

AGOA: A Hydration Procedure and Its Application to the 1-Phenyl- β -Carboline Molecule

Marcelo Z. Hernandez* , João B. P. da Silva and Ricardo L. Longo

Departamento de Química Fundamental, Universidade Federal de Pernambuco,
Av. Prof. Luiz Freire, s/n, Cidade Universitária 50740-540, Recife - PE, Brazil

Um novo procedimento, denominado AGOA, foi desenvolvido e implementado em um programa escrito em FORTRAN 77 para explorar as estruturas de hidratação ao redor de solutos polares. Este procedimento utiliza-se da análise do potencial eletrostático molecular (MEP) do soluto (ou da supermolécula) e pode ser generalizado para outros solventes polares além da água. Este procedimento foi testado para moléculas simples, bem como para moléculas complexas de interesse farmacológico, como os sistemas β -carbolínicos, derivados do indol. Este foi um teste rigoroso, e mostra que o procedimento AGOA é robusto, flexível e de baixa demanda computacional para estudos de hidratação, sendo possível a utilização de funções de onda semi-empírica ou *ab initio*. Comparações com métodos que otimizam a geometria soluto-água mostram a superioridade do procedimento AGOA no estudo da hidratação dos confôrmeros *anti* e *syn* da β -carbolina, viabilizando o seu uso para a compreensão e quantificação das interações específicas que contribuem para os efeitos do solvente.

A new procedure, named AGOA, has been developed and implemented in a computer program written in FORTRAN 77 to explore the hydration structures of polar solutes using its molecular electrostatic potential (MEP). This procedure can be generalized to polar solvents other than water. It has been tested for several small molecules, and applied to complex molecules of pharmacological interest, such as the β -carboline systems derived from indole. This is a stringent, but not general, test of the AGOA procedure and shows its robustness, flexibility and low computational costs, since either semiempirical or *ab initio* wavefunctions can be employed. The comparisons with procedures based upon the geometry optimization of the solute-water complex show the superior performance of the AGOA procedure for the *anti* and *syn* β -carboline conformers, reassuring its use to comprehend and to quantify the specific interactions involved in solvent effects.

Keywords: AGOA, hydration, solvent effects, β -carboline, electrostatic potential

Introduction

Most chemistry and biochemistry occur in condensed media, in particular, aqueous solutions. Thus, the proper simulation of these processes has to take into account the solvent effects. Consequently, since the pioneer work of Born¹ on ionic solvation, these solvent effects have been shown to be of fundamental importance for many chemical and biological processes and have then been receiving considerable attention.² There are basically three models^{3,4} to describe the solvent, namely, the continuum or dielectric model, the discrete or supermolecule model, and the discrete-continuum model, which attempts to combine the two previous ones. The continuum model treats the solvent

as a structureless dielectric medium and the solute is inserted in a cavity. Since this is a classical macroscopic description there is not a unique way of integrating it into a quantum chemical description of the solute. Thus, there are many distinct implementations of the continuum model, ranging from sophisticated Poisson equation solutions on an isoelectronic surface⁵ to simple spherical dipole reaction field.⁶ However, independent of the implementation, the continuum models are not able to describe specific solute-solvent interactions, in particular, hydrogen bonds. In addition, the definition of the solute cavity and the dielectric constant are arbitrary. The discrete model treats the solvent as individual molecules, which interact with the solute via a parametric potential⁷ (classical models) or an instantaneous Coulombic interaction between the electrons and the nuclei of the solute and the solvent molecules

