Intercharge Distance of Flexible Zwitterionic Molecules in Solution

Yves Chevalier* and Pierre Le Perchec

Laboratoire des Matériaux Organiques, CNRS, BP 24, 69390 Vernaison, France (Received: May 17, 1989)

The conformation of flexible zwitterionic molecules in solutions of polar solvents is studied by means of a NMR chemical shift method. This method makes use of chemical shift changes of NMR lines, induced by the electrostatic field caused by the electrical charges of the zwitterion. The electrostatic field at the observed nucleus is calculated, taking into account the presence of a reaction field due to the polarization of solvent molecules around the solute ones. The intercharge distances of zwitterionic (trimethylammonio)alkanoates in aqueous and methanolic solutions are then deduced from ^{13}C NMR chemical shift data as a function of the number of methylenes linking the cationic and anionic groups of the zwitterion. The comparison of experimental data with end to end distances predicted by the rotational isomerism state theory shows that the polymethylene chain has a more folded conformation than a free chain. Electrostatic attraction between the two zwitterion charges is thus important, although it could be expected that it could be offset by steric hindrance, i.e., by the bulkiness of terminal charged groups.

Introduction

The dipole moment of flexible zwitterionic molecules in solution depends mainly on the conformation of the flexible arm that links the two terminal charged groups. Because of electrostatic attraction between the two charged end groups, the dipole moment may be considerably reduced upon folding of this intercharge arm. Its value then depends strongly on arm flexibility (energy barriers between rotational trans and gauche isomers[1]), bulkiness and hydration of charged groups preventing their close approach, solvent and arm electrical properties which control electrostatic attraction between the two opposite charges, and polarization of solvent around the molecule caused by the dielectric discontinuity between solvent and solute molecule interior (image charge effect[2-4]).

Basically, electrostatic attraction by opposite charges of the zwitterion may be offset by (1) intercharge arm rigidity making trans conformations with the largest intercharge distances more likely and (2) excluded-volume effects (steric hindrance) of charged groups, where charge separation cannot be less than the closest approach distance depending on group size. With this effect, space regions where electrostatic attraction is the strongest (smallest intercharge distances) are forbidden. If charged groups are hydrated, excluded volume is enlarged by hydration water volume; closer approach requires water removal from charged groups, which is allowed if electrostatic energy is larger than dehydration energy. Lastly, electrostatic attraction is controlled by dielectric properties of the medium; if water is the solvent, its high dielectric constant should reduce electrostatic attraction. The dielectric constant of the intercharge arm is also of importance, first because it is expected to be located between the charges and also because dielectric constant differences between molecule interior and exterior (solvent) cause solvent polarization at the dielectric discontinuity and thus a supplementary electric field called reaction field.[2-4] The influence of dielectric constants is then complex, depending in particular on the location of the dielectric discontinuity, i.e., on molecular shape, solvation, etc.

Methods for direct dipole moment measurements are few. The most common one makes use of dielectric measurements, dielectric constant ϵ (for bulk), or dielectric increment $d\epsilon/dC$ (for solutions), which are related to dipole moment by electrostatic theories of Onsager,[5] Böttcher,[6] and Buckingham.[7,8] These theories only hold for dilute dipoles immersed in a continuous dielectric. When measurements are performed in a solvent (especially polar ones such as water), dielectric properties inside and outside the molecule

are different and solvent polarization around the solute molecules occurs. In calculating the electric field inside the molecule, this effect is accounted for in theories by assuming the dipolar solute molecule is located in a dielectric cavity immersed in a continuous dielectric (the solvent). This is certainly incorrect at a molecular level but was shown to work well in simple cases. These theories also account for long-range electrostatic dipole–dipole interactions which cause intermolecular correlations. Problems occur when a dipolar solvent is used and when short-range correlations occur. Thus, dipolar solvent molecules are oriented around the solute, contributing to the macroscopic dielectric constant. This effect could be accounted for with bulk liquids[7] or when solvent is of rather low polarity,[8] but it is much more difficult with highly polar solvents (water). Accounting for short-range correlations requires the knowledge of the short-range potential around the molecule which depends on many types of interactions such as steric, dispersion, and even electrostatic quadrupole[9] ones. Such correlations are obviously present in bulk liquids, so that dielectric measurements are more used for structural investigations (intermolecular correlation studies) than for dipole moment evaluation.[3,10] In order to get rid of intermolecular correlations, dilute solutions are usually chosen and extrapolation to infinite dilution provides correct interaction-free data. Here is the major problem: when molecules are strongly self-associating, the required concentration range is too low for precise measurements to be performed. It is, moreover, difficult to decide whether the concentration is low enough for dielectric increment data to be considered as interaction free. Self-association occurs in many cases: at very low concentration for solutions of dipolar molecules in nonpolar solvents because dipole–dipole attractions are so strong; it occurs also in polar solvents (especially water) because polar molecules that are investigated often contain some hydrophobic part. These limitations explain why measurements of dipole moments of α,ω-amino acids in water are so contradictory.[11-16] Salefran has, moreover, provided evidence of self-association with such compounds.[15]

It is much easier to measure a quantity A than its slope dA/dC at infinite dilution. There is thus a considerable improvement in designing a method of dipole moment evaluation that requires measurement of an observable quantity (extrapolated at infinite dilution). This is what is proposed in the present paper.

A new method based on NMR chemical shift data extrapolated to infinite dilution is used in order to overcome these difficulties.

(1) Darsey, J. A.; Rao, B. K. *Macromolecules* **1981,** *14,* 1575.
(2) Böttcher, C. J. F. *Theory of Electric Polarisation*; Elsevier: Amsterdam, 1952.
(3) Fröhlich, H. *Theory of Dielectrics*; Clarendon: Oxford, 1949.
(4) Friedman, H. L. *Mol. Phys.* **1975,** *29,* 1533.
(5) Onsager, L. *J. Am. Chem. Soc.* **1936,** *58,* 1486.
(6) Böttcher, C. J. F. *Rec. Trav. Chim.* **1943,** *62,* 119.
(7) Buckingham, A. D. *Aust. J. Chem.* **1953,** *6,* 93.
(8) Buckingham, A. D. *Aust. J. Chem.* **1953,** *6,* 323.

