Using Theoretical Descriptors in Quantitative Structure–Property Relationships: 3-Carboxybenzisoxazole Decarboxylation Kinetics

George R. Famini[a] **and Leland Y. Wilson***[,b]
[a] *US Army Edgewood Research, Development and Engineering Center, Aberdeen Proving Ground, MD 21010, USA*
[b] *Chemistry Department, La Sierra University, Riverside, CA 92515, USA*

The decarboxylation rate of 3-carboxybenzisoxazoles depends on the substituent and strongly on the solvent. Solvent dependence explanation has been based on solvent donor hydrogen bond acidity, ion pairing and dispersive interactions. Recently, to help sort the solvent effects, the empirical, Kamlet–Taft–Abraham solvatochromic parameter set has been used as a probe in a multilinear correlation analysis based on the linear solvation energy relationship (LSER) model. This paper extends the LSER work with a theoretical, computational set of molecular parameters to analyse the rate data. Good correlations and physical interpretations result and agree with empirical observations and LSER correlations. These imply a rate decrease with solvent donor hydrogen bond acidity (HBA) and increase with acceptor hydrogen bond basicity (HBB). Mechanistic interpretation can involve the way charge delocalization on the carboxylate is effected by solvent HBA interaction and by ion pairing which is, in turn, effected by solvent HBB.

Inherent to chemistry is the idea that there is a relation between microscopic (molecular structural) features and a macroscopic (empirical) property of a compound. Quantitative structure activity (property) relationships (QSAR, QSPR) go a step further in assuming this relationship can be quantified. Once such a QSAR (QSPR) is found for a particular property, the molecular structure of any related compounds can be used to predict that property.[1] The relation generally involves a correlation equation between the empirical property and molecular structural parameters; this requires an empirical data set that provides an adequate statistical sample to cover the types of compounds of interest. These QSAR (QSPR) have been used to correlate molecular structural features of compounds with their known biological, chemical and physical properties. One such equation is based on the linear free energy relationship (LFER). In 1935 Burkhardt[2] and Hammett[3] reviewed the existence of LFERs in 1935; in 1937 Hammett[4] proposed the equation that bears his name. Exner provides a recent (1988) survey of LFER and a clear discussion of the background for its use.[5]

Linear Solvation Energy Relationships

Many descriptors have been used by researchers to increase the ability to correlate biological, chemical and physical properties. A very successful set has been used in the correlations of Hansch[6] and Kamlet, Taft, Abraham and co-workers[7] who extended earlier LFER work[8] to involve solvent–solute interactions.[9] A model for a linear solvation energy relationship (LSER) is given by eqn. (1).

$$\begin{aligned}\text{property} = {} & \text{bulk/cavity terms(s)} + \\ & \text{dipolarity/polarizability term(s)} + \\ & \text{H bonding term(s)} + \text{constant} \quad (1)\end{aligned}$$

The property is often the logarithm of an empirical parameter which, in turn, can be related to a free energy (often through a rate or equilibrium constant) consistent with the LFER concept. The bulk and cavity terms model the energy to form a solute molecule sized cavity and separate the solvent molecules (endoergic); in addition, these may involve size effects such as those influencing solvent separation of reactants. The dipolarity/polarizability terms involve dipole–dipole, dipole–induced dipole and induced dipole–induced dipole (dispersive) interactions (exoergic). The H bonding terms involve both acceptor hydrogen bond basicity (HBB) and donor hydrogen bond acidity (HBA) (exoergic) and actually represent strong dipole–dipole interactions.

Kamlet, Taft, Abraham and co-workers have used the empirically based solvatochromic (LSER) descriptor set for the terms in eqn. (1) which, then, can take the form of eqn. (2). Here SSP

$$\log \text{SSP} = a\delta_{\text{H1}}{}^{2}V_{12} + b\pi^{*}{}_{1}\pi^{*}{}_{2} + c\beta_{1}\alpha_{2} + d\alpha_{1}\beta_{2} + g \quad (2)$$

represents a property influenced by solute–solvent interactions. For solutes (subscript 2) the bulk term uses the intrinsic volume, V_1; the dipolarity/polarizability terms use the parameter, π^*, and the polarizability correction, δ. The hydrogen-bonding terms employ a donor HBA descriptor, α, and an acceptor HBB descriptor, β. Early work used the molar volume, V_m, while V_I is computed; more recently, the characteristic volume of McGowan, V_x, has been substituted for V_1 and the excess molar refraction, R, has replaced δ. For solvents (subscript 1) the Hildebrand solubility parameter[10] squared, $\delta_H{}^2$ $[=(\Delta H_{vap} - RT)/V_m]$, occurs in the bulk term instead of the volume. A referee pointed out that the solute and solvent parameters for a compound are not necessarily the same. For example, β_2 (solute) for ethanol is 0.44 in tetrachloromethane whereas β_1 (solvent) is 0.77 (quite different) while for ethoxyethane they are 0.45 and 0.47, respectively (almost the same). Often not all the terms in eqn. (1) are significant.

The solvatochromic (LSER) descriptors have very successfully correlated more than 250 biological, chemical and physical properties involving solute–solvent interactions for a large number of compounds.[11] The coefficients of the descriptors in the correlation equation can provide insight into the nature of the solute–solvent interactions. However, the empirical nature of the LSER descriptors somewhat limits their ability to make *a priori* predictions. There are tables of LSER parameters and predictive relations to help in their estimation LSER values. For example, Hickey and Passino-Reader have provided 'rules of thumb' for LSER parameter estimation.[12] However, they are not as easily found for complex molecules.

Table 1 TLSER descriptors

Symbol	Name	Definition[a]	Units	Meaning
V_{mc}	Molecular volume	Molecular volume	100 Å^3	Cavity/steric
π_{I}	Polarizability index	Polarizability/V_{mc}	None	Polarizability
ε_{B}	Covalent basicity	$0.30 - \lvert \Delta E(\mathrm{h,lw}) \rvert /100$	heV	Acceptor HBB
q_-	Electrostatic HBB	Maximum $\lvert(-)$ charge$\rvert$ on an atom	acu	Acceptor HBB
ε_{A}	Covalent HBA	$0.30 - \lvert \Delta(E(\mathrm{l,hw}) \rvert /100$	heV	Donor HBA
q_+	Electrostatic HBA	Maximum (+) charge on an H atom	acu	Donor HBB

[a] $\Delta E(\mathrm{h,lw}) = E(\mathrm{h}) - E(\mathrm{lw})$; $E(\mathrm{h})$ = HOMO energy; $E(\mathrm{l})$ = LUMO energy; $E(\mathrm{lw})$ and E(hw) refer to the E(LUMO) and E(HOMO) for water, respectively; | | indicate absolute magnitudes.

Very closely related to the LSER approach is a four parameter correlation recently reported by Gajewski[13] who indicates good success for solvent rate effects using the KOPMH (Kirkwood, Onsager, Parker, Marcus, Hildebrand) equation.[14,15] The strong similarity is apparent in the use of the Hildebrand solubility parameter, the Onsager function of the permittivity, which relates to the molar refraction, and two other parameters analogous to the LSER α and β.

Attempts to correlate computationally derived structural and electronic descriptors with the solvatochromic parameters have met with moderate degrees of success by Lewis.[16] Recently Politzer has related the molecular electrostatic potential to the LSER dipolarity/polarizability[17] and solute hydrogen bonding descriptors.[18]

Theoretical Linear Solvation Energy Relationships.—In the past, theoretical chemistry has been used to provide descriptors for QSAR. Representative work is described by Loew and co-workers,[19] Pedersen,[20] and Chastrette and colleagues[21] while Lewis[22] gives a more recent summary of molecular orbital calculations applied to QSAR (MO–QSAR) for a variety of activities. An example of a theoretical descriptor is the molecular transform of Kier and Hall.[23] Ford and Livingstone[24] point out advantages of computational derived descriptors over extra-thermodynamically derived descriptors such as π and σ. They describe clearly defined molecular properties which makes the interpretation of QSAR equations more straightforward; furthermore, their values are easily obtained.

