A NEW DEVELOPMENT OF ELECTRONIC DESCRIPTORS FOR SIMULATION OF ^{13}C CHEMICAL SHIFTS

Jalali-Heravi, Mahdi *
Moosavi, Mahdi
Department of Chemistry, Shahid Bahonar University of Kerman, Kerman, Iran.

(Received : Oct. 13th 1993 , Accepted : Feb. 14th 1994)

ABSTRACT : *The ^{13}C chemical shifts for a series of compounds which includes carboxylic acids, aldehydes, ethers, ketones and hydrocarbons were simulated by using the parametric techniques. The observed chemical shifts were related to numerically encoded structural parameters called descriptors. Two new electronic doscriptors were added to the previous descriptors. Ploting of experimental versus calculated chemical shifts for 34 different carbons in the data set shows a standard deviation of 6.05 ppm and regression coefficient of 0.994.*

KEY WORDS : *Simulation, ^{13}C Chemical Shifts, Parametric Techniques, Electronic Descriptors.*

INTRODUCTION :

Development of theoretical methods are useful and seems to be necessary for interpretation of NMR spectra of large and complicated molecules. In general, there are three principal theoretical approaches for simulation of the ^{13}C chemical shifts : (1) quantum mechanical methods, (2) library shift retrieval and (3) parametric techniques.

In quantum mechanical approach both *ab initio* [1] and semiempirical [2] methods have been used for ^{13}C chemical shifts calculations.

Much work has been performed in the area of library shift retrievals [3]. This method is limited by the size and quality of the library and the versatility of the coding algorithm.

An alternative to quantum mechanical

* Corresponding author

1021-9986/93/2/73 4 / $ /2.40

methods and library shift retrieval approaches are parametric techniques. These methods are based on the construction of linear mathematical models relating the observed chemical shifts to numerically encoded structural parameters called *descriptors*. These models have the general form:

$$S = b_0 + b_1X_1 + ... + b_nX_n$$

where S is the predicted chemical shift for the carbon of interest, X_i is the numerical descriptor, b_i the coefficient determined from a multiple linear regression analysis of a set of observed chemical shifts, and n denotes the number of descriptors in the model.

The main problem in this technique is selection of suitable descriptors. In general, these descriptors are classified as topological, electronic and geometric [4].

A limitation of the parametric approach is that the calculated models are highly specific to the types of compounds in the reference set [5,6]. Therefore, the main aim of the present work is the expansion of structural diversity of the data sets, including compounds with varying functional groups. For this purpose one needs to develope new descriptors which are more general. In our previous work [4] we found that the electronic descriptors play an important role in simulation of ^{13}C chemical shifts. Therefore, in this work emphasise is placed on development of a new set of electronic descriptors. Two new electronic descriptors have been added to the previous descriptors.

EXPERIMENTAL :

Table 1 shows the compounds used in this work. A total of 22 compounds such as hydrocarbons, carboxylic acids, aldehydes, ethers, ketones and aromatic compounds were used in this study (Table 1). All compounds have been used for generation of models without any subsetting. The ^{13}C-NMR spectra were recorded under a variety of experimental conditions and were taken from the references [7] and [8]. All chemical shifts values reported in this work are relative to that of benzene.

Table 1 : Data set

1	$H_2C=CH_2$	12	H_3CCN
2	$H_2C=C=CH_2$	13	$(CH_3)_2O$
3	HCCH	14	CH_3OH
4	$H_3CCH=CHCH_3$	15	H_2CO
5	$H_3CCH=CH_2$	16	H_3CCOH
6	$H_2C=CHCH=CH_2$	17	C_6H_6
7	$H_3CCH=CHCH_3$	18	$H_3C\text{-}C_6H_5$
8	$(CH_3)_2C=CH_2$	19	$(CH_3)_2CO$
9	$H_3CCH=CHCH_3$	20	H_2NCOH
10	H_3CNC	21	HCOOH
11	HCN	22	$H_3COCOOCH_3$

The structures of all compounds used in this work were optimized by using the MOPAC program with the AM1 hamiltonian [9]. Coefficients relating the descriptors to the observed chemical shifts were obtained by using the statistical package of SPSS/PC [10].