(9) Patey, G. N.; Levesque, D.; Weis, J. J. *Mol. Phys.* **1979,** *38,* 1635.
(10) Kirkwood, J. G. *J. Chem. Phys.* **1939,** *7,* 911.
(11) Kuhn, W. *Z. Phys. Chem.* **1935,** *A175,* 1.
(12) Edward, J. T.; Farrell, P. G.; Lob, J. L. *J. Am. Chem. Soc.* **1974,** *96,* 902.
(13) Kirchnerova, J.; Farrell, P. G.; Edward, J. T. *J. Phys. Chem.* **1976,** *80,* 1974.
(14) Walder, J. A. *J. Phys. Chem.* **1976,** *80,* 2777.
(15) Salefran, J. L. *Adv. Mol. Relax. Interact. Processes* **1981,** *19,* 75.
(16) Kaatze, U.; Bieler, H.; Pottel, R. *J. Mol. Liq.* **1985,** *30,* 101.
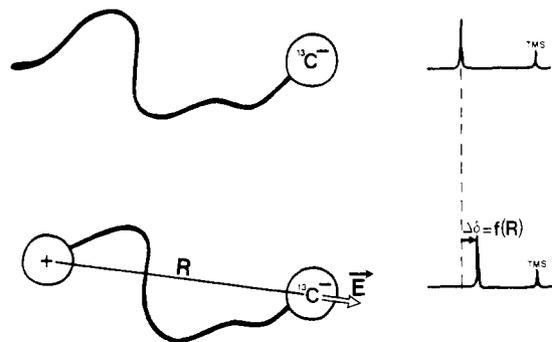
Figure 1. Principle of the NMR chemical shift method illustrated with a schematic picture of solutes and intramolecular electrical fields. Zwitterionic (trimethylammonio)alkanoate molecule and its sodium alkanoate monocharged homologue, with the corresponding NMR spectra of the $-^{13}CO_2^-$ nucleus.

It is applied to flexible zwitterionic molecules in polar solvents, namely, water and methanol. The quantity to be measured is the influence of the electrical field created by one electrical charge on the NMR chemical shift of nuclei located close to the other one. If the electric field can be evaluated, an intramolecular quantity is worked out and intercharge distances can be obtained in a rather direct way. The fact that an intramolecular property is measured, better than a macroscopic one, is also an advantage as solvent polarization only contributes to it through the reaction field.

In this paper, the principle of the method and the related basic theory are presented. They are further applied to a series of (trimethylammonio)alkanoate zwitterionic molecules (($(CH_3)_3^+$-$N(CH_2)_nCO_2^-$) of variable-length flexible polymethylene intercharge arm ($n = 1-7$), as solutions in water or methanol.

Principle of the Method

NMR chemical shift differences are due to different electron densities around the nucleus. The presence of an electrical field should displace electrons and thus influence the chemical shift value.[17] Thus, introducing an electrical charge into a molecule creates an electrostatic field that shifts all NMR resonance frequencies of nuclei in it. It is clear that the more intense the field the larger the induced shift, so that nuclei located close to the introduced electrical charge should be more shifted that those located away from it. If a quantitative estimation of the electrical field is possible, intramolecular distances can be measured. Such use of intramolecular electrical field effect for conformational analysis have been applied to some well-suited systems.[18] This is the basic feature of the method, schematically shown in Figure 1. As applied to the measurement of intercharge distance of a zwitterion, the influence of the electrical field due to one charge on the NMR resonance frequency of a nucleus located near the other one is to be measured. This is obtained by measuring the chemical shift difference between two homologous compounds, one zwitterionic having both its charges and the other one having one charge left, all the remaining part of molecule being kept constant.

Basic Theory

NMR chemical shifts depend on electrical field at the nucleus mainly because of polarization of chemical bonds around the atom. As electrons are displaced by the electric field, an additional shielding or deshielding occurs that gives rise to a shift of the resonance frequency of the nucleus. This effect is proportional to the electric field and depends on bond polarizabilities and orientations relative to the field. It is clear that large effects are expected when polarizable bonds are present and when they act all together in the same way. As a typical example of a vanishing effect due to equal contributions of bonds acting in opposite

directions in a uniform field, there is a minor effect of electric fields on ^{13}C chemical shifts of methylenes in polymethylene moieties because polarization of one C–C (or C–H) bond brings electrons to the central nucleus while the other bond abstracts the same charge because it has the same polarizability.[19,20] The distortion of p orbitals by the electric field provides a quadratic dependence on the electric field. For a uniform electric field, the total electric field dependence of the chemical shift then reads[17,21]

$$\delta = \delta_0 + \sum_i A_i E \cos \theta_i + BE^2 \qquad (1)$$

with

$$A_i = b_i \delta_e / el_i \qquad (2)$$

where the summation is over all chemical bonds adjacent to the observed atom. θ_i is the angle between bond i and E. A constant A_i contains the bond polarizability along itself b_i (denoted b_1 in Denbigh[22] or Le Fèvre[23] papers), bond length l_i and the shift increment per unit charge δ_e ($\delta_e = 160-180$ ppm/electron for ^{13}C);[24,25] e is the electronic charge. As the electric field E decreases with distance, the linear dependence is the predominant term for large separations between the charge and the observed nucleus; that is, the quadratic term is neglected when the distance is large enough. This is shown to be the case by Spiesecke:[24] while the fact that hydrogen electrons are in s orbitals makes the quadratic term negligible in ^1H NMR,[26] the ^{13}C and ^1H chemical shifts of a series of molecules have the same measured (linear) dependence on electric field. Because the electric field is not uniform (it decreases with distance), an additional contribution related to the field gradient may appear.[19,20] This effect was shown to be negligible when the observed nucleus is more than two chemical bonds from the electrical charge.[19,20] These approximations are also justified by theoretical calculations showing a linear dependence of chemical shift on intramolecular electrical field.[27] The nucleus is chosen such that the polarizability is large and that no cancellation of electron displacement can occur; the linear term is then large and predominant. In the following, the dependence will be taken as linear with respect to the electric field E. The constants A_i can in principle be calculated from bond polarizability tables,[22,23] but this procedure gave a poor agreement with experimental data in some cases.[18] A better way is thus to use a rigid compound of known structure for a measurement of the $\sum A_i \cos \theta_i$ factor which is required for evaluation of electrostatic field from chemical shift data.

For an homogeneous medium, the electric field E originating from a charge e is given by Coulomb's law

$$E = e / (4\pi\epsilon_0 \epsilon R^2) \qquad (3)$$

E then depends on inverse square distance which is to be measured, and the dielectric constant ϵ is unambiguously defined. In this case, provided the constants ϵ and A_i are known or measured from a rigid molecule, R is obtained from the chemical shift difference between the molecule and an uncharged homologue. This is what could be called the zero-order theory.

