From Hydrogen Bond to Bulk: Solvation Analysis of the n-π^* Transition of Formaldehyde in Water

SYLVIO CANUTO,[1] KALINE COUTINHO[2]

[1]*Institute de Física, Universidade de São Paulo, CP 66318, 05315-970 São Paulo, SP, Brazil*
[2]*Centro de Ciências Exatas e Tecnológicas, CCET, Universidade de Mogi das Cruzes, CP 411, 08701-970 Mogi das Cruzes, SP, Brazil*

ABSTRACT: Supermolecular calculations that treat both the solute and the solvent quantum mechanically are performed to analyze the n-π^* transition of formaldehyde in water. The liquid structures are generated by canonical (constant volume, temperature, and number of particles) (NVT) Metropolis Monte Carlo simulation. Autocorrelation function is calculated to obtain efficient ensemble average. Full quantum mechanical intermediate neglect of differential overlap/singly excited configuration interaction (INDO/CIS) calculations are then performed in the supermolecular clusters corresponding to the hydrogen bond shell and the first, second, and third solvation shells. The largest cluster, corresponding to the third solvation shell, includes 1 formaldehyde and 80 water molecules. INDO/CIS calculations are performed on a properly antisymmetric reference ground-state wave function involving all valence electrons. The results are then extrapolated to the bulk limit. The estimated limit value for the solvatochromic shift of the n-π^* transition of formaldehyde in water, compared to gas phase, is 2200 cm^{-1}.   © 2000 John Wiley & Sons, Inc. Int J Quant Chem 77: 192–198, 2000

Key words: solvent effects; hydrogen bond; formaldehyde; n-π^* transition

Introduction

The first absorption transition of gas-phase formaldehyde is known to be an n-π^* excitation located around 34,000 cm^{-1} [1]. This excitation involves a large decrease in the dipole moment of the excited state as compared to the ground state [2]. As a consequence, this transition suffers a blue shift when formaldehyde is solvated in water. The amount of this blue shift is not well determined experimentally or theoretically, but it is believed to be larger than 2000 cm^{-1} [3–5], with a very large broadening of about 4000 cm^{-1} [4]. This solvatochromic blue shift has been analyzed in several previous theoretical studies [5–12]. Most of the recent investigations have considered com-

bined quantum mechanics and simulation approach or the use of dielectric continuum. Blair et al. [5] used a combination of molecular dynamics simulation followed by ab initio Hartree–Fock calculations. Supermolecular calculations were performed for a formaldehyde surrounded by 209 water molecules represented by simple point charges. They estimated the blue shift to be 1900 cm^{-1} and noted that inclusion of a small number of water molecules could not adequately describe the formaldehyde solvatochromic shift. Later, Fukunaga and Morokuma [6] derived potential functions for the interaction between formaldehyde and water and calculated the blue shift of the n-π^* transition in solution to be around 3150 cm^{-1}. Similarly, they noted that small clusters could not represent the n-π^* transition in solution. An investigation performed by Thompson [7] in a quantum mechanics/molecular mechanics (QM/MM) study obtained a smaller value of 1150 cm^{-1}. Bader and Berne [8] used a dielectric continuum to estimate the blue shift as 3500 cm^{-1}. Very recently, Naka, Morita, and Kato [12] used a reference interaction site model self-consistent field (RISM-SCF) method and obtained a blue shift of 2000 cm^{-1}. Hence, available theoretical studies using different approaches and different sophistication levels would put the solvatochromic shift of the n-π^* excitation of formaldehyde in water, at room temperature, somewhere between 1100 and 3500 cm^{-1}, a considerably wide range. Also, they all seem to agree that a small cluster cannot adequately represent the transition in solution. Compared with acetone, where the experimental result is known, one would expect this blue shift to be around 1500–2000 cm^{-1}.

