Hydrogen Bonding. Part 9.† Solute Proton Donor and Proton Acceptor Scales for Use in Drug Design

Michael H. Abraham, Philip P. Duce, and David V. Prior
Department of Chemistry, University of Surrey, Guildford, Surrey, GU2 5XH
Derek G. Barratt, Jeffrey J. Morris,* and Peter J. Taylor*
ICI Pharmaceuticals, Mereside, Alderley Park, Macclesfield, Cheshire, SK10 4TG

Hydrogen bonding equilibrium constants have been measured for a large and varied selection of proton donors against a common acceptor (*N*-methylpyrrolidinone) and of proton acceptors against a common donor (4-nitrophenol). Together these have been used to create the log K_α and log K_β scales of proton donor and acceptor ability which are explicitly targeted to the needs of the medicinal chemist in the context of potential drug–receptor interactions. To this end they have been measured in 1,1,1-trichloroethane, a solvent never before used for hydrogen bonding studies but whose high dipolarity is considered a much better model for real biological membranes than the very non-polar solvents that have previously been employed. It is shown that this solvent imposes significant ranking changes on the solutes, since the charge transfer element in hydrogen bonding is reinforced at the expense of the purely electrostatic component. Nevertheless it is possible to scale previous data in such a way that over 80 functional group log K_α and log K_β values become available to the medicinal chemist (Table 4). In addition, data are given for a large number of parent heterocycles, most of which have never before been studied. We note that heterocycles are uniquely able to 'fine-tune' these scales, so providing at least one justification for their special interest to the medicinal chemist.

In addition to equilibrium constants we have measured the spectroscopic quantities $\Delta\nu_{C=O}$ (for donors) and β_{sm} (for acceptors). On various lines of evidence we suggest that these are enthalpy-related quantities and, following previous arguments, may function as alternative parameters suitable for use by the medicinal chemist under conditions of severe steric constraint.

Cross-comparisons of these data allow conclusions to be drawn which considerably illuminate the factors that influence hydrogen bond strength, and some of which have no precedent. A selection follows. Where a level comparison can be made, the donor order is OH > NH > CH and the acceptor order is N > O > S. However, within each category there are various sorts of family relationship. For example, phenols and alkanols lie on separate lines of log K_α *vs.* pK_a, and a similar separation for log K_β is shown by 5- and 6-membered ring heterocycles. By contrast, OH and NH donors show a single relation between log K_α and $\Delta\nu_{C=O}$, negative deviations from which are satisfactorily accounted for in terms of steric and stereoelectronic factors. The most important of the latter is lone-pair repulsion: 'α-effect' heterocycles are anomalously strong acceptors, whereas certain classes of donor, notably sulphonamides and carboxylic acids, are much weaker than would be expected from their pK_a values. More subtle anomalies attach, *inter alia*, to heterocycles as donors, CH donors generally, and amines and sulphonamides as acceptors; all however can be rationalised.

The extremes of both scales are charted. Alkyl thiols and amines are negligible as proton donors; correspondingly, π-donor hetero-atoms as *e.g.* in esters and amides are negligible acceptors. At the opposite extreme, heterocycles such as tetrazole and 4-quinolone figure prominently. Based on these results, some structural criteria are suggested that might lead to the synthesis of stronger proton acceptors than any so far known.

The importance of hydrogen bonding in biological systems has been recognised almost ever since this concept was first formulated nearly 70 years ago.[1] The helical structures of proteins[2] and DNA[3] are clamped together by hydrogen bonds; proton transfer *via* hydrogen bonds is as essential an element in enzyme catalysis as in its aqueous solution counterpart;[4] and hydrogen bonding is one of the major forces thought to be implicated in the recognition of agonists and antagonists by their receptors.[5,6] Since an authentic hormonal receptor has yet to be characterised in three-dimensional terms, this last proposition is not so far open to direct experimental test; nevertheless it reasonably follows from the many known cases of high enantiomeric selectivity[6] as is observed *e.g.* in the binding of noradrenaline and its antagonists at the β-adrenergic receptor but which is virtually abolished in analogues, such as dopamine, that lack the benzyl OH.[7] Even if some of this loss in specificity stems from more classical steric effects, it is difficult to avoid the conclusion that the anchoring of pharmacon to receptor may involve formation of a very specific hydrogen bond. Many more such examples are familiar to the medicinal chemist.[5–7]

† Part 8 is ref. 25 (*a*).

The term isosterism, originally defined[8] as pertaining to groups or molecules with a similar electronic surface, was widened by Friedman[9] to encompass similarity of biological effect. This concept of *bioisosterism* has become central to the endeavours of the medicinal chemist. Thornber's review[10] lists many cases of bioisosterism, actual or potential, in which it is clear that the nature and extent of hydrogen bonding is crucial to its definition. Neither in Thornber's review nor elsewhere, however, is any quantitative discussion of hydrogen bonding among bioisosteres to be found. This contrasts sharply with the quantification of lipophilic properties by Hansch[11] and of steric properties by Taft[12] and Verloop;[13] Hansch[14] indeed has coined the term *isolipophilic* to denote groups that possess similar lipoidal properties. No comparable treatment of hydrogen bonding ability is available to the medicinal chemist. Before jumping too quickly to the conclusion that one is needed, however, we must briefly consider the context in which it might be used.

Leaving aside infectious disease, and with major exceptions to do with correcting the body's own deficiencies (*e.g.* insulin, some steroids), most drugs are antagonists: that is, they moderate or abolish some kind of over-response. Familiar examples are the correction of hypertension and excessive stomach acidity. The purpose of an antagonist is to obstruct the binding of an agonist to the receptor. It must bind tightly, but it need not bind precisely as the agonist does; indeed it must not, or its properties will be identical. It is probable that an antagonist will use only part of the agonist's binding sites, so as to avoid triggering the agonist's response; also that it will find extra nearby binding sites of its own, unconnected with the function of the agonist. We may consider the actual mechanism of this process in the light of the Hansch equation (1), which for

$$\log (\mathrm{RBR}) = -a(\log P)^2 + b\log P + cE + dS + e \quad (1)$$

more than 20 years has formed the foundation of quantitative structure-activity relationships in biology (QSARs).[11] This attempts to express relative biological response (RBR) in terms of a set of physical variables representing hydrophobic, electronic, and steric factors. Here P is the partition coefficient, usually octanol–water, while E is commonly represented by Hammett's σ value and S is a steric interference term which is therefore, in general, negative. Where P is the only variable needed it may be presumed that all congeners bind the same way and that, therefore, differences in response depend solely on differences in rate of transport to the receptor site, of concentration in the receptor phase, or of non-specific binding in the environs of the receptor (which mechanism obtains is often very ambiguous). Many such simple relations are known.[15,16] Similarly, a small term in E may simply adjust for the failure of octanol precisely to model the discriminatory bias of the appropriate biological membrane. However the task of the medicinal chemist, once he has exhausted the possibilities represented by a given QSAR, is to search for step jumps: *i.e.* molecular modifications that may raise activity, relative to previous expectation, by an order of magnitude or more; and here equation (1) is of no direct help to him. While step jumps may arise from diverse causes, one obvious possibility lies in a new anchor point that entails a particularly well matched hydrogen bond. Bearing in mind the implied limitation imposed by S, that too large a substituent or one the wrong shape is likely to be deleterious, the search is on for hydrogen-bonding isosteres. We need scales that will help to tell us what such groups are.

When the need for such scales become evident to us nearly a decade ago, we presumed at first that we could simply go to the literature for the necessary information. It was and remains our greatest surprise that this was not the case. To be sure, the study of *solvent* properties has been comprehensively carried through by Kamlet, Taft, and their co-workers, whose α, β, and π^* scales of proton donor and proton-acceptor ability, and of dipolarity/polarisibility,[17] have rationalised a vast array of previously chaotic facts concerning the effect of solvent on spectroscopic properties, on chemical reaction rates and equilibria, and more recently[18,19] on a series of solubility-related phenomena. Nevertheless there are reasons, which we shall shortly delineate, why a solvent scale has to be greatly modified or even totally re-cast before it can perform as a *solute* scale with respect to any functional group the solvent may possess. These objections apply in even greater measure to other familiar scales such as Kosower's Z,[20] Reichardt's E_T,[21] the Y-scale of Grunwald and Winstein[22] and the acceptor A_N (though not donor D_N) numbers of Gutman,[23] since not only are all these solvent scales as well, but they have been shown[17] to represent linear combinations of the more fundamental quantities that Taft and Kamlet have derived. We found ourselves, therefore, starting virtually from scratch.

The solvatochromic[24] methodology of Taft and Kamlet entails measuring u.v. shifts for probe and reference molecules, relative to an inert standard solvent, in a series of neat organic liquids. This total immersion of the probe precludes any measurement of equilibrium constant, since every hydrogen bond that can form will, and has the additional result that any functional group which may be present in the liquid has that bulk liquid as its standard state. This may have relatively minor consequences for liquids that do not associate, though it is not inconceivable that even there some distortion can occur, the result of solvent dielectric, solvent structuring, or merely the varying number density of functional groups. Our suspicions in this respect have recently been confirmed by the demonstration[25a] of discrepancies between solvent and solute β-scales sufficient to make their cross-prediction hazardous. For associated liquids, however, the consequences in our context are disastrous. For reasons that Taft and Kamlet themselves point out,[24b] neither the behaviour of the OH group in neat alcohols, nor that of OH and NH_2 in carboxylic acids and primary carboxamides, provides the smallest guide as to how these functional groups are expected to behave when one-to-one contact is involved. Self association in the first class greatly enhances both donor and acceptor ability, while donor ability in the second is sharply reduced, for reasons that need not be elaborated but are readily explicable. Hence if conceived as applying to *solutes* all amphiprotic groups at the least are lost to the solvent β-scale,[24a] whereas since *all* proton donors (except halo-carbons) are amphiprotic, the α-scale[24b] becomes a total loss. While none of this has the slightest relevance to the original point and purpose of the α and β scales, it will sufficiently indicate why, for the purposes of drug design, a fresh start had to be made.

In attempting to devise a coherent methodology there were also some other problems to consider. One of these concerns the thermodynamic nature of the quantities we wished to derive. It has been strongly argued by Jencks and Page[26] that *incremental* free energies in the binding of their substrates to enzymes, *e.g.* from introducing a new substituent, are actually enthalpies in disguise: the entropy of binding remains substantially a constant. At the same time, those physical properties most familiar to the medicinal chemist, such as Hansch's π and Hammett's σ, as well as the biological activities themselves, are unequivocally related to Gibbs energy and it would be desirable to generate quantities that possess the same status and the increments for which are strictly comparable. That is not the case for β or α, which are scaled between nominal limits of zero and unity. In addition, the thermodynamic status of β is equivocal. The β-scale was obtained by blending data from a number of sources, some of them equilibrium constants, some of them spectral shifts;[24a] subsequent modifications[17] have

continued this process. Since i.r. and u.v. spectral shifts are subject to the Frank–Condon principle, *i.e.* their transitions occur in less time than it takes the solvent to reorganise, it is arguable that these are enthalpy related and that this blended β, therefore, is indefinable in thermodynamic terms. Possibly this is precisely why the β-scale has proved so valuable in matching such a very wide range of solvent-mediated phenomena. In our context a proper separation of ΔH from ΔG is essential, and of these, for reasons that will soon appear, ΔG must be considered the more fundamental quantity.

Finally there is the simple point that a solvatochromic methodology can deal only with solvents. This limitation excludes not only nearly all compounds of direct interest to the medicinal chemist, but in addition, important classes of functional group such as sulphones and most heterocycles. Hence a solution methodology becomes essential for our purpose.

If equilibrium constants are to be measured and compared then a standard proton donor for all acceptors, a standard proton acceptor for all donors, and a standard solvent for both sets, all have to be chosen with care. Virtually no common thread runs through previous work on this subject. The literature abounds with studies in which a handful of donors has been matched against a single acceptor, or *vice versa*, but neither the donor nor the acceptor probe has been in any way standardised, and neither has the solvent. It is possible that this does not matter: we are presently engaged in a statistical survey which will determine to what extent it is possible to combine all or most of these previous studies, from either series, into a single comprehensive linear free energy relationship.[25b] Even if this can be accomplished, however, the total of compounds covered would, in terms of functional groups, fall far short of the medicinal chemist's requirements. There is therefore ample justification for additional studies. The present study contains a multiplicity of compounds, and of functional groups, on which no such data have ever been obtained before. Others will follow in due course.

The widest ranging previous study—of the type we regard as relevant—led to the pK_{HB} scale of Taft and Schleyer,[27] which was eventually incorporated into the β-scale.[24a] Here *p*-fluorophenol was used as the donor probe and the solvent was tetrachloromethane; the pK_{HB} value for any acceptor is the logarithm of its equilibrium constant for hydrogen-bond formation. No amphiprotic compounds were studied. Our log K_β values have a similar meaning and our methodology is based essentially on theirs. One important difference, however, is that while they used an n.m.r. methodology, we have been careful, here and throughout, to avoid any technique that involves time-averaging. There are two reasons for this. Firstly, if we intend to study amphiprotic compounds, it is essential that we are able to observe all species, free or complexed, not only as a way of measuring equilibrium constants but so as to make sure that the solutions are sufficiently dilute that appreciable self-association does not occur. Secondly, any technique based on time-averaging has to assume that the signal due to the uncomplexed probe is unaffected by the presence of the substrate. While sound enough at high dilution, this can be a dangerous assumption in the relatively concentrated solutions sometimes needed for studying weak acceptors. In practice this problem has rarely been encountered to any serious extent, a notable exception (on the proton donor side) being chloroform, whose log K_α value could not have been obtained without taking this precaution.[28] Nevertheless the fact that we can be certain on this point gives us added confidence in the values in the Tables. Full details of our methodology appear below in the experimental section. Data analysis was initially carried out by the Scatchard technique, well established[29] as much the most reliable linear transformation of the observed titration curve. More recently we have employed a highly sophisticated non-linear statistical technique for fitting the titration curve directly,[30,31] and all values in the Tables are based on this. This technique allows simultaneous determination of K_β, the equilibrium constant for hydrogen bond formation, and β_{sm}, the scaled u.v. shift of the probe for total complexation, without having to classify any variable as 'dependent' or 'independent'. In being an extrapolation to infinite dilution β_{sm} differs not only from the Taft–Abboud–Kamlet (TAK)[17] solvent β but from their more recent monomer value β_m, which is in fact a transformed equilibrium constant.[32] We shall later discuss how these two quantities, K_β and β_{sm}, may be related. Our other difference is that we have used *p*-nitrophenol (PNP) as our standard donor probe. This brings us into line with Taft's more recent work[24a] and has enabled considerably weaker proton acceptors to be studied than might otherwise have been the case. It has recently been observed[33a] that the peaks of some solvatochromic indicators, but not PNP, suffer from problems of asymmetry, so that choice was possibly fortunate. Nevertheless, problems do arise following the recent discovery[33b,34] that OH and NH donors tend to rank proton acceptors in different ways. Which class should be considered as biochemically the more relevant is none too clear. As a practical point we know of no NH donor, even in the light of our own and other studies, that we might sensibly have used; most are simply far too weak. The implications of this OH/NH dichotomy are considered later.

In attempting to construct a monomer proton donor scale we were on our own. It can clearly be carried out by a similar technique to the above provided that a suitable probe with a suitable signal can be found, but this problem had not previously been addressed. We have overcome it in two ways. The first is by use of an i.r. technique based on *N*-methylpyrrolidinone (NMP) as standard acceptor probe. The choice of a carboxamide is logical in view of the ubiquity of amide groups in nature, and their strong acceptor properties,[17] while the particular property of NMP that dictated its choice was its strong and absolutely symmetrical i.r. carbonyl band. This feature has enabled a deconvolution technique to be employed that allows accurate assignment of peak position for both complexed and uncomplexed NMP down to a separation of about 10 cm^{-1}. Such a technique was necessary since NMP has almost always been used in excess so as to guard against any tendency to self-association in the donor. Details are given below. By measuring both free ν_{XH} for the donor (X = N or O) and free and complexed $\nu_{C=O}$ for NMP we are able simultaneously to obtain the association constant K_α and $\Delta\nu_{C=O}$, the carbonyl shift on complex formation. Again, because both carbonyl bands are measured together we do not have to assume that the uncomplexed frequency is unaffected by the presence of the donor; in this sense $\Delta\nu_{C=O}$, like β_{sm}, represents an extrapolation to infinite dilution. This is important in view of the suggestion, which we explore below, that $\Delta\nu_{C=O}$ may represent some form of energy from which translational and entropic terms have been removed. It also marks a sharp departure from the Taft–Kamlet solvent α scale[24b] in the derivation of which it was not possible to find probe acceptors and reference molecules structurally related in the way that, for example, *p*-nitrophenol and *p*-nitroanisole are related as probe and reference for the solvent β scale.[24a] This represents an hiatus on which there has been little comment. In the present case, NMP serves as reference molecule as well as probe.

Our second way of overcoming this problem was through the titrational calorimetric procedure that has been described.[30,31] This has enabled equilibrium constants for hydrogen bond formation, and their enthalpies, to be measured simultaneously by a mode of analysis which precisely parallels that for K_β and β_{sm} as described above. Most log K_α values in the Tables have in fact been obtained this way. A variant of this technique has even

Table 1. Solvent dielectric properties.[a]

	ε	(ε − 1)/(2ε + 1)
Vacuum	1	0
Cyclohexane	2.02	0.20
Tetrachloromethane	2.23	0.23
1,1,1-Trichloroethane	7.53	0.41
Dichloromethane	8.93	0.42
Octan-1-ol	10.3	0.43
Water	78	0.49

[a] Data on dielectric constant ε from ref. 37.

enabled K_α and ΔH_f to be obtained for donors whose propensity for dimer formation is notorious, thus enabling, for example, the proton donor ability of carboxylic acids towards another acceptor to be measured for the first time.[30] One major result of this study has been to demonstrate that, in our solvent, there is *no* unique relationship between free energies and enthalpies for hydrogen bond formation; indeed there is quite as much tendency for ΔH_f and ΔG_f to run contrary as to run together.[31] In view of the comments of Page and Jencks[26] this result is of the greatest significance, and must in our context throw the emphasis on to free energy even more decisively than before. We explore some possible consequences below.

