

## A Nuclear Magnetic Resonance and Theoretical Investigation of the Effect of Charge and Solvation on the Conformations of 1,2-Disubstituted Ethanes

Raymond J. Abraham\* and Brian D. Hudson

The School of Chemistry, University of Liverpool, P.O. Box 147, Liverpool L69 3BX

W. Anthony Thomas

Roche Products Limited, P.O. Box 8, Welwyn Garden City, Hertfordshire AL7 3AY

The conformational energies of a number of 1,2-disubstituted ethanes ( $\text{CH}_2\text{X}\cdot\text{CH}_2\text{Y}$ ) have been obtained by *ab initio* and MNDO calculations (*i.e.* for the vapour phase) and by  $^1\text{H}$  n.m.r. spectroscopy (for aqueous solutions). There is generally good agreement between theory and experiment for the neutral molecules. The 2-halogenoethylamines ( $\text{X} = \text{F}$  or  $\text{Cl}$ ,  $\text{Y} = \text{NH}_2$ ), 3-chloropropionic acid ( $\text{X} = \text{Cl}$ ,  $\text{Y} = \text{CO}_2\text{H}$ ), and succinic acid ( $\text{X} = \text{Y} = \text{CO}_2\text{H}$ ) all favour the *gauche* conformer, whereas ethylenediamine ( $\text{X} = \text{Y} = \text{NH}_2$ ) has no conformational preference.

For the charged molecules, although theory correctly predicts the direction of the conformer energy changes on ionisation of the amino and carboxy groups, it vastly overrates the extent of these changes. The most extreme case is  $\beta$ -alanine zwitterion ( $\text{X} = ^+\text{NH}_3$ ,  $\text{Y} = \text{CO}_2^-$ ), which is predicted to be wholly in the *gauche* form (by 10–30 kcal mol $^{-1}$ ), yet which shows no conformational preference in aqueous solution ( $\Delta E_{g-t}$ , -0.1 kcal mol $^{-1}$ ).

These results suggest that in carrying out molecular mechanics calculations on molecules containing charged groups exposed to aqueous solution, the dominant effect of the solvent in 'neutralising' the electrostatic interactions of the charged species must always be considered.

One of the major problems with molecular modelling studies is how to account for the effects of solvation. This is particularly important where charged species are involved, since vapour-phase energy calculations (either quantum mechanical or molecular mechanical) are dominated by the electrostatic energies associated with the charged groups. However, a conformation which is highly stabilised by electrostatic interactions in the vapour phase need not necessarily be so stabilised in aqueous solution. The conformational preferences of molecules with charged groups are of interest since such groups are present in many biologically important molecules.

Although explicit calculations of solvation effects are available, these are either prohibitively time-consuming for general use in molecular modelling (*e.g.* the quantum mechanical supermolecule approach<sup>1</sup> and the Monte Carlo methods<sup>2</sup>), or based on a simplified classical model (*e.g.* the reaction field theory<sup>3</sup>). Thus, solvation effects tend to be disregarded, or treated in a qualitative manner, as in the use of a distance-dependent dielectric constant in molecular mechanics calculations.<sup>4</sup>

A major obstacle to the quantification of the effects of solvation of charged groups is the lack of good experimental evidence on the conformation of molecules containing such groups. Richards says:<sup>5</sup> 'In order to have some data, small flexible molecules should have their conformational preferences studied in both the gas and solution phases. Accurate Hartree-Fock calculations should provide essential detail for the gas-phase conformational energy diagrams. In solution, high-resolution n.m.r. studies can give accurate population ratios. By comparing the conformer population ratios for the two environments it ought to be possible to devise an empirical scheme whereby isolated molecule calculations can be modified to incorporate solvation effects.'

Excellent examples of such small organic molecules are the 1,2-disubstituted ethanes. The conformational properties of these molecules have been extensively studied both experimentally<sup>6</sup> and theoretically.<sup>7,8</sup> This has produced some unexpected results; for example, the fact that non-bonded interactions between fluorine atoms can be attractive.<sup>7</sup> Some studies

have been done on 1,2-disubstituted ethanes with ionisable substituents. Examples are studies on the conformation of histamine,<sup>9</sup> of acetylcholine and related compounds,<sup>10,11</sup> and of the 2-halogenoethylamines.<sup>12</sup> Such studies have, however, generally been limited to a particular compound or to a closely related series of compounds; no attempt has been made to study the effect of the excess of charge itself on the conformational equilibria. Nonetheless, these compounds should be capable of providing much information about the energetics of the interactions between charged substituents.