* e-mail: zaldini@npd.ufpe.br

(quantum models). This model solves, at least partially, the problems with the continuum model, in particular, the proper description of specific solute-solvent interactions. However, the discrete models are much more computationally demanding than the continuum ones, and are highly dependent upon the positions of the solvent molecules around the solute. The most appropriate positions are obtained from statistical mechanics simulations⁷, that are not only very demanding, but also require the solute-solvent and the solvent-solvent interaction potentials, which are quite cumbersome to be obtained. Another approach⁸⁻⁹ consists in positioning the solvent molecules randomly around the solute and then use an optimization procedure to obtain the structure corresponding to the energy minimum. This procedure in addition to be very computationally demanding, is also highly dependent upon the starting structure, since the solute-solvent energy potential surface presents a large number of local minima. Thus, alternative approaches have been developed for properly positioning the solvent molecules around the solute without the need for statistical sampling techniques and/or for the explicit interaction potentials. This is the main concern of the present contribution, namely, to present a simple procedure to hydrate polar molecules, which has been denominated AGOA. This AGOA procedure is based upon the molecular electrostatic potential (MEP) of the solute molecule and the assumption that the most important interactions between the solute and the water is electrostatic, so that the positions of the water molecules are mostly defined by the solute MEP. The solute MEP is calculated with quantum chemical methods, thus limiting this approach only to very large size molecules, such as proteins or DNA. However, since the MEP is nearly localized and additive,¹⁰⁻¹² it is not very difficult to extend the present approach to treat fragments of the macromolecule and after combining these fragments to obtain adequate hydration structures for larger molecules.

The indole derivative, 1-phenyl-1,2,3,4-tetrahydro-carboline, also known as 1-phenyl- β -carboline, see Figure 1, is the polar solute chosen to apply this methodology and its implementation to obtain the hydration structures. This molecule belongs to a new class of antimicrobial compounds and is under a QSAR study in our laboratory. In addition, this compound is an example of the indole systems, which have received attention lately due to their biological importance as the chromophore of the tryptophan as well as their biological activities.¹³ Their spectroscopic and biological properties are highly dependent upon the solvation, in particular, their hydration. Therefore, several high level theoretical studies have been performed for the indole-water system,¹⁴⁻¹⁶ which suggest that the water

molecule can also strongly interact with the π -electrons of the indole ring. As a result, the 1-phenyl- β -carboline should provide a stringent test of the AGOA hydration procedure.

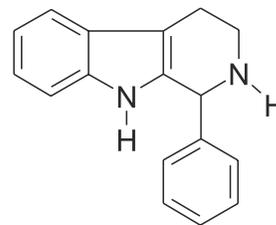


Figure 1. 1-phenyl-1,2,3,4-tetrahydro-carboline or 1-phenyl- β -carboline.

This compound presents basically two predominant conformations (*syn* and *anti*) described by the relative position of the phenyl group with respect to the hydrogen atom bonded to the nitrogen atom of the piperidine. Both conformers have been hydrated by the AGOA procedure.

Theoretical Procedure and Methodology

The AGOA hydration procedure has been implemented in FORTRAN 77, which has been tested and implemented in several machines and operating systems. This implementation of the AGOA procedure uses a file generated by the GAUSSIAN program¹⁷, named "cube" that contains the molecular electrostatic potential (MEP) of the solute molecule calculated in a 3D-grid, available for all the wavefunctions implemented in the GAUSSIAN program. This grid is chosen in such a way that the entire solute is embedded into it and the points are properly spaced so that they are kept to a minimum, but still close enough to provide a good approximation to the gradients of the MEP. The implementation of the AGOA procedure excludes the points of this grid that overlap with the solute atoms, defined by a cutoff radius properly chosen. The AGOA program has stored internally default values for the cutoff radii for H, C, N, O, P, S, F, Cl and Br atoms, but it allows the user to input cutoff radii for all the remaining periodic table.

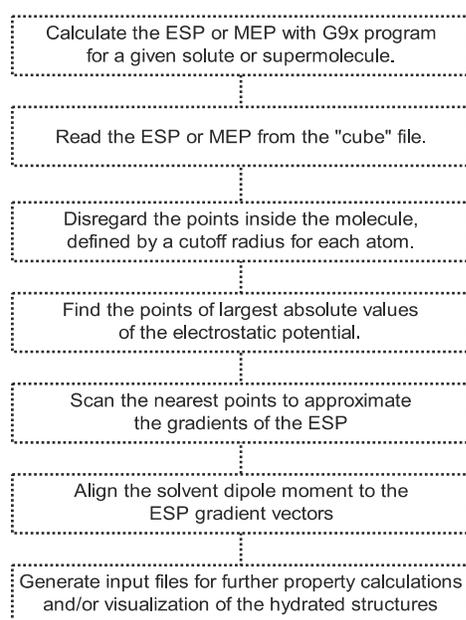
The present implementation of the AGOA procedure uses the TIP4P model for the water molecule¹⁸. This model has been widely employed in liquid water simulations and is a four-site model located at the three atoms and the additional site at 0.15 Å from the oxygen atom towards the hydrogen atoms. This site is treated as a dummy atom ("XX") in the AGOA program.

