^{13}C Nuclear Magnetic Resonance Study of Acid–Base Tautomeric Equilibria

Lowell M. Schwartz,* Robert I. Gelb, Jonathan Mumford-Zisk, and Daniel A. Laufer
Department of Chemistry, University of Massachusetts, Boston, Massachusetts 02125, U.S.A.

Acid–base tautomerisation equilibria are reported for *o*-, *m*-, and *p*-aminobenzoic acids, *p*-aminophenylacetic acid, pyridine-3-carboxylic (nicotinic) acid, and pyridine-4-carboxylic (isonicotinic) acid. These equilibria are determined by modelling the displacements of ^{13}C n.m.r. chemical shifts of each carbon due to protonation of basic sites. The models are based on the corresponding displacement due to protonation of as many as twelve related compounds. Each such model compound yields several independent estimates of the tautomeric partitioning ratio corresponding to the several carbon resonances. It is observed that the mean tautomeric partitioning ratio and its estimated uncertainty are independent of the nature or the location of ring substituents and are unaffected by heteroatom substitution in the aromatic ring. This observation has led to the development of generic models for ^{13}C n.m.r. chemical shift protonation displacements based on collections of similar model compounds. The use of generic models summarises and simplifies the determination of tautomeric equilibria.

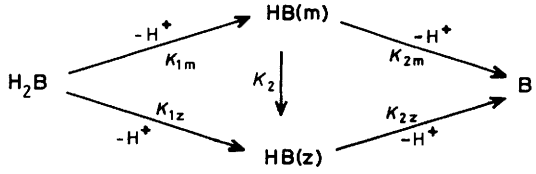
Partially protonated polyfunctional acids or bases may exist as mixtures of tautomers. For example, addition of one equivalent of acid to the glycinate anion $NH_2–CH_2–CO_2^-$ might produce either of two monoprotonated species: the zwitterionic (dipolar) $^+NH_3–CH_2–CO_2^-$ or the molecular $NH_2–CH_2–CO_2H$. Although in this case the zwitterion has long been accepted as the predominant form,[1] there are other examples where both forms exist. These include aminobenzoic acids,[2] pyridinecarboxylic acids,[3] pyridinols,[4] imidazoles,[5] and purines.[6] Tautomeric equilibria continue to be of interest in understanding such diverse phenomena as the structural and catalytic properties of biological molecules,[7,8] aqueous solvation,[9] and extrathermodynamic relationships.[10,11]

Various experimental methods have been employed for tautomeric equilibria determinations in the past. The simplest involves comparison of acidic dissociation constants of the substance in question with those of similar chemical model systems.[12] While this method indirectly relates acid dissociation $\Delta G°$ values, a more sophisticated technique compares acid dissociation $\Delta H°$ and $\Delta S°$ values.[2] Methods utilising comparisons of u.v. absorption[3,4,13] or other physical properties[14] have also been reported. In recent years n.m.r. spectrometric measurements have been widely applied to the study of acid–base tautomerism. ^1H and ^{13}C n.m.r. chemical shift displacements lead to qualitative discrimination between dissociation reactions at various sites in polyfunctional acids.[15] Reynolds *et al.*[5] estimated the tautomeric composition of neutral histidine as 80% 1-H, 20% 3-H from ^{13}C n.m.r. titration shifts of histidine and its *N*-methylimidazole derivatives. Grant and Townsend[6] determined tautomeric populations of several purines from quantitative analysis of ^{13}C n.m.r. substituent parameters. Schaal *et al.*[16] used ^{13}C n.m.r. to investigate tautomeric behaviour of nicotinic acid in $H_2O–Me_2SO$ solvents, and Bachovchin and Roberts[17] relied on ^{15}N n.m.r. to locate the tautomeric histidyl proton in the catalytic triad of α-lytic protease.

We have recently reported an application of ^{13}C n.m.r. spectrometry to the quantitative determination of the zwitterion–molecular partitioning in *p*- and *m*-aminobenzoic acids.[18] This ^{13}C n.m.r. technique has the advantage of providing several independent estimates of the partitioning from each model compound and a large number of independent estimates using several models. These multiple estimates serve as internal checks on the accuracy of the determination. We now report an extension of this methodology to investigations of tautomerism in *o*-aminobenzoic (anthranilic) acid, *p*-aminophenylacetic acid, and pyridine-3- and -4-carboxylic acids (nicotinic and isonicotinic acids, respectively.

We refer to two classes of acid–base equilibrium systems. One class consists of the systems being investigated for tautomeric behaviour, and the other of systems which serve as models from which we deduce the effects of protonation at various carbon sites. We symbolise the tautomer species as HB, the protonated tautomer as H_2B, and the unprotonated tautomer as B. 'Tautomer system' refers to any equilibrium mixture of these species. The second class of model systems involves acid–base conjugate pairs, to be symbolised by HM and M, both of which have unambiguous structures in the sense that neither HM nor M is a tautomeric mixture. Charge designations are not important and so are omitted from these symbols.

Interactions within the tautomeric system may be represented as shown in the Scheme. In any given solution we expect HB to



Scheme.

be an equilibrium mixture of the zwitterionic HB(z) and the molecular HB(m). The microscopic equilibrium constants K_{1m}, K_{1z}, K_{2m}, K_{2z} are not directly measurable but their values may be inferred from the overall acid–base dissociation constants of H_2B if the tautomeric partitioning ratio is known. One of several ways of expressing this partitioning is by the parameter X_m, the mole fraction of HB(m) relative to both forms of HB, and this will be our parameter of choice.

