

Quantitative Structure-Property Relationship (QSPR) Study of ^{17}O Carbonyl Chemical Shifts in Substituted Benzaldehydes

Rudolf Kiralj*, Márcia M. C. Ferreira rudolf@iqm.unicamp.br

Laboratório de Quimiometria Teórica e Aplicada, Instituto de Química, Universidade Estadual de Campinas, Campinas – SP, 13083-970 BRAZIL

Keywords: chemical shifts, chemometrics, benzaldehydes

Li and Li (LL) have proposed recently¹ an empirical equation for calculation of ^{17}O NMR chemical shifts in benzaldehydes. This equation uses contribution of individual *o*-, *o'*-, *m*-, *m'*- and *p*-positioned substituents and a correction constant for polar solvents. However, in spite of high predictive ability of the equation, due to its empirical nature, it cannot be validated properly as a regression model. Besides, its application is limited to benzaldehydes that include a small set of substituents. These facts have encouraged the authors of this work to develop a simple and fast Quantitative Structure-Property Relationship (QSPR) methodology for prediction of ^{17}O NMR carbonyl chemical shifts in benzaldehydes and related systems.

Geometry of fifty substituted benzaldehydes with known ^{17}O NMR shifts¹ (training set) and of ten without known shifts (prediction set) was optimized at semi-empirical PM3 level and various geometric and steric descriptors accounting for properties of the benzene ring, aldehyde group and their connecting carbon-carbon bond, were calculated. QSPR models were based on these descriptors and on partial least squares regression, principal component regression and multiple linear regression. Principal component analysis, hierarchical cluster analysis and crystal structure data retrieved from the Cambridge Structural Database (CSD) were used in order to get more insight into the chemical background of relationships between chemical shifts and molecular descriptors. All data were autoscaled prior to analysis.

Five molecular descriptors were selected: nuclear-nuclear repulsion energy for benzaldehyde (C1) and the first aromatic (C2) carbon atoms, electrostatic potential-based partial atomic charge of the carbonyl oxygen O, standard deviation of the six C-C bond lengths in the benzene ring, the C1-C2 bond length, and Mulliken partial atomic charge of C2. These descriptors exhibit high correlation with experimental ^{17}O shifts, with absolute correlation coefficients being 0.89 to 0.93. Unlike LL, the chemical shifts were averaged whenever possible (for 13 samples). The obtained PLS and PCR (two principal components) and also MLR regression model have $Q > 0.93$, $R > 0.94$, $SEV \leq 11$, $SEP < 10$, and less than 14 samples with relative error over 10%. The models were internally cross-validated (one-leave-out) and also externally validated (10 samples excluded). The models were capable to predict chemical shifts for 10 arbitrarily selected benzaldehydes, while the LL model could be applied only to four samples. Due to the simplicity and short consumed time per sample, the proposed QSPR models can be used for prediction of ^{17}O carbonyl shifts in substituted benzaldehydes and related systems.

The exploratory analysis and selected structures from the CSD show that the increase in electron density and decrease in chemical shift of the carbonyl oxygen is caused by electron donation of the benzene ring and substituents, and of the hydrogen bond established between the aldehyde and *o*-hydroxyl group. In fact, all these fragments may be considered as integral parts of the benzaldehyde π -delocalized system. The variations in the ^{17}O shifts are well described by five selected local molecular descriptors at semi-empirical level, accounting for geometric and electronic properties of the aldehyde group, the benzene ring and their connecting carbon-carbon bond. The influence of the hydrogen bonding to the chemical shifts was noticed also. The first principal component is related to cumulative electron withdrawal/donation effects felt by the carbonyl oxygen, whilst the second principal component is probably related to the variations in the benzaldehyde heteroaromatic character.

Electron withdrawal/donation effects caused directly or indirectly by substituents, hydrogen bonding and benzaldehyde non-planarity are the major factors that influence variations in electron density and ^{17}O shifts. These phenomena have been well-described by molecular descriptors from semi-empirical calculations, and the corresponding QSPR have shown to be suitable for prediction of these shifts in benzaldehydes.

Acknowledgment: FAPESP

References

- ¹ Li L.-D.; Li L.-S. *Magn. Reson. Chem.* **2004**, *42*, 977-982.