Depending upon the type and size of the substrate to be hydrated, one or more water molecules might be added simultaneously at each AGOA run. For small molecules, however, the water molecule usually perturbs significantly

the MEP of the solute. Thus, a more appropriate procedure for small solutes would be the sequential addition of the water molecules. More specifically, the MEP of the isolated solute is calculated and the AGOA program decides the best position to add a water molecule. Then, this supermolecule (solute + 1 water) is used in the calculation of the new MEP, so that, the AGOA program can place the second water molecule, that shall form a new supermolecule (solute + 2 waters), and the process is repeated until the appropriate hydration number has been reached.

Once the MEP 3D-grid has been calculated and the points that overlap with the solute (or supermolecule) have been excluded, the AGOA program performs a search to find the points in the grid corresponding to the largest negative and positive values of the MEP. The neighboring points are used to estimate the gradient of the MEP, so that the water dipole moment can be placed parallel or anti-parallel to the largest gradient vector. This procedure defines the positions of the hydrogen atoms of the water molecule, except for the dihedral angle between the three atoms of the water molecule and a given atom of the solute (or supermolecule), that is chosen randomly. As a result, the coordinates of the solute (or supermolecule), the oxygen and the dummy ("XX") atoms of the water molecule are defined in cartesian, but the hydrogen atoms of the water molecule are defined in term of internal coordinates, since it is necessary to establish the dihedral angle. Scheme 1 summarizes the flow-chart of the AGOA procedure.

Once the hydration structure has been obtained it is usually transformed into cartesian coordinates and used directly to compute the solute properties, like QSAR



Scheme 1. The algorithm for the AGOA procedure.

descriptors, in the presence of some solvent molecules, or it can be used as an initial guess for a geometry optimization procedure. In addition, this hydration structure can be employed in the determination of the solute-solvent interaction energy (E_{s-w}) evaluated approximately as,

$$E_{s-w} = E_{\text{Total}} - E_{\text{Solute}} - E_{\text{Water}} \quad (1)$$

where E_{Total} , E_{Solute} and E_{Water} correspond to the total energy of the system (solute + water), the isolated solute energy and the energy of the water cluster, respectively. In the present application of the AGOA program these energies were computed at AM1 ("Austin Model 1")¹⁹ level within the Gaussian 94 program.¹⁷ These hydration structures obtained with the AGOA program were then submitted to a partial geometry optimization, where the solute molecule was maintained frozen during the optimization procedure, allowing only the coordinates of the water molecules to be optimized. These new solute-solvent configurations are denominated AGOA-OPT.

Results and Discussion

The 3D grid of the molecular electrostatic potential (MEP) for the *anti* conformer of the 1-phenyl- β -carboline is presented in Figure 2. This MEP has been calculated with the AM1 method and the grid contains $20^3 = 8000$ points with a dimension of $20 \text{ \AA} \times 20 \text{ \AA} \times 20 \text{ \AA}$. It should be noted that this grid has successfully enclosed the whole solute molecule.