The ^{13}C n.m.r. method for determining X_m is perhaps best explained by reference to a particular example. For this purpose we will use nicotinic acid as the tautomer HB and will use benzoic acid as a model HM. Our experimental measurements consist of ^{13}C n.m.r. spectra of separate solutions of the tautomer system (typically six solutions) and of the model system (typically three or four solutions). All solutions are in 5% (v/v) $D_2O–H_2O$, made up with either HCl or NaOH in such a

way that the most acidic tautomer solution contains mostly H_2B and the most basic solution contains mostly B. Similarly, the model system solutions contain a range of compositions from mostly conjugate acid HM to mostly conjugate base M. In each tautomer solution we observe six carbon resonances corresponding to the six non-equivalent carbon atoms of *nicotinic acid*, and in each benzoic acid model solution we observe five carbon resonances. Each such observed resonance, δ_{obs}, is regarded as the mole fraction-weighted average of intrinsic chemical shifts of the acid–base conjugate species. For the tautomer system this averaging is expressed by equation (1a) and for the model system by equation (1b) where F_B or F_M is

$$F_B\delta_{obs}{}^i = [H_2B]\delta_{H_2B}{}^i + [HB]\delta_{HB}{}^i + [B]\delta_B{}^i \quad (1a)$$

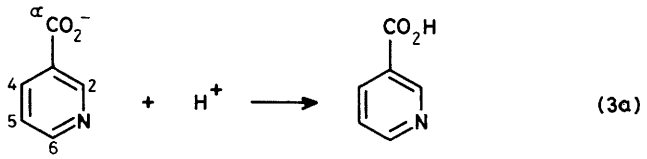
$$F_M\delta_{obs}{}^j = [HM]\delta_{HM}{}^j + [M]\delta_M{}^j \quad (1b)$$

the analytical concentration of tautomer or model, respectively. In the tautomer system δ^i refers to the intrinsic chemical shift of the *i*th carbon in the subscripted species whereas the superscript *j* refers to the *j*th carbon of the subscripted model species. Assignments of observed resonance lines to individual carbons are made by consideration of relaxation rates, symmetry factors, and approximate additivity of substituent effects.[19] The molar concentrations of species denoted by the square brackets are calculated from solution composition, from known acid dissociation constant values, and from ionic activity coefficient estimates derived from the Debye–Hückel correlation, assuming ion-size parameter values of 0.9 and 0.35 nm for H_3O^+ and OH^-, respectively, and 0.8 and 0.9 nm for tautomer and model ionic species. The intrinsic chemical shift values and their standard error estimates are extracted from the observed chemical shifts and the species concentration values by a linear multiple regression procedure.[20]

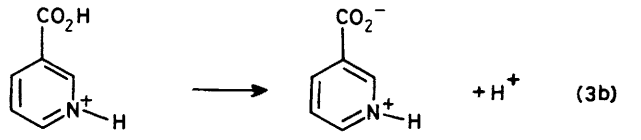
The intrinsic chemical shift $\delta_{HB}{}^i$ of the tautomer is in fact the mole fraction-weighted average of intrinsic chemical shifts, $\delta_z{}^i$ of the zwitterionic form and $\delta_m{}^i$ of the molecular form: $\delta_{HB}{}^i = X_{mi}\delta_m{}^i + (1 - X_{mi})\delta_z{}^i$. Once we know the intrinsic shifts of the tautomeric forms, we can determine the partitioning by equation (2).

$$X_{mi} = (\delta_{HB}{}^i - \delta_z{}^i)/(\delta_m{}^i - \delta_z{}^i) \quad (2)$$

We estimate values for $\delta_z{}^i$ and $\delta_m{}^i$ by hypothesising that the displacement of the chemical shift at this *i*th carbon due to protonation of a basic site on the tautomer nicotinic acid is nearly the same as the displacement of the chemical shift of the corresponding carbon due to protonation of the corresponding base site on the model benzoic acid. In other words, the molecular form of the tautomer HB(m) is regarded as the product of protonation of B at the carboxylate site [equation (3a)] and the zwitterion form is regarded as the product of deprotonation of the carboxylic acid site of H_2B [equation (3b)].



(3a)



(3b)

Thus if we represent the protonation displacement of chemical shift by $\Delta_B{}^i$, the intrinsic chemical shifts are related by equations (4a) and (4b).

$$\delta_m{}^i = \delta_B{}^i + \Delta_B{}^i \quad (4a)$$

$$\delta_z{}^i = \delta_{H_2B}{}^i - \Delta_B{}^i \quad (4b)$$

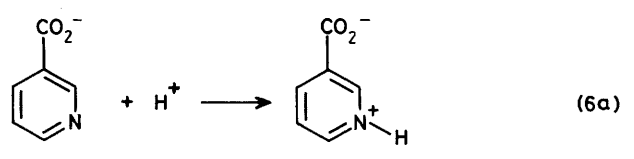
Using this hypothesis, we replace the unknown $\Delta_B{}^i$ by the measurement $\Delta_M{}^j$ from the model system HM, defined by equation (5) with the understanding that the *j*th carbon on M is

$$\Delta_M{}^j = \delta_{HM}{}^j - \delta_M{}^j \quad (5)$$

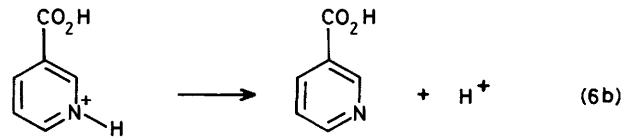
at the same position relative to the carboxylate on M as the *i*th carbon on B is relative to the carboxylate group on B.