A combined n.m.r. and theoretical study of the effect of excess of charge on conformational isomerism in simple 1,2-disubstituted ethanes has, therefore, been undertaken. The study is intended to answer three questions. (1) What is the difference in energy between the two rotational isomers in the vapour phase for both the charged and the uncharged molecules? (2) What are the corresponding energy differences in aqueous solution? (3) What is the effect of the excess of charge on the rotational isomerism both in the vapour phase and in solution? The compounds were chosen to give a good spread of energy values, ranging from the predominantly electrostatic repulsion between X and Y of succinate dianion and ethylenediammonium dication to the electrostatic attraction of  $\beta$ -alanine zwitterion.

### Experimental

Compounds were purchased from Aldrich, either as the amine hydrochloride salts or as the free carboxylic acids. 2-Fluoroethylamine hydrochloride (**1**) was purified by precipitation from ethanolic solution with diethyl ether. Succinic acid (**5**) was purified by recrystallisation from aqueous solution. All other compounds were used without further purification. Solutions were made up in  $\text{D}_2\text{O}$  and were referenced to sodium 3-trimethylsilyl[2,2,3,3- $^2\text{H}_4$ ]propionate (TSP) as internal standard. To obtain the solute in its required charge state the pH of the solution was adjusted by the addition of either 1M-DCl or solid NaOH. The  $^1\text{H}$  spectrum of methyl hydrogen succinate was also obtained in  $\text{CDCl}_3$  solution using tetramethylsilane (TMS) as the internal reference compound.

$^1\text{H}$  N.m.r. spectra for the unsymmetrical compounds (1), (2), (4), and (6) (Table 1) were obtained at 250 MHz with a Bruker WM-250 spectrometer operating in Fourier transform mode. Typical conditions were: 32 transients accumulated into 8 K data points with a pulse width of 3.0  $\mu\text{s}$  and a sweep width of 1 500 Hz, giving an acquisition time of 3.2 s. In all cases a Gaussian multiplication of the free induction decay (f.i.d.) was carried out using values of (LB) = -0.7 and (GB) = 0.5. The f.i.d. was zero-filled to 32 K data points giving a digital resolution of 0.10 Hz per point. For the symmetrical compounds (3) and (5) the  $^1\text{H}$ - $^1\text{H}$  coupling constants were obtained by observing the  $^{13}\text{C}$  satellite peaks of the  $^1\text{H}$  spectrum,<sup>13</sup> using a gated, homonuclear decoupling pulse sequence to suppress the main resonance and an automatic variable spinning rate to remove spinning side-bands. These spectra were all obtained at 400 MHz with a Bruker AM-400 spectrometer. Typical conditions were: 32 transients accumulated into 16 K data points with a pulse width of 3.0 s and a sweep width of 4 000 Hz, giving an acquisition time of 2.0 s. An exponential multiplication was performed on the f.i.d.s to increase the sensitivity using a value of 0.5 for GB. The digital resolution was 0.50 Hz per point. The  $^1\text{H}$  spectrum of methyl hydrogen succinate in  $\text{CDCl}_3$  solution was obtained at 250 MHz. A smaller sweep width of 400 Hz was used here, giving an acquisition time of 2.56 s, followed by gaussian multiplication and zero filling to 32 K. This gave a digital resolution of 0.024 Hz per point.

The coupling constants were obtained from the spectra using standard methods of spectral analysis for AA'BB' systems.<sup>14</sup> The spectra of the neutral molecules (1), (4), and (6) at 250 MHz were examples of deceptively simple AA'BB' systems, from which only the value of  $N(J + J')$  can be obtained. However, at lower applied fields (60 MHz) the spectra of each of these compounds [except (1)] was resolved into a 14-line pattern which could be analysed to give  $J$  and  $J'$ , and these results are given in Table 1. For the fluoro compound (1) and for the symmetric compounds (3) and (5) the heteronucleus was treated using the X approximation. The final refinement of the parameters was carried out using the iterative computer program PANIC.<sup>15</sup>

The observed energy differences were obtained from the coupling constants using the method of Abraham and Gatti.<sup>16</sup> The calculated values of the coupling constants of the *gauche* and *trans* forms, and hence the value of  $N(J + J')$ , were obtained from the equations (i) and (ii), where  $E_X$  and  $E_Y$  are

$$N_t = J_t^g + J_t^t = 19.42 - 0.25(E_X + E_Y) \quad (\text{i})$$

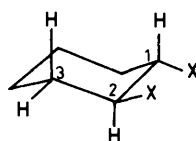
$$N_g = 0.5(J_g^t + J_g^{g'}) + J_g^g = 22.40 - 1.96(E_X + E_Y) \quad (\text{ii})$$

the Huggins electronegativities of substituents X and Y (values C, 2.6; N, 3.05; Cl, 3.15; F, 3.90). Using these values, and the observed value of  $N$ , the rotamer energy difference is then calculated using equation (iii).