RESULTS AND DISCUSSION :

This study involves 22 compounds. These compounds contain a total of 62 carbon atoms. Duplicate carbons were removed from the data set, leaving a total of 34 unique carbon atoms for the development of chemical shift models. Removing of duplicate carbon atoms prevents the introduction of statistical bias into the regression models arising from the presence of those atoms insufficiently represented in the data set. The electronic and geometric descriptors depend on the three dimentional structures of molecules. Therefore, all molecules were optimized by using the AM1 hamiltonian. SCF convergence and energy minimization criterion were limited to the program default values. Computations were restricted to *s* and *p* atomic valence orbital basis set, as set forth in the MOPAC program.

A series of numerical descriptors encoding the surrounding environment were calculated for each carbon atom. These descriptors were of three general types: topological, electronic and geometric. Definition for each type is given else-

were [4]. Two new electronic descriptors were added to the previous descriptors. First is $<r^{-3}>$ integral which is proportional to the effective nuclear charge for the atomic orbital φ_μ on atom A. This integral is given by Eq. (1).

$$<r^{-3}>_{\mu\mu A} = 1.3\,[Z_\mu / na_o]^3 \qquad (1)$$

where n and a_o are the principal quantum number of electron μ and the *Bohr* radious respectively. This integral refers to the average of the inverse cube of the separation of the valence *p* electrons from the corresponding nucleus. Second is -1/WEE parameter where WEE is the weighted average excitation energies for the excited singlet states which are mixed with the ground state by an external magnetic field.

In order to minimize chance correlation, the descriptors were screened and those descriptors containing minimal or redundant information were discarded. A total of 35 descriptors for each atom center were submitted to a stepwise multiple linear regression analysis method using the observed chemical shifts as the depended variables. The regression method were coupled to a progressive deletion procedure to help to obtain superior models.

The resulting models were evaluated based on several statistical criteria including standard error (s), correlation coefficient (R) and F value for statistical significance of the model.

Table 2 shows the best model equation generated for the data set. The labels for the descriptors are defined in Table 3. The mean effect on the predicted chemical shifts are presented for each descriptor in Table 2. The mean effect is a measure of the shielding or deshielding effects of each descriptor. It can be seen that the -1/WEE descriptor has the highest contribution to the predicted chemical shifts.

In order to illustrate the predictive ability of the selected model, the chemical shifts of 34 carbon atoms were calculated by this model. The plot of calculated versus observed chemical shifts is shown in Fig. 1. this plot shows a standard deviation of 6.05 ppm and correlation coefficient of 0.994 which confirms the high predictive ability of the model.

Table 2 : Best selected model*

descriptor	regression coefficient	mean effect
SUM1	−450.99(25.27)	85.69
−1/WEE	−3503.52(151.18)	255.25
AV1	729.29(41.03)	−51.05
N1	49.63(5.35)	5.96
O2	144.46(13.01)	8.67
HPOS4	489.78(65.68)	9.80
CCl1	−119.63(15.63)	−75.37
ASUM1	−84.96(14.83)	−39.93
HNEG1	−63.20(20.35)	11.20
ACCl3	−63.25(20.48)	−5.06
ASUM2	51.55(12.7)	18.56
ACCl1	105.60(32.77)	40.13
Constant	−276.36(17.15)	

* *Descriptors are difined in Table 3*

From the above results,one may conclude that the new electronic descriptors are very useful in simulation of ^{13}C chemical shifts. Although the data set is very diverse, the results are promising. Future research in the area will require the expansion of sizes for both data set and the molecules.

ACKNOWLEDGMENT :

Financial support from the Research Council of the Shahid Bahonar University of Kerman is gratefully acknowledged.

