

THEORETICAL SOLVATION MODELS: *AB INITIO* STUDY OF MOLECULAR AGGREGATION

AGNIESZKA SZARECKA¹, JACEK RYCHLEWSKI¹, U. RYCHLEWSKA²

¹*Quantum Chemistry Group, Faculty of Chemistry Adam Mickiewicz University
Grunwaldzka 6, 60-780 Poznań, Poland
szarecka@man.poznan.pl. rychlew@man.poznan.pl*

²*Department of Crystallography, Faculty of Chemistry Adam Mickiewicz University
Grunwaldzka 6, 60-780 Poznań, Poland*

Abstract: The results of a comparative *ab initio* study of solvation effects in a process of hydrogen bond driven dimerisation of two small organic molecules - formamide and a primary amide of the α -hydroxyacetic acid are here presented. Differences between various dielectric continuum models (i.e. Onsager, PCM, IPCM and SCIPCM), their performance and range of applicability are reported and discussed.

1. INTRODUCTION

Problems concerning phenomena occurring in solution and both qualitative and quantitative description of the influence of solvent on the structural and chemical behaviour of solute molecules are of interest to many fields of chemistry, biochemistry and physics. Solvent affects equilibrium constants and rates of reactions, conformational and tautomeric changes [1], phase transitions and crystal growth at all stages from nucleation to morphology development and polymorphic structures formation [2]. Therefore the ability to reliably model the solute-solvent interactions is of vital importance. Currently, two principal strategies are available, namely classical ensemble treatments [3] and quantum chemical models involving dielectric continuum representation of the solvent [4]. Within the latter approach so called supermolecule model [5] is also worth mentioning here, in which a cluster, containing a solute molecule and an arbitrary number of explicit solvent molecules, is immersed in a continuous dielectric medium, characterised by the macroscopic dielectric constant of the solvent. The statistical approach is mainly represented by Molecular Dynamics and Monte Carlo simulations [6-7]. The quantum chemical approach is based on the Onsager reaction field theory [8]. In this study we compare the performance of four models based on the reaction field theory and incorporated into the *ab initio* Hartree-Fock scheme (Self-Consistent Reaction Field method [9]), namely the Onsager spherical cavity approximation, the Tomasi's Polarizable Continuum Model [10-11] and its modified versions: Isodensity PCM and Self-Consistent Isodensity PCM [12], in application to the study of dimerisation of the α -hydroxyamide group (see Figure 1) and formamide (Figure 2). Here we consider two modes of dimerisation, i.e. those based on NH...OH and NH...O=C intermolecular hydrogen bonds between monomers, respectively. These and other types of bonding have been examined in detail in our previous papers [13-14], α -hydroxyamide dimer (HAD) and formamide (FMA) are shown schematically in Figure 3 and 4.

2. METHODS

2.1. Dielectric continuum methods

Solvent is represented as a continuous dielectric without discrete internal structure and described only by its macroscopic dielectric constant ϵ . Solute is embedded in a cavity of certain shape and size. The solute charge distribution interacts with the medium, polarising it and creating a reflection charge distribution on the cavity surface (the reaction field) which, in turn, will interact electrostatically with the solute leading to the net stabilisation.

This approach has both advantages and disadvantages. Inability to reproduce specific solute-solvent interactions constitutes its principal shortcoming. The advantages, however, are considerable, namely it allows the geometry and dipole moment to change under the influence of the medium, provides the possibility of calculating vibrational frequencies of solvated species quantum mechanically and does not require any empirical data as opposed to all force field methods.

In the Onsager model, solute whose charge distribution is represented by a simple dipole, is embedded in a typically spherical cavity (however, ellipsoidal cavities are also used) and interacts with the solvent. The energy of these interactions E_{int} is equal to the difference between its solvated and isolated state energies:

$$E_{int} = E_{rf} - E_0 \quad (1)$$

E_{rf} can be calculated directly from the following equation:

$$|\hat{H}^0 + \Gamma \langle \phi | \hat{\mu} | \phi \rangle \hat{\mu} | \Psi = E_{rf} \Psi \quad (2)$$

where \hat{H}^0 is the solute unperturbed Hamiltonian, Ψ is the solute total wave function and factor Γ takes the following form:

$$\Gamma = \frac{1 - \epsilon}{(2\epsilon + 1)a_0^3} \quad (3)$$

for SCaSCRF (spherical cavity approximation).