Based on the LSER philosophy a new, theoretical set of parameters for correlating a wide variety of properties has been developed.[25] These theoretical linear solvation energy relationship (TLSER) descriptors are summarized in Table 1. Good correlations and physical interpretations have resulted from the use of the parameters for a variety of biological, chemical and physical properties. Examples include the following: five non-specific toxicities;[26] activities of some local anaesthetics and the molecular transform;[27] opiate receptor activity of some fentanyl-like compounds;[28] six physicochemical properties—charcoal absorption, HPLC retention index, octanol–water partition coefficient, phosphonothiolate hydrolysis rate constant, aqueous acid equilibrium constant, electronic absorption of some ylides;[29] gas phase acidity;[30] and solubility in supercritical CO_2.[31]

The TLSER bulk/steric term for a solute is modelled by the molecular van der Waals volume, V_{mc}, in units of 100 Å^3. The TLSER bulk term for a solvent, analogous to the Hildebrand solubility parameter, δ_{H}^2, has been more difficult to model. A possibility is to use, $1/V_{\mathrm{mc}}$, analogous to δ_{H}^2, (assuming ΔH_{vap} almost constant, most likely when molecules of similar polarity and size are involved). Of course, δ_{H}^2 may be used itself since values are relatively easily available for most solvents.

The dipolarity/polarizability term uses the polarizability index, π_{I}, obtained by dividing the polarizability volume by the molecular volume to produce a unitless, size independent quantity which indicates the ease with which the electron cloud may be moved or polarized. For example, aromatics and chlorine rank high while alkanes and fluorine rank low on the scale. Based on its definition this π_{I} models the ability for dipole–induced dipole and induced dipole–induced dipole (dispersive) interactions. No adequate TLSER model for dipolarity itself has been found; the molecular dipole moment has not been statistically significant in correlations. One implication is that the gas phase dipole moment does not adequately represent the dipole moment when in a condensed phase. Of course, it should be noted that hydrogen bonding does represent a special case of dipole–dipole interactions.

The acceptor HBB and the donor HBA are composed of covalent, ε_{B} and ε_{A}, and electrostatic, q_- and q_+, terms, respectively. The covalent HBB parameters, ε_{B}, is 30 minus the magnitude of the difference between the energy of the highest occupied molecular orbital (HOMO) of the solute and the lowest unoccupied molecular orbital (LUMO) of water, $\lvert E(\mathrm{LUMO}) - E(\mathrm{HOMO,water}) \rvert = \Delta E(\mathrm{l,hw})$. Since $\Delta E(\mathrm{l,hw})$ increases with decreasing basicity, the transformation, $30 - \Delta E(\mathrm{l,hw})$, gives a parameter that increases with increasing basicity. The 30 gives a transformed variable similar in size to the untransformed one. The result is divided by 100 for convenience in presentation and comparison of coefficients; the units are in hectoelectron volts (heV). Analogously, the covalent HBA parameter, ε_{A}, is 30 minus the magnitude of the difference between the energies of the LUMO of the solute and the HOMO of water, $\Delta E(\mathrm{h,lw})$, again scaled like the covalent HBB with the same units. The water energies are included for aesthetic reasons; the smaller these differences the greater is the ability to form a hydrogen bond with water. It is important to note the difference between the ε_{a} and ε_{b} which appears in other papers involving TLSER parameters, $\varepsilon_{\mathrm{B}} = 0.30 - \varepsilon_{\mathrm{b}}$, an analogous relation for the acidity. The electrostatic contribution to the HBB is the magnitude of the largest negative partial charge, q_-, on an atom; units are atomic charge units (acu). The corresponding HBA descriptor is the partial charge, q_+, on the most positively charged H atom (acu).

The TLSER application of eqn. (2) for solvent and solute can be modelled by eqn. (3). For a given property and set of

$$\log \mathrm{SSP} = a\delta_{\mathrm{H}1}{}^2 V_{\mathrm{mc}2} + b\pi_{\mathrm{I}1}\pi_{\mathrm{mc}2} + c\varepsilon_{\mathrm{B}1}\varepsilon_{\mathrm{A}2} + d\varepsilon_{\mathrm{A}1}\varepsilon_{\mathrm{B}2} + eq_{-1}q_{+2} + fq_{+1}q_{-2} + g \quad (3)$$

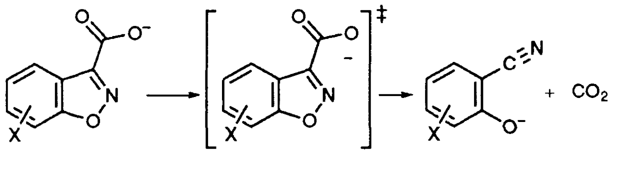
compounds, the coefficients (a–g) are determined using multilinear regression analysis to fit the data. Often not all the terms are significant. When eqn. (3) is applied to individual solutes in multiple solvents (hence, the subscript 1) it takes the form of eqn. (4). Here the coefficients contain the contribution

$$\log \mathrm{SSP} = a_2\delta_{\mathrm{H}1}{}^2 + b_2\pi_{\mathrm{I}1} + c_2\varepsilon_{\mathrm{B}1} + d_2q_{-1} + e_2\varepsilon_{\mathrm{A}1} + f_2q_{+1} + g \quad (4)$$

of the individual solute to the complementary solvent descriptors in eqn. (4). The coefficients in eqn. (4) should be related to those in eqn. (3). An analogous relation holds for the case of an individual solvent with multiple solutes. In contrast to the LSER parameters which may not be the same when referring to solvent and solvent, the TLSER parameters are taken to be the same whether the compound serves as solute or solvent. The one exception to this has been mentioned in connection with using V_{mc2} for solute and $\delta_{H1}{}^2$ for solvent. The definitions and the computational state of the art imply that the TLSER parameters apply to the gas phase.

One type of system not yet investigated with TLSER descriptors is that for rate constants for a single solute in multiple solvents. Recently, Grate, McGill and Hilvert[32] used the LSER solvatochromic parameters of Kamlet, Taft and Abraham as a chemical probe to study rate data for individual solutes in multiple solvents. This paper applies the TLSER parameters to this same set of data. Since multiple solutes and solvents were involved and their TLSER parameters are readily computed, the data makes it possible to look at multiple solute–single solvent, single solute–multiple solvent [eqn. (4)] and multiple solute–multiple solvent [eqn. (3)] systems.

Benzisoxazole Decarboxylation Kinetics.—In 1975 Kemp and Paul[33] reported an extensive study of solvent effects on the rate of decarboxylation of 3-carboxybenzisoxazoles to form 2-cyanophenolates in the presence of the base, tetramethylguanidine (TMG). In this environment these compounds are present primarily as carboxylate ($Bnzx^-$) and tetramethylguanidinium ($TMGH^+$) ions. The reaction can be represented by this mechanism.



Tables 2 and 3 contain the Kemp and Paul rate constant data and illustrate the strong dependence on solvent. For example, the rate for the 6-NO_2 compound is 10^8 faster in hexamethylphosphoramide (HMPA) than in water. Kemp and Paul proposed three primary influences to explain the solvent effects on the reaction rate: solvent donor HBA, ion pair interactions, and transition state stabilization through dispersive interactions.

Grate and colleagues (ref. 32) obtained eqn. (5) by correlating the Kemp and Paul data with the LSER parameters for 20 solvents for which parameters were available. They further removed aprotic (non-HBA), low polarity solvents to obtain eqn. (6). The statistical parameters indicate the very good fit of the data. (The statistical parameters are those commonly used except, perhaps, the VIF which is explained in the procedure section.) However, the most important aspect is the physical interpretation. These results suggest that the reaction rate decreases with solvent HBA (negative α_1 term) which is consistent with Kemp's analysis. It also suggests that the reaction rate increases with solvent HBB (positive β_1 term) and

$$\begin{array}{lllll} \log k = & -5.73\,\alpha_1 & +2.02\,\beta^1 & +0.36 & (6) \\ \pm & 0.36 & 0.70 & 0.48 & \end{array}$$

$$N = 13 \quad R = 0.983 \quad SD = 0.53$$

increases with solvent dipolarity/polarizability, $\pi_1{}^*$, and decreases with Hildebrand solubility parameter, $\delta_{H1}{}^2$.