$$\Delta E_{g-t} = RT \ln [2(N_{\text{obs}} - N_g)/(N_t - N_{\text{obs}})] \quad (\text{iii})$$

This calculation assumes that the charge (*i.e.* ionisation) on the amino and carboxy groups has no intrinsic effect on the vicinal proton-proton couplings. Although this has been shown to be the case for the average coupling in ethyl compounds (*e.g.*  $\text{Et}_3\text{N}$  versus  $\text{Et}_4\text{N}^+$ ),<sup>17</sup> this is not definitive, as any variations in  $J_g$  and  $J_t$  may cancel in  $J_{\text{av}}$ . We therefore investigated the effects of ionisation in the conformationally fixed systems *trans*-1,2-diaminocyclohexane (10) and *trans*-cyclohexane-1,2-dicarboxylic acid (11).\*

The vicinal proton-proton couplings were obtained from the first-order  $^{13}\text{C}$  satellite spectrum and are given in Table 2. These results will be considered subsequently.

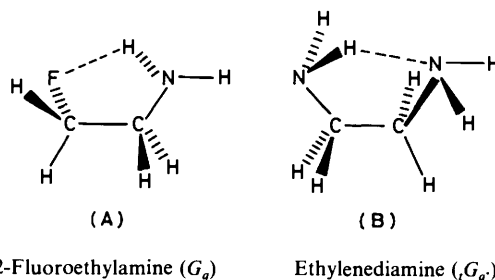


(10) X =  $\text{NH}_2$  ( $^+\text{NH}_3$ )

(11) X =  $\text{CO}_2\text{H}$  ( $\text{CO}_2^-$ )

Vapour-phase energy calculations were performed using the following procedure. Standard geometries were input (except that the carboxy group geometry was taken from a neutron diffraction study of succinic acid),<sup>19</sup> and were first optimised with respect to the XCC and CCY bond angles and the XCCY torsion angle using the MNDO method.<sup>20</sup> During this optimisation the conformation of the amine and carboxy groups was held constant. For the *gauche* form of the amines the conformations were chosen on the basis of the calculated STO-3G minimum-energy conformations of 2-fluoroethylamine ( $G_g$ ) and ethylenediamine ( $G_g$ ) shown.<sup>8</sup>

The acid groups were held in a conformation with the carboxy group orthogonal to the CCC plane. A study of the energies involved in the rotation of this group using MNDO suggested that the difference in energy between the various conformations due to rotation about the CCCO torsion angle is very small. For  $\beta$ -alanine (6) a full optimisation including the CCCO torsion angle was performed using MNDO.



These optimised geometries were then used as starting geometries for the more rigorous *ab initio* energy calculations at the STO-3G level. The XCC and CCY bond angles were further optimised at this new level of approximation, and this was followed by optimisation of the XCCY torsion angle, giving the required energy differences with full optimisation of the important internal degrees of freedom. This method of optimisation, in which a less rigorous molecular orbital method is used for the initial stages of the optimisation, leads to a great saving in computational time when working at the STO-3G level, and also has the advantage that the MNDO energy differences are available for comparison with the STO-3G results.

The optimisation of the internal degrees of freedom by STO-3G using the MNDO-optimised geometries was easily achieved in all cases except for  $\beta$ -alanine zwitterion (6), where the MNDO-optimised geometry for the *gauche* form predicted a strong hydrogen bond between one of the NH hydrogen atoms

\* A referee has pointed out that in many *trans*-1,2-disubstituted cyclohexanes the assumption of additive substituent  $A$ -values no longer holds (*e.g.* ref. 18). We take this pertinent point. Nevertheless, both the values of the couplings and their dependence on charge (Table 2) support the assumption of a large preponderance of the diequatorial conformer. Intriguingly the value of  $J_{1a-2a}$  is almost identical with that observed for the diequatorial conformer of a 1,2-dithiocyclohexane from low-temperature n.m.r. (11.0 Hz),<sup>18</sup> both supporting the  $J$  values used here and demonstrating conclusively the strong preference for the diaxial conformer in the dianion of (11).