REFERENCES :

[1] Ditchfield, R., *Mol. Phys.*, **27**, 789 (1974); pulay, P., *J. Am. chem. Soc.*, **112**, 8251 (1990); Chesnut, D.B., *Annu. Rep. NMR Spectrosc.* **21**, 51(1989).

[2] Fukui, H., *Mag. Reson. Rev.*, **11**, 205 (1987); Vauthier, E., et al, *Can. J. Chem.*, **60**, 957 (1982); Jalali- Heravi, M., and Webb, G.A., *Org. Magn. Reson.*,. **11**, 34 (1978), *ibid* **11**, 524(1978).

Table 3 : Definition of the descriptors in the selected model

descriptor	type	definition
SUM1	electronic	Sum of partial charges on atoms one bond from the carbon center
-1/WEE	electronic	Negative of inverse of weighted excitation energy
AV1	electronic	Average of partial charges on atoms one bond from the carbon center
N1	topological	Number of nitrogen atoms one bond from the carbon center
O2	topological	Number of oxygen atoms two bonds from the carbon center
HPOS4	electronic	Most positive charges on atoms four bonds from the carbon center
CCl1	topological	Corrected connectivity index for one bond from the carbon center
ASUM1	electronic	Sum of absolute charges on atoms one bond from the carbon center
HNEG1	electronic	Most negative charges on atoms one bond from the carbon center
ACCl3	topological	Average of corrected connectivity index for three bonds from the carbon center
ASUM2	electronic	Sum of absolute charges on atoms two bond from the carbon center
ACCl1	topological	Average of corrected connectivity index for one bond from the carbon center

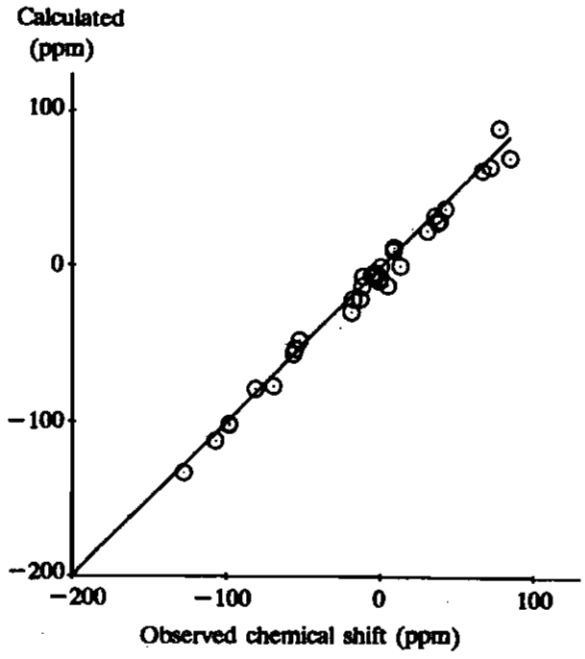

Fig. 1: Calculated versus observed chemical shifts

[3] Gray, N.A.B., Crandell, C.W., Nourse, J.G., Smith, D.H., Dagefordeand, M.L., and Djerassi, C., *J. Org. Chem.*, 46, 703(1981).

[4] Jalali-Heravi, M., and Moosavi, M., 1993, To be Published.

[5] Rance, M.L., and Jurs, P.C., *Anal. Chem.*, 61, 2489(1989).

[6] Sutton, G.P., and Jurs, P.C., *Anal. Chem.*, 62, 1884(1990).

[7] Emsley, J.W., Feeney, J., and Sutcliffe, L.H., *High Resolution NMR Spectroscopy.*, Pergamon Press, Oxford (1965).

[8] Stothers, J.B., *Carbon-13 NMR Spectroscopy*, Academic Press, New York (1972).

[9] Dewar, M.J.S., Zeobish, E.G., Healy, E.F., and Stewart, J.J.P., *J. Am. Chem. Soc.*, 107, 3902(1985).

[10] SPSS/PC, The Statistical package for IBM PC, Quiad Software Limited, Ontario, Canada (1986).