Our third major problem lay in the choice of solvent. While clearly a standard solvent must itself be devoid of proton donor or acceptor ability, especially if weak donor or acceptor solutes are of interest, the extremely non-polar nature of most solvents previously employed for hydrogen bonding studies leaves much to be desired in terms of biological realism. Without knowing how far this mattered, it seemed sensible at the outset of this work to fix on a solvent with the sort of polarity likely to obtain at the kind of receptor site that interests us. Some guidance can be obtained from recent work on the dielectric properties of biological membranes. Mehler and Eichele[35] estimate effective dielectric constants ε of 10–40 for the water-accessible surface regions of some proteins; this contrasts with ε 2–5 in the very deep hydrocarbon bilayers of phospholipids. Probably we are concerned with neither extreme. The term 'hydrophobic' as applied to receptor sites has to be interpreted with caution; it certainly does not mean 'devoid of polar groups,' or the interactions we are concerned with could not take place. Most likely some intermediate state corresponding to this intermediate range in ε, 5–10, will lie nearer to the truth. Warshel and Russell[36] define a non-polar medium as possessing $\varepsilon \ll 6$; of solvents commonly used in hydrogen bonding studies[25] this includes tetrachloromethane, tetrachloroethane, and all hydrocarbons, and while dichloromethane would classify as polar, its weak proton-donor properties[17] exclude it from consideration. We have in fact standardised on 1,1,1-trichloroethane (TCE), ε 7.53.[37] So far as we can determine this solvent has never been employed for hydrogen-bonding studies before,[25c] but it combines in unrivalled manner high polarity and chemical inertness with lack of toxicity and a total absence of hydrogen bonding properties.[17] We may put its dielectric properties in perspective by re-expressing them in terms of the Kirkwood relation[38] (ε − 1)/(2ε + 1), which for dipole-dipole interactions is much more relevant than ε itself. Table 1 lists both quantities for some relevant reaction media that range from water to a total vacuum. Like octanol, the standard solvent for partitioning studies of biological relevance,[11] TCE is clearly to be regarded as a polar medium, much closer in this respect to water than to previous standards. In addition, as a commercial dry cleaning fluid it is inexpensive and readily obtained in a high state of purity. Its excellent solvent properties have proved an invaluable bonus; many of the compounds listed in the Tables could never have been investigated in any solvent previously employed.

The purpose of this paper is therefore to present to the medicinal chemist the log K_α scale of proton donor and the log K_β scale of proton acceptor ability: the first with NMP as standard acceptor, the second with PNP as standard donor, and both in TCE as standard solvent. We lay no claim to universality. Taft and Kamlet[17,39] have many times drawn attention to the existence of family dependent (FD) properties, *i.e.*, properties that are readily correlated *via* solvatochromic parameters within compound sets of similar type but require the addition of a new, usually constant, term if dissimilar sets are to be incorporated. Laurence and his co-workers[33] go further and claim that family independent (FI) correlations scarcely exist. Recently, a quantity θ has been derived by Maria, Gal, and their co-workers[34] which can be used to classify measures of electron transfer according to the electrostatic–covalent blend that each one shows; from this it is clear that different hydrogen-bonding scales give different information. We wish to draw particular attention to the observation, by Taft and his co-workers,[40] that nitrogen acceptors gain strength relative to oxygen acceptors as solvent polarity rises. This alone would remove any pretension to universality for log K_β. We shall consider these and other complications in due course. Nevertheless K_α and K_β may still be tuned to our requirements and unexpectedly, in one respect, may perhaps have proved serendipitous. Since necessarily based on different probes, it is obvious that these scales can only fortuitously carry the same numerical weight. However, it is possible that very roughly they do. There is a common prejudice that alkanols, while amphiprotic like water, are slightly more acceptor than donor in character, at least in relative terms; this is certainly true against water in bulk,[41] is reflected by the α and β scales,[17] and would be expected for the monomer on simple inductive grounds.[42] We find log K_α *ca.* 1.2 and log K_β *ca.* 1.4 for a typical primary alkanol. These relative values are in line with the above expectation and may indicate that, to take one simple cross-comparison, phenol and thiazole as donor and acceptor, respectively, are roughly in balance. It is interesting to consider the cross-linking of peptide chains in the light of this idea. Obviously, there is no way of determining whether the main driving force is NH 'push' or carbonyl 'pull', but values of log K_α *ca.* 0.6 and log K_β *ca.* 3.1, respectively, very much suggest the latter. Wolfenden[43] has recently come to similar conclusions on entirely different grounds. Other cross-comparisons may prove equally instructive.

The most serious gap in our work to date is the whole class of charged substituents. While there is ample evidence that cations and anions form the most powerful category of proton donor and acceptor respectively,[44] these are more difficult to investigate and, as pole–dipole not dipole–dipole interactions, potentially introduce 'apple and pear' type problems. It is also arguable that, to the medicinal chemist, they are less important. Although charged groups are frequently required for agonist or antagonist activity,[6] most such interactions involve recognition sites, as in the β-adrenergic blocking agents[7] noted in passing above. As previously explained, in designing antagonists one hopes chiefly to interact with secondary binding sites. While cations and anions may certainly interact, it is also well established that charged groups greatly impede the passage of drugs across membranes.[45] Hence incorporation of extra charge into the antagonist might well be counter-productive. At present, therefore, we do not regard this omission as an urgent one to rectify.

Experimental

Materials.—The solvent 1,1,1-trichloroethane was flash distilled over fresh phosphorus pentoxide and used immediately.

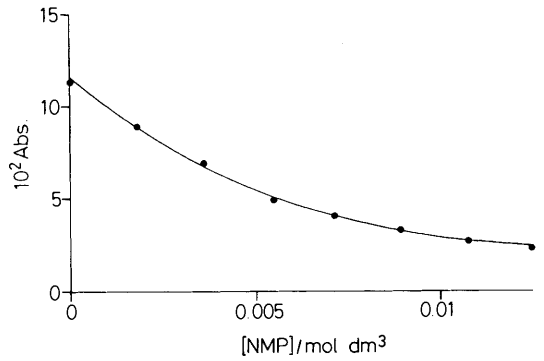


Figure 1. Representative titration curve for the donor 4-trifluoromethylphenol *vs.* NMP as acceptor. The fitted line corresponds to K = 630 dm^3 mol^{-1}.

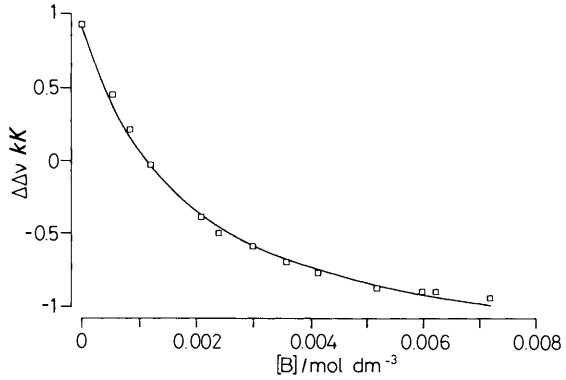


Figure 2. Representative titration curve for the acceptor 4-methylpyridine *vs.* PNP as donor. The fitted line corresponds to K = 563 dm^3 mol^{-1} and $\Delta\nu$ = 2.41 kK.

Typically, water content, measured by Karl–Fischer titration (photovolt aquatest II) was usually in the range 1–5 × 10^{-3} mol dm^{-3}. 4-Nitrophenol was recrystallised twice from water (m.p. 114–116 °C) and 4-nitroanisole from ethanol–water (m.p. 52–53 °C). NMP (Aldrich Gold Label) was used without further purification. All other compounds were either the highest grade commercially available, or were from the ICI compound collection. U.v. spectra were run on a Beckman DU8-B spectrophotometer using matched 1 cm or 0.1 cm path length cuvettes. Temperature was maintained at 25 ± 0.1 °C using a Peltier system. I.r. spectra were obtained as previously described.[31]

Measurement of K_β *and* β_{sm} *by u.v. Spectroscopy.*—The method relies on the measurement of the u.v. shift of 4-nitrophenol as a function of base concentration; β_{sm} is based on an extrapolated peak maximum for 100% complexation, not always attained. For each hydrogen-bond acceptor, some 8–12 solutions of the base were prepared in TCE, and the concentrations were adjusted to ensure that the full range (20–80%) of complexation was covered. For most solutes this is typically between 5 × 10^{-4} and 1 × 10^{-2} mol dm^{-3}. 25 μ dm^{-3} of 1 × 10^{-2} mol dm^{-3} 4-nitrophenol in TCE was then added to 2.5 cm^3 aliquots of each solution, and after allowing 15–20 min for temperature equilibration, the u.v. spectrum was recorded. The procedure was repeated for 4-nitroanisole. Peak maxima were located both by direct measurement and by first-derivative spectroscopy. Typically, the peak maximum of 4-nitroanisole remained constant up to a concentration approached only by a very few weak examples. For amphiprotic solutes such as alcohols, the solutions were checked by i.r. spectroscopy to ensure that they were sufficiently dilute to prevent significant self-association. In some cases equilibrium constants were also measured by i.r. spectroscopy, as previously described,[31] by following the decrease in OH absorbance of 4-nitrophenol with increasing base concentration.

Measurement of K_α *and* $\Delta\nu_{C=O}$.—Measurement of the equilibrium constant K_α was carried out either by titrational calorimetry or i.r. spectroscopy as previously described.[31]

Provided that the carbonyl bands were sufficiently well resolved, carbonyl shifts ($\Delta\nu_{C=O}$) could be measured directly. However, in most cases $\Delta\nu_{C=O}$ was obtained by curve fitting procedures. Two distinct approaches were used: (*a*) Fourier self-deconvolution[46] and (*b*) band synthesis using BANDFIT.[47] There are a number of well recognised problems with any form of curve fitting[48] and close attention was paid to the criteria for reliability. However, very high S/N ratios were obtained (400 scans at a nominal resolution of 2 cm^{-1}) in all spectra prior to fitting. More importantly, the fitting is carried out on a simple well-defined chemical system where there are usually only two bands to resolve (free and complexed carbonyl) and since parameters for the free band can be obtained in the absence of a hydrogen bond donor, there is a good check on the validity of any particular fit.

In general, the agreement between the two methods for estimation of $\Delta\nu_{C=O}$ was excellent (typically ±0.2 cm^{-1}). In addition, $\Delta\nu_{C=O}$ was unchanged over the full range of complexation and, typically, checks were made at about 30 and 70% complexation. The position of the free carbonyl band was found to be independent of both the nature and concentration of the hydrogen bond donor and, under the conditions used, $\nu_{C=O}$ (free) was 1 696.2 (±0.3) cm^{-1}.

Data Analysis.—The titration curves resulting either from plots of absorbance of the XH band *vs.* concentration of *N*-methylpyrrolidinone (Figure 1) for hydrogen bond donors or $\Delta\nu$ *vs.* concentration of hydrogen bond acceptor (Figure 2) were analysed by non-linear fitting techniques as described previously[31] using either SAS® or STATGRAPHICS® to yield both quantities simultaneously. The procedure for K_α and $\Delta\nu_{C=O}$ has been described.[31]

For hydrogen bond acceptors, the fitting procedure gives K_β and the $\Delta\Delta\nu$ value corresponding to full complexation, as in the procedure of Kamlet and Taft.[24a] One point of difference is that the β_{sm} scale is anchored to β_{sm} = 1.00 for the relatively non-toxic triphenylphosphine oxide (TPPO), not to hexamethylphosphoramide (HMPA) which is now known to be highly toxic. We believe that no systematic problems in cross-comparison result.

Results

Data for proton donors are assembled in Table 2 and for proton acceptors in Table 3, numbered in continuous sequence for ease of reference. Many of the former have been listed previously[31] but we have taken this opportunity to make additions and emendations while $\Delta\nu_{C=O}$ is now included; the only entries which do not contain that quantity are for compounds of particular interest that will figure in the Discussion. Standard errors for K_α have been reported;[31] for $\Delta\nu_{C=O}$ they are around ±0.2 cm^{-1} with little variation. Standard errors for log K_β appear in Table 3, calculated according to equation (2). For

$$\log(\partial x) = \partial x/x \log_e 10 \qquad (2)$$

(**101**), β_{sm} possesses the high s.e. of 0.11, which on i.r. evidence (two complexed ν_{OH}) derives from sulphonyl and carbonyl both acting as acceptors. We do not list s.e.s for β_{sm} since elsewhere

Table 2. Equilibrium and Spectroscopic Data for Proton Donors.[a]

		log K_α	pK_a[b]	$\Delta\nu_{C=O}$/cm^{-1}
(1)	Methanol	1.48	15.09	
(2)	Ethanol	1.21	15.93	15.0
(3)	Propan-1-ol	1.11	16.1	15.2
(4)	Hexan-1-ol	1.20	16.1	
(5)	Propan-2-ol	0.91	17.1	14.1
(6)	t-Butyl alcohol	0.78	19.0	13.7
(7)	$PhCH_2OH$	0.90	15.4	
(8)	$ClCH_2CH_2OH$	1.08	14.31	16.3
				21.8
(9)	$MeSCH_2CH_2OH$			16.7
				18.5
(10)	CF_3CH_2OH	2.00	12.39	21.3
(11)	$(CF_3)_2CHOH$	2.83	9.3	26.6
(12)	Phenol	2.14	10.00	22.3
(13)	2-Methylphenol	1.75		22.7
(14)	2,6-Dimethylphenol	1.08		20.1
(15)	2-Isopropylphenol	1.95		
(16)	2,6-Di-isopropylphenol	*ca.* 0		20.1
(17)	2-t-Butylphenol	1.85		23.1
(18)	2,6-Di-t-butylphenol	*ca.* 0		20.0
(19)	2-Chlorophenol	2.33		25.0
(20)	2,6-Dichlorophenol	0.98		20.0
(21)	2-Cyanophenol	2.69		28.1
(22)	3-*N*,*N*-Dimethylaminophenol	1.79	10.22	21.3
(23)	3-Methylphenol	1.89	10.09	22.3
(24)	3-Isopropylphenol	1.89	10.16	21.5
(25)	3-Chlorophenol	2.50	9.13	25.7
(26)	4-Methoxyphenol	2.18	10.21	22.7
(27)	4-Trifluoromethylphenol	2.80	8.68	25.4
(28)	4-Nitrophenol	3.12	7.15	29.0
(29)[c]		0.98	12.6[d]	18.4
(30)[c]		1.11	11.48[c]	19.1
(31)	Acetic acid	2.04	4.76	
(32)	Pivalic acid	1.77	5.04	37
(33)	Benzoic acid	2.07	4.21	39
(34)	Trifluoroacetic acid	*ca.* 3.55	0.52	
(35)[c]		0.60	(27)	14.1
(36)[c]		0.60	(17.7)	15.8
(37)	4-Nitro-*N*-methylaniline	0.73	18.37[f]	16.5
(38)[c]		1.00		17.7
(39)	2-Aminobenzothiazole	*ca.* 1.1		17.2
(40)[c]				15.4
(41)	CF_3CONH_2	1.52		18.1
(42)	$C_6H_{13}NHCOC_6H_{13}$	0.64		
(43)	$MeNHCOBu^t$	0.70		
(44)	$C_6H_{13}NHCO_2Me$			14.7
(45)	$PhNHCO_2Me$			15.8
(46)	Acetanilide	1.34		17.6
(47)	4′-*N*,*N*-Diethylacetanilide	0.48		16.5
(48)	3′-Chloro-4′-nitroacetanilide	2.48		22.9
(49)	3′-Trifluoromethyl-4′-nitroacetanilide	2.47		23.8
(50)	Thioacetanilide	1.52		
(51)	*N*,*N*′-Dicyclohexylthiourea			22.4
(52)[c]		*ca.* 1.1		19.8
(53)[c]			11.42[g]	20.5
(54)	$(CF_3CO)_2NH$	2.63		
(55)	Toluene-*p*-sulphonamide	1.15	10.17	19.3
(56)	*N*-Benzyltoluene-*p*-sulphonamide	0.90		*ca.* 22
(57)	*N*-(2-Naphthyl)toluene-*p*-sulphonamide	1.18	8.46	
(58)	$C_7H_{15}CONHSO_2Me$	*ca.* 1.0	5.39	26.0
(59)	(1-Naphthyl)$CONHSO_2Me$			30.6
(60)	Pyrrole	0.95	17.51	14.2
(61)	Indole	1.15	16.97	15.1
(62)[c]		1.20	14.5[h]	17.5
(63)[c]		1.99	10.69[h]	21.3
(64)[c]		2.18	9.51[h]	27.0
(65)[c]		2.71		34.2
(66)[c]		3.55	4.32	39.4
(67)	Chloroform	*ca.* 0.4	*ca.* 24	*ca.* 11

[a] For standard errors in log K_α see ref. 31; s.e. for $\Delta\nu_{C=O}$ is <0.2. [b] For source of data see ref. 57. [c] For structure see Scheme 1. [d] Of diethyl ketoxime. [e] Of acetophenone oxime. [f] Of 4-nitroaniline: R. Stewart and J. P. O'Donnell, *Can. J. Chem.*, 1964, **42**, 1681. [g] Of glutarimide. [h] Estimate; see ref. 57.