**Table 1.** Observed n.m.r. parameters for CH<sub>2</sub>XCH<sub>2</sub>Y

	X	Y	$\delta_A$	$\delta_B$	$^3J$	$^3J'$	R.m.s. error
(1)	F	NH <sub>2</sub> <sup>b,c</sup>	4.515	2.902	4.86 (0.10) <sup>a</sup>	2.68 (0.06)	0.04
		NH <sub>3</sub> <sup>+d,e</sup>	4.749	3.375	6.68 (0.06)	4.67 (0.07)	0.09
(2)	Cl	NH <sub>2</sub> <sup>f</sup>	3.654	2.933	6.66 (0.07)	3.70 (0.07)	0.08
		+NH <sub>3</sub> <sup>g</sup>	3.892	3.429	7.23 (0.07)	6.3 (0.05)	0.11
(3)	NH <sub>2</sub>	NH <sub>2</sub> <sup>b,h</sup>	2.70	2.70	6.4 (0.05)	7.3 (0.4)	0.1
		+NH <sub>3</sub> <sup>b,i</sup>	3.39	3.39	8.3 (0.4)	5.1 (0.2)	0.8
(4)	Cl	CO <sub>2</sub> H <sup>b</sup>	3.801	2.888	7.22 (0.15)	6.22 (0.03) <sup>a</sup>	0.08
		CO <sub>2</sub> <sup>-b</sup>	3.792	2.845	6.22 (0.03) <sup>a</sup>	5.20 (0.15)	0.09
(5)	CO <sub>2</sub> H	CO <sub>2</sub> H <sup>b,j</sup>	2.54	2.54	7.7 (0.2)	6.4 (0.7)	0.3
		CO <sub>2</sub> <sup>-k,l</sup>	2.25	2.25	8.9 (0.7)	5.1 (0.2)	0.9
(6)	+NH <sub>3</sub>	CO <sub>2</sub> <sup>-m</sup>	3.17	2.53	6.84 (0.17)	6.52 (0.17)	0.07
		CO <sub>2</sub> H	3.20	2.72	6.56 (0.32)	6.34 (0.32)	0.08
		CO <sub>2</sub> <sup>-</sup>	2.79	2.30	6.99 (0.11)	6.78 (0.11)	0.03
(7)	NH <sub>2</sub>	CO <sub>2</sub> H <sup>n,o</sup>	2.694	2.640	7.43 (0.02)	6.24 (0.02)	0.03
(8)	AcO	+NMe <sub>3</sub> <sup>p</sup>			6.93 (-)	2.35 (-)	

<sup>a</sup> Magnetically equivalent; treated as A<sub>2</sub>B<sub>2</sub>. <sup>b</sup> <sup>2</sup>J<sub>AA</sub> = <sup>2</sup>J<sub>BB</sub> = -10.0 Hz (assumed). <sup>c</sup> <sup>2</sup>J<sub>FH</sub> = 47.37 (0.02) Hz; <sup>3</sup>J<sub>FH</sub> = 29.82 (0.02) Hz. <sup>d</sup> <sup>2</sup>J<sub>AA</sub> = -13.85 (0.06) Hz; <sup>2</sup>J<sub>BB</sub> = -10.0 Hz (assumed). <sup>e</sup> <sup>2</sup>J<sub>FH</sub> = 46.61 (0.09) Hz; <sup>3</sup>J<sub>FH</sub> = 27.59 (0.05) Hz. <sup>f</sup> <sup>2</sup>J<sub>AA</sub> = -12.99 (0.07) Hz; <sup>2</sup>J<sub>BB</sub> = -10.0 Hz (assumed). <sup>g</sup> <sup>2</sup>J<sub>AA</sub> = -11.43 (0.10) Hz; <sup>2</sup>J<sub>BB</sub> = -10.0 Hz (assumed). <sup>h</sup> <sup>1</sup>J<sub>CH</sub> = 134.7 Hz. <sup>i</sup> <sup>1</sup>J<sub>CH</sub> = 145.8 Hz. <sup>j</sup> <sup>1</sup>J<sub>CH</sub> = 130.1 Hz. <sup>k</sup> <sup>2</sup>J<sub>AA</sub> = <sup>2</sup>J<sub>BB</sub> = -15.0 Hz (assumed). <sup>l</sup> <sup>1</sup>J<sub>CH</sub> = 130.1 Hz. <sup>m</sup> Values from ref. 17. <sup>n</sup> This study; CDCl<sub>3</sub> solution. <sup>o</sup> <sup>2</sup>J<sub>AA</sub> = <sup>2</sup>J<sub>BB</sub> = -12.0 Hz (assumed). <sup>p</sup> Values from ref. 10.

and one of the carboxy oxygen atoms. At the STO-3G level, the energy of the *gauche* form did not converge to a minimum, since as the hydrogen approached the carboxy group the energy became correspondingly lower, even at unreasonably short hydrogen bond lengths. Since the preferred form of β-alanine in the vapour phase is not expected to be the zwitterion, but the neutral molecule, as is the case for glycine,<sup>21</sup> it appears that the program was trying to transfer the hydrogen to the carboxy group to give the neutral molecule. A full optimisation was therefore not possible, and a single-point calculation was used at the STO-3G level, with MNDO-optimised geometry for the *gauche* form. Also for the ethylenediammonium dication (3), the succinate dianion (5), and the 3-chloropropionate anion (4) no minimum could be found for the *gauche* forms when using MNDO. At the STO-3G level again no minimum could be found for the first two, but a minimum was found for the chloropropionate anion. In cases where there is no minimum the energy was calculated by holding the XCCY torsion angle at 60°, and optimising the XCC and CCY bond angles only.