The empirical rate data and the correlation equations can be understood at the molecular level with three main points. The first point, thoroughly discussed by Grate and colleagues, is to consider the various species and their equilibria in solution: free $Bnzx^-$, solvent–ion complexes, (solvent $\cdots$ $Bnzx^-$) and (solvent $\cdots$ $TMGH^+$); tight ion–ion complexes, ($Bnzx^-$ $\cdots$ $TMGH^+$); and loose ion–ion complexes, [(solvent $\cdots$ $Bnzx^-$) $\cdots$ (solvent $\cdots$ $TMGH^+$)] which amount to solvent separated ion pairs (SSIP). The existence of ion pairs, such as ($Bnzx^-$ $\cdots$ $TMGH^+$), comes from several sources. In low polarity solvents the ($Bnzx^-$ $\cdots$ $TMGH^+$) complex would be expected, *a priori*, to predominate over the (solvent $\cdots$ $Bnzx^-$) complex. An example of empirical evidence for ion pairs comes from Smid and coworkers[34,35] who found that crown ethers and their polymers catalyse the reaction in benzene when potassium ion is the counter ion. The reaction was about a thousand times faster than with $TMGH^+$ in benzene without the ether. Apparently the crown ether can form a complex with the K^+, thus freeing the $Bnzx^-$ from the $TMGH^+$. Grate and colleagues showed that ion pairs could be appreciable in other solvents. Using dissociation constants determined by Kolthoff and co-workers[36] for analogous ion pairs in acetonitrile, they estimated the equilibrium constants on the order of 10^{-4}. This, along with the very low stoichiometric concentrations used by Kemp, indicated that there is a significant concentration of [(solvent $\cdots$ $TMGH^+$) $\cdots$ (solvent $\cdots$ Bnzx)] complexes. In more polar solvents one would expect the concentration to be decreased.

The second point is that the free $Bnzx^-$ ion and its loose (solvent $\cdots$ $Bnzx^-$) complex seem to be the forms that undergo decarboxylation. An example of empirical evidence for this is given by Smid and coworkers who found, for the 6-NO_2 compound, the rates for the free (molecular) acid, BnzxH, in dioxane (non-polar) to be slow and even slower still in the presence of stronger acids. Under these conditions, the $Bnzx^-$ concentration is small; the stronger acid shifts the equilibrium towards the BnzxH. When the $Bnzx^-$ concentration is taken into account, the rate constant becomes large thus suggesting that the free $Bnzx^-$ reacts quite rapidly.

The third point is that solvent and substituent effects on the rates can be summarized by considering their influence on the delocalization of the charge on the $Bnzx^-$ ion. Delocalizing the charge over the anion can favour the release of the neutral CO_2 molecule and, thus, an increased reaction rate.

With these three ideas the rate data can be explained at the molecular level. The decreased rate in HBA solvent and the negative sign on the α_1 term can be explained as follows. The solvent HBA would lead to tight (solvent $\cdots$ $Bnzx^-$) complexes with the hydrogen bonds near the CO_2; consequently the charge could be localized in that region. The result would be a decreased rate. Further evidence for the effects of hydrogen bonding near the carboxylate comes from the empirical result that reaction rate for the 4-OH substituent compound is slower in all solvents.[37]

$$\begin{array}{llllllll} \log k = & +5.45\,\pi_1{}^* & -1.46\,\delta^1 & -3.03\,\alpha_1 & +1.80\,\beta_1 & -1.06\,\delta_{H1}{}^2 & -2.97 & (5) \\ \pm & 0.89 & 0.55 & 0.65 & 0.66 & 0.25 & 0.52 & \\ \text{t-stat} & 6.12 & 2.66 & 4.65 & 2.74 & 2.66 & 2.66 & \\ \text{P(2-tail)} & 0.000 & 0.019 & 0.000 & 0.016 & 0.001 & 0.001 & \\ \text{VIF} & 2.29 & 1.89 & 3.26 & 2.11 & 4.78 & & \end{array}$$

$$N = 20 \quad R = 0.976 \quad SD = 0.582 \quad F = 56.2$$

Table 2 Log of rate constants, *k*, for decarboxylation of y-x-3-carboxybenzisoxazoles in 14 solvents;[a] (pseudo first-order rate constant, units = s^{-1}; T = 30 °C, base = tetramethylquanidine)

	Solute						
Solvent	6-NH_2	6-H	6-CH_3O	6-Cl	6-NO_2	5-NO_2	5,6-$(NO_2)_2$
Water	−5.700	−5.975	−5.611	−5.357	−5.134	−4.025	−2.991
Methanol	−4.959	−4.770	−4.553	−4.000	−3.602	−2.410	−1.000
Ethanol	−4.602	−3.398	−4.102	−3.509	−3.000	−1.796	−0.456
Formamide	−4.377	−4.097	−3.959	−3.456	−3.131	−1.796	
N-Methylformamide	−3.745	−3.456	−3.149	−2.553	−2.092	−0.770	
Nitromethane	−1.854	−1.509	−1.252	−0.620	−0.237	1.000	2.699
Acetonitrile	−1.071	−0.921	−0.638	0.079	0.462	1.800	
Dimethylsulfoxide	−0.700	−0.398	0.000	0.612	1.000		
Propanone	−0.398	−0.155	0.255	0.944	1.380		
N-Dimethylformamide (DMF)	−0.091	0.176	0.672	1.225	1.568		
Tetramethylene sulfone	0.079	0.477	0.653	1.431	1.806		
Dimethylacetamide (DMA)	0.279	0.544	0.903	1.845	2.204		
N-Methylpyrrolidine	0.602	1.041	1.255	2.000	2.398		
Hexamethylphosphoramide (HMPA)	0.820	1.255	1.447	2.342	2.845		

[a] From Kemp and Paul, ref. 34.