It does not hold in the case of a molecule with low internal dielectric constant ($\epsilon = 2-10$) dissolved in water of high dielectric constant ($\epsilon = 78$). Indeed, the reaction field with its complex dependence on molecular size and shape is expected to be of importance. It is, however, quite different from that of dielectric measurements because the main electric field is an internal one. The polarization of solvent molecules around the solute only influences the chemical shift through the electrostatic field it

(17) Buckingham, A. D. *Can. J. Chem.* **1960**, *38*, 300.
(18) Batchelor, J. G.; Prestegard, J. H.; Cushley, R. J.; Lipsky, S. R. *J. Am. Chem. Soc.* **1973**, *95*, 6358.
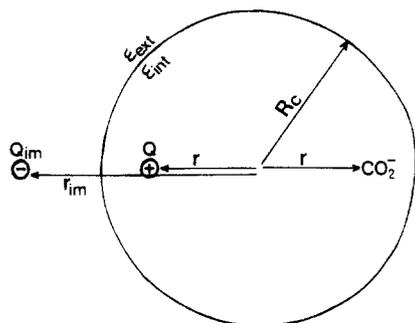
(19) Batchelor, J. G. *J. Am. Chem. Soc.* **1975**, *97*, 3410.
(20) Batchelor, J. G.; Feeney, J.; Roberts, G. C. K. *J. Magn. Reson.* **1975**, *20*, 19.
(21) Horsley, W. J.; Sternlicht, H. *J. Am. Chem. Soc.* **1968**, *90*, 3738.
(22) Denbigh, K. G. *Trans. Faraday Soc.* **1941**, *36*, 936.
(23) Le Fèvre, R. J. W. *Adv. Phys. Org. Chem.* **1965**, *3*, 1.
(24) Spiesecke, H.; Schneider, W. G. *Tetrahedron Lett.* **1961**, *14*, 468.
(25) Nelson, G. L.; Levy, G. C.; Cargioli, J. D. *J. Am. Chem. Soc.* **1972**, *94*, 3089.
(26) Del Re, G.; Pullman, B.; Yonezawa, T. *Biochim. Biophys. Acta* **1963**, *75*, 153.
(27) Seidman, K.; Maciel, G. E. *J. Am. Chem. Soc.* **1977**, *99*, 3254.

Figure 2. Electrical model for the calculation of the reaction field: a spherical cavity of low dielectric constant ϵ_{int} containing the molecule with the charge Q of the zwitterion, immersed in a continuous medium of high dielectric constant ϵ_{ext}. Image charge Q_{im} of Q is located at a distance r_{im} from the cavity center.

creates. The relevant reaction field value is that at the observed nucleus. This nucleus is close to one of the two charged end groups which is hydrophilic and solvated because of its ionic character, and thus close to solvent. The problem is the choice of the dielectric cavity model and location of the charge and observed nucleus in this cavity. This is a difficult task, especially with flexible molecules, but must be done correctly since reaction field and consequently chemical shifts are very sensitive to it.[28,29] The cavity of low dielectric constant contains the molecule and may also contain some water that is immobilized around the solute (hydration water). Indeed, when water is not allowed to move as freely as in bulk, its response to an electric field is altered, which can be accounted for by reducing its dielectric constant. The dielectric constant of unrotating water is that measured at frequencies higher than the dielectric relaxation frequency of bulk water; its value of 4.5[30] is actually very low, the polarization mechanism then being distortion of the molecule.[31] Once this choice has been made, a mathematical problem remains and a judicious choice of molecular geometric parameters is advisable for the computation difficulty to be reasonable. Actually, this choice is difficult enough to decide for only simple geometric shapes to be retained, so that only ellipsoids have been used for calculations. The expected strong dependence of electric field on molecular shape and volume makes the choice of the uncharged homologue crucial; it has to be of the same volume and shape.

A first order of magnitude of the reaction field can be obtained by modeling the system as a spherical dielectric cavity of radius R_c in a continuous dielectric (Figure 2). When the charge is not close to the dielectric discontinuity, the reaction field is well approximated as that created by a fictitious image charge whose location r_{im} and charge Q_{im} are given by[2,4]

$$r_{im} = R_c^2/r \qquad (4)$$

$$Q_{im} = -Q\frac{\epsilon_{ext} - \epsilon_{int}}{\epsilon_{ext} + \epsilon_{int}}\frac{R_c}{r} \qquad (5)$$

Applying eq 4 and 5 to a model system, in which both charge Q and the observed nucleus are located on a diameter, symmetrically from the sphere center (at r), the total electrical field originating from Q, as experienced at the observed nucleus, is

$$E = \frac{e}{4\pi\epsilon_0\epsilon_{int}4r^2}\left[1 - \frac{\epsilon_{ext} - \epsilon_{int}}{\epsilon_{ext} + \epsilon_{int}}\frac{4R_c/r}{(1 + R_c^2/r^2)^2}\right] \qquad (6)$$

The first simple dependence to be noticed is that of dielectric constants. Taking $\epsilon_{int} = 2$ as for hydrocarbons, E is unsensitive to ϵ_{ext} as long as $\epsilon_{ext} \gg \epsilon_{int}$. In the present case, the reaction field is only reduced by 7% upon substitution of deuterated water ($\epsilon = 76$) by CD_3OD ($\epsilon = 32$). The exact value of ϵ_{int} is also of minor importance for the $(\epsilon_{ext} - \epsilon_{int})/(\epsilon_{ext} + \epsilon_{int})$ term when $\epsilon_{int} \ll \epsilon_{ext}$,

so that the calculated electric field is inversely proportional to ϵ_{int}. $\epsilon_{int} = 2$ has been chosen in the following, this value being very important and difficult to estimate. Only relative variations were worked out (a rigid compound of known structure was used for calibration of the $\sum A_i \cos \theta_i$ factor).

The reaction field exclusively depends on the ratio R_c/r. The dependence on R_c/r is very sharp when R_c is close to r, which means that the choice of $r - R_c$ is then more crucial than that of R_c itself (the cavity volume). Equations 4–6 are not valid when r is close to R_c, but the same trend is observed with an exact theory.[2,4] In our case the trimethylammonio group is bulky, and its central positive charge is not close to the cavity edge. A close approach of the charge and the cavity edge has thus to be avoided in the choice of the cavity model. If R_c is chosen in such a way that the cavity volume is equal to the molecular volume, r may be found larger than R_c for fully extended zwitterions; that is, the charged groups may be located outside the model dielectric cavity. The consequence is that the intercharge distance $R = 2r$, as deduced from fitting the experimental data to this model calculation, is small for the charged groups to sit inside the cavity. This result is an artifact due to an unrealistic situation arising from the choice of the spherical shape and cavity volume. An alternative way is to place the charges at a fixed distance (d) from the cavity edge, d being estimated from bond lengths and van der Waals radii; thus, R_c is set as $r + d$. The cavity volume then largely exceeds the molecular one.