The MNDO calculations were all performed with a VAX 11/780 computer. The STO-3G calculations were performed by using the program GAUSSIAN 76 (ref. 22) with a CDC 7 600 machine.

## Results

The observed n.m.r. parameters for the full set of compounds are given in Tables 1 and 2, together with the value for acetylcholine. This exists predominantly in the *gauche* form in aqueous solution and is given for comparison. The n.m.r. parameters for (1) and (2) are in agreement with previous literature values.<sup>12</sup> The results in Table 2 show clearly that protonation of the amino groups in (10) causes a significant increase in all the neighbouring vicinal couplings, whereas ionisation of the carboxy groups of (11) results in no general effect, the *ax-ax* couplings decreasing slightly and the *eq-ax* coupling increasing. The alternative possibility that ionisation is changing the conformer equilibrium of (10) may be eliminated on the basis of the known conformer preference of (10) (assuming additive *A* values, the *eq*, *eq*-conformer is stabilised

**Table 2.** Vicinal <sup>3</sup>J<sub>HH</sub> couplings in *trans*-1,2-disubstituted cyclohexanes

		$J_{1a-2a}$	$J_{2a-3a}$	$J_{2a-3c}$
(10)	X = NH <sub>2</sub> <sup>a</sup>	9.25	11.09	4.00
	X = +NH <sub>3</sub>	10.25	11.45	4.34
(11)	X = CO <sub>2</sub> H <sup>b</sup>	11.4	11.4	3.4
	X = CO <sub>2</sub> <sup>-</sup>	11.1	11.1	3.7

<sup>a</sup> <sup>1</sup>J<sub>CH</sub> 135.0 Hz. <sup>b</sup> <sup>1</sup>J<sub>CH</sub> 132.7 Hz.

by *ca.* 2.8 kcal mol<sup>-1</sup>),\* also this would not explain the smaller (*ca.* 50%) increase in  $J_{2a-3a}$  relative to  $J_{1a-2a}$ , which is exactly what would be expected on the basis of additive substituent effects.

These results may be utilised in equations (i) and (ii) by simply assuming a proportionate increase in the rotamer couplings ( $N_i$  and  $N_g$ ). Table 3 gives the observed values of  $N$  together with the calculated values  $N_g$  and  $N_i$  and the derived rotamer energies ( $\Delta E_{g-i}$ ) from equations (i)–(iii). Also given in this Table are the calculated rotamer energies of the compounds at both the STO-3G and MNDO levels. Table 4 gives the STO-3G-calculated geometries of the compounds.

Table 3 reveals that in most cases the rotamer energies given by MNDO and STO-3G are in accord with predictions based on the electrostatic interactions between the two substituents. The uncharged amines have a preference for the *gauche* form, which may be due to the formation of an intramolecular hydrogen bond.<sup>8</sup> The charged 2-halogenoethylamines show a much stronger preference for the *gauche* form due to the electrostatic attraction between the negative halogen atoms and the +NH<sub>3</sub> group. For ethylenediammonium dication both MNDO and STO-3G predict a strong preference for the *trans* form, due to the large electrostatic repulsion of the two +NH<sub>3</sub> groups.

3-Chloropropionic acid is predicted to favour the *gauche*

\* 1 kcal = 4.184 kJ.

Table 3. Rotamer energies  $\Delta E_{g-t}$  for  $\text{CH}_2\text{XCH}_2\text{Y}$ 