E_{int} is therefore calculated from the equation (4) [9,15]:

$$E_{int} = \frac{1}{2} \Gamma \langle \bar{\mu}_{tot} \rangle^2 = -\frac{1}{2} \Gamma \langle \langle \phi | \hat{\mu}_{el} | \phi \rangle + \bar{\mu}_{nuc} \rangle^2 \quad (4)$$

where:

$$\hat{\mu}_{el} = \sum_{i=1}^n e_i \hat{r}_i \quad (5)$$

and

$$\bar{\mu}_{nuc} = \sum_{k=1}^N Z_k \vec{R}_k \quad (6)$$

correspond to the electronic component of the dipole moment operator and the nuclear component of the dipole moment, respectively.

Despite Onsager model's obvious deficiencies, i.e. its applicability only to solutes of approximately spherical shape and those whose interactions with the medium may be justifiably represented by dipole-dipole ones, it has proved considerably successful [16].

Polarisable Continuum Model constitutes a significant improvement in the description of solute-solvent interactions, provides us with a means of taking into account a mutual polarisation of solute and solvent in a self-consistent way and allows us to study molecules of complicated, non-spherical shape and charge distribution. In this case a cavity has a much more realistic shape and is constructed of certain number of interlocking spheres (in most cases the number and radii of the spheres correspond to the number of atoms in the solute molecule and their van der Waals radii). Subsequently, the cavity surface is divided into certain number of surface elements ΔS_i with the charges q_i at the centres of them. The surface charge density $\sigma^0(s)$ is derived in the first step from the electric field generated by the unperturbed solute charge distribution $\rho^0(\mathbf{r})$. The surface charges are calculated according to the equation (7):

$$q_i^0 = \sigma(s_i) \Delta S_i \quad (7)$$

These charges constitute a contribution to the electric field generating new $\sigma^1(s)$ and a new set of q_i^1 and these steps are repeated until self-consistence is achieved. Thus we have the reaction field potential V_o^f , due to the final set of charges q_i^f , which is subsequently added to the unperturbed solute hamiltonian:

$$(\hat{H}^0 + \hat{V}_o^f) \Psi = E \Psi \quad (8)$$

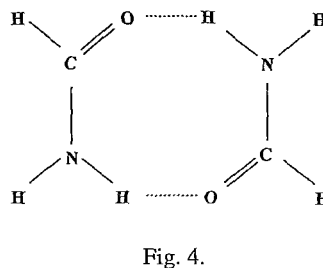
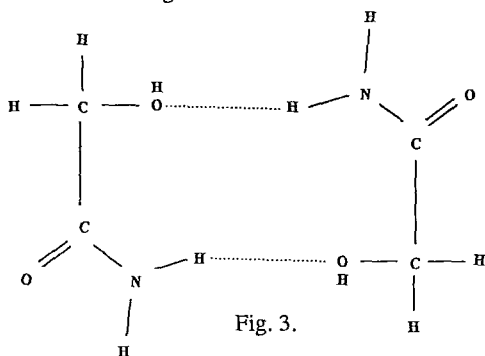
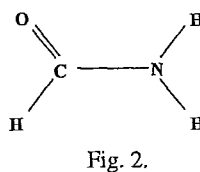
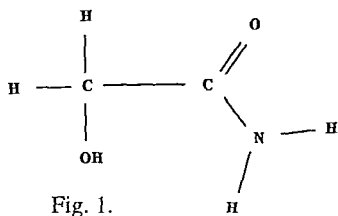
and after having solved the Schrödinger equation, we obtain new, polarised, $\rho^1(\mathbf{r})$. The above steps are repeated until self-consistence of $\rho(\mathbf{r})$ and $\sigma(s)$ is achieved.

There is a range of problems in PCM. First of all, the cost of calculations is significantly higher because going from a spherical to realistic cavity shapes results in slow convergence of numerical integrations, which in turn is due to discontinuities at the intersections of the spheres. It is also worth emphasising that special parameters such as size (volume) & shape of the cavity and surface grid, which affect substantially the results of calculations, are subject to an arbitrary choice.