Table 3. Equilibrium and spectroscopic data for proton acceptors.

		log K_β	±s.e.	pK_a[b]	β_{sm}
(2)	Ethanol	1.41	0.10	−1.94	0.62
(5)	Propan-2-ol	1.36	0.03		0.66
(6)	t-Butyl alcohol	1.45	0.04		0.66
(68)	Dibutyl ether	1.28	0.02	−2.39[c]	0.50
(69)	t-Butyl methyl ether	1.46	0.03		0.55
(70)	Tetrahydrofuran	1.69	0.03	−2.32	0.54
(71)	Anisole	0.30	0.04	−5.4	0.23
(72)	$MeO(CH_2)_2OMe$	1.69	0.02		0.53
(73)	1,4-Dioxane	1.28	0.02	−3.42	0.40
(74)	1.4-Thioxane	1.06	0.04		0.38
(75)	1,3-Dioxolane	0.70	0.13		0.23
(76)	Acetone	1.61	0.01	−2.85	0.45
(77)	Pentan-3-one	1.50	0.01		0.48
(78)	$MeCOPr^i$	1.52	0.05	−3.57	0.48
(79)	$MeCOBu^t$	1.44	0.02	−3.50	0.48
(80)	Pr^iCOPr^i	1.39	0.03		0.51
(81)	Cyclohexanone	1.70	0.03		0.49
(82)	Acetophenone	1.46	0.05	−4.36	0.38
(83)	Ethyl acetate	1.43	0.01	−4.61	0.44
(84)	γ-Butyrolacetone	1.67	0.03		0.48
(85)	Dihydro-2(3*H*)-thiophenone	1.32	0.04		0.40
(86)	Dimethylformamide	2.81	0.03	−1.60	0.68
(87)	Diethylformamide	2.73	0.04		0.77
(88)	$Bu^tCON(Me)Bu^t$	2.53	0.02		0.66
(89)	Dimethylthioacetamide	1.76	0.08		
(90)	*N*-Methylpyrrolidinone	3.12	0.02	−0.36[d]	0.79
(91)	*N*-Dimethylbenzamide	2.82	0.03	−1.20	0.76
(92)	Tetramethylurea	3.19	0.02	0.10	0.74
(93)	Tetramethylthiourea	1.96	0.14		
(94)[a]		2.38	0.03		0.70
(95)	$PhOCONMe_2$	2.09	0.02		0.64
(96)	*N*-Methylmaleimide	1.67	0.05		0.42
(97)	*N*-Methylquinol-4-one	>4			
(98)	Dimethyl sulphoxide	3.06	0.02	−1.54	0.77
(99)	Tetramethylenesulphone	1.61	0.02		0.42
(100)	$PhSO_2N(Me)CH_2Ph$	1.36	0.13		0.38
(101)[a]		0.99	0.20		0.46
(102)	Triphenylphosphine oxide	3.85	0.01		1.00
(103)	Triethyl phosphate	3.17	0.03		0.76
(104)	Isopropylamine	2.84	0.06	10.63	*ca.* 0.90
(105)	Benzylamine	2.36	0.14	9.62	0.75
(106)	Allylamine	2.63	0.04	9.49	0.79
(107)	$CN(CH_2)_2NH_2$	1.74	0.05	7.80	0.57
(108)	$CF_3CH_2NH_2$	1.01	0.01	5.59	0.55
(109)	Pyridine	2.52	0.01	5.22	0.77
(110)	2-Methoxypyridine	1.28	0.04	3.06	0.64
(111)	2-Fluoropyridine	1.41	0.05	−0.44	0.49
(112)	2-Chloropyridine	1.48	0.02	0.72	0.59
(113)	2-Cyanopyridine	1.00	0.04	−0.26	0.55
(114)	3-Methylpyridine	2.65	0.03	5.52	0.82
(115)	3-Fluoropyridine	1.82	0.04	2.97	0.66
(116)	3-Chloropyridine	1.77	0.06	2.84	0.67
(117)	3-Bromopyridine	1.76	0.03	2.84	0.65
(118)	3-Cyanopyridine	1.41	0.02	1.39	0.52
(119)	3-*N*,*N*-Diethylcarbamoylpyridine	2.76	0.03	3.35[e]	0.71
(120)	4-Methylpyridine	2.78	0.03	6.03	0.85
(121)	3,4-Dimethylpyridine	3.06	0.03	6.46	0.84
(122)	4-Methoxypyridine	2.87	0.06	6.47	0.92
(123)	4-*N*,*N*-Dimethylaminopyridine	3.54	0.05	9.70	1.25
(124)	4-Acetylpyridine	2.20	0.09	3.51	0.59
(125)	Pyrazine	1.46	0.05	0.65	0.55
(126)	Pyrimidine	1.67	0.01	1.23	0.77
(127)	Pyridazine	2.53	0.02	2.24	0.74
(128)	Isoxazole	1.06	0.05	−2.03	0.51
(129)	Oxazole	1.67	0.03	0.80	0.58
(130)	2,4,5-Trimethyloxazole	2.65	0.03	3.51	0.75
(131)	Thiazole	1.90	0.08	2.52	0.72
(132)	Benzothiazole	1.76	0.05	1.2	0.62
(133)	1-Methylpyrazole	2.22	0.08	2.09	0.70
(134)	1-Methylimidazole	3.68	0.04	7.25	0.92
(135)	1-Benzyl-1,2,4-triazole	2.38	0.03		0.70

Table 3 (*continued*)

		log K_β	±s.e.	pK_a[b]	β_{sm}
(136)	1-Phenethyl-1,2,3-triazole	2.56	0.03	1.25[f]	0.73
(137)	1-Methylbenzotriazole	2.17	0.06	*ca.* 1.6[g]	0.70
(138)[a]				2.05[h]	
(139)[a]		3.37		2.64[i]	0.85
(140)[a]		0.57			
(141)[a]		2.51	0.04		0.70
(142)[a]		1.98	0.04		0.77
(143)[a]		0.79			
(144)[a]		1.99	0.05	*ca.* −1.9[j]	0.60
(145)[a]		1.51	0.02	0.99[k]	0.59
(146)	Me_2C=NOPh	1.10	0.04		0.60
(147)[a]		2.90		−1.05[l]	0.75
(148)	Me_2NCN	2.00	0.03	1.2	0.59
(149)	Acetonitrile	1.23	0.03	−10.1[m]	0.41
(150)	$MeOCH_2CN$	1.04	0.02		0.44
(151)	$MeO(CH_2)_2CN$	1.28	0.02		0.42
(152)	$ClCH_2CN$	0.61	0.03	−12.8[m]	0.42
(153)	PhCN	1.06	0.02	−10.4[m]	0.40
(154)	4-Methoxybenzonitrile	1.32	0.06		0.42
(155)	4-Chlorobenzonitrile	0.92	0.03		0.40

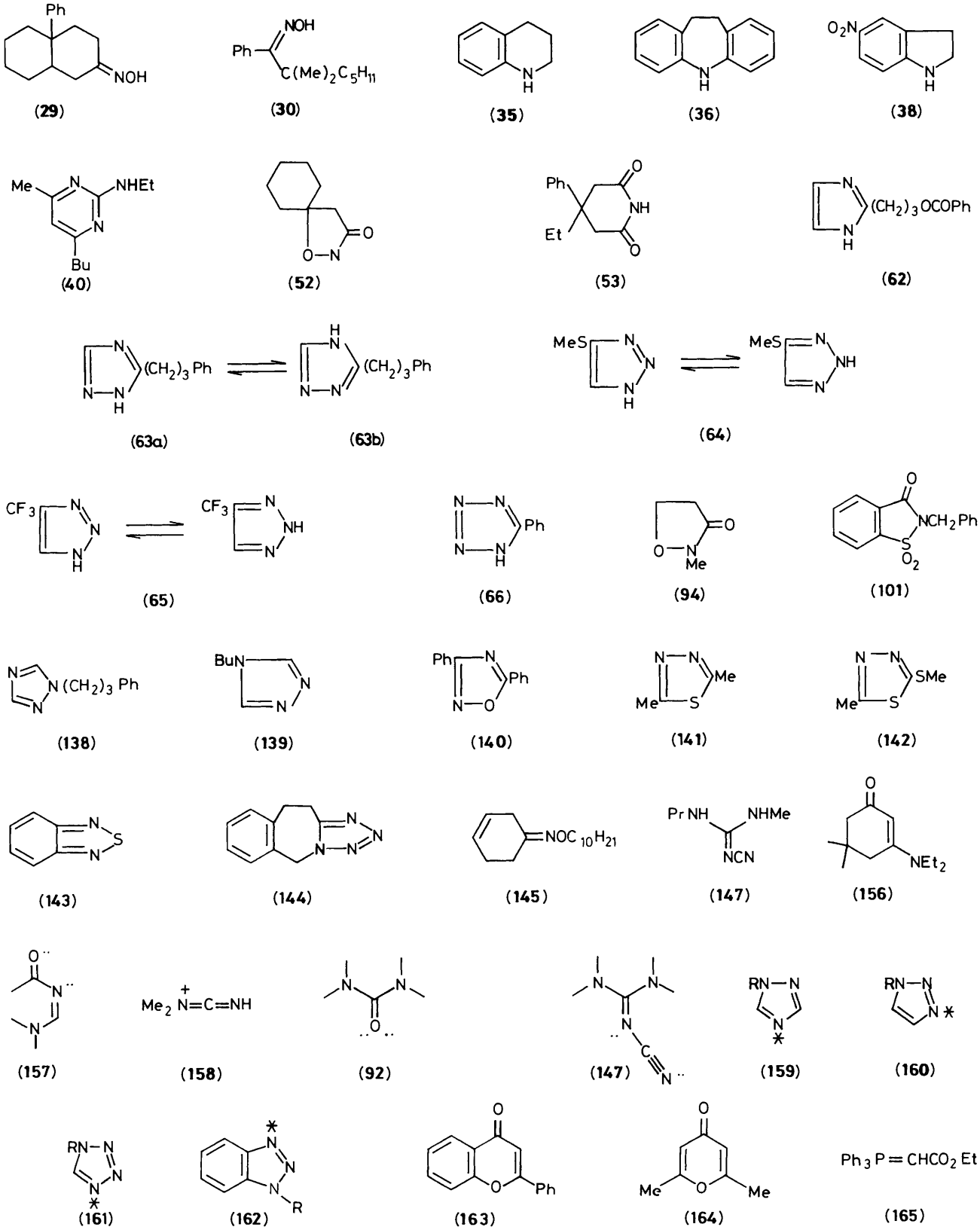
[a] For structure see Scheme 1. [b] Main sources: D. D. Perrin, 'Dissociation Constants of Bases in Aqueous Solution,' Butterworths, London, 1970; A. R. Katritzky and C. W. Rees, eds., 'Comprehensive Heterocyclic Chemistry,' Pergamon, Oxford, 1984; J. Elguero, C. Marzin, A. R. Katritzky, and P. Linda, 'The Tautomerism of Heterocycles,' Academic Press, New York, 1976; E. M. Arnett and G. Scorrano, *Adv. Phys. Org. Chem.*, 1976, **13**, 83; ref. 39. [c] For Et_2O. [d] For *N*,*N*-dimethylacetamide. [e] For pyridine-3-carboxamide. [f] For 1-methyl-1,2,3-triazole. [g] For benzotriazole. [h] For 1-methyl-1,2,4-triazole. [i] For 4-methyl-1,2,4-triazole. [j] Estimate for tetrazole. [k] For acetoxime: J. W. Smith, in 'The Chemistry of the Carbon–Nitrogen Double Bond,' ed. S. Patai, Interscience, New York, 1970, p. 235. [l] G. A. Cockayne, unpublished data for 1,2-dimethyl-3-cyanoguanidine. [m] J. Grundnes and P. Klaboe, in 'The Chemistry of the Cyano Group,' ed. Z. Rappoport, Interscience, New York, 1970, p. 123.

these are remarkably small: 0.016 ± 0.011 for 73 compounds of which only eight exceed 0.03. The high precision obtained for all classes of data allows us to place considerable confidence in the results of the cross-comparisons to be described later.

Some individual results require comment. Simple aliphatic amines have proved much the most difficult general class for which to obtain consistent results. Values of K_β have covered the range 400–900 not only for all three classes of amine but for each individual amine examined. Isopropylamine (**104**) was for some reason the least 'ill-behaved' and its log K_β value is quoted as about average for the set and a convenient mark for the remainder. The cause of this extraordinary variability is unknown; it is not due to full proton transfer since the characteristic u.v. peak of 4-nitrophenoxide anion was never seen. Nevertheless in this quite polar solvent any hydrogen bond between a strong base and a strong acid must lie quite close to the border with salt formation and that factor may contribute. It almost certainly contributes to the anomalous β_{sm} of 1.25 for 4-dimethylaminopyridine (**123**) (see Figure 7). Kamlet *et al.*[39] have pointed out that anomalies are expected for compounds, specifically including this one, in which strong conjugative π-electron donation is present; here the effect is still further enhanced by solvent polarity. Once again, u.v. evidence rules out full proton-transfer. Note however that K_β is not anomalous, *i.e.* distortions of this type do not contribute to Gibbs energy. Since NMP is only a reluctant proton transfer base, similar anomalies are not to be found among proton donors.

Conformational or similar ambiguities attach to some compounds. The 2-substituted ethanols (**8**) and (**9**) each show, as donors, two $\Delta\nu_{C=O}$ bands which are presumably due to complexation by *trans* and *gauche* conformers. For (**8**) the *gauche* form is known to be favoured by about a factor of two in tetrachloromethane[49] and since the smaller $\Delta\nu_{C=O}$ is similar to (**8**) and (**9**) while the larger one is very different, the *gauche* conformer is most likely responsible for the latter band. The primary amino group of (**39**) and (**41**) could bond to acceptors *via* either NH; considerations both of molecular dipole and of stereoelectronic repulsion (see later) suggest the *E*-NH as in secondary amides to be the dominant contributor, both here and generally. Potential tautomeric problems exist for the azoles (**63**)–(**66**). While for tetrazole and 1,2,4-triazole the dominant form is well established[50] as (**66**) and (**63a**), there remain ambiguities for 1,2,3-triazole[50] so both forms are displayed for (**64**) and (**65**). It has been seen above that the acylsulphonamide (**101**) probably forms hydrogen bonds to both acceptor moieties and potential ambiguities of a similar sort attach to (**74**), (**107**), (**119**), (**124**), (**147**), (**148**), (**150**), (**151**), and especially the heterocycles (**135**)–(**138**), (**140**), and (**144**), all of which will be discussed below.

Some of these last examples raise in acute form the question of statistical correction. It might seem obvious that K_β for pyrazine (**125**) with equivalent nitrogens ought to be halved, but nevertheless there are complications even in such apparently straightforward cases. For example, if log K_β is corrected for pyrazine the same should be done for its pK_a, though this correction is not normal. In fact we correct both values on Figure 8. Near-equivalence poses further problems: while the contiguous nitrogens of pyridazine (**127**) are clearly equivalent those of (**135**)–(**137**) are probably not, whereas for the tetrazole (**144**) we do not know whether its acceptor properties are confined to one nitrogen or a function of all three. Another dimension is added to the problem by the varying tally of lone pairs. Amines possess one lone pair but ketones two: should K_β for the latter be halved? If so, what does one then do about esters and amides, whose lone pairs are certainly not equivalent? This problem is not trivial: there is evidence for esters that bonding to the *Z* lone pair is preferred despite apparent steric hindrance,[51] as indeed expected on σ-resonance grounds,[52] and the same is likely for amides. More surprisingly, there is recent evidence[53] that ethers



Scheme 1.

* Preferred site of interaction (see the text).

behave like nitrogen bases and quite unlike carbonyl groups in their inability to form hydrogen bonds simultaneously with two proton donors. It was already known from crystal structure studies that ether lone-pairs—unlike those of carbonyl—are remarkably non-directional [54] (an observation partially contradicted by studies in the gas phase [55]), but this new result has

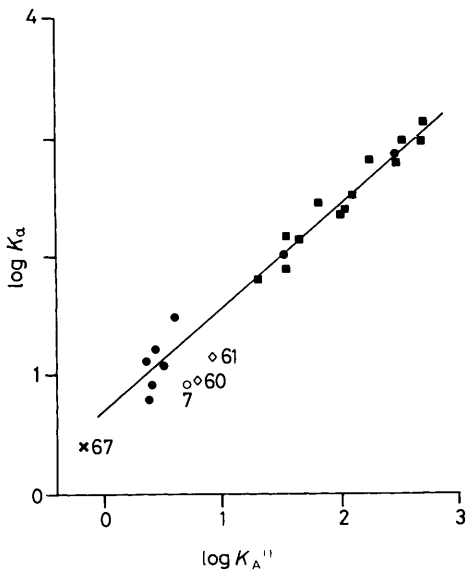


Figure 3. Relation between log K_α and log K_A^H. The correlation line of equation (3) is for the filled points. Key: circles, alkanols; squares, phenols; diamonds, NH donors; cross, CH donor.

profound implications some of which are explored below. Even the apparently fail-safe position, that statistical correction might be reasonable for the two oxygen atoms *e.g.* of SO_2 and NO_2, is thrown in jeopardy by recent crystal structure evidence[56] that hydrogen bonding to the latter prefers to lie along its axis, *i.e.* the bond to XH is bifurcated. Sulphones apparently behave in a more orthodox manner.[55] With so many potential traps for the unwary we have preferred to incorporate no statistical corrections into Tables 2 and 3; these will be performed in an *ad hoc* manner whenever appropriate and signalled clearly in the accompanying discussion. Since $\Delta\nu_{C=O}$ and β_{sm} are not equilibria, statistical correction is of course inapplicable.