For example, on the basis of ^{13}C n.m.r. measurements of six solutions of nicotinic acid ranging in pH from 1.5 to 11.7, and applying multiple linear regression in the form of equation (1a), we calculate that at the C-6 (*para* to the carboxylate), the intrinsic chemical shifts of B, HB, and H_2B are 151.45, 143.89, and 145.48 p.p.m. respectively. The corresponding carbon of the benzoic acid model system is at the C-4 or *para* position, and we calculate intrinsic chemical shifts there of 132.24 and 134.72 p.p.m. for M and HM, respectively. Thus, the protonation displacement $\Delta_M{}^{C4}$ is 134.72 − 132.24 or 2.48 p.p.m. By using this value in place of Δ^{C6} in equations (4a) and (4b), we estimate that $\delta_m{}^{C6} = 151.45 + 2.48 = 153.93$ p.p.m. and $\delta_z{}^{C6} = 145.48 - 2.48 = 143.00$ p.p.m. Then, by equation (2), $X_{mC6} = (143.89 - 143.00)/(153.93 - 143.00) = 0.081$. Thus, by using the protonation displacement of the chemical shift at the C-4 position of the benzoic acid–benzoate system as a model for the protonation displacement at the C-6 position of nicotinic acid, we estimate that nicotinic acid HB exists as 8.1% in the molecular form and 91.9% in the zwitterionic form. Each of the six non-equivalent carbons of nicotinic acid may be modelled in this way from corresponding carbons in benzoic acid, and each such calculation yields an independent estimate of the partitioning parameter X_m for nicotinic acid. However, these independent estimates are not equally reliable and this topic is discussed (see Uncertainty Considerations).

The number of X_{mi} estimates that can be derived by applying a particular model system to a particular tautomer system depends on the symmetries of the rings of both systems. For example, in the less symmetric model system *m*-hydroxybenzoic acid, C-2 and C-6 are non-equivalent and both are *ortho* to the carboxylate group. Each of these may serve as a model for the two non-equivalent C-2 and C-4 which are both *ortho* to the carboxylate group in nicotinic acid. Thus C-2 on nicotinic acid generates two estimates of X_m based on the two *ortho* carbons in the model, and C-4 on nicotinic acid also generates two estimates. Two pairs within the four estimates generated in this way at the *ortho* positions are clearly statistically interdependent. In a more symmetric tautomer such as isonicotinic acid, there is



(6a)



(6b)

Table 1. Sample calculation of tautomeric partitioning in nicotinic acid as modelled with benzoic acid—benzoate

		Tautomer system — nicotinic acid			
C-3	C-2	C-6	C-5	C-4	C-α
		B intrinsic chemical shifts (p.p.m.)			
133.47 (0.22)[a]	150.09 (0.22)	151.45 (0.22)	124.96 (0.22)	138.68 (0.23)	174.25 (0.22)
		HB intrinsic chemical shifts (p.p.m.)			
136.68 (0.21)	143.37 (0.21)	143.89 (0.21)	127.99 (0.21)	147.00 (0.22)	169.20 (0.21)
		H_2B intrinsic chemical shifts (p.p.m.)			
131.59 (0.23)	143.88 (0.23)	145.48 (0.23)	128.75 (0.23)	148.31 (0.24)	165.84 (0.23)
		α-Carboxylate protonation model system — benzoic acid/benzoate			
ipso	*ortho*	*para*	*meta*	*ortho*	C-α
		M intrinsic chemical shifts (p.p.m.)			
137.33 (0.05)	129.84 (0.05)	132.24 (0.05)	129.34 (0.05)	129.84 (0.05)	176.62 (0.05)
		HM intrinsic chemical shifts (p.p.m.)			
130.91 (0.06)	130.64 (0.06)	134.72 (0.06)	129.74 (0.06)	130.64 (0.06)	171.97 (0.06)
		Protonation displacements $\Delta_M{}^j$ (p.p.m.)			
−6.42 (0.27)	0.080 (0.10)	2.48 (0.15)	0.040 (0.09)	0.80 (0.10)	−4.65 (0.22)
		Tautomeric partitioning			
		Mole fraction molecular form X_{mi}			
0.121 (0.023)	0.037 (0.024)	0.081 (0.019)	0.120 (0.057)	0.064 (0.024)	1.449 (0.52)
	Mean $\overline{X_m} = 0.079$		$SE(\overline{X_m}) = 0.011$		$\Delta X_m = 0.041$

[a] Values in parentheses are standard error estimates.

only a single *ortho* resonance. This generates only a single extimate when modelled with a symmetric model such as *p*-hydroxybenzoic acid, but generates two estimates when modelled with *m*-hydroxybenzoic acid.

Not only are estimates of tautomeric partitioning derivable by modelling the protonation of the carboxylate group of the tautomer, but also by modelling the protonation of the nitrogen atom. In place of the reactions in equations (3a) and (3b), the alternative reactions are those shown in equations (6a) and (6b). These lead to the alternative relationships in equations (7a) and (7b).

$$\delta_z{}^i = \delta_B{}^i + \Delta_B{}^i \quad (7a)$$

$$\delta_m{}^i = \delta_{H_2B}{}^i - \Delta_B{}^i \quad (7b)$$

The protonation displacements $\Delta_B{}^i$ are replaced by protonation displacements $\Delta_M{}^j$ derived from corresponding carbons of model systems which feature a pyridine ring. For example, pyridine itself can serve as a model. We calculate that the intrinsic chemical shifts at C-2 of the nicotinic acid system are 150.09, 143.37, and 143.88 p.p.m. for **B**, **HB**, and H_2B, respectively, while those at C-2 of pyridine are 149.57 and 142.16 p.p.m. for **M** and **HM**, respectively. Thus, the protonation displacement $\Delta_M{}^{C2}$ is −7.41 p.p.m. Using this value in place of $\Delta_B{}^i$ in equations (7a) and (7b) leads to $\delta_z{}^{C2} = 142.68$ p.p.m. and $\delta_m{}^{C2} = 151.29$ p.p.m. Using these results together with $\delta_{HB}{}^{C2} = 143.37$ p.p.m. in equation (2) yields $X_{mC2} = 0.080$. Each of the five non-equivalent carbons (excluding the carboxylate carbon) of nicotinic acid can be modelled by five corresponding carbons of pyridine.