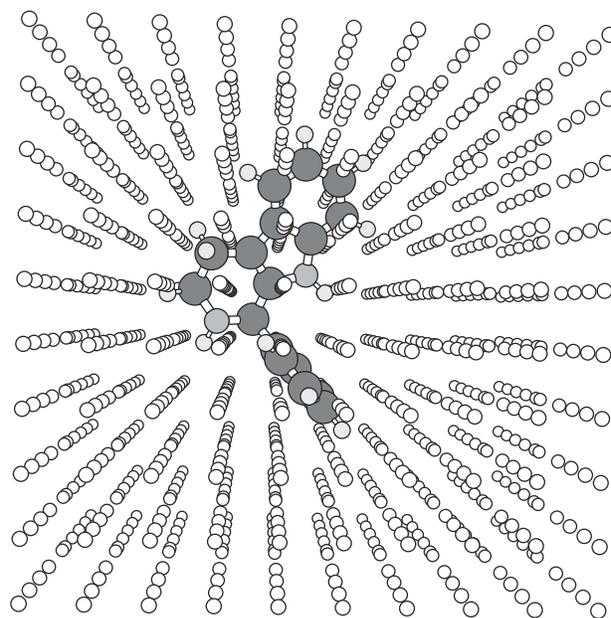


Figure 2. 3D grid of the calculated electrostatic potential for the *anti* conformer.

The AGOA program excludes the points of this grid that are inside the volume defined by the solute (or supermolecule). It has been verified that the implemented procedure properly excluded all points inside the solute according to the default cutoff radii.

Following, a search within this new grid is then performed and for the same conformer the points with the largest negative and positive values are selected. In addition, the gradient vectors of the MEP at each of these selected points are estimated and are displayed in Figure 3.

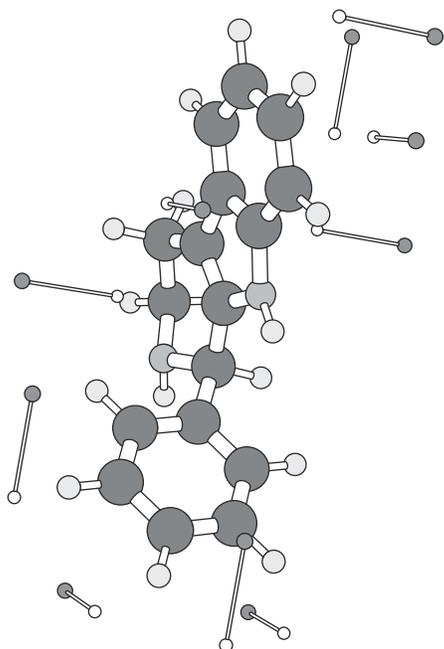


Figure 3. *Anti* conformer with 10 vectors indicating the orientations of the dipole moment for the water molecules.

It is worthwhile noticing that this procedure readily yields hydration structures that include the interaction of the water molecule with the π -electrons of the indolic ring. This result is very encouraging since previous studies of the indole-water systems required high level *ab initio* calculations to yield such a hydration structures.¹⁴⁻¹⁶

For each hydration structure the solute-water (E_{s-w}) interaction energy has been estimated according to equation (1). These hydration structures were also used as the initial guess for a partial geometry optimization of the water molecules (AGOA-OPT) with the AM1 method. Comparisons between the results without (AGOA) and with (AGOA-OPT) the geometry optimization are presented in Table 1 and illustrated in Figure 4.

As expected, the solute-water interaction energies are larger with the AGOA-OPT procedure, since it allows for an inter and intramolecular relaxation of the water molecules. However, this result seems to be an artifact of

the AM1 method, which usually yields unrealistic hydration structures and hydration energies.²⁰ In addition, this AGOA-OPT procedure lacks convergence of the interaction energy with respect to the number of water molecules added. It also lacks a consistent *anti/syn* energy relationship according to the hydration of each conformer. In contrast, the AGOA procedure yielded monotonic convergent results for the solute-water interaction energy as well as a consistent relationship between the energies of the *anti* and

Table 1. Solute-water intermolecular energies in kJ mol^{-1} (see Equation 1)^{a)}

number of water molecules	AGOA		AGOA-OPT	
	<i>anti</i>	<i>syn</i>	<i>anti</i>	<i>syn</i>
1	-2.46	-9.80	-15.52	-13.20
2	-6.06	-12.55	-26.37	-22.80
3	-11.93	-16.09	-32.94	-35.73
4	-14.41	-26.58	-22.97	-41.60
5	-16.18	-29.25	-28.46	-33.31
6	-17.76	-30.00	-42.06	-41.27
7	-18.77	-29.66	-27.29	-47.77

^{a)}The enthalpies of formation of the isolated *anti* and *syn* conformers are 354.18 and 348.44 kJ mol^{-1} , respectively.