	X	Y	$N_g$	$N_t$	$N_{\text{obs.}}$	$\Delta E_{\text{soln.}}$	$\Delta E_{\text{vap.}}$	
							MNDO	STO-3G
(1)	$\left\{ \begin{array}{l} \text{F} \\ \text{F} \end{array} \right.$	$\text{NH}_2$	8.78	17.68	9.72	-0.86	-0.49	-0.68
		$^+\text{NH}_3$	9.22	18.56	9.36	-2.1	-3.60	-3.82
(2)	$\left\{ \begin{array}{l} \text{Cl} \\ \text{Cl} \end{array} \right.$	$\text{NH}_2$	10.25	17.87	11.33	-0.66	-0.11	-0.53
		$^+\text{NH}_3$	10.76	18.76	10.93	-1.87	-2.55	-4.81
(3)	$\left\{ \begin{array}{l} \text{NH}_2 \\ ^+\text{NH}_3 \end{array} \right.$	$\text{NH}_2$	10.44	17.90	12.7	-0.1	-0.28	-1.29
		$^+\text{NH}_3$	11.48	19.69	15.6	0.42	7.60	10.82
(4)	$\left\{ \begin{array}{l} \text{Cl} \\ \text{Cl} \end{array} \right.$	$\text{CO}_2\text{H}$	11.13	17.98	12.40	-0.47	-0.11	-0.66
		$\text{CO}_2^-$	11.13	17.98	12.68	-0.32	4.87	5.85
(5)	$\left\{ \begin{array}{l} \text{CO}_2\text{H} \\ \text{CO}_2^- \end{array} \right.$	$\text{CO}_2\text{H}$	12.21	18.12	12.8	-0.9	0.86	-0.29
		$\text{CO}_2^-$	12.21	18.12	15.3	0.5	12.80	13.65
(6)	$\left\{ \begin{array}{l} ^+\text{NH}_3 \\ ^+\text{NH}_3 \\ \text{NH}_2 \end{array} \right.$	$\text{CO}_2^-$	11.33	18.01	13.36	-0.10	-13.03	-31.13
		$\text{CO}_2\text{H}$	11.33	18.01	12.90	-0.3		
		$\text{CO}_2^-$	11.89	18.9	13.77	-0.18		

Table 4. STO-3G-optimised geometries for  $\text{CH}_2\text{XCH}_2\text{Y}$ 

	X	Y	<i>gauche</i>			<i>trans</i>	
			XCC	CCY	XCCY	XCC	CCY
(1)	$\left\{ \begin{array}{l} \text{F} \\ \text{F} \end{array} \right.$	$\text{NH}_2$	110.0	110.3	63.5	109.6	110.5
		$^+\text{NH}_3$	109.7	110.1	52.7	110.4	109.3
(2)	$\left\{ \begin{array}{l} \text{Cl} \\ \text{Cl} \end{array} \right.$	$\text{NH}_2$	110.8	112.1	65.4	109.0	111.8
		$^+\text{NH}_3$	109.8	110.1	53.3	110.4	109.1
(3)	$\left\{ \begin{array}{l} \text{NH}_2 \\ ^+\text{NH}_3 \end{array} \right.$	$\text{NH}_2$	113.1	110.0	60.7	110.1	110.1
		$^+\text{NH}_3$	114.7	114.7	60 <sup>b</sup>	110.5	110.5
(4)	$\left\{ \begin{array}{l} \text{Cl} \\ \text{Cl} \end{array} \right.$	$\text{CO}_2\text{H}$	113.3	112.7	66.0	110.6	112.1
		$\text{CO}_2^-$	115.3	115.1	71.0	110.5	114.8
(5)	$\left\{ \begin{array}{l} \text{CO}_2\text{H} \\ \text{CO}_2^- \end{array} \right.$	$\text{CO}_2\text{H}$	112.9	112.9	56.8	111.3	111.3
		$\text{CO}_2^-$	117.9	117.9	60 <sup>b</sup>	114.0	114.0

<sup>a</sup> XCCY 180° for all *trans* compounds. <sup>b</sup> Did not converge to a minimum.

form by both MNDO and STO-3G. Surprisingly, succinic acid is also predicted to favour slightly the *gauche* form by STO-3G, whilst MNDO predicts a preference for the *trans* form, as expected on the basis of electrostatic and steric interactions. For the charged acids both methods predict the expected destabilisation of the *gauche* form due to electrostatic repulsion between the two negatively charged groups.

For  $\beta$ -alanine there is predicted to be a large electrostatic attraction between the  $\text{NH}_3^+$  group and the  $\text{CO}_2^-$  group. Not surprisingly therefore, both MNDO and STO-3G predict a strong preference for the *gauche* form. In particular, STO-3G predicts stabilisation of the *gauche* form by some 30 kcal, suggesting that this is the only form which needs to be considered in terms of the conformation of this molecule.

The optimised geometries shown in Table 4 are also interesting. For the compounds where the *gauche* form is favoured the XCC and CCY bond angles are smaller than the 'normal' value of 111.0° for C-C-C angles, due to the large electrostatic attraction between X and Y. In those cases where excess of charge adds to this interaction the XCCY torsion angles are also smaller than 60°. Conversely, in the cases where there is a large electrostatic repulsion between X and Y the XCC and CCY bond angles have been 'opened out' in both the *trans* and (especially) the *gauche* forms. Note that in the cases where there is a very large electrostatic repulsion due to the addition of excess of charge, the XCCY torsion angle does not converge to a minimum. These calculations therefore suggest

that the *gauche* form of these compounds does not even exist as a local minimum in the vapour phase.