Discussion

Proton Donors.—These cover three decades in K_α with NH donors generally weaker than OH but spanning a greater range and responsible, indeed, for the highest as well as almost the lowest entries. Our single CH entry, for chloroform (**67**), represents the weakest proton donor in this set and this value for K_α, with the associated $\Delta\nu_{C=O}$, had to be obtained by special techniques.[28] We have no results for aliphatic SH or amine NH donors but the evidence from other sources[57] is that these are almost immeasurably weak, their complexes even with strong proton acceptors lying at the limit where hydrogen bonding *per se* fades into a generalised weakly dipolar interaction. Despite striking recent evidence for this in the case of ammonia[58] it seems so little generally appreciated as to be worth re-emphasis here. This conclusion may need some modification where aliphatic amines are able to take part in co-operative bonding, which for alkanols as solutes is known to occur,[59] and indeed we have evidence for this from partitioning studies,[60] but as unactivated proton donors *e.g.* at the biological receptor, simple alkylamines have little to commend them. Also, while such activation might be provided by water, a more likely consequence would be full proton transfer to yield the cation (amine pK_a values are remarkably insensitive to their environment[61]). Cation formation is beyond the scope of this paper, as previously explained; in addition, while a large favourable enthalpy of formation is expected,[62] the corresponding Gibbs energy even for cation–anion association rapidly reduces as water is added until, in fully aqueous solution, it becomes quite small.[63] Hence water, while stabilising the cation, will also tend to reduce its binding energy.

In the course of the present work we have investigated the possibility that, for a given solvent, all scales of proton donor ability *vs.* a given acceptor (or *vice versa*) may be interrelated. For proton donors in solvent tetrachloromethane we have derived equation (3), in which K^i is a measured equilibrium

$$\log K^i = L_B \log K_A^{H_i} + D_B \qquad (3)$$

constant and where the slope L_B and intercept D_B terms characterise a given acceptor while log K_A^H is the proton donor term.[57] This equation is 'reasonably general'[57] in that log K^i values for most donor–acceptor combinations are predictable with a mean s.e. of 0.093. Hence log K_A^H is a 'reasonably general' scale of solute proton-donor ability for solvent tetrachloromethane (it is not a scale for TCE). It is only this, and not universal, since certain specific combinations of donors and acceptors have to be excluded. Most such exclusions concern weak donors and acceptors although, as we shall see, the pattern is not so simple. The basic reason for all these exclusions goes back to the varying blend of electrostatic and charge transfer forces that is involved in any donor–acceptor combination. This blend can now be characterised by the Maria–Gal[34] angle θ but was long foreshadowed by Drago's E-C treatment (for ΔH_f),[64] has analogies with the 'co-ordinate covalency' approach of Kamlet and Taft,[39] and was already implicit in the demonstration[65] that hydrogen bonding of acceptors with some weak proton donors is adequately correlated only when π* is added to β. It is only because most proton donor–acceptor combinations possess rather similar θ values that a 'reasonably general' scale can be constructed. On Figure 3 we plot all available data for log K_α *vs.* log K_A^H. With the exception of benzyl alcohol (**7**) where the data may be suspect, all phenols and alkanols fit the relation of equation (4). The alkanols show

$$\log K_\alpha = 0.870(0.038)\log K_A^H + 0.70(0.06) \qquad (4)$$
$$(n = 21,\ r^2 = 0.972,\ s = 0.13,\ F = 667)$$

some scatter, perhaps for steric reasons; if methanol (K_α too high) and t-butyl alcohol (K_α too low) are omitted, the statistics are improved but the equation itself is not appreciably affected. The NH donors pyrrole (**60**) and indole (**61**) lie Δlog K_α *ca.* 0.4 below this line: while these are known 'unreliables' in terms of equation (3), their combination with NMP is 'allowable' and the probable explanation is a solvent effect. Pyrrole and indole typify the class of weak donor whose specific hydrogen-bonding ability can be enhanced by a superimposed non-specific dipolar interaction.[65] Increasing solvent polarity will devalue this factor relative to hydrogen bonding *per se*, so that donor strength apparently falls. Hence the ability of pyrrole-like NH (*e.g.* in tryptophan) to form hydrogen bonds in a biological context, relative to phenolic OH (*e.g.* in tyrosine) is likely to vary quite sharply with the polarity of its surroundings. Since CH donors in general show this effect even more,[65] it is rather surprising that the point for chloroform (**67**) should lie so close to the regression line. It must be remembered however that these very low K_α and K_A^H values are hard to determine and more subject to error than most.

We turn to the possible significance of our second experimental quantity, $\Delta\nu_{C=O}$. The Badger–Bauer relation,[66] that $\Delta\nu_{OH}$ for a proton donor opposite a series of acceptors should be linearly related to $-\Delta H_f$ for hydrogen bond formation, was later discounted by Rao *et al.*[67] who, on detailed examination of the evidence, preferred a parabolic relationship (if any). Nevertheless Drago and co-workers have demonstrated a number of

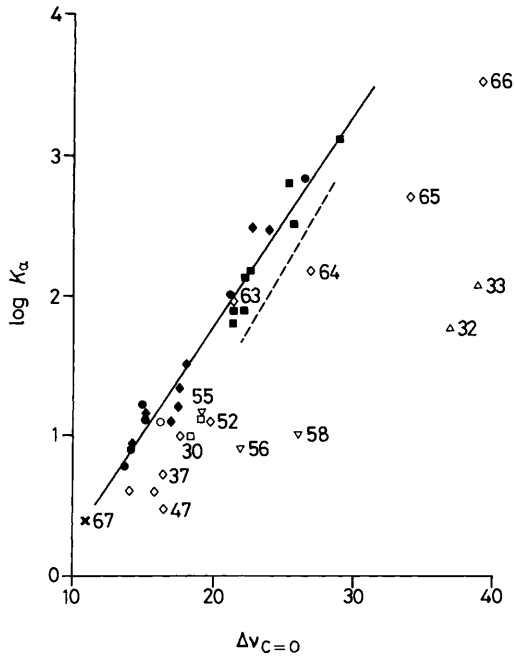


Figure 4. Relation between log K_α and $\Delta\nu_{C=O}$. The correlation line of equation (6) is for the filled points; for dashed line see text. Key: circles, alkanols; squares, phenols; diamonds, NH donors; upright triangles, carboxylic acids; reversed triangles, sulphonamides; cross, CH donor.

such relations involving both $\Delta\nu_{OH}$[68] and $\Delta\nu_{C=O}$[69] with $-\Delta H_f$ while Thijs and Zeegers-Huyskens[70] have recently produced extensive further data of the latter sort. A start in the resolution of this conflict was made by Kamlet *et al.*[71] who showed that ν_{OD} for MeOD with a set of mixed acceptors gave family lines against solvent β; it is possible with hindsight to see some such families in the data assembled by Rao *et al.*[67] Parallel results had previously been obtained by Bellamy and Pace,[72] who demonstrated that $\Delta\nu_{C=O}$ for carbonyl acceptors of various sorts splits into lines when plotted against $\Delta\nu_{OH}$ for the common proton donor. Nevertheless the same authors showed that this does not apply in reverse: $\Delta\nu_{OH}$ *vs.* $\Delta\nu_{C=O}$ gives a single relation when a set of OH donors is examined opposite a single carbonyl acceptor. Part of the confusion may arise from thermodynamic ambiguity: some of these relations are with enthalpy, some with free energy, and for some, such as β as is noted above, it is not clear which or precisely what blend is concerned.

When we started this investigation we had hoped, *contra* Rao *et al.*,[67] to establish one or a few simple relations between $\Delta\nu_{C=O}$ and $-\Delta H_f$ such that the first could be used to predict the second, or *vice versa*. As noted above this aim seemed reasonable in the light of the Frank–Condon principle, while in addition Bellamy's observation[72] suggests that $\Delta\nu_{C=O}$ may be a better behaved parameter than $\Delta\nu_{OH}$. If the relation between $-\Delta H_f$ and Gibbs energy turned out to be tolerably regular, it would then have been possible to use $\Delta\nu_{C=O}$, an easy quantity to measure, to obtain at least a rough estimate of log K_α, experimentally a much more difficult one.* We were therefore somewhat disconcerted to discover that, while relations between ΔG_f and ΔH_f do exist, they not only show no parallelism but many are reversed in slope: it is actually more common than not to find that the bond appears to weaken (ΔH_f less negative) as it becomes more favourable (ΔG_f more negative).[31] Carboxylic acids and anilides, for example, show this phenomenon. This extraordinary paradox has now been traced to an unfavourable enthalpy of desolvation which, in this quite polar solvent, can approach or exceed the favourable enthalpy of hydrogen bond formation.[73] This phenomenon is much attenuated in the less polar solvent tetrachloromethane.[73] It follows that measured ΔH_f values for hydrogen bonding may entirely mislead: the warning of Jencks and Page[26] is all too apposite. The quantity we actually require is one from which translational and entropic terms have been removed. Two recent observations suggest that, after all, $\Delta\nu_{C=O}$ for the common acceptor may reasonably approximate this quantity. There is firstly the semi-theoretical treatment of Bürgi and Dunitz[74] which postulates that bond length, bond strength, and force constant are all related linearly. Secondly, Hillier and his co-workers[75] have recently established, by theoretical studies at the STO-3G level, the relationship of equation (5) for several OH donors against

$$-\Delta H \text{ Kcal mol}^{-1} = 0.09(0.80) + 0.397(0.055)\Delta\nu_{C=O} \quad (5)$$
$$(n = 4, r^2 = 0.964, s = 0.63, F = 53)$$

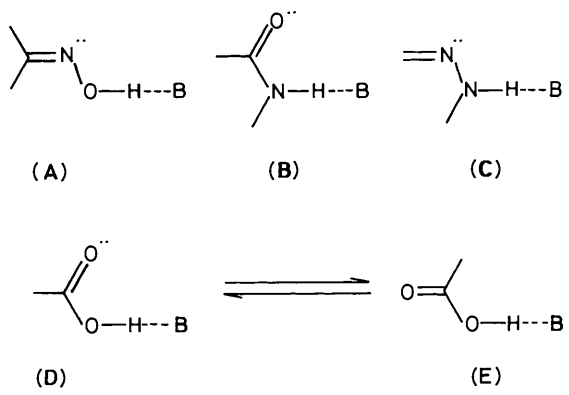
formaldehyde as common acceptor, where $-\Delta H$ is the standard enthalpy change *in vacuo* at 0 K. This relation is linear with a virtually zero intercept. Its effect is essentially to re-establish the Badger–Bauer relationship, but with the proviso that the appropriate enthalpy term is related to ΔH_0°, and this may not be accessible to experiment. We shall refer to this term (the Jencks–Page quantity[26]) as $\Delta H_0'$ in the subsequent discussion, without prejudice to its exact thermodynamic status. For any equation such as (5) to work, a constant probe in a constant solvent is prerequisite.

A linear relation between ΔG_f and $\Delta H_0'$ has been postulated by Jencks and Page[26] for enzyme binding though it is not, of course, required in free solution. Figure 4 shows log K_α as a function of $\Delta\nu_{C=O}$. This demonstrates the single relation of equation (6) from which some points fall short but which

$$\log K_\alpha = 0.152(0.007)\Delta\nu_{C=O} - 1.26(0.14) \quad (6)$$
$$(n = 22, r^2 = 0.960, s = 0.15, F = 479)$$

no point exceeds. It is fit by alkanols, phenols, anilides, carboxamides, and some heterocycles. Even the solitary CH donor (**67**) lies on this line. If $\Delta\nu_{C=O}$ is a measure of $\Delta H_0'$ then equation (6) defines a relationship that will only hold in the absence of some extra entropic constraint. The most obvious such constraints are steric or stereoelectronic. Virtually every deviant point is readily explained in one of these ways [the sole exception is (**47**) where K_α is small and may be at fault]. Most obviously, the category of 2-substituted phenols gives its own correlation line (– – – –) below that of equation (6) (the actual points are omitted to avoid congestion). These lines are separated by $\Delta\log K_\alpha$ *ca.* −0.3, *i.e.* exactly that factor expected on statistical grounds. Most other effects are stereoelectronic. The preferred conformation of oximes (**A**)[76] is shown in Scheme 2; here the nitrogen lone-pair will tend to repel that of the incoming proton acceptor, hence $\Delta\log K_\alpha$ *ca.* −0.55 for (**29**) and (**30**). The change in lone-pair orientation should make this greater for lactams (**B**); hence $\Delta\log K_\alpha$ *ca.* −0.7 for (**52**), whereas primary and secondary carboxamides are normal, is not a surprise. Pyrrole, indole, and imidazole lie on the line; however, the triazoles (**64**) and (**65**) and the tetrazole (**66**) fall away from it very sharply, plausibly because of lone-pair repulsion as shown for (**C**). For (**66**), the strongest proton donor we have yet encountered, this amounts even so to a 15-fold reduction in potential hydrogen bonding strength ('theoretical' log K_α 4.73). It is therefore surprising that (**63**) falls on the line, unless it

* A referee points out that, in principle, the enthalpic term in log K_α relates to the variation in the X–H and C=O bond energies and the energy of the H · · · O bond. Unfortunately, the solvation problems here discussed mean that, at least in TCE, these quantities are inaccessible to experiment. We only wish we were able to resolve this dilemma.



Scheme 2.

prefers to act as donor through the normally disfavoured[50] tautomer (**63b**). No tautomer of (**64**)–(**66**) can escape the constraint illustrated as (**C**).

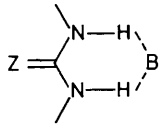
The most spectacular deviant points are for carboxylic acids and sulphonamides. Sulphone is much bigger than carbonyl, contains two oxygens, and is virtually incapable of adopting any conformation in which all lone pairs are shielded from the incoming proton acceptor. The preferred conformation in the crystal state[77] is not the worst in this respect that might have been envisaged, but must still dictate a very tight angle of approach. In fact the sulphonamides (**55**) and (**56**) fall short of the line by $\Delta\log K_\alpha$ *ca.* −0.5 and −1.2 respectively whereas the acylsulphonamide (**58**), with its extra acceptor group, deviates by *ca.* −1.7, *i.e.* falls short by a factor of 50. Carboxylic acids represent a special and intriguing case. With the lactam-like conformation (**D**) a similar shortfall to (**52**) might have been expected, but the actual $\Delta\log K_\alpha$ *ca.* −2.6 is clearly excessive. Since this is around the margin by which the *E*-conformer is disfavoured,[76,78] bonding to NMP as in (**E**) becomes a possibility. This is certainly not the way in which bonding takes place to amides in the solid state, however.[79] Alternatively, $\Delta\nu_{C=O}$ is here distorted by an unusually high degree of proton transfer, though the evidence to be presented below suggests that this is no greater than for phenols. As a further possibility, electrostatic attraction between OH and carbonyl in the preferred *Z*-conformer (**D**) may allow a rival proton acceptor, in solution, only a minute window in which to operate. For the present we have to leave this question open.

2-Chloroethanol (**8**) poses a special problem. Its putative *trans*-conformer ($\Delta\nu_{C=O}$ 16.3 cm^{-1}) lies near to the line and so presumably accounts for most of the bonding (as seen above, the conformational balance is probably quite even). However, we have now to explain both why the gauche-conformer should apparently be capable of an intrinsically stronger bond ($\Delta\nu_{C=O}$ 21.8 cm^{-1}), and also why this is not then the dominant form. Probably this greater strength represents an increase in the effective OH dipole, since now reinforced by C–Cl (in the *trans*-form these are opposed), more than offset by lone-pair repulsion from the same source. Such a potential conflict may often be realised in binding to biological receptors, and elsewhere, and it is instructive to observe this example.

The final class of deviants comprises the aromatic amines (**35**)–(**38**), for which $\Delta\log K_\alpha$ lies in the range −0.3 to −0.5. At these very low K_α values it is difficult to be certain that this difference is real, but if so it probably derives from a ponderal effect on the rapid inversion that is as much a feature of aromatic as of aliphatic amines.[80] A far greater constraint of related origin operates on $\log K_\beta$ (see below).

* 1 eu = 4.184 J k^{-1} mol^{-1}.

If equation (6) is taken seriously we may use it, within limits, to predict $\log K_\alpha$. For (**9**) with similar $\Delta\nu_{C=O}$ we expect about the same value as for (**8**), *ca.* 1.1. A similar value is predicted for the carbamates (**44**) and (**45**) if these, like carboxamides, fall on the line (lone-pair repulsion from sp^2 oxygen may reduce this value). *N*,*N*-Dicyclohexylthiourea (**51**) forms an interesting special case. Here $\log K_\alpha$ *ca.* 2.1 is predicted, far higher than for a simple amide or thioamide and not to be explained in electronic terms. Since ureas and thioureas are dominated by the *Z*,*Z*-conformer[76] we suspect the bifurcated hydrogen bond shown as Scheme 3 (Z = O or S); this is known to have a large enhancing effect on *K* for OH donors.[53] If so, ureas in general may show enhanced donor bonding, an unexpected bonus that once again may carry implications for drug design.



Scheme 3.

Some comment is required on the different slopes of equations (5) and (6). Translated into $-\Delta G_f$, that of (6) becomes 0.207; for NMP against a set of phenols in tetrachloromethane this rises to 0.331 while the slope *vs.* $-\Delta H_f$ is virtually identical, *i.e.* ΔS_f is virtually a constant (*ca.* −10 eu*).[73] So the difference is a solvent effect. Polarisation of carbonyl is disfavoured in the less polar solvent, hence smaller carbonyl shifts and a different slope. The gas phase, equation (5), continues this trend, though here the different acceptor may cause complications. The differing intercepts of these equations are also of interest. Equation (6) predicts zero carbonyl shift at $\log K_\alpha$ −1.26, remarkably close to the value of −1.1 known to represent an effective zero for the $\log K_A^H$ scale.[57] However, this may be misleading. Equation (4) substituted with this minimum value of $\log K_A^H$ yields a putative minimum value for $\log K_\alpha$ of about −0.3. This calculation is of questionable validity and needs to be checked with further work in TCE on a range of proton acceptors. Nevertheless, it is interesting to pursue one possible implication. Equation (6) can be solved in an alternative manner to yield $\Delta\nu_{C=O}$ *ca.* 6 cm^{-1} at $\log K_\alpha$ − 0.3. Given that $\nu_{C=O}$ for NMP moves by more than 17 cm^{-1} between hexane (π^* − 0.08) and TCE (π^* 0.49), a shift of the above order is plausible for a degree of dipolar attraction short of actual hydrogen-bond formation. The theoretical calculations giving rise to equation (5) are specifically tuned to hydrogen bonding so are not expected to throw light on this question. Whatever its precise value, a finite $\Delta\nu$ at zero hydrogen bond strength is not implausible.