Uncertainty Considerations.—Each estimate of the partitioning parameter X_m is subject to both random and systematic errors. Random errors in the measured carbon resonances give statistical uncertainties in the X_m values. We can identify three sources of random error. First, the precision of the spectrometer itself is *ca.* 0.02 p.p.m. for sharp well separated lines. This means that if duplicate spectra are recorded with a single sample, a standard deviation of *ca.* 0.02 p.p.m. is observed for a particular carbon resonance. These uncertainties in observed chemical shifts are propagated through the multiple linear regression calculation and yield standard error estimates for the intrinsic chemical shifts for both the tautomer and model system species. In cases where broad resonance lines are observed, we use the residual variance about the regression to estimate standard error estimates for intrinsic chemical shifts.

A source of random error results from our use of different sample tubes for different tautomer and model systems. Any given sample tube affects the observed resonances by a small displacement in one direction or the other. We estimate the effect of these small displacements as a random error of *ca.* 0.1 p.p.m. difference between tautomer and model chemical shifts. In order to estimate the effect of this error on the X_m determinations, we add 0.1 p.p.m. to the standard error estimate for the intrinsic chemical shift of each tautomer system species.

Thirdly, there is a random error associated with the preparation of each solution. We estimate that the precision of weighing solid reagents and the precision of pipetting appropriate amounts of stock solution and HCl or NaOH into the sample tubes to achieve the desired distribution of conjugate acid–base species together amount to an average statistical uncertainty of *ca.* 1.5% in the concentrations of individual species. The principal effect of these uncertainties is to propagate *ca.* 3% random errors into each Δ_M protonation displacement. These random error estimates contribute to a standard error estimate $SE(X_m)$ for each X_m. In ref. 18 we derive a propagation-of-variance formula by which $SE(X_m)$ values are calculated from standard errors of intrinsic chemical shifts.

Table 1 shows the results of these calculations for the nicotinic acid tautomer system with carboxylate protonation displacements modelled with benzoic acid–benzoate. In a given column we align the intrinsic chemical shift of a particular carbon of nicotinic acid, the intrinsic shift of the corresponding carbon of benzoic acid, and the X_{mi} value generated by these data. Standard error estimates for each entry are given in

parentheses. We note that the standard error estimates $SE(X_{mi})$ differ from carbon to carbon. In particular, the X_{mi} estimate generated by the carboxylate C_α is substantially different from the other values, but the precision of the C_α value is considerably less than the others. In order to arrive at a 'best' estimate of X_m as predicted for nicotinic acid using benzoic acid–benzoate as a protonation model, we calculate a weighted mean with a weighting factor W_i for each X_{mi} taken as the reciprocal of the squared standard error estimate, as shown in equation (8)

$$\overline{X_m} = \Sigma_i W_i X_{mi} \Sigma^i W_i \quad (8)$$

where $W_i = [SE(X_{mi})]^{-2}$. The standard error estimate $SE(\overline{X_m})$ of this weighted mean $\overline{X_m}$ is $(\Sigma_i W_i)^{-1/2}$.[21] Applying these formulae to the six entries in the last row of Table 1, we calculate a weighted mean $\overline{X_m}$ of 0.079 with a corresponding standard error estimate of 0.011. This uncertainty of 0.011 refers to the precision of the weighted mean value $\overline{X_m} = 0.079$ as estimated by propagating known sources of random error in the experimental measurements. This result implies that if we repeated the entire experiment by preparing new solutions of nicotinic and benzoic acids, filling various sample tubes, recording ^{13}C n.m.r. spectra, and calculating new X_{mi} estimates, the standard deviation of these duplicate estimates would be near the values appearing in parentheses in the last row of Table 1, and the standard deviation of the weighted mean $\overline{X_m}$ would be *ca.* 0.011.

In addition to these random errors there are systematic errors which affect the accuracies of the X_m estimates. We expect the major source of systematic error to be the hypothesis that the protonation displacement at a particular carbon in the tautomer system equals the protonation displacement observed at a corresponding carbon of a model system. The protonation displacement at each carbon of the tautomer system must differ somewhat from that of the corresponding carbon in a model. These errors are responsible for the scatter of the X_{mi} values derived from the six different carbons of nicotinic acid and shown in Table 1. This scatter is much greater than can be explained by statistical variation of experimental data alone. However, we are unable to make *a priori* estimates of these errors, but can only observe their manifestation in the scatter of the X_{mi} values. We summarise the effects of these systematic errors by calculating a weighted root (RMS) deviation of each X_{mi} from the mean, according to equation (9).

$$\Delta X_m = [\Sigma_i W_i (X_{mi} - \overline{X_m})/\Sigma_i W_i]^{1/2} \quad (9)$$

On the basis of the six X_{mi} values shown in Table 1, we calculate a root mean error (r.m.s.) of 0.041. Because we have predicted a scatter of *ca.* 0.011 in $\overline{X_m}$ due to random errors and this component is included in the overall scatter of 0.041, we conclude that systematic errors average to *ca.* 0.03 in $\overline{X_m}$ for nicotinic acid modelled with benzoic acid–benzoate.

Tautomeric Partitioning Results.—We have calculated tautomeric partitioning for several systems each based on from six to twelve model systems. The results are shown in Table 2. Listings of $\overline{X_m}$ values and their standard error estimates at each carbon of each tautomer–model system, as well as intrinsic chemical shift data on which these results are based, are available on request from the authors.