Further improvements have therefore been necessary. The significant improvements, mainly towards both more realistic shape of the cavity and its varying during the computation, are IPCM and SCIPCM models. In both of them the cavity is defined based upon the isosurface of the total electron density. „Normal" PCM procedure is then followed with one exception, namely the isosurface changes at each iteration since each iteration updates the solute density. The IPCM process, however, is not fully self-consistent. SCIPCM overcomes this problem as it contains a suitable term which couples the isodensity and solute Hamiltonian.

2.2. Computational details

We have employed Gaussian94 [17] package to perform all the computations. Dimers HAD (built of α -hydroxyamide moieties, (Figures 1 and 3)) and FMA (cyclic formamide dimer (Figures 2 and 4)) have been completely optimised in the gas phase at the RHF/6-311++G** [18-20] level and in solution with Onsager and SCIPCM methods. PCM and IPCM energies have been calculated for the gas-phase geometries of HAD and FMA without further reoptimisation. We have considered two values of dielectric solvent: 2.00 and 80.00 corresponding to non-polar (\sim cyclohexane) and polar (\sim water) solvent. In the case of FMA we have additionally carried out a detailed SCIPCM study on the dependence of solvation energy on the dielectric constant of the solvent, namely we have reoptimised the FMA gas-phase structure with ϵ varying from 2.00 to 80.00, with a particularly small step in the region of low dielectric constants.



In Onsager and PCM calculations additional special parameters are necessary : cavity radius and number of points per sphere, respectively. In the Onsager calculations we have used the cavity radii derived from molecular volume calculations (so called *Volume* algorithm in Gaussian94 [21]). In the PCM calculations - for comparison - we have used 80 points per sphere with tightened SCF convergence criteria and 100/150 points per sphere with default SCF criteria for HAD and FMA, respectively. For isodensity models the default value of 0.0004 e/bohr^3 was used to build the cavities.

3. RESULTS AND DISCUSSION

Data concerning the optimised gas-phase and solvated structures of HAD, i.e. inter- and intramolecular hydrogen bonding geometries, α -hydroxyamide group conformation and dipole moments, are collected in Table 1. Solvation energies (i.e. calculated with respect to the gas-phase energy $E_{\text{sol}} = E_{\text{M}} - E_{\text{RHF}}$) as predicted by various methods are given in Table 2. In Tables

3 and 4 we have collected the analogous data for FMA, i.e. NH...O=C hydrogen bonding characteristics and dipole moments for several selected values of dielectric constant ϵ (Table 3) and solvation energies (Table 4). Plot 1 presents the dependence of $E_{\text{sol}}(\text{SCIPCM})$, in kcal/mol, of FMA on dielectric constant of the medium. Figures 5, 6 and 7 show the optimised geometries of HAD from the gas-phase, Onsager/ $\epsilon=80$ and SCIPCM/ $\epsilon=80$ calculations, respectively.

Table 1. HAD - geometrical data

| Geometrical parameters ^a | gas-phase | Onsager model | | SCIPCM model | |
|-------------------------------------|------------|---------------|------------|----------------|------------|
| | | 2.00 | 80.00 | 2.00 | 80.00 |
| PA(inter) | 2.1838(45) | 2.2036 | 2.3639 | 2.1654(71) | 2.1540(2) |
| DA(inter) | 3.1531(6) | 3.1525(6) | 3.2351 | 3.1315(34) | 3.1132(4) |
| DPA(inter) | 164.10(08) | 158.90(1) | 145.74 | 163.02(7) | 161.05(9) |
| PA(intra) | 2.4010(3) | 2.3723(2) | 2.2716 | 2.3894(2) | 2.3658(63) |
| DA(intra) | 2.7352(3) | 2.7255 | 2.6559 | 2.7306(5) | 2.7231(3) |
| DPA(intra) | 98.73(2) | 99.88 | 101.53 | 99.13 | 100.08(7) |
| HOCC | -79.52(69) | -86.09(8) | -146.85(4) | -82.30 | -85.52(46) |
| NCCO | -13.96(88) | -7.99(8) | 18.94(5) | -13.02(-12.98) | -9.88(2) |
| CCCC | -21.48 | -36.25 | -82.18 | -22.80 | -28.44 |
| dipole moment | 4.3328 | 5.8529 | 9.2672 | 4.9313 | 6.0476 |

^adistances given in Å, torsion angles in degrees and dipole moments in Debyes; PA:proton-acceptor distance, DA:donor-acceptor distance, DPA-hydrogen bond angle; two NH...OH bonds are present - the values given in parenthesis next to the values for one of the bonds correspond to the other one.