Partial proton transfer as in hydrogen bonding is expected to relate to full proton transfer as in aqueous acidity. For proton acceptors, Taft *et al.*[27] found family lines that were later to be rationalised through the co-ordinate covalency parameter ξ. It is of interest whether proton donors show the same pattern. Figure 5 displays our results. There are four main families, summarised as in equations (7)–(10), plus many deviant points,

For carboxylic acids:

$$\log K_\alpha = 3.79(0.09) - 0.39(0.03)pK_a \quad (n = 6,\ r^2 = 0.980,\ s = 0.13,\ F = 201) \qquad (7)$$

For phenols:

$$\log K_\alpha = 6.25(0.57) - 0.40(0.06)pK_a \quad (n = 9,\ r^2 = 0.859,\ s = 0.11,\ F = 42) \qquad (8)$$

For alkanols:

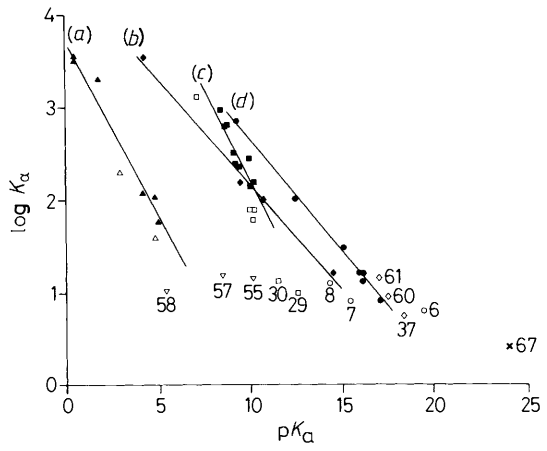


Figure 5. Relation between log K_α and pK_a. Correlation lines are for the filled points, as follows: (*a*), equation (7); (*b*), equation (10); (*c*), equation (8); (*d*), equation (9). Key: circles, alkanols; squares, phenols; diamonds, NH donors; upright triangles, carboxylic acids; reversed triangles, sulphonamides; cross, CH donor.

$$\log K_\alpha = 5.06(0.11) - 0.24(0.01)pK_a \quad (9)$$
$$(n = 7,\ r^2 = 0.995,\ s = 0.05,\ F = 972)$$

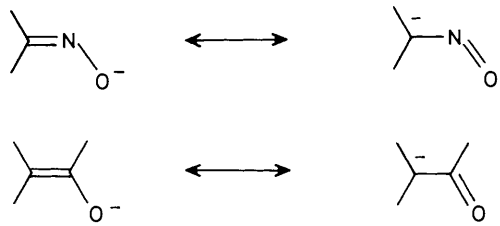
For azoles:

$$\log K_\alpha = 4.49(0.14) - 0.23(0.01)pK_a \quad (10)$$
$$(n = 4,\ r^2 = 0.993,\ s = 0.10,\ F = 278)$$

most of them negative as on Figure 4. These correlations supplement those for log K_A^H,[57] with the important addition of the azoles. They include compounds absent from Table 2 since lacking $\Delta\nu_{C=O}$ but whose K_α has been reported.[31] From the carboxylic acids we have omitted the point for hexanoic acid, which is dubious, and 2-bromobenzoic acid, where additional stereoelectronic repulsion appears to occur. From the alkanols we omit (**7**) which is dubious and (**8**) which is ambiguous (K_α and K_a may reflect different conformers), and also t-butyl alcohol (**6**), whose ionisation is clearly disfavoured more than its proton donor ability by lack of solvation.[57] The phenol set includes no 2-substituents but, even so, scatter is significantly greater than elsewhere. The reason lies in the inductive-resonance blend which is different for partial and full proton transfer;[57] hence the points which reflect this most have been omitted from the correlation, though in fact their inclusion would scarcely affect its slope. The omission of pyrrole and indole from the line for the azoles is considered below.

In discussing the similar equations in log K_A^H we have noted that their slopes may be considered as pseudo-Brönsted coefficients: pseudo, in that hydrogen bonding in some other solvent is being compared with full proton transfer in water.[57] Hence while their absolute values are of unknown significance, their relative values are strictly comparable. The high 'pseudo-α' value of 0.39 for carboxylic acids plausibly reflects the ease with which the developing charge can delocalise into the resonance-stabilised carboxylate ion. At first sight, therefore, the identical value for phenols is unexpected. However, employing the much larger log K_A^H database, we have used a dual substituent parameter (DSP) treatment to demonstrate a degree of resonance involvement in hydrogen bonding by phenols not far short of that for full proton transfer.[57] (There is still enough difference to explain the scatter on Figure 5, however). Its most likely explanation goes back to the observation by Taft and Topsom[81] that the solvation of lone pairs by proton donor solvents restricts their delocalisation, so that removal of this constraint will tend to raise the ρ_R/ρ_I ratio. However that may be, the concordance of these two criteria helps confirm our interpretation of these similar 'pseudo-α' slopes. Consistently, this slope is much less for alkanols where the charge cannot be delocalised, and also for the azoles, from imidazole to tetrazole, where considerable delocalisation in the anion is offset by the much greater reluctance of N than of O to carry a negative charge. This probably explains why pyrrole (**60**) and indole (**61**) do not lie on the azole line: there is no other hetero-atom to help in its dispersal, so their proton transfer acidity is much less in comparative terms than otherwise expected. The still greater reluctance of carbon acids to support anionic charge will explain the position of chloroform (**67**), which continues the trend set by pyrrole and indole and, however poor as a proton donor, is clearly much better than its high aqueous pK_a would predict.

The oximate anion can delocalise charge in a similar manner to phenoxide (Scheme 4) and it is significant, therefore, that (**29**) and (**30**) if corrected according to their



Scheme 4.

deviations from equation (6) would fall on the phenol line. Similarly, the nearly parallel phenol and carboxylic acid lines of Figure 5 are separated by almost the same margin, $\Delta\log K_\alpha$ *ca.* −2.4, as obtains on Figure 4. Such simplicities will only occur when the steric and stereoelectronic factors operating to restrict K_α are absent from ionisation. This is not so for the azoles, where the relation of log K_α with pK_a possesses a regularity not found for $\Delta\nu_{C=O}$. Here the azoles are less effective proton donors than the alkanols by a constant factor of $\Delta\log K_\alpha$ *ca.* −0.4 at any given pK_a, whereas against $\Delta\nu_{C=O}$ they become less effective as the nitrogen content builds up (range 0 to −1.2). The origin of this surprising regularity may lie in enhanced repulsion within the anion as its lone-pair content increases, so that the two resulting forms of lone-pair repulsion cancel out in this comparison. No comparable phenomenon is possible for any other set of compounds studied.

Few other regularities are apparent. Simple carboxamides, for which pK_a 17.7 is estimated,[82] lie on this basis remarkably close to the azole line; log K_α 0.4 is predicted where for (**42**) and (**43**), *ca.* 0.6 is found. If the imide (**53**) with a similar $\Delta\nu_{C=O}$ to the lactam (**52**) possesses a similar value of log K_α, then this is displaced $\Delta\log K_\alpha$ *ca.* −0.8 from the azole line, comparable to the displacement shown by (**52**) on Figure 4. If so, lactams and imides show similar stereoelectronic effects. Sulphonamides as NH donors are displaced from the azole line of Figure 5 even more than from the common line of Figure 4. As for the azoles some lone-pair repulsion in the anion is anticipated, but on crystal structure evidence this is not pronounced,[77] so this difference may originate in the ionisation process. *N*-Methyl-4-nitroaniline (**37**) lies by contrast above the azole line; perhaps the motional restriction which we have suggested may lower log K_α has an even larger effect on ionisation, so that the anomaly really lies in pK_a.

Proton Acceptors.—Whereas proton donors are virtually confined to various sorts of OH and NH group, acceptors vary much more in character: they comprise sp- and sp^2-oxygen and sulphur; sp-, sp^2-, and sp^3-nitrogen; and a few miscellaneous

categories such as aliphatic fluorine and the π-cloud. Table 3 includes no results in the last two categories or for sp^2-sulphur, but otherwise our coverage is comprehensive and in particular includes many heterocycles never before investigated, along with certain new structural types such as lactones, oximes, and an imide. Alongside some extrapolations from the log K_B^H scale[25b] to be found in Table 4, this gives the medicinal chemist data for a range of functionalities not remotely approached by any previous compilation. Conveniently, it will be found that log K_α and log K_β span very similar ranges in quantitative terms.

Table 4. Functional group proton donor and acceptor values.[a]

	log K_α	log K_β
Alkene[b]		(−0.4 to 0.0)
Alkyne	(0.1)	(0.3)
$(Hal)_2CH_2$	(0.1)	
$(Hal)_3CH$	(0.3 to 0.5)	
$(O_2N)_2$ CH Alk	(1.1)	
Alk F		(−0.2)
Alk X[c]		(−0.4)
Alk NH_2	*d*	2.8[e]
Ar NH_2	0.6	(1.0)
Ar NH Ar	(0.6)	(0.4)
Het NH_2	(0.6 to 1.0)	
Alk NHO Alk	(0.8)	
Alk $NHNO_2$	(1.6)	
Alk C=NO Alk		1.5
Alk C=NO Ar		1.1
Amidine C=N		(3.4)
Guanidine C=N		(4.2)
Alk C≡N		1.2
Ar C≡N		1.0
Alk SCN		(0.9)
Cyanamide[f]		2.0
Cyanoguanidine[f]		2.9
Water	(1.2)	(1.2)
Alk OH	1.2	1.4)
Alk O Alk		1.5
Ar OH	2.1	0.2[g]
Ar OD	(2.0)	
Ar O Alk		0.3
Si OH	(1.3)	
Alk NOH	(1.6)	
Alk C=NOH	1.0	1.5
Ar C=NOH	1.1	
Alk CO_2H	2.0	*f*
Ar CO_2H	2.0	*f*
Alk CHO		(1.1)
Ar CHO		(1.2)
Alk CO Alk		1.6
Ar CO Alk		1.4
Ar CO Ar		(1.4)
Alk CO_2 Alk		1.4
Ar CO_2 Alk		(1.2)
Lactone		1.7
Alk CONH Alk	0.7	3.0[e]
Ar CONH Alk	0.7[g]	(2.8)[e]
Alk CONH Ar	1.3	2.5[e]
Lactam	*f*	3.1[e]
Cyclic imide	1.1	1.4[h]
Alk NHCONH Alk	2.1[f]	3.2[e]
Alk OCONH Alk		2.4[e]
Ar OCONH Alk		2.1[e]
Alk OCONH Ar	1.1	(2.0)[e]
Ar NO_2		(0.7)
Het N-Oxide		(3.5)
Alk SO Alk		3.0
Ar SO Alk		(2.7)
Ar SO Ar		(2.4)
Alk SO_2 Alk		1.6
Ar SO_2 Alk		(1.4)
Sulphonamide	1.2	1.4[e]
Acylsulphonamide	1.0	1.0[e,f]
Ar PO		3.9
Alk O_3PO		3.2
Alk SH	*d*	
Alk S Alk		(0.4)
Alk CSNH Alk	(1.0)	1.8[e]
Ar CSNH Alk	(1.0)[g]	(1.6)[e]
Alk CSNH Ar	1.5	
Alk NHCSNH Alk	2.1[f]	2.0[e]
Alk OCSNH Alk		(1.2)[e]
Alk SCSNH Alk		(1.1)[e]
Alk NCS		(0.0)
Ar PS		(1.5)
Alk O_3PS		(1.0)

[a] Alk = alkyl, Ar = aryl, Het = heterocyclic; values in parentheses are scaled from log K_A^H or log K_B^H (see text). [b] Includes aromatic. [c] X = Cl, Br, or I. [d] Negligible hydrogen-bonding ability: see text. [e] Value assumed to be unchanged by replacement of H by alkyl. [f] See text. [g] Estimate. [h] Statistically corrected.

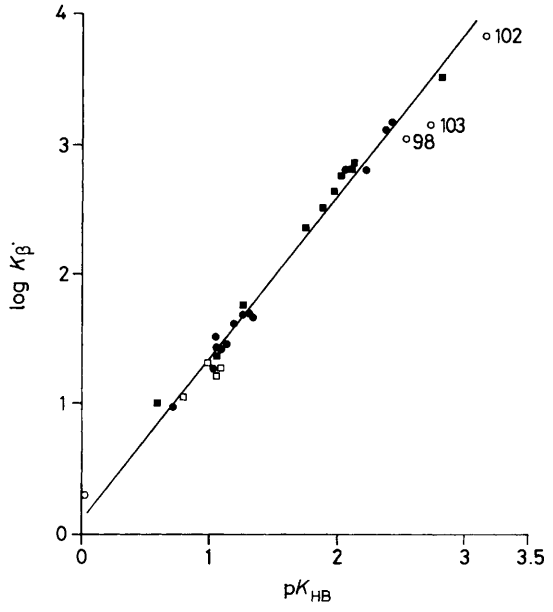


Figure 6. Relation between log K_β and pK_{HB}; the correlation line is for the filled points. Key: circles, oxygen acceptors; squares, nitrogen acceptors.

Figure 6 displays the relation between log K_β and the Taft–Schleyer[27] pK_{HB}. Since both are based on phenol probes a fair correspondence is not unreasonable and, for 24 compounds, this is expressed as equation (11). Two categories, nitriles (two out of four) and X=O bases (X = P or S), lie off the line. It is known[83] that the latter (but not N=O) are best expressed as the dipolar form $\overset{+}{X}$–$\overset{-}{O}$, hence in the more polar solvent TCE they are likely to lose acceptor ability relative *e.g.* to carbonyl where this is less important. This effect, while fairly small, is shown

$$\log K_\beta = 1.27(0.03)pK_{HB} + 0.11(0.05) \quad (11)$$
$$(n = 24, r^2 = 0.990, s = 0.08, F = 2\,141)$$

$$\log K_\beta = 1.30(0.05)\log K_B^H + 0.11(0.07) \quad (12)$$
$$(n = 39, r^2 = 0.958, s = 0.18, F = 843)$$

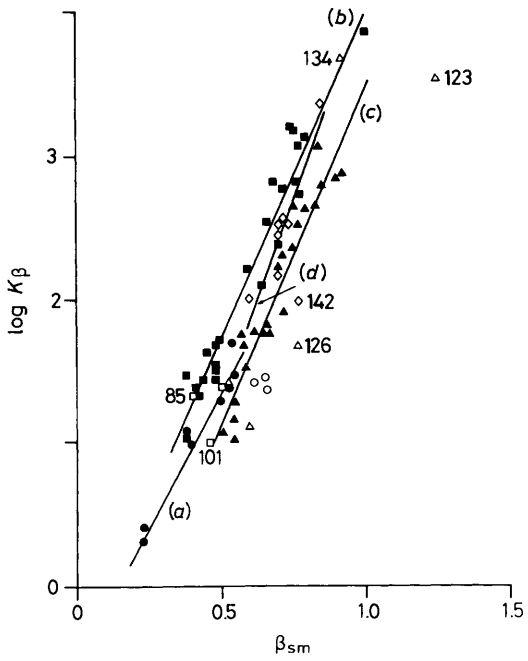


Figure 7. Relation between log K_β and β_{sm}. Correlation lines are for the filled points only, except for equation (16), as follows: (*a*), equation (13); (*b*), equation (14); (*c*), equation (15); (*d*), equation (16), from which (**142**) is excluded. Key; filled circles, ethers; open circles, alkanols; squares, sp-oxygen acceptors; diamonds, potential α-effect nitrogen bases; triangles, all other nitrogen bases.

equally by comparison with our comprehensive log K_B^H scale of proton acceptor ability,[25b] where again both categories are outliers from equation (12). [This scale is defined in an analogous manner to log K_A^H; *cf.* equation (3) and the accompanying discussion.] For nitriles this is also reasonable since these, while weak acceptors, are strongly dipolar. The behaviour of $\overset{+}{P}\text{–}\overset{-}{O}$ and $\overset{+}{S}\text{–}\overset{-}{O}$ illustrates the less common situation in which high charge-transfer ability is overshadowed by a still more pronounced electrostatic term; for donors, this effect has been encountered with 3,5-dinitrophenol.[84] Otherwise, oxygen and nitrogen bases lie on the same line. Since log K_B^H has been scaled to 4-fluorophenol as donor (in tetrachloromethane), it is unsurprising that equations (11) and (12) are identical to within their combined error limits. Their non-unit slope probably mostly reflects the difference in the probe. The only other outlier is anisole (**71**) where K is small and the data may be in error.

We observed near the start of this paper that, while our interest is primarily in Gibbs energy, there may be contexts in which enthalpy is the more appropriate quantity. For proton donors we developed $\Delta\nu_{C=O}$ in an attempt to fill this need and, as demonstrated above, it probably does. For acceptors the equivalent quantity is β_{sm}, plausibly an enthalpy-related quantity on account of the Frank–Condon principle. We may investigate this possibility along the lines of Maria *et al.*[34] For the derivation of θ there are 22 defining compounds of which 12 appear in Table 3 while three more can be added if the reasonable substitutions of (**104**) for Et_3N, (**68**) for Et_2O, and (**103**) for the trimethyl homologue are allowed. For log K_β, θ lies in the range 65–69° according to the compound set employed, which is firmly that expected (66–70°) for the Gibbs energy of hydrogen bonding in 'well-behaved' series of compounds,[34] *i.e.* those corresponding to our 'reasonably general' scales.[25,27] For β_{sm}, we similarly find θ at 40–42°. This is at the top of the range for experimentally determined hydrogen-bond enthalpies (−12 to 42°)[34] as is consistent with our postulate that β_{sm}, like $\Delta\nu_{C=O}$, represents a form of energy from which entropic and translational terms have been largely removed. Hence we postulate β_{sm} as the corresponding measure of $\Delta H_0'$ for proton acceptors.