In a realistic choice of cavity shape and location of charges in it, the first parameter that should be controlled is the distance d between the charges and the cavity edge, in order to avoid their too close proximity. The cavity can then be chosen in such a way that both its volume and shape satisfactorily fit the molecule. Electric field calculations for ellipsoidal cavity shapes with the charge and observed nucleus being on the major axis (but not necessarily at the foci) have thus been carried out according to the method of Westheimer and Kirkwood[32] originally designed for calculating pK's of diacids. Calculation procedures described by Ehrenson[33] were used.

Experimental Section

Synthesis. The general synthesis of (trimethylammonio)alkanoates is quaternization of trimethylamine with ethyl ω-bromoalkanoates, followed by basic hydrolysis of the ester function.[34] Trimethylamine (25% aqueous solution) and ethyl ω-bromoalkanoates were refluxed for 15 h in methanol. After solvent evaporation, Br^- ions of the resulting ethyl ω-(trimethylammonio)alkanoates were exchanged for OH^- by percolation of their aqueous solutions through an anion-exchange resin (Amberlite IRA400, OH^- form). Basic hydrolysis of the ester group was then carried out by heating (50–70 °C) the aqueous solution in vacuo (in a rotative evaporator). Compounds were recrystallized first in dimethylformamide and then in acetone–ethanol (90/10) mixtures, yielding pure and salt-free compounds as white crystals.

(Trimethylammonio)propionate could not be obtained by this way because of the Hoffmann elimination reaction occurring in basic media, yielding trimethylammonium acrylate. Thus, a solution of ethyl 3-bromopropionate and dimethylamine (40% aqueous solution) was stirred for 3 days at room temperature in the presence of K_2CO_3. After filtration and rough acetone evaporation, ethyl (dimethylamino)propionate was extracted from water with ether. Quaternization with methyl iodide in acetone rapidly yielded ethyl (trimethylammonio)propionate iodide as a white precipitate which was hydrolyzed with excess HBr (refluxed for 12 h in 48 wt% HBr aqueous solution). The solution was evaporated and redissolved three times in water for efficient elimination of excess HBr, and (trimethylammonio)propionic acid iodide is precipitated from its concentrated aqueous solution by addition of acetone. (Trimethylammonio)propionate was then obtained by fast percolation of the aqueous solution on an an-

(28) Scholte, T. *Physica* **1949**, *15*, 437.
(29) Diehl, P.; Freeman, R. *Mol. Phys.* **1961**, *4*, 39.
(30) Grant, E. H.; Shack, R. *B. J. Appl. Phys.* **1967**, *18*, 1807.
(31) Hill, N. E. *Trans. Faraday Soc.* **1963**, *59*, 344.

(32) Westheimer, F. H.; Kirkwood, J. G. *J. Chem. Phys.* **1938**, *6*, 513.
(33) Ehrenson, S. *J. Am. Chem. Soc.* **1976**, *98*, 7510.
(34) Laughlin, R. G. US Patent 4,287,174, 1981.

TABLE I: Distance between Nitrogen Atom and Carboxylate Carbon Atom for Fully Extended Molecules R_{max}, θ_{C-C} Angle, Molecular Volume of (Trimethylammonio)alkanoates, and ^{13}C NMR Chemical Shift Data in Solvents D_2O and CD_3OD

$$(CH_3)_3{}^+N(CH_2)_nCO_2{}^-$$

			molecular	δ, ppm	
n	R_{max}, Å	θ_{C-C}, deg	vol, Å³	in D_2O	in CD_3OD
1	2.48	33.45	156.6	−15.24	−14.38
2	3.81	22.16	182.8	−7.64	−7.15
3	4.98	34.93	209.1	−3.65	−3.44
4	6.30	27.59	235.3	−1.83	−1.64
5	7.51	35.04	261.6	−0.94	−0.92
7	10.04	35.10	314.0	−0.36	−0.31

ion-exchange resin in its OH^- form and evaporation to dryness.

Distances, Angles, and Volumes. Interatomic distances and bond angles were estimated from crystallographic data on glycine[35] and geometry of polymethylene chains (1.265 Å/CH_2 group for a fully extended chain[36]). An odd and even effect is not negligible for small intercharge distances; it influences both R and θ_{C-C} (see text) values and is taken into account. Volumes are calculated from group increments according to the method of Cabani.[37] As the tabulated data of Cabani only deal with uncharged groups, data from volumes of electrolytes[38] were introduced into it. The volume of the trimethylammonio group is then obtained by subtracting to the volume of the tetramethylammonium cation, the methyl group increment, and the Traube covolume[39] A_z. The carboxylate group volume is calculated in the same way from that of acetate ion. All these molecular data are summarized in Table I.

NMR Measurements. ^{13}C NMR chemical shift measurements were performed with a Bruker WP80 working at 20.1 MHz under ^1H broad-band decoupling, and the temperature was 30 ± 1 °C. Two deuterated solvents of different dielectric constant were used: water (ϵ = 76) and methanol (ϵ = 32) in which internal chemical shift references were dioxane and tetramethylsilane (TMS), respectively. Spectra were run for all compounds as a function of concentration, and data were extrapolated to infinite dilution for chemical shifts to be independent of any correlation between solute molecules. Experimental data are chemical shift differences between the carboxylate ^{13}C of zwitterionic (trimethylammonio)alkanoates and the sodium alkanoates of the same volume. The volume increment of a trimethylammonio group in a molecule (95 Å³) best corresponds to that of a propyl group (84 Å³), so the chemical shift variation upon replacement of the terminal propyl group of sodium alkanoate by the charged trimethylammonio group was used in this study. The carboxylate chemical shifts of sodium pentanoate and its higher homologues are identical (117.80 ppm from dioxane in D_2O and 183.17 ppm from TMS in CD_3OD); the choice of the reference compound is thus not crucial in this particular case. Uncertainty in chemical shift differences is 0.02 ppm. Measurements have also been performed with (trimethylammonio)undecanoate (n = 10), but the concentration dependence of chemical shift is steep, so that extrapolation to infinite dilution is difficult; the chemical shift difference with the corresponding sodium alkanoate being small, the estimated uncertainty is then found of the order of 100%. This compound was thus discarded for the analysis. The reason for this difficulty may be self-association of a hydrophobic origin, occurring at low concentration. Experimental data are shown in Table I and Figure 3.