In this study we use computer simulation, but in the subsequent quantum mechanical calculations the solvent water molecules are considered with all the valence electrons. We perform a systematic study of the n-π^* blue shift of formaldehyde in water where the solvent molecules are explicitly included in the quantum mechanical calculations. This includes not only the electrostatic interaction between the solute and the solvent but also the corresponding induced polarizations both in the solute and in the solvent. We use Monte Carlo simulations to generate structures of the liquid and supermolecular quantum mechanical calculations, with all-valence electrons, to obtain the separate contributions of the different solvation shells to the blue shift. Starting from the hydrogen-bonded water, we analyze the contribution of the first, second, and third solvation shells. The largest calculation involves the ensemble average of many quantum mechanical results obtained with the formaldehyde solute surrounded by 80 water solvent molecules. This latter starts with a self-consistent field INDO calculation with a properly antisymmetric wave function with 652 valence electrons. The excitation energies are obtained next using CIS. The final result is estimated after extrapolation to the bulk limit. Thus we not only investigate how the solvation shells influence the blue shift but we also extrapolate our results to obtain our best estimate of the solvatochromic shift of the n-π^* transition of formaldehyde in water, compared to gas phase. This approach has been successfully used to obtain solvatochromic shifts in both polar and nonpolar solutes and solvents [13].

Methods of Calculation

MONTE CARLO SIMULATION

The Monte Carlo (MC) simulation is performed using standard procedures for the Metropolis sampling technique [14] in the canonical ensemble, where the number of molecules N, the volume V, and the temperature T are fixed. As usual, periodic boundary conditions in a cubic box [15] are used. In our simulation, we use one formaldehyde molecule plus 343 molecules of water. The volume of the cubic box is determined by the experimental density of the water [16]; $\rho = 0.9966$ g/cm^3 at $T = 25°$C. The molecules interact by the Lennard-Jones plus Coulomb potential with 3 parameters for each atom i (ε_i, σ_i, and q_i). For water molecules, we used the simple point charges (SPC) potential developed by van Gunsteren et al. [17]. For formaldehyde molecule, we used the same potential of Blair et al. [5]. It is known that upon solvation the C=O distance of formaldehyde is slightly increased. This is because of the hydrogen bonds formed in the carbonyl side. Also, this increase influences the solvatochromic shift contributing to a red shift [5, 18]. Thus for the geometry of formaldehyde we have used a lengthened C=O distance of 1.2173 Å as obtained from an ab initio second-order Møller–Plesset (MP2) full geometry optimization with a 6-311++G** basis set of the cluster formaldehyde–water. The other geometric parameters used were $R_{CH} = 1.102$ Å and $H\hat{C}O = 121.4°$.

In the calculation of the pairwise energy, each molecule interacts with all other molecules within a center-of-mass separation that is smaller than the

cutoff radius $r_c = 10.9$ Å. For separations larger than r_c, we use the long-range correction of the potential energy [15]. The initial configuration is generated randomly, considering the position and the orientation of each molecule. A new configuration is generated after 344 MC steps, i.e., after selecting all molecules sequentially and trying to translate it in all the Cartesian directions and also rotate it around a randomly chosen axis. The maximum allowed displacement of the molecules is autoadjusted after 17,000 MC steps to give an acceptance rate of new configurations around 50%. The maximum rotation angle is fixed during the simulation in $\delta\theta = \pm15°$. The full simulation [19] involves a thermalization stage of about 1.7×10^6 MC step followed by an averaging stage of 16.5×10^6 MC steps. Thus, the total number of MC configurations generated in the simulation is 48,000. During the averaging stage some thermodynamical properties, as the internal energy and the heat capacity at constant volume, are calculated, and they are in agreement with that calculated for liquid water. The radial distribution function is also calculated during the averaging stage in the simulation.

As quantum mechanical calculations will be performed on the configurations generated by the MC simulation, it is important to optimize the statistics. Configurations that are statistically correlated will not give important additional information. Therefore we calculate the autocorrelation function [20]. For Markovian processes this follows an exponentially decaying function [21, 22]:

$$C(t) = \sum_i^n c_i e^{-t/\tau_i},$$

where t is the interval of MC configurations. The correlation step is

$$\tau = \int_0^\infty C(t)\, dt.$$

Configurations that are separated by $t > 3\tau$ have a statistical correlation that is less than 5%. Figure 1 shows the calculated autocorrelation function. From this and the equation above it can be seen that the correlation step is ~280. Therefore, in calculating the averages we select configurations separated by 800 MC steps. As the total number of MC configurations generated in the simulation is 48,000, the averages are taken over 60 configurations, separated by 800 MC steps [22].
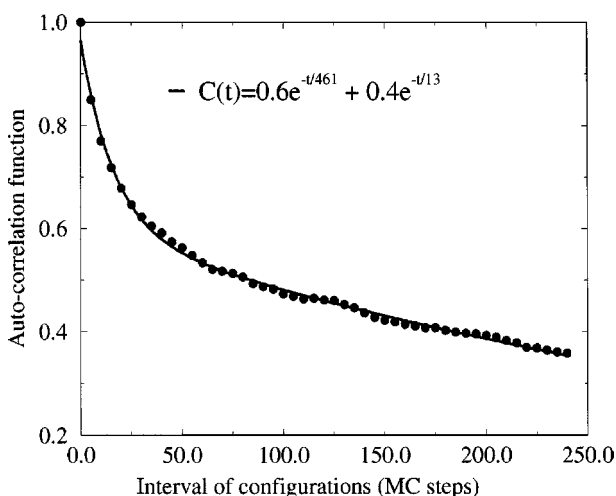

FIGURE 1. Calculated autocorrelation function for formaldehyde in water and best exponential fit.

QUANTUM MECHANICAL CALCULATIONS

The excitations energies are calculated using the ZINDO program [23] within the INDO/CIS model in the supermolecules generated by the MC simulations. The quantum mechanical calculations are then performed for the supermolecular cluster composed of one formaldehyde and all solvent molecules within a particular solvation shell. Each water molecule includes 8 valence electrons, and the Hartree–Fock wave function is antisymmetric with respect to the entire solute–solvent system. Next, CIS calculations are performed to obtain the excitation energies. The solvatochromic shift is obtained as the average over a chain of L transition energy values:

$$\langle\Delta E\rangle = \frac{1}{L}\sum_i^L \Delta E_i.$$

For the first, second, and third hydration shells the value of L is 60, obtained previously from the autocorrelation function. For the hydrogen bonds a specific statistics will be performed, where we analyze the number of hydrogen bonds in each configuration, using both the geometric and energetic criteria [24]. In total, nearly 250 quantum mechanical calculations are performed. This is a computational task that would be very difficult to accomplish if programs like Zerner's INDO/CIS were not available.

Results and Discussions

HYDROGEN BOND

Solvation shells are defined from the radial distribution function as obtained from the MC simulation. We discuss separately the hydrogen bond formed between formaldehyde and water. Figure 2 shows the radial distribution function between the oxygen of formaldehyde and the oxygen of water, $g_{OO}(r)$. As can be seen there is a hydrogen bond (HB) peak that starts at around 2.4 Å and has a minimum at 3.2 Å. Integration of this peak gives 2.6 water molecules. As expected this number is the same as that obtained by Blair et al. [5]. We call this the HB solvation, to separate it from the other solvation shells.

Hydrogen bonds are also obtained using a geometric criterium (radial $R_{00} \leq 4$ Å and angular $O\hat{O}H \leq 30°$) [24]. With this in the 60 MC configurations we find 123 hydrogen bonds. This gives an average of 2.1 bonds. This is close but lower than the average value obtained from the radial distribution function. Figure 3 shows the pairwise formaldehyde–water energy distribution. It can be seen that for pairs with distance lower than 4 Å, some have interaction energy either too low or even positive. Therefore, it is quite appropriate to use also an energetic criterion to identify the hydrogen bonds. Thus only those structures with a binding energy higher than 3.0 kcal/mol are included. Thus using both the geometric and energetic criteria we find a total of 114 hydrogen bonds in the separated 60 MC configurations. This gives an average of 1.9 hydrogen bonds. Table I gives the statistics obtained for the hydrogen bonds formed. We find that 23.3% of the configurations have one hydrogen bond. Similarly, 63.3%
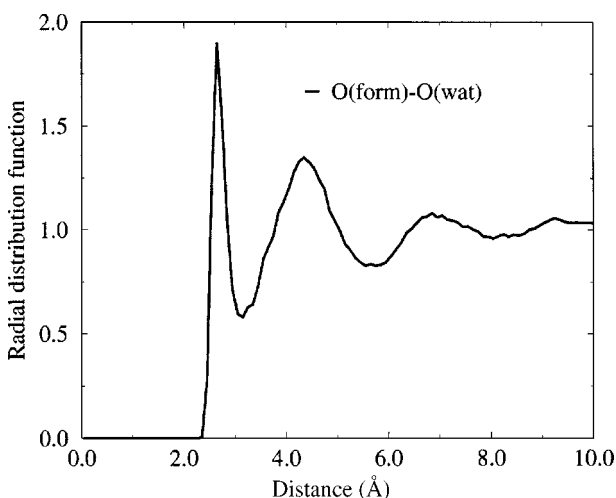