In the light of this presumption we may examine the relation between log K_β and β_{sm}. Figure 7 displays an overall trend in which four main family relationships may be distinguished. Ethers obey equation (13); there are no outliers but we have

$$\log K_\beta = 3.65(0.33)\beta_{sm} - 0.46(0.14) \quad (13)$$
$$(n = 8,\ r^2 = 0.995,\ s = 0.11,\ F = 125)$$

$$\log K_\beta = 4.58(0.22)\beta_{sm} - 0.56(0.14) \quad (14)$$
$$(n = 24,\ r^2 = 0.950,\ s = 0.18,\ F = 414)$$

$$\log K_\beta = 4.78(0.35)\beta_{sm} - 1.27(0.25) \quad (15)$$
$$(n = 22,\ r^2 = 0.904,\ s = 0.20,\ F = 188)$$

$$\log K_\beta = 5.60(0.83)\beta_{sm} - 1.51(0.60) \quad (16)$$
$$(n = 7,\ r^2 = 0.901,\ s = 0.15,\ F = 45)$$

applied statistical corrections to the diethers (**72**), (**73**), and (**75**). sp-Oxygen acceptors follow equation (14), with statistical corrections applied to the imide (**96**) and the sulphones (**99**) and (**100**). Here the very hindered ketone (**80**) has been omitted from the correlation while (**101**) is also rejected because of ambiguity in β_{sm} (see Results section). Interestingly, the thione (**85**) lies close to this line; other thiones (Table 3) have u.v. spectra too strong for β_{sm} to be determined accurately. Equation (14) is also fit by the acylpyridines (**119**) and (**124**), so demonstrating that, under these circumstances, the major site for hydrogen bonding is carbonyl oxygen not pyridine nitrogen. This agrees with previous conclusions,[85] based on i.r. evidence, and illustrates the way in which the relation between log K_β and β_{sm} may be used to elucidate problems of structure as well as providing quantitative information. (It should be noted, however, that the i.r. method is superior in its ability to detect more than one form of hydrogen bonding when several are present.[85,86]) Amines and most nitrogen heterocycles, with the oxime (**145**) [but not (**146**), which is an outlier], obey equation (15) [pyrazine (**125**) is statistically corrected]. We have excluded all 2-substituted pyridines from this set; these tend to possess lower log K_β values than β_{sm} would predict, *i.e.* behaviour exactly parallel to that of the 2-substituted phenols discussed above. For the strong base (**123**) β_{sm} is anomalous, possibly (as noted in the Results section) through an abnormal degree of proton transfer. However the point for pyrimidine (**126**) is equally anomalous and that explanation is implausible here. The case of 1-methylimidazole (**134**), an outlier in the opposite sense, is discussed below. Nitriles as a set are anomalous, with virtually no change in β_{sm} for considerable variation in log K_β, so all have been omitted from this correlation (and from Figure 7). It has been seen above that their electrostatic-charge transfer balance is abnormal too.

Our final category comprises the potential 'α-effect'[87,88] bases by pyridazine (**127**) and the azoles (**135**)–(**137**), (**139**), (**141**), and (**144**) [(**142**) is rejected as an outlier]. These possess lone pairs on adjacent atoms and, for statistical reasons alone, might be favoured over other nitrogen acceptors. Here we find equation (16) to which, deliberately, statistical corrections have not been applied. Clearly the line for these compounds lies close to that for sp-oxygen where, again, two lone pairs are present. [The fit of 1-methylimidazole (**134**) to this line may indicate an exceptional degree of lone-pair availability almost as if this compound were to possess a truly dipolar structure. Its acceptor ability in the context of solvent–water partitioning is indeed remarkable.[60]] In fact this data set may be combined with that for sp-oxygen to give equation (17) without very much loss in

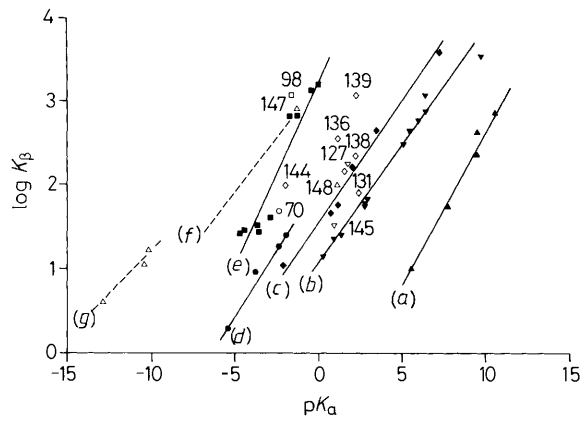


Figure 8. Relation between log K_β and pK_a. Correlation lines are for the filled points, as follows: (*a*), equation (19); (*b*), equation (20); (*c*), equation (21); (*d*), equation (22); (*e*), equation (23); (*f*) and (*g*), see text. Key: circles, alkanols, and ethers; squares, carbonyl compounds; triangles, alkylamines; inverted triangles, 6-membered nitrogen heterocycles; diamonds, 5-membered nitrogen heterocycles; open triangles, nitriles.

$$\log K_\beta = 4.44(0.23)\beta_{sm} - 0.52(0.15) \quad (17)$$
$$(n = 31, r^2 = 0.930, s = 0.20, F = 382)$$

$$\log K_\beta = 4.03(0.22)\beta_{sm} - 0.72(0.14) \quad (18)$$
$$(n = 30, r^2 = 0.925, s = 0.20, F = 347)$$

statistical rigour. Similarly, (13) and (15) may be combined to give (18) to which the same applies. These new equations, (17) and (18), are nearly parallel. They correspond to two (or more) lone pairs, or only one, respectively, and in the region of greatest interest are separated by $\Delta\log K_\beta$ *ca.* 0.5, which at $\beta_{sm} > 0.5$ is also about the separation between the major data sets that fit (14) and (15). It is plausible to regard this separation as essentially entropic in origin, the result of a fundamental difference in steric demand. Millen and his co-workers[58b] have shown, for one class of acceptor in the gas phase, that a variation of $\pm 30°$ in angle of approach entails a loss of 3.8 kJ mol^{-1} in binding energy, which roughly translates into the vertical spread of Figure 7. A recent crystal structure study[89] shows, *inter alia*, that hydrogen bonds formed by nitrogen heterocycles are highly directional whereas those to sulphonamide possess almost no directionality at all. Carbonyl possesses a high degree of directionality[54,90] but with an available angle of approach much greater than that of nitrogen heterocycles.[89] Crystal structures do, of course, represent a compromise between packing forces of different kinds.[91] Our evidence suggests that a single lone pair is effectively disadvantaged, relative to more than one, by about a factor of three in free solution. This is not a large factor on a scale of four decades but the constraints imposed by the biological receptor might very well amplify it greatly.

The near-equivalence of the lines for amines and ethers on Figure 7, and their successful combination in equation (18), re-emphasise Hine's evidence[53] that ethers in solution possess effectively one lone pair. This restriction must clearly apply to OH as well, accounting for the observation[92] that alkanols in bulk never form more than two hydrogen bonds. Interestingly, the three alkanol points on Figure 7 (open circles) fall below even the ether line. A probable reason for this is the H–H repulsion term which is present in OH···OH hydrogen bonding[93,94] but, of course, is absent when ether is the acceptor. The result, paradoxically, is an improved enthalpic term since the highly restricted angle of approach is itself the most energetically favourable,[94] so that ΔH is not diluted by contributions from less favoured conformers. These effects cancel in Gibbs energy so that ethers and alkanols possess essentially identical log K_β values. Other forms of enthalpy–entropy compensation may be hidden in these data; they will show themselves, if present, as varying β_{sm} values at a near constant log K_β. These variations are much more pronounced for acceptors *vs.* a common donor (than *vice-versa*) since acceptors vary so much more in character. However, one factor that does not appear to be important in this respect is the extra rigidity imparted by cyclisation; we have looked for, and failed to find, any systematic trend attaching to cyclic *vs.* open-chain compounds. Nevertheless there is some slight evidence for an effect on ΔG: log K_β is somewhat enhanced in the ether (**70**), the ketone (**81**), and possibly the amide (**90**) by the side of their nearest open-chain equivalents, and the same may be true for the sulphone (**99**) though here we have no comparative data by which to judge. This is also true for γ-butyrolactone (**84**) though its origin here may be the 'α-effect'. Hybridisation changes at the ring atoms may be responsible, though this is uncertain and the effect itself is marginal. The magnitude of any such effect in nitrogen ring heterocycles cannot, of course, be judged (but see below).

Equations (17) and (18) extrapolate to a mean value of log K_β *ca.* −0.6 for zero β_{sm}. It has been seen above that log $K_A{}^H$ and log $K_B{}^H$ are scaled to a common effective zero in the region of −1.1,[25,27] while log K_α probably possesses an effective zero somewhere between −1.3 and −0.3. These figures are scarcely precise but appear to indicate that the same types of relation are likely to hold. This 'effective zero' represents the point at which a hydrogen bond (directional) becomes indistinguishable from a vaguely dipolar attraction (non-directional).[57] For proton donors, we have seen that alkylamines and alkyl thiols enter this category. Probably none of the proton acceptors we have studied is quite so weak; the nearest likely categories in a more general survey are halogens (except fluorine), π-donor heteroatoms, and some π-acceptors (see the next section).

In their original derivation of the pK_{HB} scale, Taft and his co-workers[27] demonstrated family relationships between pK_{HB} and pK_a as a series of roughly parallel lines each of slope *ca.* 0.2. The most prominent such families were primary amines, 3- and 4-substituted pyridines, and carbonyl compounds, with fragmentary evidence for others. These families have since been characterised by the 'co-ordinate covalency' parameter ξ[38] of which values exist for P=O, C=O plus S=O, ether, pyridine, and amine bases. Each of these should therefore give families in any plot of log K_β *vs.* pK_a. This plot is shown as Figure 8; results are necessarily more fragmentary than *e.g. vs.* β_{sm} since many relevant pK_a values are unknown (*cf.* Table 3). Most of the above families do appear, along with limited data for nitriles and clear evidence, for the first time, that 5- and 6-membered ring nitrogen heterocycles lie on different lines.

The equations that describe these families appear as (19)–(23). All the alkylamines have been used to define (19). For (20)

For aliphatic amines:

$$\log K_\beta = 0.37(0.04)pK_a - 1.08(0.31) \quad (19)$$
$$(n = 5, r^2 = 0.974, s = 0.14, F = 110)$$

For 6-membered ring heterocycles:

$$\log K_\beta = 0.27(0.01)pK_a + 1.07(0.06) \quad (20)$$
$$(n = 12, r^2 = 0.983, s = 0.11, F = 561)$$

For 5-membered ring heterocycles:

$$\log K_\beta = 0.29(0.02)pK_a + 1.55(0.06) \quad (21)$$
$$(n = 6, r^2 = 0.987, s = 0.11, F = 312)$$

For alcohols and ethers:

$$\log K_\beta = 0.32(0.03)pK_a + 2.05(0.11) \quad (22)$$
$$(n = 4, r^2 = 0.981, s = 0.08, F = 104)$$

For carbonyl compounds:

$$\log K_\beta = 0.44(0.05)pK_a + 3.24(0.14) \quad (23)$$
$$(n = 9, r^2 = 0.923, s = 0.24, F = 83)$$

we have omitted the 2-substituted pyridines, since there is clear evidence in Table 3 that short-range forces affect log K_β and pK_a very differently, and the acylpyridines (**119**) and (**124**) which have been shown above to bond predominantly at carbonyl. Otherwise, the only omission is pyridazine (see below). Pyrazine (**125**) and pyrimidine (**126**) have been statistically corrected for pK_a as well as log K_β. All the 'α-effect' bases have been omitted from equation (21) but the only outliers otherwise are the triazole (**138**) and thiazole (see below). The necessary omission of tetrahydrofuran (**70**) from equation (22) may point to some effect of bond hybridisation or rigidity on partial proton transfer that is not reflected in pK_a; this could be relevant to the difference between 5- and 6-membered ring heterocycles (again, see below). Here dioxane, again, has been statistically corrected in both quantities. The least satisfactory relation is that for the carbonyl compounds of line (*e*). Here no compound has been omitted but there is considerable dispute as to what such pK_a values actually mean (the same to a lesser extent is true of ethers). In accordance with consensus opinion at the present time, we have taken ours from the very slim 'approved' list of Arnett and Scorrano[95] in which the raw data have been properly corrected for activity effects by the Bunnett–Olsen procedure. However, this debate has recently been re-opened by Johnson and Stratton[96] who have shown that the expected variation in pK_a with substituent for several classes of weak organic base correlates better with the original Hammett values than with those derived by these more sophisticated procedures. If the original Hammett values are used, so that, *e.g.* pK_a −4.36 for acetophenone (**83**) becomes *ca.* −7, the slope [line (*f*)] of equation (23) falls to *ca.* 0.25. This is the value obtained if only the four most basic carbonyl compounds are considered, for which there is little difference between pK_a as derived from different scales. If these alone are used, or 'raw' pK_a values are taken throughout, the slopes of equations (20)–(23) become almost identical. It should be noted that the parallelism reported by Taft *et al.*[27] was based on the original Hammett values. The dashed line (*g*) shown for nitriles—it cannot be dignified as a correlation—is drawn on this basis also.

The separation between the lines of Figure 8 may be discussed in either of two mutually exclusive ways. We may enquire why certain classes of compound should be stronger or weaker as bases at a given degree of proton acceptor ability, or why these same classes are better or worse as proton acceptors at a given pK_a. Concerning the first, Kamlet *et al.*[39] discuss their ξ-values in terms of atom electronegativity: as this increases, a full charge will become steadily more disfavoured relative to a partial one. Similarly, Arnett and Scorrano[95] draw attention to the much greater solvation requirements on protonation of oxygen than of nitrogen bases. These factors, which are not unconnected, will satisfactorily explain why lines (*d*) and (*e*) on Figure 8 lie to the left of lines (*a*)–(*c*). From line (*g*) it seems probable that nitriles, which are still more electronegative than any class of carbonyl,[97] will lie further to the left again.

The second argument is also worth pursuing. To take any line as standard has to be arbitrary, but by selecting (*b*) we standardise on a series of rigid compounds the hybridisation of whose acceptor atom is unaffected by protonation, whether partial or complete. By contrast alkylamines, which rapidly invert in the free-base form,[80] are converted to a rigid tetrahedral geometry not only in the cation but merely by accepting a hydrogen bond. Hence almost all this motional constraint is forced in the initial step; the energy required for partial proton transfer is greatly enhanced, hence their sharply reduced log K_β values from those otherwise expected. This argument predicts a considerably elevated log K_β for very rigid amine bases. No data appear in Table 3, but by using equation (12) we predict log K_β *ca.* 3.5 for quinuclidine, which places it well above line (*a*). Since aromatic amines invert just as fast or faster,[80] they may be expected to lie on or near line (*a*) as well, and once again a high degree of rigidity should confer advantages.

The difference between 5- and 6-membered ring nitrogen heterocycles was unexpected but may be rationalised. The internal bond angle at sp^2-nitrogen is[98] about 110°; on protonation this opens out to 120° or so. In rigid 5-membered rings this opening cannot take place, so an energetic penalty results and line (*c*) lies to the left of line (*b*). Greater lone-pair accessibility in the 5-membered ring may also help. Hence these are relatively favoured by Δlog K_β *ca.* 0.5 at a given pK_a. Possibly thiazole (**131**) lies between these lines since its sulphur atom makes the ring intermediate in size (but benzothiazole. whose pK_a is approximate, seems normal). The position of the oxime (**145**) is also indeterminate. The triazole (**138**) which lies above the line possesses two acceptor sites (see below).

The 'α-effect' brings about some quite spectacular exaltations in log K_β. Pyridazine (**127**) deviates from line (*b*) by Δlog K_β 0.64, *i.e.* to a position above line (*c*). The 1,2,3-triazole (**136**), 1,2,4-triazole (**139**), and tetrazole (**144**) lie above line (*c*) by margins of Δlog K_β 0.65, 0.84, and 1.0 respectively (pK_a for the last is estimated so this margin may be in error). Only the benzotriazole (**137**) fails to show the expected exaltation, but its quoted pK_a is doubtful. [Statistical corrections for pK_a and log K_β have been applied to (**127**) and (**139**)]. Since we are comparing equilibria, this exaltation monitors the difference in the magnitude of the effect and not its absolute value, which may be even greater than appears. The 'α-effect' is well documented as a phenomenon in aqueous nucleophilic and general acid–base catalysis,[99] though for pyridazine it is not always found.[100] However it shows itself in the higher aqueous pK_a of pyridazine than of its isomeric diazines (Table 3) and is clearly evident in some related solvent–water partitioning phenomena,[60] so its presence while more marked here is not without precedent. In fact, the use of 'α-effect' heterocycles is one of the ways in which medicinal chemists may optimise proton-acceptor ability at the expense of full protonation, a highly desirable aim if biological membranes are to be crossed.[45] Taft *et al.*[101] have drawn attention to the potential use of exceptionally electron-rich carbonyl compounds for just this purpose, notably the cyclohexenone (**156**) for which, from line (*f*), we may estimate log K_β *ca.* 4. Because of its strength we have been unable to obtain an accurate value for the quinolone (**97**) for which about the same value would be expected. Other oxoheterocycles may fall in this category. It may even be that one may combine these principles in structures of type (**157**), where a kind of α-effect could operate, to generate compounds of quite remarkable proton acceptor ability, but so far we possess no data.