Table 2. Solvation energies of HAD in kcal/mol

| | Onsager | PCM ^{a,b} | IPCM | SCIPCM |
|--------------------|---------|-------------------------|--------|--------|
| $\epsilon = 2.00$ | -1.05 | -10.68 / -10.40 | -9.59 | -6.88 |
| $\epsilon = 80.00$ | -4.81 | - ^c / -32.42 | -25.44 | -18.32 |

^a 80 points per sphere and tightened SCF convergence criteria

^b 100 points per sphere and default SCF convergence criteria

^c lack of convergence of the first stage of PCM cycle (self-consistence of the electric field and surface charges q_i)

Table 3. FMA - geometrical data

| SCIPCM ϵ | E_{sol} | PA | DA | DPA | NCNC | dipole moment |
|-------------------|------------------|------------|-------------|------------|-------|---------------|
| 2.00 | -3.58 | 2.0497(6) | 3.0399(8) | 169.02(3) | 0.10 | 0.005 |
| 10.00 | -8.20 | 2.0781(56) | 3.0701(678) | 170.23(35) | 0.01 | 0.007 |
| 50.00 | -9.46 | 2.0845(29) | 3.0769(54) | 170.53(62) | 0.03 | 0.005 |
| 80.00 | -9.58 | 2.0874(68) | 3.0796(2) | 170.55(61) | -0.05 | 0.004 |
| gas-phase | 0 | 2.0299 | 3.0179 | 168.86(4) | -0.03 | 0.000 |

distances given in Å, torsion angles in degrees, energies in kcal/mol and dipole moments in Debyes; PA: proton-acceptor distance, DA: donor-acceptor distance, DPA-hydrogen bond angle; two NH...O=C bonds are present - the values given in parentheses next to the values for one of the bonds correspond to the other one

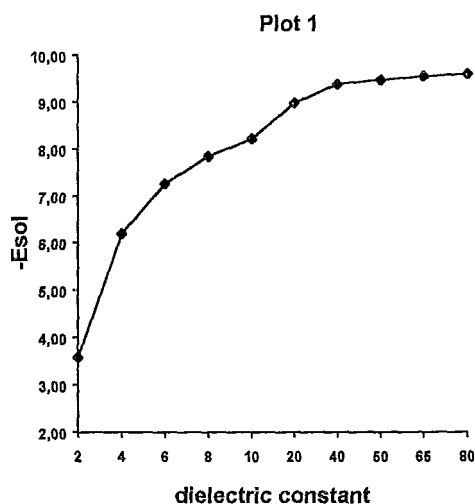
Table 4. Solvation energies of FMA in kcal/mol

| | Onsager | PCM ^a | IPCM | SCIPCM |
|--------------------|---------|---------------------|--------------------|--------|
| $\epsilon = 2.00$ | 0.00 | -3.95 ^b | -3.50 ^d | -3.58 |
| $\epsilon = 80.00$ | 0.00 | -14.42 ^c | 10.99 ^e | -9.58 |

^a 150 points per sphere and default SCF convergence criteria

^b dipole moment $\mu = 0.031$

^c $\mu = 0.027$; ^d $\mu = 0.034$; ^e $\mu = 0.035$



As one can see from the data collected in Tables 1 and 2 there are significant differences between the results produced by the methods employed. The Onsager model predicts much larger changes in the dipole moment, intermolecular hydrogen bond angles and also in the

$\text{CH}_2(\text{OH})\text{COOH}$ group conformation. The latter one is also, according to the Onsager results, less planar while in the SCIPCM the degree of planarity increases under the rising polarity. Interestingly, Onsager model predicts elongation of PA (proton-acceptor) and DA (donor-acceptor) distances with the increase in solvent polarity while in SCIPCM we observe shortening of these distances. Also the differences between both $\text{NH}\dots\text{OH}$ bonds are larger according to the SCIPCM model.