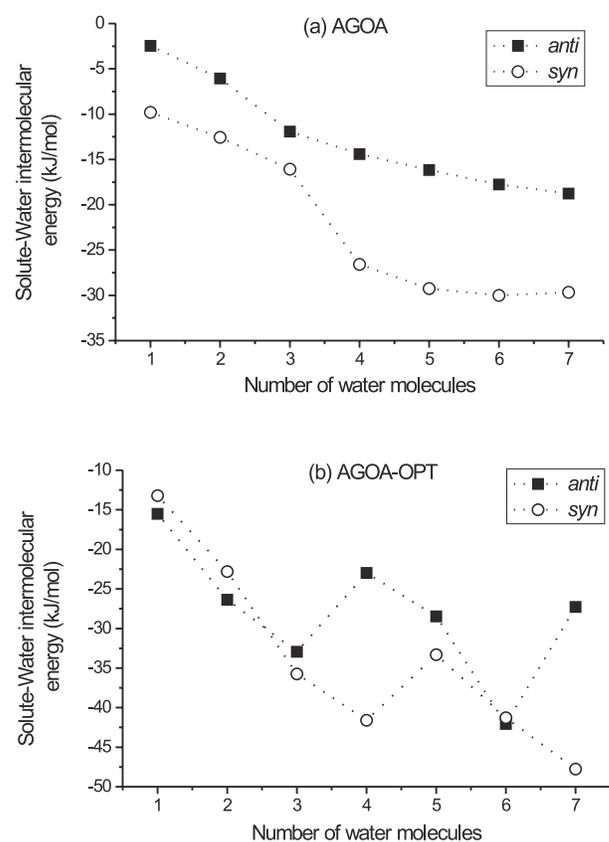


Figure 4. The solute-water interaction energy as a function of the number of water molecules.

syn conformers. As a result, it can be seen that the hydration is important for the relative stabilization of the conformers, and that the *syn* conformer is consistently more stable than the *anti* one. In attempting to establish the origin for the preferential hydration of the *syn* conformer the hydration structures of both conformers are displayed in Figure 5.

For comparison, the hydration structures of these conformers obtained with the AGOA-OPT procedure are illustrated in Figure 6.

It can be seen that the geometry optimization with the AM1 leads to the clustering of the water molecules as well as to the formation of bidentate hydrogen bonds, which is

an artifact of the AM1 method. It should also be noted that this optimization procedure yields hydration structures where the interactions of the water with the indole π -electrons are absent. These unrealistic results yielded by the optimization with the AM1 are also a consequence of simulating the solvent effects by a finite cluster, and since in the bulk these solvent molecules would be strongly interacting with the remaining water molecules present in the liquid, it should decrease the solute-solvent interaction energy. These results are also reflected into the convergence of the interaction energy with respect the number of water molecules added in the hydration process (see Figure 7).