FIGURE 2. Pairwise radial distribution function for the oxygen atom of formaldehyde and oxygen atom of water.
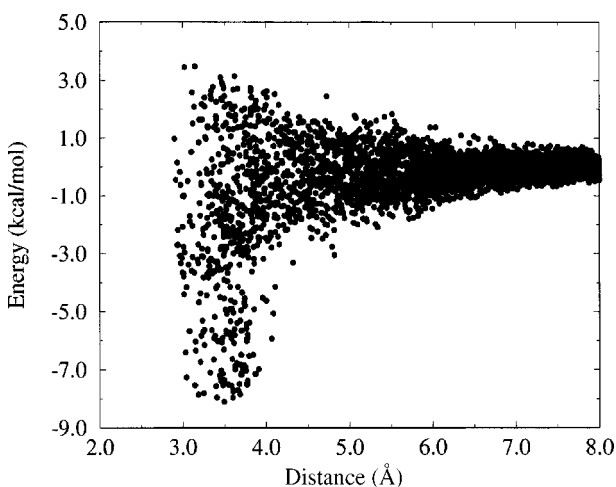

FIGURE 3. Calculated pairwise energy interaction for formaldehyde and water.

TABLE I

Statistics of the hydrogen bonds formed between formaldehyde and water and their contribution to the blue shift of the n-π* transition.

Number of HB	Geometric		Geometric and energetic	
	Occurrence (%)	Shift (cm^{-1})	Occurrence (%)	Shift (cm^{-1})
1	18.3	152	23.3	146
2	60.0	335	63.3	313
3	20.0	262	13.4	360
4	1.7	988	—	
Total	100 (123)	305 ± 65	100 (114)	281 ± 65

[a] First result uses the geometric criterion to obtain a hydrogen bond. The second uses in addition the energetic criterion [24].

forms two and 13.4% forms three hydrogen bonds. In the subsequent quantum mechanical calculations these structures are analyzed separately. As shown, the total blue shift contribution of the HB solvation is 281 cm^{-1}. But those structures with one, two, and three hydrogen bonds contribute, respectively, 146,313, and 360 cm^{-1}. The blue shift of 281 cm^{-1} due to the HB shell is lower than that obtained by Blair et al. [5], showing that in this case the induced polarization of the hydrogen-bonded water is important. Also, we include here only those water molecules that satisfy the energetic and geometric criteria. As a result the average number of hydrogen bonds reduce from 2.6 to 1.9. Indeed, using the radial distribution to identify hydrogen bonds is not a good criterion [24].

SOLVATION SHELLS

Figure 4 shows the radial distribution function between the C atom of formaldehyde and the O atom of water. Three solvation shells are discernible. The first shell starts at 2.8 Å ending at 4.75 Å. The second ends at 6.45 Å and the third at 8.35 Å. Integration of the radial distribution function gives a total of 15, 35, and 80 water molecules in these shells, respectively. Table II thus gives the calculated solvatochromic shift and summarizes the results, including the spread of calculated values, obtained from the lowest and the highest transition energies calculated for each solvation shell. The largest calculation, obtained for the third solvation shell, is composed of the formaldehyde and 80 water molecules and gives a calculated blue shift of 1942 cm^{-1}. This result is in very good agreement with that obtained before by many others [5, 10–12]. The analysis of the variation of the shift with the solvation shells (Fig. 5) shows that this is not a stable value with respect to the number of solvent
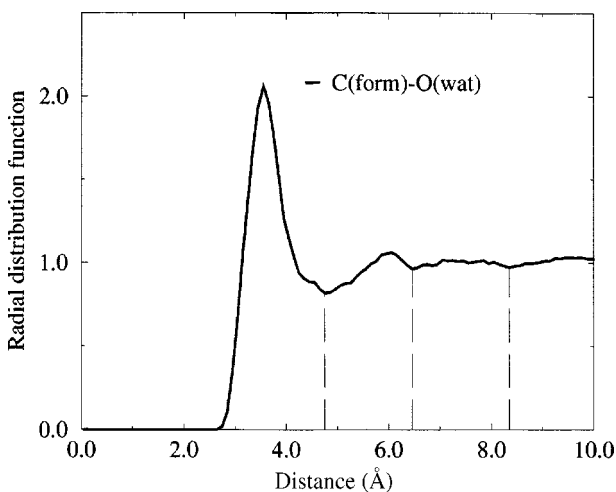