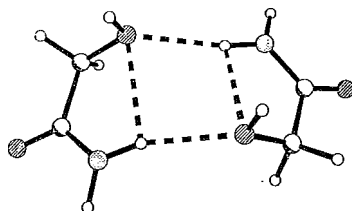


Fig. 5.

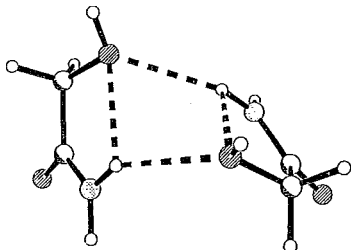


Fig. 6.

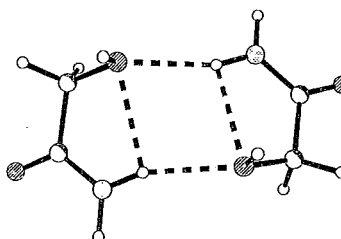


Fig. 7.

Intramolecular bonds in the Onsager and SCIPCM optimised structures are similar and the direction of changes when increasing the polarity is also similar. It is worth noticing that in Onsager (polar solvent calculations the mutual arrangement of α -hydroxycarboxylic planes is much different from all the rest of the resulting structures (the angle between the planes is about 90° in Onsager) $\epsilon = 80$ and about 120° in the rest of the structures).

Degree of stabilisation of the HAD dimer also varies significantly with the model used. In general Onsager model predicts very low solvation energies, i.e.

- 1 kcal/mol in a non-polar solvent versus approximately 10 and 7 kcal/mol in PCM/IPCM and SCIPCM, respectively)
- about 5 kcal/mol in a polar solvent compared to about 30 and 18 kcal/mol in PCM/IPCM and SCIPCM, respectively.

It appears important to report a failure in obtaining convergence of the PCM scheme in the case of 80 points per sphere calculation at the stage of self-polarisation, i.e. obtaining self-consistent set of reaction field surface charges. Convergence has been obtained when we employed 100 points per sphere.

In the case of FMA, because of its very low dipole moment, the Onsager model cannot be applied at all since in such a case it produces the same results as those obtained from the gas-phase

calculations. PCM/IPCM/SCIPCM predict stabilisation of FMA dimer in both non-polar and polar media. The decrease in the solute energy is observed - by about 4 kcal/mol in a non-polar solvent and between 10-14 kcal/mol in a polar one. Dependence of hydrogen bonding geometrical parameters on dielectric constant shows a consistent elongation of both PA and DA distances as the polarity increases. This is accompanied by a slight but also systematic increase in the hydrogen bond angle. PA increases by about 0.02Å in a non-polar surrounding and by 0.06Å in a polar one. Analysis of Plot 1, which shows the SCIPCM E_{sol} versus ϵ for FMA, points to the interesting fact, namely that the largest changes in E_{sol} , for relatively small increase in the solvent dielectric constant, occur in the non-polar region, while for highly polar media (beginning with $\epsilon \approx 40.00$) we observe a kind of saturation and E_{sol} changes only by tenths of kcal/mol with the increase in polarity from 65 to 80.

CONCLUSIONS

In conclusion we would like to emphasise the fact that all models, despite the differences, predict a considerable degree of stabilisation of HAD and FMA (except for the Onsager model), further increasing with the increase in the polarity of the medium. Both optimised structures predicted by Onsager model and SCIPCM, although noticeably different in the case of HAD, retain the same scheme of intermolecular and intramolecular hydrogen bonding. The similarities between the solvation energies predicted by PCM/IPCM/SCIPCM group of models, for non-polar solvents in particular, and the difference between them and that of the Onsager model suggest that dipolar representation of the molecular charge distribution is a considerable oversimplification in the case of HAD and cannot be applied at all in the case of FMA. Inapplicability of Onsager model to the FMA dimer stems from the dimer's very low dipole moment.