Results

Estimation of the $A = \sum A_i \cos \theta_i$ Factor. A is first estimated from a rigid compound of known geometry. (Trimethyl-

(35) Marsh, R. E. *Acta Crystallogr.* **1958**, *11*, 654.
(36) Tanford, C. *The Hydrophobic Effect*; Wiley: New York, 1973.
(37) Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. *J. Solution Chem.* **1981**, *10*, 563.
(38) Millero, F. J. *Chem. Rev.* **1971**, *71*, 147.
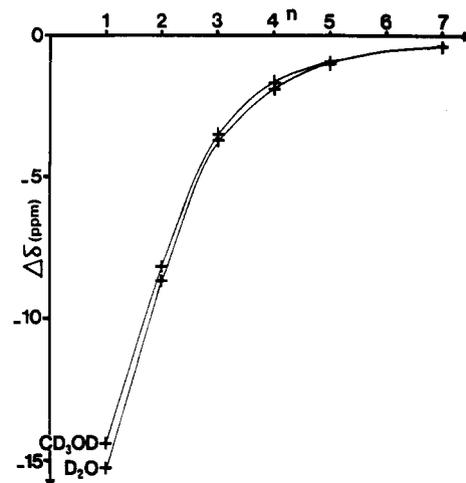(39) Traube, J. *Samm. Chem. Chem. Tech. Vortr.* **1899**, *4*, 255.

Figure 3. ^{13}C NMR chemical shift difference of carboxylate line between (trimethylammonio)alkanoates and sodium alkanoates as a function of the number of intercharge methylenes n in the zwitterion. Data are for D_2O and CD_3OD as solvents.
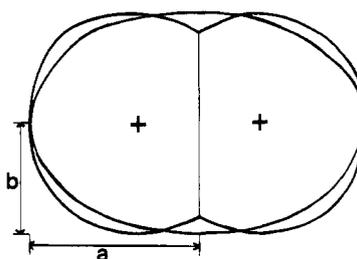

Figure 4. Comparison of molecular shape of $(CH_3)_3{}^+NCH_2CH_2CO_2{}^-\cdot 2H_2O$ and an ellipsoid having the same internal volume, its major axis being a = 5.15 Å and its minor axis being b = 3.36 Å. Positions of charge and observed nucleus are shown as crosses.

TABLE II: Geometrical Parameters and Longitudinal Polarizabilities around the Carboxylate Carbon Atom

bond	l, Å	θ, deg	$10^{24}b$, esu	$10^{40}b$, SI units
C=O	1.265		2.3	2.56
C-O⁻	1.261		0.89	0.99
C-C	1.523	22.16	0.99	1.10
-$CO_2{}^-$	1.263	157.84	1.46	1.62

ammonio)acetate (n = 1) is a good candidate, but there is some doubt about the validity of the approximations presented in the theoretical part of the paper. The observed ^{13}C nucleus is two bonds apart from the charged nitrogen atom, so that some contribution of the electric field gradient to the chemical shift may be expected,[19,20] through the bonds polarization (inductive effect) may also occur. This compound was thus eliminated and the used rigid reference is (trimethylammonio)propionate (n = 2). Although some gauche conformations are possible as observed on a sticks and balls model, steric hindrance due to the bulky trimethylammonio and hydrated carboxylate groups makes the trans conformation the only allowed one. In deciding the cavity shape and volume, the charge to edge distance d = 3.3 Å is set equal to the radius of a sphere having the volume of tetramethylammonium ion.[38] The observed ^{13}C nucleus is set symmetrically at 3.3 Å from the other side (Figure 4); a carboxylate groups is not bulky enough (13.3 Å³) to fill a end cap with such a radius (75 Å³ for a hemisphere). The excess volume is 62 Å³, which corresponds to two water molecules. The dielectric cavity is thus chosen as containing the molecule and two hydration water molecules bound to the carboxylate groups. It should be noted that two water molecules are indeed strongly bound to the carboxylate group of glycine.[40] The ellipsoid having such a volume and its major axis equal to $R + 2d$ has a minor axis of 6.7 Å, which

(40) Kokpol, S. U.; Doungdee, P. B.; Hannongbua, S. V.; Rode, B. M.; Limtrakul, J. P. *J. Chem. Soc., Faraday Trans. 2* **1988**, *84*, 1789.
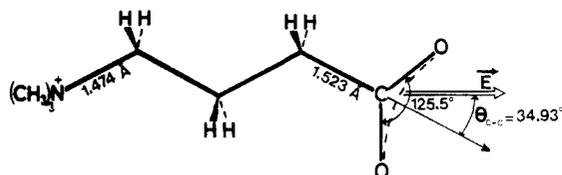
Figure 5. Geometry around the carboxylate group for the calculation of the A factor from longitudinal polarizabilities.

agrees quite well with the expected molecular dimension $(2d)$. The cavity fits well the molecule as shown in Figure 4.

From the electric field calculated at the carboxylate ^{13}C nucleus in such a cavity and the experimental chemical shift of -7.64 ppm, the value of A is found as -6.2×10^{-15} SI units in D_2O. $A = -5.8 \times 10^{-15}$ SI units is found in CD_3OD, which is close to the value obtained in D_2O.

A is then calculated from bond longitudinal polarizabilities and the geometry of the carboxylate group. Figure 5 and Table II summarize the values of bond lengths and angles[35] and longitudinal polarizabilities b_i from Le Fèvre.[23] δ_e was taken as 180 ppm/electron.[25] C=O and C–O$^-$ bonds are actually identical because of electron delocalization in the carboxylate group, so that the average value b_{C-O} was taken for each bond. The rotational motion around the symmetry axis of the CO_2^- group (the C–C bond) partially averages further the $\cos \theta_i$ factor. This geometrical average is performed by replacing the CO_2^- group by a single fictitious bond with an angle $\theta = \pi - \theta_{C-C}$ and of longitudinal polarizability