Table 3 TLSER parameters[a] for solvents, solutes and residuals for solvents

Solvent[e]	V_{mc}	π_I	ε_B	q_-	ε_A	q_+	$\delta_H{}^2$	log *k*[b]	diff.[c]
Water	0.1933	0.0581	0.1237	0.3255	0.1237	0.1628	2.2970	−5.134	0.000
Methanol	0.3647	0.0860	0.1314	0.3291	0.1402	0.1803	0.8586	−3.602	0.152
Formamide	0.4090	0.0965	0.1371	0.3553	0.1667	0.1572	1.5134	−3.131	−0.068
Ethanol	0.5423	0.0927	0.1326	0.3235	0.1429	0.1800	0.6782	−3.000	−0.105
N-Methylformamide	0.5790	0.1006	0.1381	0.3437	0.1682	0.1557	0.6150	−2.092	0.027
Nitromethane	0.4740	0.1093	0.1302	0.3342	0.1817	0.0500	0.6632	−0.237	0.405
Acetonitrile	0.4529	0.0937	0.1177	0.1145	0.1622	0.0209	0.5766	0.462	−0.221
Dimethylsulfoxide	0.7209	0.1046	0.1475	0.7196	0.1734	0.0525	0.7063	1.000	0.230
Propanone	0.6441	0.0972	0.1381	0.2867	0.1715	0.0232	0.3791	1.380	0.362
DMF	0.7693	0.1042	0.1441	0.4698	0.1649	0.0576	0.5812	1.568	−0.583
(Methylene)$_4$sulfoxide	1.0129	0.1192	0.1368	0.6788	0.1942	0.0814	0.7489	1.806	−2.104
DMA	0.9634	0.1026	0.1452	0.4656	0.1658	0.0285	0.4879	2.204	−0.279
N-Methylpyrrolidine	1.0311	0.1053	0.1516	0.4541	0.1495	0.0169	0.5339	2.398	0.250
HMPA	1.8436	0.1107	0.1456	0.6518	0.1874	0.0096	0.3071	2.845	−0.098
Trichloromethane[f]	0.7540	0.1114	0.1165	0.1130	0.1849	0.0876	0.3711	−3.090	1.656
Tetrachloromethane	0.9058	0.1172	0.1132	0.0704	0.1912	0.0000	0.3088	−2.820	3.596
Benzene	0.8463	0.1204	0.1517	0.0594	0.1744	0.0593	0.3506	−2.320	4.068
Dimethoxymethane	0.8076	0.0978	0.1369	0.3304	0.1474	0.0190	0.5905	−1.440	3.198
Dioxane	0.8598	0.1045	0.1379	0.3275	0.1480	0.0364	0.4184	−1.390	2.742
Dichloromethane	0.6045	0.1036	0.1207	0.1605	0.1773	0.0555	0.4088	−1.330	1.033
Ethoxyethane	0.9035	0.0996	0.1365	0.3423	0.1455	0.0072	0.2351	−1.050	3.318
Benzonitrile	0.9984	0.1274	0.1474	0.0865	0.1833	0.0699	0.5142	0.400	0.535
THF	0.7889	0.1021	0.1378	0.3270	0.1471	0.0217	0.3615	0.600	1.190
$(CH_3OCH_2CH_3)_2O$	1.4171	0.1024	0.1157	0.3572	0.1502	0.0128	0.4000[d]	0.700	1.151
y-z-Benzisoxazoles, anion									
6-NH_2	1.3385	0.1356	0.1982	0.5807	0.1531	0.0971			
6-H	1.2249	0.1328	0.1980	0.5709	0.1518	0.0784			
6-CH_3O	1.4657	0.1329	0.1978	0.5741	0.1529	0.0840			
6-Cl	1.3791	0.1348	0.1959	0.5688	0.1565	0.0885			
6-NO_2	1.3795	0.1430	0.1935	0.5683	0.1648	0.0920			
5-NO_2	1.3888	0.1406	0.1940	0.5740	0.1628	0.1129			
5,6-$(NO_2)_2$	1.5638	0.1460	0.1900	0.5635	0.1715	0.1172			

[a] $\delta_H{}^2$ units are J m^{-3}, other units in Table 2. [b] For the 6-NO_2 compound. [c] diff. = (log *k* calculated from equation for 6-NO_2 in Table 5) − (empirical log *k*). [d] Estimated. [e] These 14 solvents were used for majority of solutes, see Table 6. [f] These 10 solvents in addition to the previous 14 were used for the 6-NO_2 compound.

The increased rate in HBB solvents, based on a positive β_1 term, can be explained as follows. The solvent HBB can lead to tight (solvent···$TMGH^+$) complexes through hydrogen bonding. Grate had pointed out the HBA nature of $TMGH^+$. This would diminish the ($Bnzx^-$ ···$TMGH^+$) complexes and result in more free $Bnzx^-$ and loose (solvent···$Bnzx^-$) ions and an increased rate. This is consistent with the Kemp's conclusion that ion pairs can be involved.

The increased rate in non-HBA (aprotic) solvents relative to water and the catalysis by extraction from water into these solvents can be explained as follows. Many of these solvents have acceptor HBB so that the explanation in those cases is given in the previous paragraph. However, Kemp assumed that the primary solvent–$Bnzx^-$ interaction in these solvents to be dispersive. The highly polarizable $Bnzx^-$ can have dispersive interactions with polarizable solvents. Conclusions regarding dispersive interactions must come from data for non-polar, non-HBA solvents. Most of the non-HBA solvents with increased rate relative to water were also polar and their rate increased with polarity. This is contrary to expectation for ion-dipole interactions which would favour charge localization and, thus, decreased rate.

Table 4 Correlations for solvents with anionic y-x-3-carboxybenzisoxazoles

$\log k = a_1V_{mc2} + b_1\pi_{12} + c_1\varepsilon_{B2} + d_1q_{-2} + e_1\varepsilon_{A2} + f_1q_{+2} + g$; $N = 7$; VIF = 1.05

Coeff. ± t-stat. *P*(2-tail)	a_1	b_1	c_1	d_1	e_1	f_1	g	*R*	SD	*F*
Water	n/s	n/s	n/s	−73.33 14.89 4.92 0.008	n/s	27.58 5.62 11.1 0.000 75.49	30.99 8.64 3.58 0.023	0.989	0.195	89.6
Methanol[a]	n/s	n/s	n/s	−128.68 19.42 6.63 0.003	n/s	7.33 10.3 0.001 n/s	62.70 11.27 5.56 0.005	0.990	0.254	94.7
Ethanol	n/s	n/s	n/s	n/s	n/s	81.37	n/s	n/s	n/s	n/s
NO_2CH_3[a]	n/s	n/s	n/s	−153.64 22.45 6.84 0.002	n/s	8.47 9.61 0.001	79.76 13.03 6.12 0.004	0.989	0.293	88.4

[a] Outlier is 6-NH_2 compound. n/s, not significant at 0.95 level.

The positive π_1* term can be explained as follows. This parameter is related to the ability of the solvent to undergo dipole and induced dipole (dispersive) interactions. Since eqn. (6), which applies to polar solvents (the low polarity ones have been removed) does not have π_1*, Grate and colleagues associated this solvent dipolarity/polarizability parameter with the ion pairs expected in low polarity solvents. Solvent dipolarity/polarizability might influence the separation of ion pairs to form (solvent ··· Bnzx$^-$) complexes more than it would localize the charge through dipole interaction. When polarity is low the induced dipole (dispersive) interactions could be significant and result in delocalized charge on the (solvent ··· Bnzx$^-$) complex. This could be viewed as favouring the move into the transition state and, thus, relate to Kemp's proposal that dispersion interaction helps stabilize the transition state. Practically concurrent with the work of Grate and colleagues, a study on the carbon kinetic isotope effects on the 5-NO_2 compound gave strong evidence that the transition state structure changes very little with changes in solvent.[38] This implied that the transition state structure might not depend strongly on an interaction with the solvent particularly with regard to solvation; consequently, it would suggest that the proposed role of transition stabilization through dispersive interaction might not be important.

The negative sign on the $\delta_{H1}{}^2$ term can be explained as follows (largely due to Grate). Increased solvent cohesion would make it more difficult for ion pairs to separate; ion pair dissociation would involve more surface area contact between ions and solvent. Hence, maintaining ion pairs would favour a decreased rate.

The role of the Bnzx$^-$ substituents can be explained as follows. For example, the 5,6-$(NO_2)_2$ compound increases the polarizability of the Bnzx$^-$ and leads to more charge delocalization and an increased rate. This agrees with experiment. NH_2, less electron donating, gives a decreased rate.

Procedure

Rate constant data was taken from the work of Kemp and Paul. Structure entry and visualization was done with the in-house developed molecular modelling package, MMADS[39] and PCMODEL (Serena Software, Bloomington, IN, USA). Molecular geometries were optimized and TLSER descriptors were calculated using the MNDO algorithm contained in MOPAC v6.0.[40,41] The q_- and q_+ values were taken from the Mulliken population analysis (in MOPAC). The molecular volume for the optimized geometry was determined using the algorithm of Hopfinger.[42] Multilinear regression analysis [using MYSTAT (Systat, Evanston, Il, USA)] was used to obtain the coefficients in the correlation equation.

The correlation equations were selected based on the coefficients being significant at the 0.95 level ('large' t-statistic) or higher, the correlation coefficient, *R*, as large as possible, small cross correlation [small variance inflation factor (VIF)] and a minimum number of outliers. The VIF is defined as $1/(1-R^2)$ where *R* is the correlation coefficient of one variable against the others; small (closer to one) values imply small cross correlation.[43] Outliers were taken as compounds whose calculated values were three or more standard deviations from the mean.