Table 2. Tautomeric partitioning predictions

Tautomeric system	Model system	X_m	$SE(X_m)$	ΔX_m
o-Aminobenzoic acid	α-Carboxylate protonation models			
(anthranilic acid)	Benzoic acid	0.667	0.008	0.066
	Salicylic acid	0.641	0.009	0.100
	m-Hydroxybenzoic acid	0.671	0.006	0.077
	p-Hydroxybenzoic acid	0.674	0.008	0.064
	m-Nitrobenzoic acid	0.663	0.006	0.078
	p-Aminobenzoic acid	0.671	0.008	0.068
	Nicotinic acid *N*-oxide	0.708	0.008	0.076
	Isonicotinic acid *N*-oxide	0.669	0.008	0.061
	Amine protonation models			
	Aniline	0.741	0.009	0.073
	p-Aminobenzoic acid	0.673	0.010	0.065
	o-Aminosulphonic acid	0.654	0.009	0.073
	Sulphanilic acid	0.673	0.011	0.061
	Grand mean (84 X_{mi} values)	0.673		0.076
m-Aminobenzoic acid[a]	α-Carboxylate protonation models			
	Benzoic acid[a]	0.400	0.011	0.021
	Salicylic acid	0.416	0.008	0.036
	m-Hydroxybenzoic acid	0.401	0.008	0.023
	p-Hydroxybenzoic acid[a]	0.401	0.011	0.021
	m-Nitrobenzoic acid	0.405	0.008	0.027
	p-Aminobenzoic acid	0.403	0.011	0.027
	Nicotinic acid *N*-oxide	0.426	0.013	0.022
	Isonicotinic acid *N*-oxide	0.403	0.011	0.020
	Amine protonation models			
	Aniline[a]	0.400	0.012	0.018
	p-Aminobenzoic acid	0.422	0.010	0.033
	o-Aminosulphonic acid	0.438	0.008	0.031
	Sulphanilic acid[a]	0.425	0.010	0.031
	Grand mean (93 X_{mi} values)	0.412		0.030

Table 2 (*continued*)

Tautomeric system	Model system	X_m	$SE(X_m)$	ΔX_m
p-Aminobenzoic acid[a]	α-Carboxylate protonation models			
	Benzoic acid[a]	0.973	0.012	0.035
	Salicylic acid	0.981	0.020	0.118
	m-Hydroxybenzoic acid	0.971	0.011	0.030
	p-Hydroxybenzoic acid[a]	0.979	0.012	0.023
	m-Nitrobenzoic acid	0.969	0.011	0.017
	Nicotinic acid *N*-oxide	0.986	0.013	0.043
	Isonicotinic acid *N*-oxide	0.985	0.018	0.045
	Amine protonation models			
	Aniline[a]	1.079	0.024	0.103
	o-Aminosulphonic acid	0.924	0.025	0.139
	Sulphanilic acid[a]	0.998	0.026	0.014
	Grand mean (48 X_{mi} values)	0.979		0.056
p-Aminophenylacetic acid	β-Carboxylate protonation models			
	Phenylacetic acid	0.295	0.021	0.044
	p-Hydroxyphenylacetic acid	0.293	0.021	0.046
	Amine protonation models			
	Aniline	0.291	0.026	0.012
	p-Aminobenzoic acid	0.342	0.021	0.046
	o-Aminosulphonic acid	0.355	0.018	0.035
	Sulphanilic acid	0.334	0.022	0.027
	Grand mean (29 X_{mi} values)	0.322		0.047
Nicotinic acid	α-Carboxylate protonation models			
	Benzoic acid	0.079	0.011	0.041
	m-Hydroxybenzoic acid	0.079	0.008	0.052
	p-Hydroxybenzoic acid	0.077	0.011	0.051
	m-Nitrobenzoic acid	0.092	0.008	0.079
	p-Aminobenzoic acid	0.065	0.013	0.064
	Nicotinic acid *N*-oxide	0.057	0.009	0.037
	Isonicotinic acid *N*-oxide	0.104	0.011	0.065
	Pyridine nitrogen protonation models			
	Pyridine	0.035	0.015	0.067
	Nicotinamide	0.073	0.012	0.078
	Isonicotinamide	0.040	0.016	0.062
	Pyridine-3-sulphonic acid	0.072	0.012	0.075
	Grand mean (74 X_{mi} values)	0.075		0.064
Isonicotinic acid	α-Carboxylate protonation models			
	Benzoic acid	0.001	0.017	0.071
	m-Hydroxybenzoic acid	0.017	0.013	0.085
	p-Hydroxybenzoic acid	0.011	0.017	0.065
	m-Nitrobenzoic acid	0.038	0.012	0.112
	p-Aminobenzoic acid	0.054	0.017	0.059
	Nicotinic acid *N*-oxide	0.038	0.021	0.135
	Isonicotinic acid *N*-oxide	0.038	0.014	0.045
	Pyridine nitrogen protonation models			
	Pyridine	0.062	0.018	0.013
	Nicotinamide	0.094	0.017	0.101
	Isonicotinamide	0.038	0.020	0.088
	Pyridine-3-sulphonic acid	0.094	0.016	0.095
	Grand mean (47 X_{mi} values)	0.042		0.090

[a] Results of these tautomer–model systems were published previously.[18]

In Table 2 we show for each tautomer–model system the mean partitioning parameter $\overline{X}_m$, its standard error estimate SE($\overline{X}_m$), and an overall uncertainty estimate in the form of the weighted r.m.s. error ΔX_m as calculated from equation (9). For each tautomer system we have summarised the results from all the models into a single grand mean X_m and its weighted r.m.s. error. These quantities, which are entered as grand means in Table 2, result from equations (8) and (9), which combine all the $\overline{X}_m$ estimates from all the models.