$$b_{CO_2^-} = (b_{C=O} + b_{C-O}) \cos (\angle(O-C-O)/2) \qquad (7)$$

This way of averaging is incorrect but does not introduce significant errors as compared to uncertainties on b_i and δ_e values. A is then found equal to -5.9×10^{-16} SI units. This disagrees with the experimental value by a factor of 10. Batchelor[18] also observed large discrepancies between A values as obtained from experiment, from longitudinal polarizabilities, or from CNDO calculations. The origin of this may be ascribed to uncertainties on estimation of ϵ_{int}, δ_e, and especially b_i values which are indeed controversial.[41] b_i values may, however, be correct. Allen[42] has indeed recently found $b_{C-C} = 1.00 \pm 0.09 \times 10^{-40}$ SI units in halogenated alkanes, which compares well with the tabulated value of Le Fèvre $(b_{C-C} = 1.1 \times 10^{-40}$ SI units).[23] What is here incorrect is probably the use of these longitudinal polarizabilities in estimating NMR shifts: ^{13}C NMR shift are much more sensitive to the p electronic charge distribution than to that of s electrons; the significant polarizability would then be the p one, i.e., the ability for electrons of the p electronic orbitals to move in an electric field. The choice of reference compounds in assessing b_{CO}'s can also be criticized: The present A value is calculated from longitudinal polarizabilities $b_{C=O}$ and b_{C-O} available from tables,[23] as determined from acetone and trioxane, which are not representative of those of a carboxylate group. The C–O bond of an ether group is especially different in polarizability from a carboxylate C–O$^-$ bond. The presence of a negative charge delocalized along the whole CO_2^- group is expected to make this group a very polarizable one and may be the main reason for the failure in using tabulated data of Le Fèvre. Calibration with a rigid compound is thus necessary.

Intercharge Distances. Once A is known, and a cavity geometry is decided, intercharge distances are obtained. The two parameters that completely describe the ellipsoidal cavity are its volume and the charge to edge distance d. In order to check for the sensitivity of the method on these two parameters, a systematic variation of their values has been made in a fitting procedure to experimental data. The results for (trimethylammonio)octanoate $(n = 7)$, which is the most sensitive case, are shown in Figure 6.   In these calculations, the charge to edge distance d is the same for the rigid reference $(n = 2)$ compound and the flexible $(n = 7)$ one, and
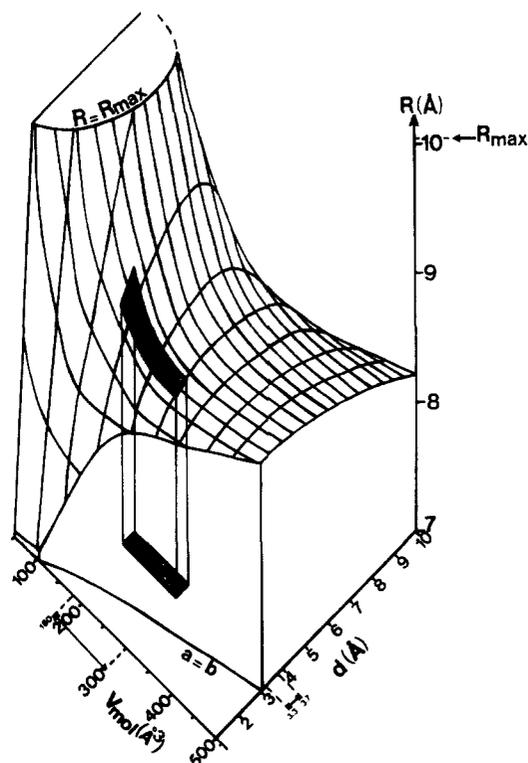

Figure 6. Best fit to experimental data for (trimethylammonio)octanoate $(n = 7)$ as a function of input parameters describing the dielectric cavity shape (molecular volume V_{mol} of $n = 2$ compound, and charge to edge distance d). The domain of investigation is limited toward the low d values by the requirement of prolate ellipsoidal shape (solid line $a = b$ at the bottom) and toward the low volume values because the intercharge distance cannot exceed the length of a fully extended arm R_{max} (10.04 Å). The "realistic domain" of volume and d is shaded. It is in a region of low sensitivity to cavity parameters.

the volume difference between these two compounds is their molecular volume difference (160 Å^3). It is observed that the method is very sensitive to the choice of the cavity geometry when the volume is low. The low d values for which the method is expected to be very shape sensitive cannot be reached because in that case, a prolate ellipsoid cannot fit the volume; the investigation stops when a spherical shape is obtained for the $n = 2$ compound.

The realistic values of volume are those close to the molecular one, with some hydration water molecule. The "best" value with two hydration water molecules has been selected, the error being estimated from the calculation for the dry molecule and for the tetrahydrate. The d value may be estimated by several means: $d = 3.3$ Å is obtained from the radius of the tetramethyl-ammonium ion, while $d = 3.5$ Å or $d = 3.7$ Å is estimated from the sum of a N–C bond length (1.474 Å) and the van der Waals radius of a methyl group (2.0 or 2.27 Å).[43] These values define the hatched area in Figure 6. It is observed that the deduced intercharge distance is rather insensitive to the shape of the cavity in the "realistic domain" of d and molecular volume values. The error or resolution of the method is estimated from this kind of plot and is found of the same order as the error coming from uncertainties on experimental chemical shift data. Thus, in the present case, the method is reliable because of the favorable molecular geometry: large molecular volume and large charge to molecular edge separation.

Intercharge distances (nitrogen to ^{13}C distances) are shown in Table III and Figure 7 and compared with the length of a fully extended intercharge arm. Data collected in methanol as a solvent have been treated in the same way as for aqueous solution, and the results are also shown in Table III. Coiling of the intercharge arm increases with its number of methylenes (n), as expected. The values of the ellipsoid minor axis b are also realistic as compared

(41) Piccinni-Leopardi, C.; Reisse, J. *J. Magn. Reson.* **1981**, *42*, 60.
(42) Allen, G. W.; Armstrong, R. S.; Aroney, M. J.; Pierens, R. K.; Williams, A. J. *J. Chem. Soc., Faraday Trans. 2* **1988**, *84*, 1775.

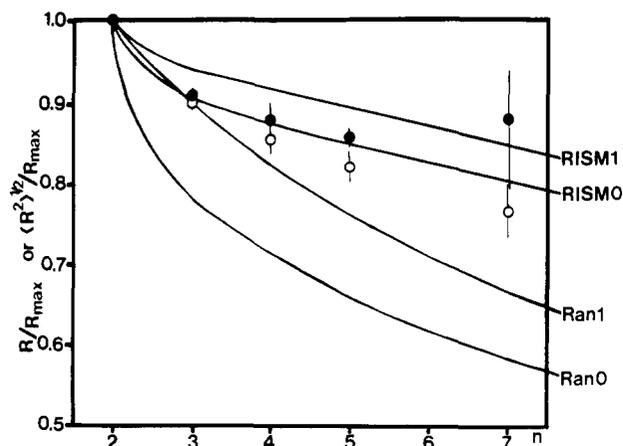(43) Edward, J. T. *Chem. Ind.* **1956**, 774.