Results

The analysis of the data essentially follows the organization of Kemp and Paul while their data referred to 30 °C and Hildebrand solubility parameter data to 25 °C, the 5 °C difference represents less than 2% error based on the free energy relation, $\ln k = -E_a/RT +$ constant, model. Table 2 contains their data for seven solutes and 14 solvents while Table 3 contains their data for the 6-NO_2 compound over 24 solvents (which include the 14 in Table 2) as well as the TLSER parameters for the compounds used in this work. The 14 member set solvents are quite polar while the 10 member set solvents have low polarity. The diff(erence) column in Table 3 shows that the calculated values for the 10 solute set are all higher than experiment. Tetramethylene sulfone, an outlier in the set of 14, is quite different in that its calculated value is much less than experiment.

Tables 4–6 summarize the correlation equations; each coefficient is accompanied by its standard error (±), t-statistic (t-stat.), 2-tail probability *P*(2-tail) and the variance inflation factor (VIF) to indicate the quality of the 'fit'. There is some redundancy since small standard error accompanies a large t-statistic and small probability. To help judge its quality, each equation is accompanied by its correlation coefficient, *R*, standard deviation(error), SD and its Fisher index of statistical significance, *F*.

Because of the systematic arrangement of the data in Tables 2 and 3, it is conceivable to do three types of analysis. (*i*) A single solvent ··· multiple Bnzx$^-$ study across rows in Table 2. Four solvents have seven solutes, three solvents have six solutes while the remaining seven have only five solutes. There are two problems; there are not enough data points to give reliable results and the full set of Bnzx$^-$ descriptors are strongly cross

Table 5 Correlations for 6-x-3-carboxybenzisoxazoles with solvents, δ_{H1}^2, ε_{B1} and q_{+1}[a]

$\log k = a_2\delta_{H1}^2 + b_2\pi_{11} + c_2\varepsilon_{B1} + d_2q_{-1} + e_2\varepsilon_{A1} + f_2q_{+1} + g$; $N = 13$; VIFs[b]

Coeff. ± t-stat. P(2-tail)	a_2	b_2	c_2	d_2	e_2	f_2	g	R	SD	F
NH_2[c]	−0.617	n/s	57.22	n/s	n/s	−26.16	−7.144	0.993	0.320	214
	0.224		10.58			1.68	1.532			
	2.75		5.41			15.59	4.66			
	0.022		0.000			0.000	0.001			
H	−0.831	n/s	64.49	n/s	n/s	−26.00	−7.746	0.994	0.300	269
	0.210		9.92			1.57	1.437			
	3.95		6.50			16.52	5.10			
	0.003		0.000			0.000	0.001			
CH_3O[c]	−0.819	n/s	64.84	n/s	n/s	−26.32	−7.474	0.993	0.336	218
	0.235		11.11			1.76	1.609			
	3.48		5.84			14.94	4.64			
	0.007		0.000			0.000	0.001			
Cl	−1.009	n/s	66.60	n/s	n/s	−27.11	−6.802	0.994	0.342	235
	0.239		11.30			1.79	1.637			
	4.21		5.89			15.12	4.19			
	0.002		0.000			0.000	0.002			
NO_2	−1.161	n/s	66.34	n/s	n/s	−26.55	−6.362	0.994	0.322	268
	0.237		10.67			1.69	1.546			
	5.14		6.22			15.68	4.12			
	0.003		0.000			0.000	0.003			

[a] Tetramethylene sulfone, dropped from set. [b] VIF = 1.73, 1.27 and 1.62, respectively. [c] *N*,*N*-dimethylformamide an outlier, retained.

Table 6 Correlations for anionic y-x-3-carboxybenzisoxazoles and solvents

$\log k = aV_{mc2}\delta_{H1}^2 + b\pi_{11}\pi_{12} + c\varepsilon_{B1}\varepsilon_{A2} + d\varepsilon_{A1}\varepsilon_{B2} + eq_{-1}q_{+2} + fq_{+1}q_{-2} + g'$

Coeff. ± t-stat. P(2-tail) VIF	a	b	c	d	e	f	g	R	SD	F
$N = 81$	n/s	319.0	412.1	n/s	n/s	−40.89	−12.04	0.945	0.809	214
		61.1	70.1			2.82	1.364			
all solutes		5.22	5.88			14.49	8.83			
		0.000	0.000			0.000	0.000			
		1.84	1.41			1.41				
N = 70	n/s	341.1	356.0	n/s	n/s	−44.56	−11.16	0.969	0.628	340
		53.1	58.5			2.41	1.151			
6-x solutes		6.42	6.08			18.84	9.69			
		0.000	0.000			0.000	0.000			
		1.84	1.51			1.42				
$N = 81$[a]	n/s	199.9	276.0	n/s	29.92	−39.77	−8.837	0.901	1.033	92.9
		70.0	89.5		8.36	3.46	1.701			
all solutes		2.85	3.08		3.58	11.50	5.20			
and solvents		0.005	0.003		0.001	0.000	0.000			
		1.76	1.74		1.37	1.42				

[a] Benzene and 6-NO_2 compound outlier, kept.

correlated. However, smaller sets are not strongly correlated. Table 4 is presented to suggest possible correlations, only. The case for ethanol is interesting in that while q_{-2} and q_{+2} are the most significant, they do not reach the 0.95 significance level.

(*ii*) A multiple solvent • • • single Bnzx study based on eqn. (3) down columns in Tables 2 and 3. This was the type of study done in the LSER study of Grate and colleagues. This can be done for the 6-NO_2 compound over the 24 solvents in Table 3 as well for the five 6-x compounds (including the 6-NO_2) over the 14 solvents in Table 2. The correlation equations are found in Table 5. The 24 solvent case did not correlate well; R was 0.0857 with ε_{B1}, q_{-1} and q_{+1} significant [and close to it for ε_{B1} since P(2-tail) = 0.055]. The 14 solvent set correlated very well for the 6-x compounds but when the other 10 solvents are analysed, the best equation has $R = 0.871$ with ε_{B1}, q_{-1} and ε_{A1} significant with tetrahydrofuran (THF) an outlier. There is great cross correlation for the latter two. Removing the outlier improves R to 0.952. Interestingly, analysis of the 10 member subset with the LSER parameters also showed strong cross correlation. The 5-NO_2 compound has data for only 7 solvents so the correlation equation can be taken as suggestive only and its best fit is listed here in eqn. (7). Formamide was an outlier

$\log k =$	$-0.626\,\delta_H^2$	$-21.20\,q_{+1}$	$+3.14$	(7)
±	0.130	3.06	0.44	
t-stat	4.82	6.93	7.17	
P(2-tail)	0.009	0.002	0.002	
VIF	1.17	1.17		

$N = 7$ $R = 0.992$ SD = 0.455 $F = 65.8$

but was retained. The 5,6-NO_2 compound has only four solvents which precludes attaching any significance to a correlation.