FIGURE 4. Pairwise radial distribution function for the carbon atom of formaldehyde and the oxygen atom of water.

molecules included. It is somewhat surprising that one apparently has to go beyond the third solvation shell. Longer range effects could in principle be included by a reaction field with this largest supermolecule enclosed in a cavity [25]. But as the reaction field depends on the inverse third power of the cavity radius, its contribution is rather negligible here. The result shown in Figure 5 clearly suggests that the monotonic behavior of the calculated energy shift permits an extrapolation of the results to the bulk limit. In doing so we obtain a limiting value of ∼2200 cm^{-1}. This would be our best estimate of the blue shift of the n-π^* transition of formaldehyde in water. This final value agrees with some previous calculations, in particular those by Blair et al. [5] and the very recent (RISM-SCF) result of Naka, Morita, and Kato [12]. Compared to the case of acetone the blue shift of formaldehyde

TABLE II _____

Variation of the calculated (INDO/CIS) shift and spread (in cm^{-1}) of the n-π^* transition of formaldehyde in water with the solvation shells.

Solvation shell	N	M	L	Distance (Å)	Blue shift	Spread
First	15	132	60	4.75	1237 ± 75	2500
Second	35	292	60	6.45	1667 ± 90	3100
Third	80	652	60	8.35	1942 ± 90	3300
Limit	Bulk				∼2200	

[a] N is the number of water molecules included. M is the total number of valence electrons. L is the number of MC configurations used for ensemble average.
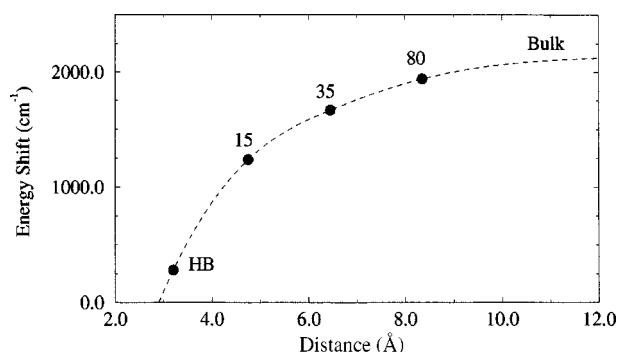
FIGURE 5. Calculated blue shift of the n-π^* transition of formaldehyde in water, as a function of the solvation shell. Also shown are the number of molecules included in the quantum mechanical calculations. HB is the hydrogen bond shell. The others are the first, second, and third solvation shells. See Figure 4.

is predict to be larger by 500–700 cm^{-1}. Another important aspect of the n-π^* transition of formaldehyde in water is the broadening of the absorption line shape. Figure 6 shows the calculated bandwidth for the largest supermolecular calculations, namely, the third solvation shell. This result is also in very good agreement with both experiment [4] and recent theoretical estimates [5, 8, 12].

Summary and Conclusions

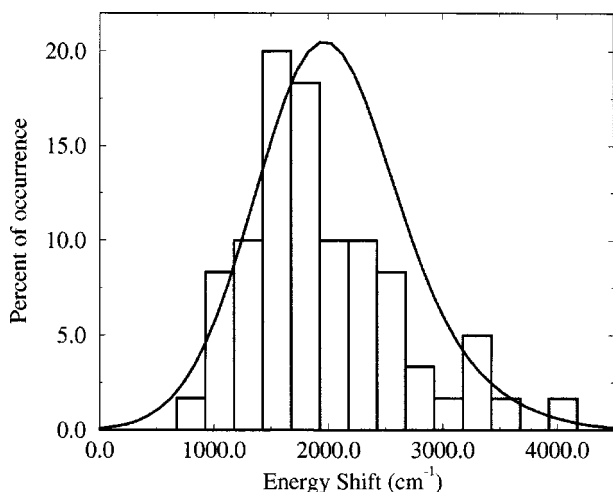A sequential Monte Carlo–quantum mechanics study of the influence of the hydration shells to the