Figure 7. Nitrogen to observed ^{13}C nucleus distance R, relative to that of a fully trans intercharge arm R_{max}, as a function of the number of intercharge methylenes n, for solutions in water (O) and methanol (●). Bars show errors which are sum of experimental ones and those coming from different choices of the dielectric cavity. Lines are comparisons with end to end distances $\langle R^2 \rangle^{1/2}/R_{max}$ from model predictions: random walk on a tetrahedral lattice (Ran) and RISM models, without (Ran0 and RISM0) or with one rigid bond (Ran1 or RISM1).

TABLE III: Nitrogen to Carboxylate Distances R As Deduced from Experimental Data as a Function of Number n of Intercharge Methylenes, for Solvents D_2O and CD_3OD

		n			
	2	3	4	5	7
R, Å, in D_2O	3.81	4.50	5.40	6.15	7.7
R, Å, in CD_3OD	3.81	4.54	5.63	6.45	8.8

to expected ones. Although it is the complex result of a fit to experimental data, b increases as the number of methylenes n increases, that is, as coiling increases, which is the expected trend (Figure 8).

Discussion

Comparison with Conformation Models of $-(CH_2)_n-$ *Moieties.* It has been shown that coiling of flexible zwitterionic molecules increases as the number of methylenes contained in the flexible intercharge arm increases. This result has to be compared with theoretical predictions of end to end distances of flexible molecules. Models usually predict mean-square end to end distances $\langle R^2 \rangle$, and $\langle R^2 \rangle^{1/2}$ can be compared to intercharge distances as deduced from experimental data. The main problem in doing such a comparison is that the way of averaging over all conformations is different in experiment interpretation and theoretical prediction. This is, however, made this way in order to get a first qualitative sight of phenomena.

The simpler theoretical model is the random walk. Bonds are free without any constraint from the others, which leads to a Gaussian chain: the mean-square end to end distance $\langle R^2 \rangle$ then increases linearly with the number n of bonds. More realistic is the random walk on a lattice. The chain is allowed to adopt any conformation on a lattice (bond lengths and angles are thus respected); a tetrahedral lattice with angles β such as $\cos \beta = 1/3$ fits well a polymethylene chain. The prediction is also a linear variation of $\langle R^2 \rangle$ with n in the limit of large n values; otherwise, $\langle R^2 \rangle$ reads[44-46]

$$\langle R^2 \rangle = l_{C-C}^2(2n - 1.5(1 - 3^{-n})) \qquad (8)$$

where l_{C-C} is the carbon–carbon bond length (1.54 Å). As seen in Figure 7, the experimental conformation is more extended (trans or rigid) than predicted by the model. The $n = 2$ compound is rigid because of end-group bulkiness and so is the $(CH_3)_3^+$-NCH_2CH_2- part of the molecule because of steric hindrance; a
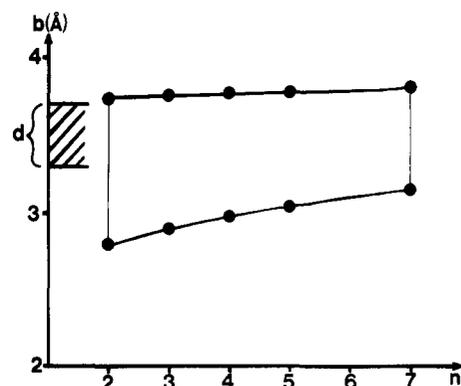

Figure 8. Variation of the minor diameter of ellipsoids (b) with n for aqueous solutions. The hatched region covers all b values found when the molecular volume and charge to edge distance d are varied between their limiting values (the shaded area of Figure 6).

modification of the model that rigidifies the intercharge arm is then to set one bond rigid. In order to account for that fact, $\langle R^2 \rangle^{1/2}$ is calculated for a $n - 1$ bond chain and the contribution of one bond to a fully extended (all-trans) chain is added as

$$\langle R^2 \rangle^{1/2}(n) = \langle R^2 \rangle^{1/2}(n - 1) + l_{C-C} \cos (\beta/2) \qquad (9)$$

This is an approximate procedure. The underlying assumption is that the orientation of the rigid bond is the same as in the all-trans compound; this is, however, considered as the best simple one as the conformation of zwitterions is found rather extended. As seen in Figure 7, the actual conformation is still more extended than in these random-walk models.

A more refined and realistic model is the rotational isomerism state model (RISM) which accounts for the larger probability of trans conformations by assigning to trans conformations a lower potential energy.[46-48] An energy difference of 2.3 kJ/mol between trans and gauche conformations in a polymethylene chain is the currently accepted value. One bond can be set rigid in the same way as above, in order to account for the rigidity of the $n = 2$ compound; this gives the RISM1 model (RISM0 with all bonds flexible). RISM0 model fits well experimental data, but it is sterically impossible. RISM1 (with one bond set rigid) is preferred and predicts a too much extended mean conformation.

A further supplementary refinement is accounting for the fact that a position cannot be occupied twice, that is, accounting for internal steric constraints. The prediction of a self-avoiding random walk on a tetrahedral lattice for the limiting case of an infinite polymer is that the mean-square end to end distance $\langle R^2 \rangle$ varies as the power $6/5$ of n instead of 1 for a random walk;[49,50] internal steric constraints thus make the conformation more rigid. Such a model is not available in a simple way for a RISM model of a short chain, but the qualitative trend is that it will predict a more extended conformation than the original RISM theory. As RISM theory is predicting a more extended conformation than the experimental one, this latest refinement will not allow a good fit to experimental data. The conclusion is thus that folding of the intercharge arm occurs, which is more pronounced in water than in methanol. The origin of such folding is ascribed to electrostatic attraction between the two charges.

Discussion. The picture that emerges from this study is still qualitative. Comparison with end to end distance predictions of conformation models is not the best way for interpretation of our data because the averaging process is not the same in experiments and theory. The ingredients for a quantitative interpretation are, however, known: (1) calculate end to end distance for each conformation from purely geometrical considerations; (2) calculate the electrostatic field at the observed nucleus and then the chemical shift for each conformation, from end to end distance and mo-

(44) Eyring, H. *Phys. Rev.* **1932**, *39*, 746.
(45) Kuhn, W. *Kolloid Z.* **1934**, *68*, 2.
(46) Flory, P. J. *Statistical Mechanics of Chain Molecules*; Wiley: New York, 1969.