Table 7 Residuals for log *k* for the $N = 81$ combined solute-sample set based on equation in Table 6

	log *k*				log *k*		
Combination	Obs.	Calc.	Resid.	Combination	Obs.	Calc.	Resid.
Water 6-NH_2	−5.700	−5.590	−0.110	DMF[a]	−0.091	0.195	−0.286
6-H	−5.975	−5.630	−0.345		0.176	0.060	0.116
6-MeO	−5.611	−5.599	−0.012		0.672	0.112	0.560
6-Cl	−5.357	−5.325	−0.032		1.255	0.428	0.827
6-NO_2	−5.134	−4.762	−0.372		1.568	1.171	0.397
Methanol[a]	−4.959	−4.310	−0.649	Tetramethylene sulfone[a]	0.079	−0.180	0.259
	−4.770	−4.372	−0.398		0.477	−0.316	0.793
	−4.553	−4.342	−0.211		0.653	−0.270	0.923
	−4.000	−4.032	0.032		1.431	0.048	1.383
	−3.602	−3.372	−0.230		1.806	0.806	1.000
Ethanol[a]	−4.602	−3.936	−0.666	DMA[a]	0.279	0.887	−0.608
	−4.398	−3.955	−0.443		0.544	0.740	−0.196
	−4.102	−4.028	−0.074		0.903	0.797	0.106
	−3.509	−3.878	0.369		1.845	1.107	0.738
	−3.000	−3.415	0.415		2.204	1.849	0.355
Formamide[a]	−4.377	−2.946	−1.431	Methylpyrrolidine[a]	0.602	1.683	−1.081
	−4.097	−3.031	−1.066		1.041	1.526	−0.485
	−3.959	−2.993	−0.966		1.255	1.587	−0.332
	−3.456	−2.674	−0.782		2.000	1.907	0.093
	−3.131	−1.969	−1.162		2.398	2.677	−0.279
Meformamide[a]	−3.745	−2.670	−1.075	HMPA[a]	0.820	1.712	−0.892
	−3.456	−2.756	−0.700		1.255	1.550	−0.295
	−3.149	−2.721	−0.428		1.447	1.609	−0.162
	−2.553	−2.398	−0.155		2.342	1.921	0.421
	−2.092	−1.679	−0.413		2.845	2.684	0.161
Nitromethane[a]	−1.854	−0.279	−1.575	Water 5-NO_2	−4.025	−4.941	0.916
	−1.509	−0.417	−1.092	5,6-$(NO_2)_2$	−2.991	−4.374	1.383
	−1.250	−0.369	−0.881	Methanol[b]	−2.410	−3.580	1.170
	−0.620	−0.075	−0.545		−1.000	−2.936	1.936
	−0.237	0.635	−0.872	Ethanol[b]	−1.796	−2.182	0.386
Acetonitrile[a]	−1.071	−1.053	−0.018		−0.456	−1.512	1.056
	−0.921	−1.183	0.262	Formamide 5-NO_2	−1.796	−3.190	1.394
	−0.638	−1.137	0.499	Meformamide 5-NO_2	−0.770	−1.896	1.126
	0.008	−0.878	0.886	Nitromethane[b]	1.000	0.446	0.554
	0.462	−0.251	0.713		2.699	1.070	1.629
Dimethylsulfoxide[a]	−0.700	0.549	−1.249	Acetonitrile 5-NO_2	1.800	−0.412	2.212
	−0.398	0.409	−0.807				
	0.000	0.463	−0.463				
	0.612	0.784	−0.172				
	1.000	1.539	−0.539				
Propanone[a]	−0.398	0.330	−0.728				
	−0.155	0.190	−0.345				
	0.255	0.244	0.011				
	0.944	0.539	0.405				
	1.380	1.243	0.137				

[a] Same sequence as the first water case. [b] Same sequence as the just preceding water case.

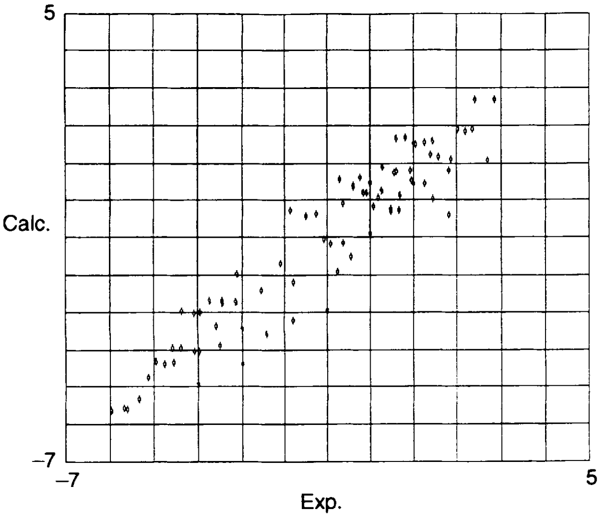


Fig. 1 Plot of log *k* (calc.) *vs.* log *k* (exp.) for $N = 81$ equation in Table 6

(*iii*) Combined solute and solvent study based on eqn. (3) over whole of Tables 2 and 3. Table 2 has 81 sample points and Table 3 provides 10 more. Table 6 contains the three equations corresponding to the rectangular array in Table 2 ($N = 70$), the irregular array in Table 2 ($N = 81$) and the array from combining Tables 2 and 3 ($N = 91$). The greater number of sample points can suggest statistically better interpretive and predictive qualities for these equations.

Finally, Table 7 contains the residuals for the 81 point case and Fig. 1 shows the corresponding plot of calculated *versus* empirical values for log *k*.

Discussion

The results show that not all of the descriptors are significant. Furthermore, the correlations are good to very good. Correlation coefficient values show that more than 80% of the variance is accounted for with $R > 0.945$ for most cases. Again, SD values indicate that most of the equations provide rate constant, *k*, values within a factor of 2 to 6 of the empirical values over a range as high as 10^8. As a fraction of the range for log k, the SD values correspond to $0.04 < \mathrm{SD/range} < 0.1$; that is, 4 to 10%. Most terms are significant at well over the 0.96 level, the worst cases are for the intercepts of solvent–multiple solute equations which have a very small set of data points, anyway. The VIF values (1.05 to 1.84) indicate that the parameter sets have low cross correlation. Examination of the parameters, their coefficients and t-scores makes it possible to infer something about the chemistry of this reaction.

Solute Sets.—The effect of solute structure on rate can be explained qualitatively. Examining any row (a given solvent) in Table 2 shows that the rate tends to increase going from the 6-NH_2 to the 5,6-$(NO_2)_2$ compound. Apparently, the nitro group (electron withdrawing) helps delocalize the charge over the whole ion through participation in the extensive pi resonance; there is a concomitant increase in polarizability. Charge delocalization favours the loss of the neutral CO_2. Conversely, the NH_2 group (electron donating) would help localize the charge favouring the anion and slowing the rate. Due to the small sample size, information in Table 4 can suggest some features, only. The positive sign on the q_{+2} term implies that the rate increases with increasing solute HBA parameter. In terms of eqn. (4), this can imply that the rate would increase with increasing solvent HBB, the complementary solvent parameter. The q_{+2} descriptor is most significant. The negative sign on the q_{-2} term implies that the rate decreases with increasing solute HBB; again, this can imply, eqn. (4), that increasing the complementary solvent parameter, HBA, accompanies a rate decrease.

Solvent Sets.—Table 5 shows two prominent features: (*i*) q_{+1}, the solvent HBA, is the most statistically significant parameter in each equation; (*ii*) the equations for each of the 6-x compounds are quite similar. As a rough guide there are trends across the coefficients with some increasing and others decreasing algebraically. The similarity of the equations implies a similar mechanism for each 6-x compound. The trends of the coefficients are consistent with the increase in rate from the 6-NH_2 to the 6-NO_2 compound. Eqn. (7) for the 5-NO_2 compound has two of the same parameters, $\delta_{H1}{}^2$ and q_{+1} (with the same signs), that the 6-x equations in Table 5 have. The small sample size prevents a conclusion regarding the difference in the mechanisms for the 5- and 6-x compounds.

The negative sign for q_{+1} term indicates that rate decreases with increasing solvent HBA which is consistent with empirical and LSER results. This is consistent with the idea that the solvent HBA interacts with the Bnzx^- to produce a complex with localized charge near the CO_2. It should be noted that this diminishes ion pairs. In keeping with eqn. (4) it suggests that the rate decreases with increase in solute HBB parameter and, indeed is consistent with the solute results. The positive sign on the ε_{B1} term suggests that the rate increases with increasing solvent HBB; again, this agrees with the LSER result and with the trends in the 10 set solvent set (low polarity) where HBB solvents tend to provide increased rates. This is consistent with the idea that the solvent HBB can produce a complex with the TMGH^+ which releases a free Bnzx^- in which the charge can be delocalized. This last term suggests, eqn. (4) that the rate may increase with increase in solute HBA and, again, is consistent with the solute results. The negative sign on the $\delta_{H1}{}^2$ term agrees with the LSER result, eqn. (5). The implication is that the rate would decrease with increasing solvent cohesion. As pointed out in the LSER discussion increased solvent cohesion could favour retaining ion pairs with a resulting decrease in rate. Pair dissociation would involve more solute surface area and thus more (solvent$\cdots$solvent) cohesion to overcome in order to solvate the ions. Increased solvent cohesion would be expected to accompany increased solvent dipolarity (HB included) and, to a lesser extent, polarizability (dispersive interactions). However, VIF values show that, despite this expected correlation, $\delta_{H1}{}^2$ is not cross-correlated strongly to the other parameters. It is to be noted that there is no direct evidence of a role for solvent polarizability.