FIGURE 6. Histogram and Gaussian convolution of the calculated blue shifts obtained for the third solvation shell. See text.

n-π^* transition of formaldehyde in water was made. The solvatochromic shift was analyzed as a function of the solvation shells, starting from the hydrogen bonds and extending to the third solvation shell. All the solvent molecules are treated quantum mechanically. The INDO self-consistent field equations are first solved for the entire solute–solvent supermolecule. Then INDO/CIS calculations are performed to obtain the transition energies. To obtain the average, calculations are made on the configurations, generated by the Monte Carlo simulation that are statistically nearly uncorrelated. Indeed the autocorrelation function is used to obtain structures that are only ~5% correlated. A detailed analysis is made of the contribution of the hydrogen bonds. Next, 60 quantum mechanical calculations were made for each of the first, second, and third solvation shells that are composed, respectively, of one formaldehyde molecule surrounded by 15, 35, and 80 water molecules. The corresponding calculated blue shifts are, respectively, 1237, 1667, and 1942 cm^{-1}. Extrapolating these results to the bulk limit gives our best estimate of this solvatochromic shift. This value is found to be close to 2200 cm^{-1}.

References

1. Brand, J. C. D. J Chem Soc 1956, 858.
2. Jones, V. T.; Coon, J. B. J Mol Spectr 1969, 31, 137; Buckingham, A. D.; Ramsey, D. A. Can J Phys 1970, 48, 1242.
3. Bercovici, T.; King, J.; Becker, R. S. J Chem Phys 1972, 56, 3956.
4. Beiber, R.; Truempler, G. Helv Chim Acta 1947, 30, 1860.
5. Blair, J. T.; Krogh-Jespersen, K.; Levy, R. M. J Am Chem Soc 1989, 111, 6948; Levy, M.; Kitchen, D. B.; Blair, J. T.; Krogh-Jespersen, K. J Phys Chem 1990, 94, 4470.

6. Fukunaga, H.; Morokuma, K. J Phys Chem 1993, 97, 59.

7. Thompson, M. A. J Phys Chem 1996, 100, 14492.

8. Bader, J. S.; Berne, B. J. J Chem Phys 1996, 104, 1293.

9. Iwata, S.; Morokuma, K. J Am Chem Soc 1973, 95, 7563; Del Bene, J. E. J Am Chem Soc 1973, 95, 6517; Debolt, S. E.; Kollman, P. A. J Am Chem Soc 1990, 112, 7515.

10. Bader, J. S.; Berne, B. J. J Chem Phys 1997, 106, 2372.

11. Mennucci, R.; Cammi, R.; Tomasi, J. J Chem Phys 1998, 109, 2798.

12. Naka, K.; Morita, A.; Kato, S. J Chem Phys 1999, 110, 3484.

13. Canuto, S.; Coutinho, K.; Zerner, M. C., submitted.

14. Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. J Chem Phys 1953, 21, 1087.

15. Allen, M. P.; Tildesley, D. J. Computer Simulation of Liquids; Clarendon: Oxford, 1987.

16. Lide, D. R. (ed.) CRC Handbook of Chemistry and Physics, 64th ed.; CRC Press: 1983–1984.

17. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Intermolecular Forces; Pullmam, B., Ed.; Reidel: Dordrecht, 1981, p. 331.

18. Coutinho, K.; Saavedra, N.; Canuto, S. J Mol Struct (Theochem) 1999, 466, 69.

19. Coutinho, K.; Canuto, S. DICE: A Monte Carlo Program for Molecular Liquid Simulation; University of São Paulo: São Paulo, 1997.

20. Coutinho, K.; Canuto, S. Adv Quant Chem 1997, 28, 89.

21. Chatfield, C. The Analysis of Time Series. An Introduction, 3rd ed.; Chapman and Hall: London, 1984; Krätschmer, R.; Binder, K.; Stauffer, D. J Stat Phys 1976, 15, 267.

22. Coutinho, K.; Oliveira, M. J.; Canuto, S. Int J Quant Chem 1998, 66, 249.

23. Zerner, M. C. A Semi-empirical Program Package; University of Florida: Gainesville, 1997.

24. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. J Chem Phys 1983, 79, 926.

25. Karelson, M.; Zerner, M. C. J Phys Chem 1996, 96, 6949.