We have included in Table 2 *p*-aminobenzoic acid, both as a tautomer and as a protonation displacement model for other tautomers. We are able to use *p*-aminobenzoic acid as a protonation displacement model because we[18] and others[2] have shown that the monoprotonated species exists virtually

Table 3. Generic protonation displacements

Protonation site	Carbon	Count	Δ_{gen}[a] p.p.m.	SE(Δ_{gen})[b] p.p.m.	Unweighted standard deviation, p.p.m.
α-Carboxylate[c]	*ipso*	8	−6.76	0.099	0.67
	ortho	12	0.67	0.015	0.58
	meta	11	0.26	0.014	0.27
	para	7	2.14	0.045	0.47
	C-α	8	−4.33	0.069	0.50
	(all)	46	0.36		3.1
α-Amine[d]	*ipso*	4	−16.59	0.35	0.63
	ortho	5	7.08	0.15	0.86
	meta	5	0.38	0.029	0.33
	para	4	10.40	0.23	0.94
	(all)	19	0.70		10.4
Pyridine N[e]	*ortho*	6	−6.62	0.11	0.49
	meta	6	3.59	0.065	0.40
	para	4	9.07	0.18	0.49
	(all)	16	1.75		6.6

[a] Weighted mean over a number (count) of protonation displacements at carbons situated relative to the protonation site. [b] Standard error estimate of Δ_{gen}. [c] Based on eight systems: benzoic, salicylic, *m*-hydroxybenzoic, *p*-hydroxybenzoic, *m*-nitrobenzoic, and *p*-aminobenzoic acids, and nicotinic and isonicotinic acid *N*-oxides. [d] Based on four systems: aniline, *p*-aminobenzoic, *o*-aminosulphonic, and sulphanilic acids. [e] Based on four systems: pyridine, nicotinamide, isonicotinamide, and pyridine-3-sulphonic acid.

completely in the molecular form. This conclusion is reflected in the grand mean $\overline{X}_m$ value of 0.979 ± 0.056, found for the partitioning parameter. Thus, the first and second protonations of *p*-aminobenzoate ion occur unambiguously at the carboxylate and amine sites, respectively, and these two protonations can serve as legitimate models for other systems. We observe that where *p*-aminobenzoic acid is used as a protonation model, the resulting mean $\overline{X}_m$ agrees with all other models used for that tautomer system. This observation further validates the deduction that *p*-aminobenzoic acid is all or nearly all in the molecular form.

We observe that there is remarkable agreement among the several mean $\overline{X}_m$ values predicted by the various models for each tautomeric system. This agreement is observed for six different tautomer systems, each modelled with six to twelve models. This serves to confirm our hypothesis that the modelling of ^{13}C n.m.r. protonation displacements as described is a valid and useful procedure for determining tautomeric partitioning. The entries in the column labelled ΔX_m are the r.m.s. errors associated with the corresponding mean $\overline{X}_m$ for each tautomer–model system. The magnitude of this r.m.s. error reflects the degree to which protonation displacements at individual carbons of the model system are a successful approximation for the protonation displacements of corresponding carbons of the tautomer system. We are struck by the fact that these errors seem to be independent of the substituent locations around the aromatic ring. For example, the tautomer system *o*-aminobenzoic acid features *ortho* substituency. Conventional wisdom would predict that *ortho*-substituted models would serve as better protonation displacement models for tautomeric prediction than would *meta*- or *para*-substituted models. Yet, we observe that the r.m.s. errors for $\overline{X}_m$ predicted by the two *ortho*-substituted models, salicylic and *o*-aminosulphonic acid, are no less than the r.m.s. errors of the *meta*- or *para*- substituted systems, and no less than the r.m.s. errors of the carboxylate models with pyridine ring systems, nicotinic and isonicotinic acid *N*-oxides. This general invariance of modelling success to ring substituency pattern is true for all six tautomeric systems examined to date.

However, we notice different magnitudes of ΔX_m values in the different tautomer systems. For example, all the model systems, used either for carboxylate or amine protonation displacement for the tautomer *m*-aminobenzoic acid, yield ΔX_m values in the range 0.018—0.036, whereas for anthranilic acid the range is 0.061—0.10. Apparently, the average systematic error associated with using the model systems for estimating protonation displacements in *m*-aminobenzoic acid is less than the average systematic error associated with using these models with anthranilic acid. Therefore, by implication, the two protonation reactions of *m*-aminobenzoate ion affect the magnetic environments of all the carbons in a manner which is more typical of the general class of aromatic ring systems than is true with anthranilic acid. This reasoning, together with the observed invariance of modelling success with ring substituents, leads us to attempt to construct generic models for protonation displacement.

Generic Modelling of Protonation Displacements.—We have constructed generic models for protonation displacement about three types of protonation sites, α-substituted carboxylate and amine groups and the pyridine nitrogen. The results are shown in Table 3. We have not yet examined a sufficient number of models to construct a generic model for the β-substituted carboxylate group in *p*-aminophenylacetic acid. Under the heading Δ_{gen} we show the weighted mean of all non-equivalent protonation displacements at the specified carbon as positioned relative to protonation site. The weighting factors W_j reflect the standard error estimate of each displacement and are the same as those defined for use in equation (8). Averages are taken over all models having the specified type of protonation site. The column labelled SE(Δ_{gen}) is the standard error estimate for Δ_{gen} and is given by SE(Δ_{gen}) = $(\Sigma_j W_j)^{-1/2}$. Because these standard errors reflect the precision of the weighted mean rather than the scatter of the component displacements within the mean, we also list the conventional unweighted standard deviation of the several protonation displacements. Entries in rows marked 'all' summarise the results of all protonation displacements regardless of carbon location in the molecules of a given type. We notice that the unweighted standard deviation of each 'all' entry is much larger than the Δ_{gen} value itself. This indicates that the average of all protonation displacements is within a statistical uncertainty of 0 p.p.m. In other words, the algebraic sum is about zero for all protonation displacements taken over all systems and all carbons. Also, the standard deviation over all