Acknowledgements

KBN is gratefully acknowledged for supporting us with the grants number 8T11F 029 15 and SPUB/COST/T-9/D-9. We are also grateful to Poznań Supercomputing and Networking Center for access to Cray Y MP EL and Cray J916.

References

- [1] B.Ya. Simkin, I.I. Sheikhet *Quantum Chemical and Statistical Theory of Solutions: a Computational Approach* Ed. T. J. Kemp, Ellis Horwood Ltd. 1995.
- [2] a) A. Gavezzotti, G. Filippini, Chem. Commun. **278**; (1998) b) A. Gavezzotti, G. Filippini, J. Kroon, B. P. van Eijck, P. Klewinghaus, Chem. Eur. J. **3**, 893 (1997); c) N. Blagden, R. J. Davey, H. F. Lieberman, L. Williams, R. Payne, R. Roberts, R. Rowe, R. Docheity, J. Chem. Soc. Faraday Trans. **94**, 1035 (1998).
- [3] J. P. Hansen, I. R. McDonald *Theory of Simple Liquids*, Academic Press, Harcourt Brace Jovanovich 1986.
- [4] P. Claverie, J. P. Daudey, J. Langlet, B. Pullman, D. Piazzola, J. Phys. Chem. **82**, 405 (1978).
- [5] G. W. Schnuelle, D. L. Beveridge, J. Phys. Chem. **79**, 2566 (1975).
- [6] M. P. Allen, D. J. Tildesley *Computer Simulation of Liquids*, Clarendon Press, Oxford 1996.
- [7] A. R. Leech *Molecular Modelling Principles and Applications* Addison Wesley Longman Ltd., London 1996.
- [8] L. Onsager, J. Am. Chem. Soc. **58**, 1486 (1936).

-
- [9] M. W. Wong, M. J. Frisch, K. B. Wiberg, J. Am. Chem. Soc. **113**, 4776 (1991).
- [10] S. Miertus, E. Scrocco, J. Tomasi, Chem. Phys. **55**, 117 (1981).
- [11] S. Miertus, J. Tomasi, Chem. Phys. **65**, 239 (1982).
- [12] J. B. Foresman, T. A. Keith, K. B. Wiberg, J. Snoonian, M. J. Frisch, J. Phys. Chem. **100**, 16098 (1996).
- [13] A. Szarecka, U. Rychlewska, J. Rychlewski, J. Mol. Struct, in press.
- [14] A. Szarecka, J. Rychlewski, U. Rychlewska, Progress in Theoretical Chemistry, in press.
- [15] a) M.W. Wong, M. Frisch, J. Chem. Phys. **95**, 8991 (1991); b) O. Tapia, O. Goscinski, Mol. Phys. **29**, 1653 (1975).
- [16] M. W. Wong, K. B. Wiberg, M. J. Frisch J. Am. Chem. Soc. **92**, 523 and 1645 (1992).
- [17] Gaussian94 (Revision D.1), M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. A. Keith, G. A. Patersson, J. A. Montgomery, K. Raghavachari, M. A. Al.-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. Stefanow, A. Nanayakkara, M. Challacombe, C. Y. Peng,, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin. D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 1995.
- [18] Pople, J. A.; Beveridge, D. L.; *Approximate Molecular Orbital Theory*, McGraw-Hill; New York, 1970.
- [19] Clark T., Chandrasekhar J., Spitznagel G. W., Schlayer von P. R. (1983) J. Comput. Chem. **4**, 294.
- [20] a) M. J. Frisch, J. A. Pople, J. S. Binkley J. Chem. Phys. **80**, 3265; (1984) b) R. Krishnan, J. S. Binkley, R. Seeger, J. A. Pople J. Chem. Phys. **80**, 3265 (1980); c) M. W. Wong, M. J. Frisch, K. B. Wiberg, J. Am. Chem. Soc. **113**, 4776 (1991).
- [21] a) Gaussian94 User's Guide; b) J. B. Frisch, A. Frisch, *Exploring Chemistry with Electronic Structure Methods*, Sec. Edition, Gaussian, Inc. Pittsburgh, PA, 1996.