(47) Flory, P. J.; Jernigan, R. L. *J. Chem. Phys.* **1965**, *42*, 3509.
(48) Baram, A.; Gelbart, W. M. *J. Chem. Phys.* **1977**, *66*, 617.
(49) Domb, C. *J. Chem. Phys.* **1963**, *38*, 2957.
(50) Wall, F. T.; Whittington, S. G. *J. Phys. Chem.* **1969**, *73*, 3953.

lecular volume and shape (which depend on conformation); (3) calculate the potential energy of each conformation from energy differences between trans and gauche conformation, electrostatic energy, and steric constraints; and (4) calculate the chemical shift by averaging all chemical shifts according to their probability. Note that this last averaging is only correct when the exchange rate between conformations is faster than their chemical shift difference, which is the actual case since only one unique NMR line is experimentally observed (fast exchange).

The best agreement of theoretical predictions with experimental values is obtained for RISM theory when all bonds are assumed flexible. $n = 2$ compound is set rigid, however. It is of course expected that the RISM model works better than the random-walk one, since its validity and predictive power are well established.[1] Experimental values of R/R_{max} are slightly less than theory, especially with methanol as a solvent, and this observation suggests that folded conformations are more likely than expected from the RISM1 model. Electrostatic attraction between the two charges of the zwitterion is indeed stronger for folded conformations of low intercharge distance. The slight discrepancy between experimental data and theory is thus evidence of the importance of electrostatics in the conformation of flexible zwitterion. The problem is to know whether such a difference is significant; a correct average should be performed in order to get a definitive answer to this question.

Experimental data also show that intercharge arm folding is more pronounced in water than in methanol. This is quite unexpected as the dielectric constant of methanol is less than that of water: the electric field should then be stronger in methanol than in water. The origin of that could be the hydrophobic nature of the intercharge arm. When n is large enough, some alkane-like material is exposed to solvent in the case of extended conformation, while coiling reduces the contact area with solvent. Hydrophobic effects, which are mainly due to solvent structure,[36] are much more effective in water than in methanol. The ability for water to bind with both its hydrogens is the key for promoting solvent structuring

and thus hydrophobic effect. Methanol cannot show this behavior: as an example, micelles that do form in water from hydrophobic effect cannot form in methanol; molecular dispersion is preferred.[51–53] Such coiling of alkanes in aqueous solution has also been calculated from Monte Carlo simulations.[54–56]

Conclusion

Mean intercharge distances of flexible zwitterions as solutions in polar solvents have been obtained from NMR chemical shift measurements. The NMR method is quite general, provided a rigid reference compound is available. Experimental quantities extrapolated to infinite dilution are free of any influence of self-association, which is an improvement over methods based on dielectric increment measurements. Data are, however, still dependent on molecular shape in a rather complex way. Measured intercharge distances as a function of the length of the flexible intercharge arm show that coiling occurs. A first qualitative interpretation of the intercharge arm conformation provides experimental evidence of folding originating in electrostatic attraction between the two charges of the zwitterion and from hydrophobic coiling when water is the solvent.

(51) Ray, A. *Nature* **1971**, *231*, 313.
(52) Magid, L. *Solution Chemistry of Surfactants*; Mittal, K. L., Ed.; Plenum: New York, 1979; Vol. 1, p 427.
(53) Beesley, A. H.; Evans, D. F.; Laughlin, R. G. *J. Phys. Chem.* **1988**, *92*, 791.
(54) Rosenberg, R. O.; Mikkilineni, R.; Berne, B. J. *J. Am. Chem. Soc.* **1982**, *104*, 7647.
(55) Jorgensen, W. L. *J. Chem. Phys.* **1982**, *77*, 5757.
(56) Jorgensen, W. L. *J. Phys. Chem.* **1983**, *87*, 5304.

Theoretical Study of the Ionization Potentials of B₂

Pablo J. Bruna and James S. Wright*

Ottawa-Carleton Chemistry Institute, Carleton University, Ottawa, Canada K1S 5B6 (Received: May 24, 1989; In Final Form: August 16, 1989)

Extensive multireference configuration interaction (MRD-CI) calculations carried out with a contracted 6s5p2d1f Gaussian basis set indicate that B_2^+ has a $X^2\Sigma_g^+(\sigma_g^2\sigma_u^2\sigma_g)$ ground and a low-lying $1^2\Pi_u(\sigma_g^2\sigma_u^2\pi_u)$ state ($T_e = 0.25$ eV). The computed dissociation energies D_e are 1.90 eV for $X^2\Sigma_g^+$ and 1.65 eV for $1^2\Pi_u$. The adiabatic ionization potentials (IP) of B_2 from $X^3\Sigma_g^-(\sigma_g^2\sigma_u^2\pi_u^2)$ into $X^2\Sigma_g^+(\pi_u^2 \rightarrow \sigma_g, \infty)$ and into $1^2\Pi_u(\pi_u \rightarrow \infty)$ are 8.99 and 9.24 eV, respectively; both results are expected to underestimate the true values by about 0.15 eV. The present data do not support recent experimental and ab initio results from Hanley, Whitten, and Anderson (*J. Phys. Chem.* **1988**, *92*, 5803). Those authors reported a $^2\Pi_u$ ground state, with a D_e of 0.8 ± 0.6 eV (0.7 eV) and an adiabatic IP ($X^3\Sigma_g^- \rightarrow {}^2\Pi_u$) of 10.3 eV (8.9 eV), with values in parentheses corresponding to their ab initio results. Since the photoelectron spectrum of B_2 (not yet recorded to our knowledge) represents a valuable and alternative experimental technique for the spectroscopic characterization of low-lying states of B_2^+, the IP's are investigated here for one-electron ionization processes having $X^3\Sigma_g^-$, $1^5\Sigma_u^-$, and $1^3\Pi_u$ of B_2 as lower state. The corresponding Franck–Condon factors are also reported.

1. Introduction

According to recent collision-induced dissociation (CID) experiments on boron cluster ions,[1,2] the dissociation energy (D_0) of B_2^+ is as low as 0.8 ± 0.6 eV, a bond strength significantly below

a D_0 of 3.00 ± 0.02 eV for neutral $B_2(X^3\Sigma_g^-)$.[3] Based on the appearance potentials (AP), Hanley et al.[2] estimated for the process $B_2 \rightarrow B_2^+$ an adiabatic ionization potential (IP) of 10.3 eV, about 2.0 eV higher than the IP of B atom. Furthermore,

(1) Hanley, L.; Anderson, S. L. *J. Phys. Chem.* **1987**, *91*, 5161.
(2) Hanley, L.; Whitten, J. L.; Anderson, S. L. *J. Phys. Chem.* **1988**, *92*, 5803.

(3) Huber, K. P.; Herzberg, G. In *Molecular Spectra and Molecular Structure. IV. Constants of Diatomic Molecules*; Van Nostrand: New York, 1979.