Table 4. Tautomeric partitioning using generic protonation displacements

	Protonation site					
	α-Carboxylate		α-Amine or Pyridine N		Both sites	
Tautomer system	X_m	ΔX_m	X_m	ΔX_m	X_m	ΔX_m
o-Aminobenzoic acid	0.669	0.042	0.713	0.058	0.677	0.048
m-Aminobenzoic acid	0.404	0.025	0.412	0.024	0.408	0.024
p-Aminobenzoic acid	0.980	0.024	1.043	0.052	0.991	0.039
Nicotinic acid	0.067	0.046	0.044	0.070	0.064	0.050
Isonicotinic acid	0.044	0.093	0.083	0.103	0.054	0.098

carbons is much greater than the standard deviation at any given carbon. Standard deviations of protonation displacements at individual carbon locations are typically *ca.* 0.6 p.p.m. and do not exceed 0.94 p.p.m. These relatively small standard deviation values explain why the tautomeric partitioning estimates are independent of substituent orientation of the model system. Apparently, the presence or absence of a substituent group bonded to a ring carbon has little effect on the protonation displacement at that carbon. In other words, protonation displacements of ^{13}C n.m.r. chemical shifts primarily reflect symmetry properties of the delocalised π-electron system rather than specific resonance, field, or inductive interactions of ring substituents or even of heteroatom substitution in the ring. Thus the position of a given carbon atom relative to the protonation site is of dominant importance in determining the ^{13}C n.m.r. chemical shift displacement.

We expect that if we calculate tautomeric partitioning ratios using the appropriate generic protonation displacements, Δ_{gen}, in place of specific model protonation displacements, we should obtain similar partitioning results to those listed as 'Grand means' in Table 2 for each tautomer. In these calculations we weight each generic carbon by the reciprocal of the square of the corresponding $SE(\Delta_{gen})$ value. Our expectation is confirmed by the results given in Table 4 where we show mean X_m values and associated weighted r.m.s. errors as predicted on the basis of the generic displacements due to protonation separately at the α-carboxylate sites and the nitrogen sites, and averaging over both sites. We observe that for each tautomer system the separate predictions agree within the r.m.s. error bounds and also that the overall results agree with the 'Grand means' within r.m.s. error bounds.

Discussion

Our estimation of tautomeric partitioning in nicotinic and isonicotinic acids (Table 4) indicates that monoprotonated forms of both acids are largely if not completely zwitterionic. These results appear to be in agreement with those reported by Green and Tong[3] and by Schaal *et al.*[16]

Previous investigations of the tautomeric behaviour of aminobenzoic acid gave values between 0.8 and 0.94 for the *ortho* isomer, between 0.3 and 0.5 for the *meta* isomer, and between 0.83 and 0.99 for the *para* isomer.[2] From Table 4, we report corresponding values of 0.68 ± 0.05, 0.41 ± 0.02, and 0.99 ± 0.04 for *o*-, *m*-, and *p*- amino benzoic acids, respectively, as estimated from generic modelling of protonation displacements at both the carboxylate and amine sites. Our results for *meta* and *para* isomers are in good agreement with values of 0.40 for *m*-aminobenzoic acid and 0.98 for *p*-aminobenzoic acid which are found by comparing acid dissociation ΔH° and ΔS° values of each of these isomers with corresponding thermodynamic data for chemical model systems.[2] However, our finding of $X_m = 0.68 \pm 0.05$ for *o*-aminobenzoic acid must be compared with earlier estimates of *ca.* 0.6,[14] 0.91,[2] and 0.94.[2] We shall focus our discussion on the previously reported X_m values close to 0.9, which are based on the modelling of ΔH° and ΔS° values for the protonation of methyl *o*-aminobenzoate. We note that ΔH° and ΔS° for protonation reactions in aqueous media reflect primarily the difference in interaction of the solvent with successively protonated solute species rather than intramolecular bonding interaction differences of the solute species. For example, if we survey a large number of *meta*- and *para*-substituted benzoic acids having widely differing intramolecular bonding interactions, we find that the acid dissociation ΔH° values are all in the range 0 ± 1 kcal mol^{-1} * and ΔS° values are all near −20 cal mol^{-1} K^{-1}.[22] The uniformity of these values reflects two effects which are common to all these systems: the solvent interaction and structuring about the dissociated proton, and the similar solvation properties of the carboxylate anions. However, in many *ortho*-substituted acids there exists the possibility of intramolecular interaction which may affect solvation properties in such specific ways that ΔH° and ΔS° values could not be modelled on the basis of related chemical systems. For example, in *o*-aminobenzoic acid we can envisage a monoprotonated structure in which the proton is bound mutually by adjacent amine and carboxylate groups. If this structure were regarded as a zwitterion, the overall solvent interaction with this ammonium group would be substantially less than expected for a typical ammonium group. The acid dissociation ΔH° and ΔS° values would reflect this *ortho* interaction and these values could not be readily interpreted by comparison with chemical model systems having differing intramolecular interactions. In other words, while chemical model systems such as methyl *o*-aminobenzoate may accurately reflect tautomeric behaviour in *meta*- and *para*-substituted acids, specific interactions which are possible in *ortho*-substituted acids do not permit reliable chemical modelling.