The 6-NO_2 compound with 24 solvents case needs to be considered. As mentioned in the results section the full 24 solvent set does not give as good a correlation as the 14 solvent subset does. This differs from the LSER results in eqn. (5) where 20 solvent subset gave very good correlation. The separation by Kemp and Paul into subsets of 10 (low polarity) and 14 (polar) solvents is borne out quantitatively by the TLSER results. The diff. column in Table 3 shows results of the application of the correlation equation for the 6-NO_2 compound, 14 solvent set in Table 5 to all 24 solvents. The 10 solvent subset has all positive values indicating that the calculated values are too high; the 14 solvent subset shows much smaller and more random residuals. Tetramethylene sulfone is conspicuously different in that its calculated value is much lower than empirical one. The absence of q_{+1}, so prominent in the other subset, from the correlation equation seems to be due to its small variability across the set as seen in Table 3.

Solvent–Solute Combined Sets.—Table 6 contains equations all 81 points in Table 2 plus 10 more from Table 6; these larger sets are more significant statistically. For the $N = 81$ equations the 5-NO_2 and 5,6-$(NO_2)_2$ compounds contribute only 11 out of a possible 28 points. Kemp and Paul point out that the larger rate constants of these compounds coupled with equipment limitations prevented measurement for these solvent–solute combinations. The inclusion of these compounds lowered the quality of the statistical parameters but not the qualitative content. The $N = 91$ point equation includes the 10 solvent subset combined with the 6-NO_2 compound. The subset was extensively discussed earlier in the solvent section. Inclusion of these compounds lowered the quality of the fit and brought in one more term.

The physical content of the three equations is essentially the same. The positive sign on the π_{11} term, polarizability, shows that the rate increases with increasing solvent polarizability and is consistent with the LSER relation, eqn. (5). The positive sign on the solvent acceptor HBB, ε_{B1} and q_{-1}, also implies a rate increase while the negative sign on the solvent HBA, ε_{A1} and q_{+1}, implies a rate decrease. The HBB and HBA results are consistent with the LSER correlations and the HBA result is consistent with experiment. Some of the complementary solute parameters suggested by the solute and solvent equations appear here with the proper signs.

Summary for Interpretation in Terms of the Reaction Mechanism.—The correlation equations in Table 6 can be examined in terms of the three proposed mechanisms for the solvent sensitivity. The results may be best summarized by considering the two main influences mentioned in the introduction. These are the localization of charge near the CO_2 in the Bnzx^- and the relative concentrations of the solvated ions and ion pairs.

Support for the role of the solvent donor HBA comes from the presence of the HBA term, q_{+1} and/or ε_{A1}, with a negative sign. This indicates that the rate constant would decrease with increasing solvent HBA which is consistent with experiment. This may be explained by the formation of the (solvent$\cdots\mathrm{Bnzx}^-$) complex by the interaction of solvent HBA and Bnzx^- HBB with the bonding site near the CO_2 on the Bnzx^-. The large localization would hinder the release of the neutral CO_2.

Support for the role of ion pairs comes from the positive signs on the solvent acceptor HBB terms, ε_{B1} and q_{-1}, and the solvent polarizability term, π_{11}. This indicates that the rate constant would increase with increasing HBB and polarizability. The solvent HBB role can be explained by an increase in the (solvent $\cdots$TMGH$^+$) complex concentration from the interaction of the solvent HBB and the TMGH$^+$ HBA. This would decrease the (Bnzx$\cdots$TMGH$^+$) and [(solvent$\cdots$Bnzx$^-$)$\cdots$(solvent$\cdots$TMGH$^+$)] complex concentrations and increase the (solvent$\cdots$Bnzx$^-$) complex concentration and, hence, be associated with a rate increase. Greater polarizability, often associated with increased size, could be associated with increased separation of the (solvent$\cdots$Bnzx$^-$) and (solvent$\cdots$TMGH$^+$) complexes so that the Bnzx$^-$ complex concentration is increased. This increased separation of complexes could decrease charge localization on the anion and be associated with increased rate. Support for the role of size comes by noting that very high correlations are obtained with no outliers in the 14 solvent-single solute cases, for each the five 6-x compounds, when V_{mc1} and V_{mc1}^2 terms (along with q_{+1}) are included. The equations suggest that volume up to a maximum of about 130 Å^3 is associated with increased rate.

The role of solute structure can be explained by noting that substituents in 5 and 6 positions that are electron withdrawing tend to delocalize the charge away from the CO_2 and facilitate loss of the neutral CO_2. Conversely, electron-donating groups tend to decrease the rate.

Transition state stabilization by dispersion is more difficult to support. First, the correlation equations cannot distinguish between the ground and transition states. In fact, the molecular calculations apply to the ground state. In the TLSER case the polarizability term, involving π_1, occurs only in the combined solute–solvent case. This is positive thus indicating a rate increase with increased solvent and solute polarizability. This could be explained by the interaction of the solvent polarizability with the highly polar (ionic) Bnzx$^-$. Increased solvent polarizability could facilitate charge delocalization through dispersive interactions with the very polarizable Bnzx$^-$. This dispersive interaction could enhance the delocalization of the charge over the Bnzx$^-$ and thus favour loss of the neutral CO_2. Another possibility is the greater polarizability of the solvent permits stronger dispersive interactions with the Bnzx$^-$ and TMGH$^+$. This would permit greater separation of the ion pairs and, thus, an increased rate. Consequently one can see that it is not necessary to invoke the concept of transition state stabilization through dispersive interactions.

TLSER Comments.—The theoretical model for solvent cohesive forces, $1/V_{mc1}$, was not significant. The Hildebrand solubility parameter provides a much better model. Using the reciprocal of the volume implies that the cohesive interactions between solvent molecules do not change much over the solvent sample set. That is an unrealistic assumption. In fact, the polarizability and HBA and HBB should contribute to these cohesive terms in some way. Perhaps a parameter involving these descriptors divided by the volume could work as a cohesion model.

Outliers.—The few outliers are listed in the Tables. Outliers can often be correlated ('explained') by looking for unique features. For Table 6 and $N = 9$ the point corresponding to benzene and the 6-NO_2 compound is an outlier with its calculated value being too high. The 6-NO_2 compound and benzene have the highest ε_{A2} and ε_{B1} values, respectively in their sets. The term involving these descriptors does make a positive contribution to log k. Table 5 shows that tetramethylene sulfone was dropped in that set, it had a very large standardized deviation (*ca.* 6). Its calculated value was much lower than the empirical one. In the case of the 6-NH_2 and 6-CH_3 equations, *N*,*N*-dimethylformamide was an outlier; but, its inclusion still gives very good correlations. The reason for this being an outlier is not apparent. Table 4 shows that the 6-NH_2 compound is an outlier for the methanol and nitromethane cases. Its inclusion still gave good correlations. Again, the reason for this being an outlier is not apparent.

Conclusions

The TLSER descriptors provide good to very good correlation equations for the rate constants for the decarboxylation of y-x-3-carboxybenzisoxazoles. The relative ease with which the descriptors can be obtained for solvent and solute made it possible to use a large number of sample points. More importantly, the results make physical sense. There is strong agreement with the LSER correlations and experiment. The good correlations and physical interpretations also support the fundamental LSER model of eqn. (1) and, more specifically, that of eqn. (2).