The acylpyridines (**119**) and (**124**) do not appear on Figure 8 since, as seen above, protonation takes place at one site but hydrogen bonding at another; if plotted, they would fall at arbitrary positions between lines (*b*) and (*e*). Since dimethylcyanamide (**148**), which is known to protonate on the amino group,[102] falls between lines (*a*) and (*g*), this is evidence that in hydrogen-bonding terms it behaves as a nitrile. In fact, we may go further. Extrapolation of the nitrile line to log K_β *ca.* 2 leads

to a prediction of pK_a *ca.* −6 for the disfavoured cation (**158**). From pK_a *ca.* zero for a 'regular' carbodi-imide,[102] this places the tautomeric ratio between these species as *ca.* 10^6 in favour of the cyanamide form. Given the enormous approximations involved, including the use of Hammett pK_a values, this ratio must be regarded as indicative only, but nevertheless it is not unreasonable and it certainly lies in the right direction. The result for the cyanoguanidine (**147**) is even more intriguing. With its very low pK_a and high log K_β this behaves as no sort of imino-compound and, in fact, it plots almost along an extrapolation of the nitrile line. Protonation on nitrile is improbable in water[103] and high level MO calculation shows it to be highly disfavoured for the isolated molecule;[104] however, these same calculations reveal nothing to choose between hydrogen bonding to sp- or sp^2-nitrogen. There is precedent for preferential bonding to nitrile when both possibilities are open.[105] The evidence is best reconciled if we suppose that protonation is predominantly, but perhaps not exclusively, at imino-nitrogen, whereas the reverse (in TCE) applies to hydrogen bonding. The cyanoguanidine moiety has been proposed as a urea isostere.[106] If the above picture is correct, the analogy is perhaps even closer than its originators had imagined; not only is log K_β similar—compare (**147**) with (**92**)—but their distribution of lone-pair density may be very much the same.

It was noted above that the slopes of equations (7)–(10) may be regarded as 'pseudo-α' values in Brönsted terms. In a similar way, the slopes of equations (19)–(23) become 'pseudo-β' values. While these have no significance in the conventional sense, they are strictly on the same scale since aqueous pK_a cancels out in the comparison. Hence it is interesting that the latter should be relatively constant while the former are so much more variable. It is even more significant that, at least for strong donors, $\alpha > \beta$. This suggests a fundamental asymmetry in hydrogen bond formation even at equilibrium which does not appear to have been considered. In fact, it is confirmed by MO calculations which show the developing positive charge to be shared between the proton in flight and its acceptor.[94] The possible implications of this asymmetry to the interpretation of acid–base catalysis will be explored on another occasion.

General Survey of Donors and Acceptors.—Table 4 comprises a general survey of those functional groups of most use to the medicinal chemist. Many of these are scaled from log K_A^H or log K_B^H and where this has been done, as seen above, certain cautions are necessary. Since NH donors tend to be over-valued by equation (4) we have carried out a survey of available data and, as a result, the derived log K_α values of Table 4 have been scaled down by as much as 0.4. A rather smaller adjustment appears needed for CH and here the mean reduction is 0.2. Similarly, scaling adjustments of *ca.* −0.3 have been applied to S=O and about half this to C=S. Values for halogen and the π-cloud have had to remain unadjusted but in any case these values are small. One or two estimates are by close analogy; for example, since average log K_β for alkanols and alkyl ethers differs by *ca.* 0.1, the same is assumed for their aryl analogues. It would be tempting to assume the same log K_β values for carboxylic acids and their esters, but we are forced to hesitate since log K_α is so abnormal and the same factors might affect log K_β; here we await further evidence. For amines and amides we assume that alkylation will have substantially no effect on log K_β, a view that receives some support.[25c] It has been seen above that the proton donor ability of simple alkylamines and alkyl thiols is essentially zero. An important consequence, therefore, is that these are not bioisosteres for OH.

A superficial reading of Table 4 might suggest that log K_β varies much more, and covers a wider range, than log K_α. This is misleading, as Table 2 will demonstrate: most types of substitution will increase log K_α, so that the ranges actually observed are very similar. Equations (7)–(10) and (19)–(23) should be useful for generating further values where pK_a is known. In the special case of log K_β for heterocycles, the parallelism between equations (20) and (21) allows generalisation to equation (24) [provided of course that bonding is to

$$\log K_\beta(RX) = \log K_\beta(RH) + 0.28\Delta pK_a \qquad (24)$$

ring N: *cf.* (**119**) and (**124**) as exceptions]. Here ΔpK_a is that of RX relative to RH (R = heterocycle, X = substituent) and there seems no reason to restrict its use to heterocycles that fall on lines (*b*) and (*c*); it should be equally applicable to outliers such as, *e.g.*, pyridazine and tetrazole.

Many cross-comparisons have been discussed and more will be evident on inspection, but a few general trends may be noted. Where a level comparison can be made, OH donors are stronger than NH typically by $\Delta\log K_\alpha$ 1–1.5. The comparison for acceptors is more difficult to apply but, typically, N scores over O by a similar margin for each hybridisation type, with S much weaker again. (A valuable comparison of O and S acceptors has recently appeared.[32]) In fact hybridisation has little effect except for nitrile, which is much weakened relative to other forms of N, and alkyne, which on rather limited evidence[25c] is a surprisingly good acceptor. Substitution of aryl for alkyl sharply lowers log K_β and raises log K_α where attachment is directly to a hetero-atom, but otherwise its effect is more muted. Most heterocycles behave in this respect as 'super-aryls'. Inside many functional series the expected electronic effects are clearly apparent; thus amides and ureas are much stronger acceptors than esters or ketones, while sulphone is much weaker than sulphoxide. In view of the known difference in protonation site for carboxamides[107] and sulphonamides,[108] *i.e.* O and N respectively, it is of interest that the latter are even weaker acceptors than sulphones. Both differences occur since sulphone is ineffective as a resonance acceptor by the side of carbonyl,[97] so that only the inductive effect of nitrogen counts.

We may also comment on competitive proton-acceptor sites within a single functional group. Electronic arguments will dispose of the nitrogen atom of amides, or the corresponding oxygen of esters, as likely sites for a hydrogen bond in free solution (the crystal state of course is another matter). The latter conclusion extrapolates from log K_β for alkyl and aryl ether (Table 4) and is of special importance as the contrary assumption is sometimes made. The same conclusion applies to urethanes, ureas, their thio-analogues, and indeed to π-donor heteroatoms generally. Simple π-excessive heterocycles are extremely poor proton acceptors[25c] and what bonding there is probably attaches to the π-cloud.

One other general trend, especially among acceptors, is a tendency to bunching. Esters, ethers, and ketones, all at log K_β *ca.* 1.5, are to that extent rough isosteres. Then at log K_β *ca.* 3.0 we have urea, carboxamide, and sulphoxide (but not sulphone). Relatively little falls in this gap. However, it is readily bridged by the heterocycles, which not only span a very wide range in log K_β but are much better able to 'fine-tune' this range than most of the conventional substituents in Table 4. In this respect at least, the common predeliction of medicinal chemists for heterocycles seems amply justified.

Heterocycles as acceptors have one problem, however. Except in the simplest cases, as imino-nitrogens are added it is not immediately evident which is likely to be the preferred acceptor site. Some light is thrown on this by a ^{15}N n.m.r. study in which chemical shifts were assigned unequivocally to individual nitrogen atoms.[109] On the assumption that the chemical shift for sp^2-N at highest field will identify the preferred acceptor site, these are as shown starred for (**159**)–(**162**). All seem chemically reasonable. Of course we do not know by what

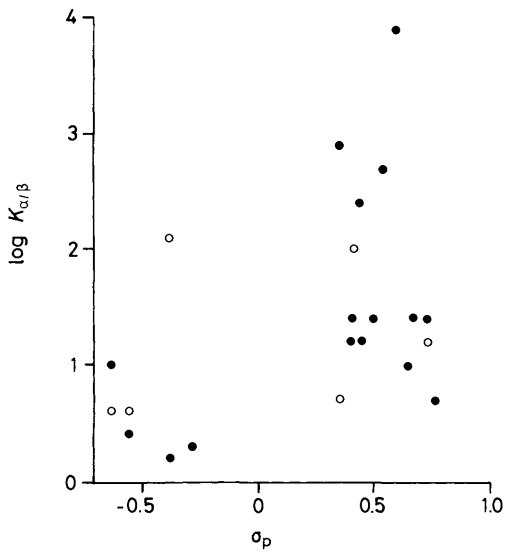


Figure 9. Relation between log K_α (open circles) and log K_β (filled circles) as a function of σ_p.

margin each site will be favoured. In the case of (**162**), while crystal structure evidence confirms N-3 as the sole protonation site, it also demonstrates a much lower degree of bond localisation than the conventional representation would suggest,[110] and similar ambiguities may be present in other cases.

In an attempt to extend the present range we draw special attention to the extreme class of proton acceptor represented by aminoenones such as (**156**) and some related oxoheterocycles since, with one exception,[101] their potential in this respect does not seem to have been recognised till now. Nitrogen need not always be the donor atom: from equation (12), flavone (**163**) and the pyrone (**164**) calculate to give log K_β 2.6 and 3.3 respectively. This enormous gain in proton-acceptor ability by vinylogous amides and esters over their 'standard' equivalents (Δlog K_β 1–1.5) is probably due, in the former class, to avoidance of the σ-resonance effects[52] that, in the latter, remove lone-pair electron density from carbonyl. We have previously classified aminoenones as 'superamides' on spectroscopic grounds[111] and a similar potential exists among putative acylguanidines.[112] Gramstad[113] has drawn attention to the extreme proton-acceptor ability of phosphorus ylides such as (**165**). The 'principle of vinylogy'[114] can probably be used to construct some exceedingly powerful proton acceptors as and when the need arises.

There are also the quantities $\Delta\nu_{C=O}$ and β_{sm} which, as we have seen, are probably enthalpy-related. For work in unconstrained circumstances, *e.g.* free solution, Gibbs energy is unquestionably the quantity that the medicinal chemist requires, but that is not always the relevant situation. At the receptor surface, if all substituents are forced along a narrow trajectory, $\Delta H_0'$ might turn out to be more relevant.[26] It may then be worth trying one of these quantities instead.

Finally we draw attention to an aspect of equation (1) on which there has been no comment. Use of an electronic term in the Hansch equation[11] presupposes that any electrical effect due to a change in substitution is adequately covered by Hammett's σ or some equivalent. For its influence transmitted to some remote group, that may be adequate. For a direct influence on the biological receptor, it plainly is not. Figure 9 shows log K_α and log K_β as a function of σ_p (hence only aromatic substituents appear in this comparison). The result is merely a scattergram and, while σ_p was chosen for its blend of σ_I with σ_R, it would have made little difference what form of σ were used. It seems scarcely surprising that the electronic term has so rarely proved of major importance in correlations according to equation (1). We hope in this paper to provide the medicinal chemist with a more viable alternative.

Overview and Rationale.—The present results have all been obtained in TCE, a solvent conveniently situated half way along the polarity scale (π^* 0.49; *cf.* zero for cyclohexane, 0.29 for tetrachloromethane, and 1.09 for water[17]). Biological surroundings of relevance can range all the way from deep lipid bilayers to the nearly aqueous. We need to enquire how our ranking orders may be affected by such variations.

While we can give no quantitative answers, the underlying principles are fortunately quite simple. We have seen that NH donors lose importance relative to OH as solvent polarity rises since the electrostatic component diminishes relative to that of charge transfer.[65] Very strong NH donors may fail to show this trend but, in that case, their behaviour will simply become more 'regular' and so of less concern. Proton acceptors behave in the opposite way: nitrogen (except nitrile) gains ground relative to oxygen since its lone pair is more polarisible and so becomes more accessible as the surroundings become more polar.[40] Also, as seen, dipolar species such as P=O and S=O will lose out *e.g.* relative to carbonyl as solvent polarity rises. Hence, if we arbitrarily take the OH ··· O bond as standard, then as solvent polarity rises, OH ··· N bonds will strengthen, NH ··· O will weaken, and NH ··· N may perhaps remain about the same. It is impossible so far to quantify this picture over the whole solvent range. We know nothing concerning other species but it seems plausible that the relative importance of the charge-transfer component should generally fall in the order OH > NH > CH and N > O > S. Hence, for example, the volatile anaesthetics, the best of which are halo-carbon CH donors, will probably exert their effects in regions of extremely low polarity, which indeed is supported by the evidence.[115]

Finally we need to consider whether quantitative hydrogen-bonding scales are relevant to medicinal chemistry at all. It has been strongly argued that hydrogen bonding can be treated quite adequately as an indicator variable,[116] and it has recently been demonstrated by Fersht and his co-workers[117] that the excision by protein engineering of a single hydrogen bond from an enzyme-substrate complex involves a penalty of only 0.5–1.5 kcal mol^{-1} in binding energy. The common rationale is that of Jencks:[118] since enzyme-substrate binding simply replaces two hydrogen bonds to water by two others, and water itself forms strong hydrogen bonds, the nett change in energy in this respect is likely to be rather small.

We shall consider this and the counter-argument fully at another time, but a summary is as follows. The resemblance between drug-receptor and enzyme–substrate interactions is considerable but it can be overdrawn.[119] Most receptors of the type that concern us here exist not on protein surfaces but deep in the interior, protected indeed from unwanted influences by their hydrophobic environment.[5–7] Water is not present in bulk (if at all). The properties of bulk and monomeric water are different to a degree that is rarely appreciated. Bulk water is an exceptional proton donor;[17] as a monomer it is no better than ordinary, whether donor or acceptor (Table 4). In saturated TCE solution it is monomeric to *ca.* 90% on our i.r. evidence and in nonpolar solvents this behaviour is quite normal.[25c] There is a confusion here between its fugacity in a hydrophobic environment, which is very high, and its ability to form hydrogen bonds in such surroundings. The latter is in no way unusual: we have demonstrated that, in TCE, water does not compete with other donors for NMP even when present in excess.[31] And since the difference between strong and weak donors (or acceptors) increases as solvent polarity falls,[25c] the ability of water to compete will get worse as this happens, not better. The effect of

a hydrophobic environment is to squeeze water out, not to enhance its (relative) binding ability. Hence water will exert no effective competition to other polar groups and their differential efficacy will show itself. We possess QSAR studies that can only be rationalised if this argument is valid.[120] The special properties of water in biological media is a myth.

Conclusions.—While the present study is still far from complete, we have judged it sufficiently mature for release at this stage despite the annoying omissions a different selection of which will occur to each reader. Much remains to be done. Nevertheless we intend these scales as the start of a data base which the medicinal chemist can use to quantify ideas that till now have been able to receive only qualitative expression. If $\log K_{\alpha}$, $\log K_{\beta}$, $\Delta\nu_{C=O}$, and β_{sm} join the pantheon of variables that currently include[11] $\log P$, π, MR, σ, and E_s, amongst others, we shall have achieved our objective.

Acknowledgements

We thank the SERC for grants to Mr. D. V. Prior and Dr. P. P. Duce. This paper is dedicated to the memory of Mortimer J. Kamlet, without whose pioneering endeavours none of this work would have been possible.