## Discussion

Most of the major trends in this set of compounds are reproduced by both the experimental solution-phase energies and the vapour-phase energy calculations. Thus, the 2-halogenoethylamines have a preference for the *gauche* conformation in both neutral and charged states. The effect of the addition of excess of charge is to increase the electrostatic attraction between the halogen and the amino group, and hence increase the population of the *gauche* conformation. Interestingly, STO-3G predicts that this increase in energy due to the addition of excess of charge is greater for the chloro-substituted amine than for the fluoro-substituted case, whereas MNDO predicts the opposite trend. In solution the increased stabilisation of the *gauche* form is about 1.2 kcal mol<sup>-1</sup> in both cases. The increased stabilisation of the *gauche* conformation is much smaller in solution than it is in the vapour phase, presumably owing to the much greater solvation of the  $^+\text{NH}_3$  group effectively shielding the excess of charge.

Ethylenediamine is predicted to form a strong intramolecular hydrogen bond by STO-3G, giving a marked preference for the *gauche* conformation. This hydrogen bond is not predicted by the MNDO method and is not observed in solution where the

energy difference between the two conformations is only very slightly in favour of the *gauche* form. On going to the charged amine the dominant feature is expected to be the electrostatic repulsion between the two  $^+NH_3$  groups, leading to a much greater population of the *trans* form. This is indeed the case, the vapour-phase calculations showing a very large shift to the *trans* conformation, and, likewise, the solution-phase energy difference showing an increased preference for the *trans* conformation of about 0.5 kcal mol<sup>-1</sup>. Once again, although the solution-phase energy differences show the expected change on going to the charged species, this is very much smaller than that given by the vapour-phase calculations. For the acids a shift towards a higher population of the *trans* conformation on going to the anion is predicted by the vapour-phase calculations. In solution, succinic acid shows this trend, the *trans* conformer gaining in energy on ionisation, but to a much smaller extent than predicted. 3-Chloropropanoic acid shows a preference for the *gauche* form, possibly due to interaction between the negative halogen atom and the positive carbon of the carboxy group. Surprisingly however there is no evidence of any significant increase in the population of the *trans* conformation in the anion. Again the vapour-phase calculations considerably exaggerate the extra stability of the *trans* form observed in the anions in aqueous solution.

The stability of the *gauche* conformation of un-ionised succinic acid in solution is interesting in that this suggests that there is an attractive interaction between the two carboxy groups. However, the vapour-phase energy differences show no evidence of such an interaction, STO-3G giving a much smaller preference for the *gauche* form and MNDO a preference for the *trans* form. Furthermore, the geometry of the CO<sub>2</sub>H group precludes the formation of any intramolecular hydrogen bond between the acid groups. It is possible that the large experimental error in this result (the r.m.s. error of coupling constants = 0.9 Hz, giving an error of 0.5 kcal in the rotamer energy difference) may be responsible.

To provide further information on this subject, it was decided to look at the <sup>1</sup>H n.m.r. spectrum of methyl hydrogen succinate. In this compound there is only one carboxy group, and hence any intramolecular hydrogen bonding could only involve a single hydrogen bond in both the *gauche* and the *trans* forms. Unfortunately, in D<sub>2</sub>O the spectrum of methyl hydrogen succinate appears as a singlet for both the neutral acid and the carboxylate anion, owing to accidental magnetic equivalence of the two CH<sub>2</sub> groups. In CDCl<sub>3</sub> solution the chemical-shift separation of the two resonances is sufficient to allow analysis of the spectrum.

The spectral parameters obtained (Table 1) correspond to a value of -0.25 kcal mol<sup>-1</sup> for  $\Delta E$ . Thus, the preference for the *gauche* form is smaller than that for succinic acid, and is close to the STO-3G energy difference for neutral succinic acid. This observation is consistent with both hypotheses for the high population of the *gauche* form in succinic acid.

## Conclusions

The energy differences obtained in this study confirm and illustrate the premise that in theoretical calculations of the conformer energies of polar and charged molecules (*i.e.* in the vapour phase) electrostatic effects are predominant.

The calculated conformer energy differences for the uncharged molecules at the STO-3G level are in generally good agreement ( $\leq$  ca. 0.5 kcal mol<sup>-1</sup>) with the corresponding observed values for aqueous solution, even though no account has been taken of solvation effects. The only exception to this is ethylenediamine (calculated -1.3 versus observed -0.1 kcal mol<sup>-1</sup>); competitive hydrogen bonding with water could well be the reason for the discrepancy.