On the other hand, the ^{13}C n.m.r. technique described here has been found to yield consistent X_m results for *o*-aminobenzoic acid (Table 3) regardless of whether the protonation displacement models are derived from *ortho*-, *meta*-, or *para*-substituted compounds and whether the protonation site is the amine or carboxylate group. We conclude that the ^{13}C n.m.r. method is insensitive to solvation or desolvation effects resulting from direct intramolecular interaction between *ortho* substituents. If this were not true, we would have observed different X_m estimates with *ortho*-substituted models than with *meta*- and *para*-substituted models. We conclude that this ^{13}C n.m.r. technique provides accurate estimates of proton distribution even where classical chemical modelling fails.

Experimental

Commercial samples of reagents for ^{13}C n.m.r. measurements were assayed by acid–base titration. If appreciable impurity

* 1 kcal = 4.184 kJ.

concentrations were detected, the reagents were recrystallised until impurity levels were satisfactorily reduced. Stock solutions in 5% D_2O–H_2O (v/v) were made up from weighed portions of the assayed samples. In a typical experiment, a 5-ml aliquot portion of the stock was pipetted into a 10-mm n.m.r. sample tube and the ^{13}C n.m.r. spectrum then recorded. The spectrum was recorded again after the addition of each of two half- or three third-equivalents of 0.95M–NaOH or HCl solution in 5% D_2O–H_2O (v/v); these were added to the n.m.r. sample tube with a 200-μl micropipet.

^{13}C N.m.r. data were acquired at 30 ± 2 °C with a Bruker HX-270 spectrometer operating at 67.89 MHz for ^{13}C detection. Typical instrument settings were: 11 μs pulse width (30° tip angle); 14 kHz spectral width; 0.3 s acquisition time; 1.5 s pulse delay time; and 4—8 K transients. Chemical shifts are quoted downfield from external $(CH_3)_4Si$.

The calculation of intrinsic chemical shift values from observed resonances and solution composition data required acid dissociation pK values at 30 °C. If reliable values could not be found in published tables, they were obtained by pH potentiometric titration experiments using an Orion model 801 pH meter and conventional glass and reference electrodes. pK Values were calculated from titration data by a nonlinear regression program described elsewhere.[23]

Acknowledgements

The Bruker HX-270 experiments were performed at the n.m.r. facilities of the Francis Bitter National Magnet Laboratory, Massachusetts Institute of Technology. The n.m.r. facility is supported by Grant 00995 from the Division of Research of the National Institutes of Health and by the National Science Foundation under Contract C-76D.

References

1 J. P. Greenstein and M. Winitz, 'Chemistry of the Amino Acids,' Wiley, New York, 1961, vol. 1, p. 435.
2 J. J. Christensen, D. P. Wrathall, R. M. Izatt, and D. O. Tolman, *J. Phys. Chem.*, 1967, **71**, 3001.
3 R. W. Green and H. K. Tong, *J. Am. Chem. Soc.*, 1956, **78**, 4896.
4 D. E. Metzler, C. M. Harris, R. J. Johnson, D. B. Siano, and J. A. Thomson, *Biochemistry*, 1973, **12**, 5377.
5 W. F. Reynolds, I. R. Peat, M. H. Freedman, and J. R. Lyerla, *J. Am. Chem. Soc.*, 1973, **95**, 328.
6 M.-T. Chenon, R. J. Pugmire, D. M. Grant, R. P. Panzica, P. Townsend, and L. B. Townsend, *J. Am. Chem. Soc.*, 1975, **97**, 4636.
7 D. E. Metzler, 'Biochemistry. The Chemical Reactions of Living Cells,' Academic Press, New York, 1977, pp. 54, 191, 333.
8 P. R. Rony, *J. Am. Chem. Soc.*, 1969, **91**, 6090.
9 J. A. Sordo, M. Klobukowski, and S. Fraga, *J. Am. Chem. Soc.*, 1985, **107**, 7569.
10 H. H. Jaffe, *J. Am. Chem. Soc.*, 1955, **77**, 4445.
11 B. Van de Graaf, A. J. Hoefnagel, and B. M. Wepster, *J. Org. Chem.*, 1981, **46**, 653.
12 R. Wegscheider, *Monatsh. Chem.*, 1902, **23**, 287; L. Ebert, *Z. Phys. Chem. (Leipzig)*, 1926, **121**, 385.
13 I. M. Klotz and D. M. Gruen, *J. Am. Chem. Soc.*, 1945, **67**, 843.
14 G. Devoto, *Gazz. Chim. Ital.*, 1934, **64**, 371.
15 K. Wutrich, 'NMR in Biological Research: Peptides and Proteins,' American Elsevier, New York, 1976.
16 T. Khan, J. C. Halle, M. P. Simmonin, and R. Schaal, *J. Phys. Chem.*, 1977, **81**, 587.
17 W. W. Bachovchin and J. D. Roberts, *J. Am. Chem. Soc.*, 1978, **100**, 8041.
18 D. A. Laufer, R. I. Gelb, and L. M. Schwartz, *J. Org. Chem.*, 1984, **49**, 691.
19 S. Gould and D. A. Laufer, *J. Magn. Reson.*, 1979, **34**, 37.
20 R. I. Gelb, L. M. Schwartz, and D. A. Laufer, *J. Am. Chem. Soc.*, 1981, **103**, 5664.
21 P. R. Bevington, 'Data Reduction and Error Analysis for the Physical Sciences,' McGraw-Hill, New York, 1969, ch. 5.
22 J. W. Larson and L. G. Hepler, in 'Solute–Solvent Interactions,' eds. J. F. Coetzee and C. D. Ritchie, Marcel Dekker, New York, 1969, ch. 1.
23 L. M. Schwartz and R. I. Gelb, *Anal. Chem.*, 1978, **50**, 1571.

Received 21*st March* 1986; *Paper* 6/560