References

1 W. M. Latimer and W. H. Rodebush, *J. Am. Chem. Soc.*, 1920, **42**, 1419.
2 L. Pauling, R. B. Corey, and H. R. Branson, *Proc. Natl. Acad. Sci. U.S.A.*, 1951, **37**, 205.
3 J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 737.
4 W. P. Jencks, 'Catalysis in Chemistry and Enzymology,' McGraw–Hill, New York, 1970, *passim.*
5 C. J. Cavallito, 'Structure-Activity Relationships,' Section 5 of 'International Encyclopaedia of Pharmacology and Therapeutics,' ed. G. Peters, Pergamon, Oxford, 1973, vol, I, p. 1.
6 A. Albert, 'Selective Toxicity,' 7th edn., Chapman and Hall, 1985, p. 498.
7 E. J. Ariëns, in 'Stereochemistry and Biological Activity of Drugs,' eds. E. J. Ariëns, W. Soudijm, and P. B. M. W. M. Timmermans, Blackwell, Oxford, 1982, p. 16.
8 H. Erlenmayer and M. Leo, *Helv. Chim. Acta*, 1932, **15**, 1171.
9 H. L. Friedman, *N.A.S.N.R.C., Publ.*, 206, 1951, p. 295.
10 C. W. Thornber, *Chem. Soc. Rev.*, 1979, **8**. 563.
11 C. Hansch, *Acc. Chem. Res.*, 1969, **2**, 232.
12 R. W. Taft, in 'Steric Effects in Organic Chemistry,' ed. M. S. Newman, Wiley, New York, 1956, p. 556.
13 A. Verloop, W. Hoogestraaten, and J. Tipker, in 'Drug Design,' ed. E. J. Ariëns, Academic Press, New York, vol. 7, 1977, p. 165.
14 C. Hansch, *Intra-Sci. Chem. Rep.*, 1974, **8**, 17.
15 C. Hansch and W. J. Dunn, *J. Pharm. Sci.*, 1972, **61**, 1.
16 C. Hansch and J. M. Clayton, *J. Pharm. Sci.*, 1973, **62**, 1.
17 M. J. Kamlet, J.-L. M. Abboud, and R. W. Taft, *Prog. Phys. Org. Chem.*, 1981, **13**, 485.
18 M. H. Abraham, P. L. Grellier, J.-L. M. Abboud, R. M. Doherty, and R. W. Taft, *Can. J. Chem.*, 1988, **66**, 2673.
19 M. J. Kamlet, R. M. Doherty, J.-L. M. Abboud, M. H. Abraham, and R. W. Taft, *Chem. Tech. (Leipzig)*, 1986, 566.
20 E. M. Kosower, *J. Am. Chem. Soc.*, 1958, **80**, 3253.
21 C. Reichardt, *Angew. Chem., Int. Ed. Engl.*, 1965, **4**, 29.
22 E. Grunwald and S. Winstein, *J. Am. Chem. Soc.*, 1948, **70**, 846.
23 V. Gutmann, 'The Donor-Acceptor Approach to Molecular Interactions,' Plenum Press, New York, 1978.
24 M. J. Kamlet and R. W. Taft, *J. Am. Chem. Soc.*, 1976, **98**, (*a*) p. 377; (*b*) p. 2886; R. W. Taft and M. J. Kamlet, *J. Chem. Soc., Perkin Trans. 2*, 1979, 1723.
25 (*a*) M. H. Abraham, G. J. Buist, P. L. Grellier, J. J. Morris, P. J. Taylor, C. Laurence, P. Nicolet, P.-C. Maria, J.-F. Gal, J.-L. M. Abboud, R. M. Doherty, M. J. Kamlet, W. J. Shuely, and R. W. Taft, *J. Phys. Org. Chem.*, 1989, in the press; (*b*) M. H. Abraham, P. L. Grellier, D. V. Prior, J. J. Morris, P. J. Taylor, C. Laurence, and M. Berthelot, *Tetrahedron Lett.*, 1989, **30**, 2571; M. H. Abraham, P. L. Grellier, D. V. Prior, J. J. Morris, and P. J. Taylor, *J. Chem. Soc., Perkin Trans. 2*, submitted; M. H. Abraham, P. L. Grellier, D. V. Prior, R. W. Taft, J. J. Morris, P. J. Taylor, C. Laurence, M. Berthelot, R. M. Doherty, M. J. Kamlet, J.-L. M. Abboud, K. Sraidi, and G. Guihéneuf, *J. Am. Chem. Soc.*, 1988, **110**, 8534; (*c*) literature data accumulated by M. H. A. in preparation for the preceding analyses.
26 W. P. Jencks and M. I. Page, *Proc. Eighth FEBS Meeting*, Amsterdam, 1972, **29**, 45; W. P. Jencks, *Adv. Enzymol.*, 1975, **43**, 219; M. I. Page, *Angew. Chem., Int. Ed. Engl.*, 1977, **16**, 449.
27 R. W. Taft, D. Gurka, L. Joris, P. von R. Schleyer, and J. W. Rakshys, *J. Am. Chem. Soc.*, 1969, **91**, 4801.
28 D. G. Barratt and J. J. Morris, unpublished observations.
29 D. A. Deranleau, *J. Am. Chem. Soc.*, 1969, **91**, 4044.
30 M. H. Abraham, P. P. Duce, R. A. Schulz, J. J. Morris, P. J. Taylor, and D. G. Barratt, *J. Chem. Soc., Faraday Trans. 1*, 1986, **82**, 3501.
31 M. H. Abraham, P. P. Duce, J. J. Morris, and P. J. Taylor, *J. Chem. Soc., Faraday Trans. 1*, 1987, **83**, 2867.
32 J.-L. M. Abboud, C. Roussel, E. Gentric, K. Sraidi, J. Lauransan, G. Guihéneuf, M. J. Kamlet, and R. W. Taft, *J. Org. Chem.*, 1988, **53**, 1545.
33 (*a*) P. Nicolet and C. Laurence, *J. Chem. Soc., Perkin Trans. 2*, 1986, 1071; (*b*) C. Laurence, P. Nicolet, and M. Helbert, *ibid.*, p. 1081.
34 (*a*) P. C. Maria, J.-F. Gal, J. De Franceshi, and E. Fargia, *J. Am. Chem. Soc.*, 1987, **109**, 483; (*b*) M. H. Abraham, P. L. Grellier, D. V. Prior, J. J. Morris, P. J. Taylor, P. C. Maria, and J.-F. Gal, *J. Phys. Org. Chem.*, 1989, **2**, 243; (*c*) J. Lauranson, *J. Chim. Phys.*, 1968, **65**, 1937.
35 E. L. Mehler and G. Eichele, *Biochemistry*, 1984, **23**, 3887.
36 A. Warshel and S. T. Russell, *Quart. Rev. Biophys.*, 1984, **17**, 283.
37 J. A. Riddick and W. B. Bunger, 'Organic Solvents,' vol. 2, in 'Techniques of Chemistry,' ed. A. Weissberger, 3rd edn., Wiley–Interscience, New York, 1970.
38 J. G. Kirkwood, *J. Chem. Phys.*, 1934, **2**, 351.
39 M. J. Kamlet, J. L. M. Abboud, M. H. Abraham, and R. W. Taft, *J. Org. Chem.*, 1983, **48**, 2877; M. J. Kamlet, J.-F. Gal, P. C. Maria, and R. W. Taft, *J. Chem. Soc., Perkin Trans. 2*, 1985, 1583.
40 L. Joris, J. Mitsky, and R. W. Taft, *J. Am. Chem. Soc.*, 1972, **94**, 3438.
41 R. N. Smith, C. Hansch, and M. M. Ames, *J. Pharm. Sci.*, 1975, **64**, 599.
42 R. W. Taft, M. Taagepera, J.-L. M. Abboud, J. F. Wolf, D. J. De Frees, W. J. Hehre, J. E. Bartmess, and R. T. McIver, Jr., *J. Am. Chem. Soc.*, 1978, **100**, 7765.
43 R. Wolfenden (*a*) *Biochemistry*, 1978, **17**, 201; (*b*) *Science*, 1983, **222**, 1087.
44 For aqueous solution data see, *e.g.*, E. M. Arnett, B. Chawla, L. Bell, M. Taagepera, W. J. Hehre, and R. W. Taft, *J. Am. Chem. Soc.*, 1977, **99**, 5729; R. W. Taft, M. H. Abraham, R. M. Doherty, and M. J. Kamlet, *J. Am. Chem. Soc.*, 1985, **107**, 3105. For gas-phase data see, *e.g.*, J. E. Bartmess, G. Caldwell, M. D. Rozeboom, and J. P. Kiplinger, *J. Am. Chem. Soc.*, 1984, **106**, 4660; P. Kebarle and G. Caldwell, *ibid.*, p. 967; P. Kebarle and Y. C. Lau, *Can. J. Chem.*, 1981, **59**, 151; W. R. Davidson, J. Sumner, and P. Kebarle, *J. Am. Chem. Soc.*, 1979, **101**, 1675.
45 See, *e.g.*, C. A. M. Hogben, L. A. Schanker, D. J. Tocco, and B. B. Brodie, *J. Pharmacol. Exp. Ther.*, 1957, **120**, 540; J. K. Seydel and K. J. Schaper, *Pharmac. Ther.*, 1982, **15**, 131; J. Gutknecht and A. Walter, *Biochem. Biophys. Acta*, 1981, **649**, 149; J. L. De Young, H. G. H. Tan, H. E. Huber, and M. A. Zoglio, *J. Pharm. Sci.*, 1978, **67**, 320.
46 Program contributed to Digilab Inc. by D. J. Moffatt and co-workers, Molecular Spectroscopy Section, National Research Council of Canada.
47 E. Suzuki, *Anal. Chem.*, 1969, **41**, 37; program adapted as for ref. 46.
48 See, *e.g.*, W. F. Maddams, *Appl. Spectroscp.*, 1980, **34**, 245.
49 P. J. Krueger and H. D. Mettee, *Can. J. Chem.*, 1964, **42**, 326.
50 J. Elguero, C. Marzin, A. R. Katritzky, and P. Linda, 'The Tautomerism of Heterocycles,' Academic Press, New York, 1976.
51 P. L. Huyskens, H. Marshal, and Th. Zeegers-Huyskens, *J. Mol. Struct.*, 1987, **158**, 379.
52 A. J. Kirby, 'The Anomeric Effect and Related Stereoelectronic Effects at Oxygen,' Springer–Verlag, Berlin, 1983.
53 J. Hine, S. Hahn, and D. E. Miles, *J. Org. Chem.*, 1986, **51**, 577.
54 P. Murray-Rust and J. P. Glusker, *J. Am. Chem. Soc.*, 1984, **106**, 1018.
55 (*a*) A. C. Legon and D. J. Millen, *Acc. Chem. Res.*, 1987, **20**, 39; (*b*) J. W. Bevan, A. C. Legon, D. J. Millen, and S. C. Rogers, *Proc. R. Soc. London, Ser. A*, 1980, **370**, 239.

56 T. W. Panunto, Z. Urbáncyak-Lipíowska, R. Johnson, and M. C. Etter, *J. Am. Chem. Soc.*, 1987, **109**, 7786.
57 M. H. Abraham, P. L. Grellier, D. V. Prior, P. P. Duce, J. J. Morris, and P. J. Taylor, *J. Chem. Soc., Perkin Trans. 2*, 1989, 699.
58 D. D. Nelson, Jr., G. T. Fraser, and W. Klemperer, *Science*, 1987, 238, 1670.
59 H. Kleeberg, D. Klein, and W. A. P. Luck, *J. Phys. Chem.*, 1987, **91**, 3200.
60 D. E. Leahy, P. J. Taylor, and A. R. Wait, *Quant. Struct. Act. Relat.*, 1989, **8**, 17.
61 S. E. Odinokov, A. A. Mashkowsky, and A. A. Nabiullin, *Spectrochim. Acta, Part A*, 1983, **39**, 1065.
62 R. G. Bates, in 'Hydrogen-Bonded Solvent Systems,' ed. A. K. Covington and P. Jones, Taylor & Francis, London, 1968, p. 49.
63 N. Stahl and W. P. Jencks, *J. Am. Chem. Soc.*, 1986, **108**, 4196.
64 R. S. Drago, *Struct. Bonding (Berlin)*, 1973, **15**, 72; *Coord. Chem. Rev.*, 1980, **33**, 251.
65 M. J. Kamlet, C. Dickinson, T. Gramstad, and R. W. Taft, *J. Org. Chem.*, 1982, **47**, 4971.
66 R. M. Badger and S. H. Bauer, *J. Chem. Phys.*, 1939, **5**, 839.
67 C. N. R. Rao, P. C. Dwivedi, H. Ratajczak, and W. J. Orville-Thomas, *J. Chem. Soc., Faraday Trans. 2*, 1975, **71**, 955.
68 T. D. Epley and R. S. Drago, *J. Am. Chem. Soc.*, 1967, **89**, 5770; R. S. Drago and T. D. Epley, *J. Am. Chem. Soc.*, 1969, **91**, 2883.
69 D. G. Brown, R. S. Drago, and T. F. Bolles, *J. Am. Chem. Soc.*, 1968, **90**, 5706.
70 R. Thijs and Th. Zeegers-Huyskens, *Spectrochim. Acta, Part A*, 1984, **40**, 307.
71 M. J. Kamlet, A. Solomonovici, and R. W. Taft, *J. Am. Chem. Soc.*, 1979, **101**, 3734.
72 L. J. Bellamy and R. J. Pace, *Spectrochim. Acta, Part A*, 1971, **27**, 705.
73 M. H. Abraham, P. P. Duce, D. V. Prior, R. A. Schulz, J. J. Morris, and P. J. Taylor, *J. Chem. Soc., Faraday Trans. 1*, 1988, **84**, 864.
74 H. B. Bürgi and J. D. Dunitz, *J. Am. Chem. Soc.*, 1987, **109**, 2924.
75 X. Q. Lewell, I. H. Hillier, M. J. Field, J. J. Morris, and P. J. Taylor, *J. Chem. Soc., Faraday Trans. 2*, 1988, **84**, 893.
76 O. Exner, in 'The Chemistry of Double-Bond Functional Groups,' Supplement A, Part 1, ed. S. Patai, John Wiley, New York, 1977, p. 1.
77 R. L. Beddoes, L. Dalton, J. A. Joule, O. S. Mills, J. D. Street, and C. I. F. Watt, *J. Chem. Soc., Perkin Trans. 2*, 1986, 787.
78 L. Eberson, in 'The Chemistry of Carboxylic Acids and Esters,' ed. S. Patai, Interscience, New York, 1969, p. 211.
79 Z. Berkovitch-Yellin, S. Ariel, and L. Leiserowitz, *J. Am. Chem. Soc.*, 1983, **105**, 765.
80 S. Sorriso, in 'The Chemistry of Functional Groups,' Suppl. F, Part 1, ed. S. Patai, John Wiley, New York, 1982, p. 1.
81 R. W. Taft and R. D. Topson, *Prog. Phys. Org. Chem.*, 1987, **16**, 1.
82 R. S. Moldey and R. G. Kallen, *J. Am. Chem. Soc.*, 1972, **94**, 6739.
83 F. Bernardi, H. B. Schlegel, M.-H. Whangbo, and S. Wolfe, *J. Am. Chem. Soc.*, 1977, **99**, 5633.
84 J.-L. M. Abboud, K. Sraidi, G. Guihéneuf, A. Negro, M. J. Kamlet, and R. W. Taft, *J. Org. Chem.*, 1985, **50**, 2870.
85 J. De Taeye, G. Maes, and Th. Zeegers-Huyskens, *Bull. Soc. Chim. Belg.*, 1983, **92**, 917.
86 C. Laurence, personal communication.
87 J. D. Aubort and R. F. Hudson, *Chem. Commun.*, 1970, 937.
88 R. W. Taft, F. Anvin, M. Tangepera, J. Catalan, and J. Elguero, *J. Am. Chem. Soc.*, 1986, **108**, 3237.
89 A. Vedani and J. D. Dunitz, *J. Am. Chem. Soc.*, 1985, **107**, 7653.
90 R. Taylor, O. Kennard, and W. Versichel, *J. Am. Chem. Soc.*, 1983, **105**, 5761.
91 R. Taylor and O. Kennard, *Acc. Chem. Res.*, 1984, **17**, 320.
92 F. Franks and D. J. G. Ives, *Q. Rev. Chem. Soc.*, 1966, **20**, 1.
93 H. Umeyama and K. Morokuma, *J. Am. Chem. Soc.*, 1977, **99**, 1316.
94 R. H. Davies, R. C. Mason, D. A. Smith, D. J. McNeillie, and R. James, *Int. J. Quantum Chem., Quantum Biol. Symp.* 5, 1978, 221.
95 E. M. Arnett and G. Scorrano, *Adv. Phys. Org. Chem.*, 1976, **13**, 83.
96 C. D. Johnson and B. Stratton, *J. Org. Chem.*, 1987, **52**, 4798.
97 M. Charton, *Prog. Phys. Org. Chem.*, 1981, **13**, 119.
98 J. V. Greenhill, M. J. Ismail, P. N. Edwards, and P. J. Taylor, *J. Chem. Soc., Perkin Trans. 2*, 1985, 1255.
99 See, *e.g.* T. C. Bruice, A. Donzel, R. W. Hufmann, and A. R. Butler, *J. Am. Chem. Soc.*, 1967, **89**, 2106.
100 J. A. Zoltewicz and L. W. Deady, *J. Am. Chem. Soc.*, 1972, **94**, 2765; P. M. Bond, E. A. Castro, and R. B. Moodie, *J. Chem. Soc., Perkin Trans. 2*, 1976, 68.
101 R. W. Taft, J. F. Gal, S. Geribaldi, and P. C. Maria, *J. Am. Chem. Soc.*, 1986, **108**, 861.
102 S. V. Hill, A. Williams, and J. L. Longridge, *J. Chem. Soc., Perkin Trans. 2*, 1984, 1009.
103 R. G. Button, J. P. Cairns, and P. J. Taylor, *J. Chem. Soc., Perkin Trans. 2*, 1985, 1555; P. J. Taylor and A. R. Wait, *ibid.*, 1986, 1765.
104 I. H. Hillier, unpublished observations
105 C. Laureys and Th. Zeegers-Huyskens, *J. Mol. Struct.*, 1987, **158**, 301.
106 C. R. Ganellin, *Annu. Rep. Med. Chem.*, 1979, **14**, 91.
107 R. B. Homer and C. D. Johnson, in 'The Chemistry of Amides,' ed. J. Zabicky, Interscience, London, 1970, p. 187.
108 T. Birchall and R. J. Gillespie, *Can. J. Chem.*, 1963, **41**, 2642.
109 D. S. Wofford, D. M. Forkey, and J. G. Russell, *J. Org. Chem.*, 1982, **47**, 5132.
110 F. Giordano and A. Zagari, *J. Chem. Soc., Perkin Trans. 2*, 1978, 312.
111 D. Smith and P. J. Taylor, *Spectrochim. Acta, Part A*, 1976, **32**, 1489.
112 J. V. Greenhill, M. J. Ismail, G. R. Bedford, P. N. Edwards, and P. J. Taylor, *J. Chem. Soc., Perkin Trans. 2*, 1985, 1265.
113 T. Gramstad, *Spectrochim. Acta*, 1965, **21**, 741.
114 J. March, 'Advanced Organic Chemistry,' John Wiley, New York, 1985, p. 418.
115 R. H. Davies, R. D. Bagnall, B. Crooks, and W. G. M. Jones, *Int. J. Quantum. Chem., Quantum Biol. Symp.* 3, 1986, 53.
116 T. Fujita, T. Nishioka, and M. Nakajima, *J. Med. Chem.*, 1977, **20**, 1071.
117 A. R. Fersht, J. P. Shi, J. Knill-Jones, D. M. Lowe, A. J. Wilkinson, D. M. Blow, P. Brick, P. Carter, M. M. Y. Waye, and G. Winter, *Nature*, 1985, **314**, 235.
118 See, *e.g.* ref. 4, pp. 393–394 and ch. 8 *passim*.
119 M. I. Page, in 'Quantitative Approaches to Drug Design,' ed. J. C. Dearden, Elsevier, Amsterdam, 1983, p. 109.
120 J. J. Morris, L. R. Huges, and P. J. Taylor, in 'QSAR in Drug Design and Toxicology,' ed. D. Hadži and B. Jerman-Blažič, Elsevier, Amsterdam, 1987, p. 204.

Received 20*th June* 1988; *Paper* 8/02460J