However, when the charged species are considered the results are very different. Even though in most cases the calculated conformer energy differences are in the same sense as the observed values, the actual energy differences are calculated to be so much greater than the observed values that one must question the appropriateness of attempting theoretical calculations of the conformer energies of charged species without an explicit account of solvation effects. This is particularly noteworthy for those molecules with two charged groups. For example in ethylenediammonium dication, succinate di-anion, and  $\beta$ -alanine zwitterion the calculated conformer energy differences are between 10 and 20 kcal mol<sup>-1</sup>, whereas the observed ones are an order of magnitude smaller (0.4, 0.5, and -0.1 kcal mol<sup>-1</sup>, respectively). The ability of the water solvent to 'neutralise' the charged groups' electrostatic effects by solvation is thus seen to be a dominant factor in determining the observed conformer energies.

This is an important point, because even a qualitative reliance on the calculated conformer energies in such cases would lead to incorrect conclusions. The calculated conformer energies would suggest that all these molecules only exist in one conformation in solution (the *trans* conformer for ethylenediammonium dication and succinate dianion, and the *gauche* conformer for  $\beta$ -alanine zwitterion), whereas in fact in all three cases the conformer energy differences are so small in aqueous solution that a significant fraction of both conformers is present. This may be considered a general phenomenon for any molecule with two or more charged groups exposed to the aqueous solvent, and illustrates the considerable problems yet to be solved in molecular modelling studies of compounds containing ionisable groups.

## Acknowledgements

We thank Mr. I. W. A. Whitcombe for the <sup>13</sup>C satellite spectra at 400 MHz, Miss J. Fisher for assistance with the WM 250 spectra, Roche Products Limited for the use of their computer graphics and n.m.r. facilities, and the University of Liverpool Computer Department for computational facilities, and acknowledge an S.E.R.C. CASE award (to B. D. H.) and grant to purchase the WM 250 spectrometer.

## References

- 1 A. Pullman and B. Pullman, *Quart. Rev. Biophys.*, 1975, **7**, 505.
- 2 J. M. Hammersley and D. C. Handscomb, 'Monte Carlo Methods,' Methuen, London, 1964.
- 3 R. J. Abraham and E. Bretschneider, 'Internal Rotation in Molecules,' ed. W. J. Orville-Thomas, Academic Press, London, 1976, ch. 13.
- 4 G. Wipff, A. Dearing, P. K. Weiner, J. M. Blaney, and P. A. Kollman, *J. Am. Chem. Soc.*, 1983, **105**, 997.
- 5 W. G. Richards, 'Quantum Pharmacology,' Butterworths, London, 1979.
- 6 R. J. Abraham and K. G. R. Pachler, *Mol. Phys.*, 1963, **7**, 165.
- 7 R. J. Abraham and R. Stolevik, *Chem. Phys. Lett.*, 1981, **77**, 181.
- 8 L. Radom, W. A. Lathan, W. J. Hehre, and J. A. Pople, *J. Am. Chem. Soc.*, 1973, **95**, 693.
- 9 R. J. Abraham and D. Birch, *Mol. Pharmacol.*, 1975, **11**, 663.
- 10 P. Partington, J. Feeney, and A. S. V. Burgen, *Mol. Pharmacol.*, 1972, **8**, 269.
- 11 A. Pullman and G. N. J. Port, *Theor. Chim. Acta (Berlin)*, 1973, **32**, 77.
- 12 P. Levins and Z. B. Papanastassiou, *J. Am. Chem. Soc.*, 1985, **87**, 826.
- 13 N. Sheppard and J. J. Turner, *Proc. Roy. Soc., A*, 1959, **252**, 506.
- 14 R. J. Abraham, 'Analysis of High Resolution N.m.r. Spectra,' Elsevier, Amsterdam, 1971.
- 15 PANIC: Bruker n.m.r. software. Version for ASPECT 2000.
- 16 R. J. Abraham and G. Gatti, *J. Chem. Soc. B*, 1969, 961.

- 17 A. G. Massey, E. W. Randall, and D. Shaw, *Spectrochim. Acta*, 1964, **20**, 379.
- 18 N. S. Zefirov, L. G. Gurvich, A. S. Shashkov, M. Z. Krimer, and E. A. Vorobeva, *Tetrahedron*, 1976, **32**, 1211.
- 19 J. L. Leviel, G. Auvert, and J. M. Savariault, *Acta Crystallogr., Sect. B*, 1981, **37**, 2185.
- 20 M. J. S. Dewar and W. Thiel, *J. Am. Chem. Soc.*, 1977, **99**, 4899, (Quantum Chemistry Program Exchange no. 428).
- 21 R. D. Suenram and F. J. Lovas, *J. Mol. Spectrosc.*, 1978, **72**, 372.
- 22 GAUSSIAN 76 (Quantum Chemistry Program Exchange no. 368).

*Received 10th January 1986; Paper 6/080*