The TLSER results support empirical and LSER observations. The rate decreases with increasing solvent donor HBA. The rate increases with increasing solvent acceptor HBB. There is some evidence that rate increases with increasing dispersive intractions.

Molecular explanation for the solvent HBA effect can be based on the idea that solvent HBA helps form a (solvent$\cdots$Bnzx$^-$) complex in which the charge is localized near the CO_2 thus hindering the loss of the neutral molecule. Explanation for the solvent HBB effect can be based on the idea that the solvent HBB helps form a (solvent$\cdots$TMGH$^+$) complex and, thus, release relatively free and/or solvated Bnzx$^-$ ions from (TMGH$^+$ $\cdots$ Bnzx$^-$) ion pairs. The charge on these relatively free anions is delocalized and, thus, the neutral CO_2 can be more readily released. This supports a role for ion pairing. However, there is not strong support for the rate being increased through stabilization of the transition state by dispersive interaction with the solvent. A possible explanation is that solvent dispersive interactions with the Bnzx$^-$ could contribute to the separation of the ion pairs and/or delocalization of charge on the Bnzx$^-$.

The similarity of the equations for individual 6-x compounds over the solvents suggests a similar interaction of the solvents with each compound and essentially agree with the solvent descriptor significance of the combined solvent–solute set. The effect of the solute substituent on reaction rate can be associated with the ability of the charge to be distributed over the Bnzx$^-$ enhancing the release of the neutral molecule.

Also important was inclusion of the empirical, reasonably readily available Hildebrand solubility parameter as a solvent cohesive parameter with the TLSER descriptors. However, the applicability of a reciprocal term for the solvent volume to represent this cohesive effect was not significant in this case.

Acknowledgements

The authors gratefully acknowledge the significant contribution of Jay W. Grate of Pacific Northwest Laboratories toward the understanding of the benzisoxazole chemistry.

References

1 S. Gupta, *Chem. Rev.*, 1987, **87**, 1183.
2 G. N. Burkhardt, *Nature (London)*, 1935, **17**, 684.
3 L. P. Hammett, *Chem. Rev.*, 1935, **17**, 125.
4 L. P. Hammett, *J. Am. Chem. Soc.*, 1937, **59**, 125.
5 O. Exner, *Correlation Analysis of Chemical Data*, Plenum Press, New York, 1988, p. 25.

6 C. Hansch, *Acc. Chem. Res.*, 1969, **B2**, 232.

7 M. J. Kamlet, R. W. Taft and J.-L. M. Abboud, *J. Am. Chem. Soc.*, 1977, **91**, 8325.

8 I. A. Koppel and V. A. Palm, *Advances in Linear Free Energy Relationships*, ed. N. B. Chapman and J. Shorter, Plenum Press, London, 1972.

9 M. J. Kamlet, R. M. Doherty, M. H. Abraham and R. W. Taft, *Quant. Struct.-Act. Relat.*, 1988, **7**, 71.

10 J. H. Hildebrand, J. M. Prausnitz and R. L. Scott, *Regular and Related Solutions*; Van Nostrand-Reinhold, Princeton, 1970.

11 M. J. Kamlet, R. W. Taft, G. R. Famini and R. M. Doherty, *Acta Chem. Scand.*, 1987, **41**, 589.

12 J. P. Hickey and D. R. Passino-Reader, *Environ. Sci. Technol.*, 1991, **25**, 1753.

13 J. J. Gajewski, *J. Org. Chem.*, 1992, **57**, 5500.

14 J. G. Kirkwood, *J. Chem. Phys.*, 1934, **2**, 351; L. Onsager, *J. Am. Chem. Soc.*, 1936, **58**, 1486.

15 B. G. Cox, G. R. Hedwig, A. J. Parker and D. W. Watts, *Aust. J. Chem.*, 1974, **27**, 477; Y. Marcus, *Pure Appl. Chem.*, 1983, **55**, 977.

16 D. V. F. Lewis, *J. Comput. Chem.*, 1987, **8**, 1084.

17 T. Brinck, J. S. Murray and P. Politzer, *Mol. Phys.*, 1992, **76**, 609.

18 J. S. Murray and P. Politzer, *J. Chem. Res.*, 1992, **5**, 110.

19 G. H. Loew, M. Poulsen, E. Kirkjian, J. Ferrell, B. S. Sudhindra and M. Rebagliati, *Environ. Health Perspect.*, 1985, **61**, 69.

20 L. Pedersen, *Environ. Health. Perspect.*, 1985, **61**, 185.

21 M. Chastrette, M. Rajzmann, M. Chanon and K. F. Purcell, *J. Am. Chem. Soc.*, 1985, **107**, 1.

22 D. F. V. Lewis, in *Progress in Drug Metabolism*, ed. J. W. Bridges and L. F. Chasseaud, John Wiley, London, 1990, p. 205.

23 L. B. Kier and L. Hall, *Molecular Connectivity in Structure-Activity Analysis*, Research Studies Press, Letchworth, 1986.

24 M. G. Ford and D. J. Livingstone, *Quant. Struct.-Act. Relat.*, 1990, **9**, 107.

25 G. R. Famini, *Using Theoretical Descriptors in Quantitative Structure-Activity Relationships* V. CRDEC-TR-085, US Army Chemical, Research, Development and Engineering Center, Aberdeen Proving Ground, MD, 1989.

26 L. Y. Wilson and G. R. Famini, *J. Med. Chem.*, 1991, **34**, 1668.

27 G. R. Famini, R. J. Kassel, J. W. King and L. Y. Wilson, *Quant. Struct.-Act. Relat.*, 1991, **10**, 344.

28 G. R. Famini, W. P. Ashman, A. P. Mickiewicz and L. Y. Wilson, *Quant. Struct.-Act. Relat.*, 1992, **11**, 162.

29 G. R. Famini, C. E. Penski and L. Y. Wilson, *J. Phys. Org. Chem.*, 1992, **5**, 395.

30 G. R. Famini, B. C. Marquez and L. Y. Wilson, *J. Chem. Soc., Perkin Trans 2*, 1993, 773.

31 G. R. Famini and L. Y. Wilson, *J. Phys. Org. Chem.*, 1993, **9**, 539.

32 J. W. Grate, R. A. McGill and D. Hilvert, *J. Am. Chem. Soc.*, 1993, **115**, 8577.

33 D. S. Kemp and K. G. Paul, *J. Am. Chem. Soc.*, 1975, **97**, 7305.

34 J. Smid, A. Varma and S. C. Shah, *J. Am. Chem. Soc.*, 1979, **101**, 5764.

35 M. Shirai and J. Smid, *J. Am. Chem. Soc.*, 1980, **102**, 2863.

36 I. M. Kolthoff, *Anal. Chem.*, 1974, **46**, 1992; M. K. Chantooni and I. M. Kolthoff, *J. Phys. Chem.*, 1976, **80**, 1307.

37 D. S. Kemp, D. D. Cox and K. G. Paul, *J. Am. Chem. Soc.*, 1975, **97**, 7312.

38 C. Lewis, P. Paneth, M. H. O'leary and D. Hilvert, *J. Am. Chem. Soc.*, 1993, **115**, 1410.

39 J. M. Leonard and G. R. Famini, *A User's Guide to the Molecular Modeling Analysis and Display System*, CRDEC-TR-030, US Army Chemical Research, Development and Engineering Center, Aberdeen Proving Ground, MD, 1989.

40 M. J. K. Dewar and W. Thiel, *J. Am. Chem. Soc.*, 1977, **99**, 4899.

41 J. J. P. Stewart, *Mopac Manual* FJSRL-TR-88-007, Frank J. Seiler Research Laboratory, US Air Force Academy, Colorado Springs, CO, 1988.

42 A. J. Hopfinger, *J. Am. Chem. Soc.*, 1980, **102**, 7126.

43 D. A. Belesley, E. Kuh and R. E. Welsh, *Regression Diagnostics*, Wiley, New York, 1980.

Paper 3/07604K
Received 30*th December* 1993
Accepted 25*th March* 1994