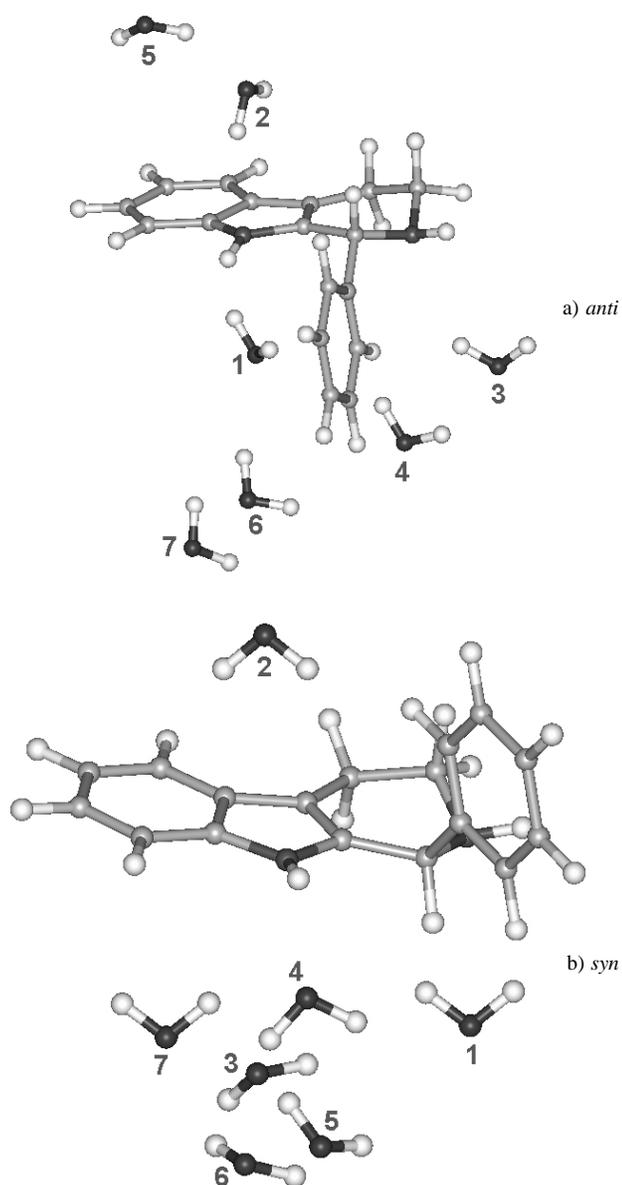


Figure 5. The hydration structures of the (a) *anti* and (b) *syn* conformers using the AGOA procedure.

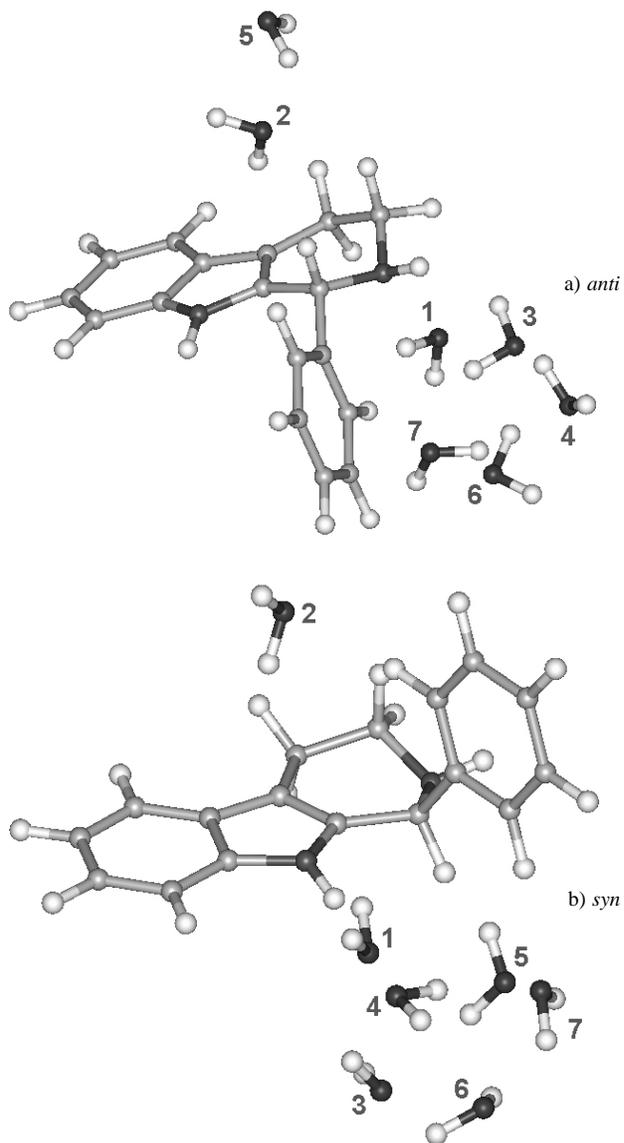


Figure 6. Hydration structures of the (a) *anti* and (b) *syn* conformers using the AGOA-OPT procedure.

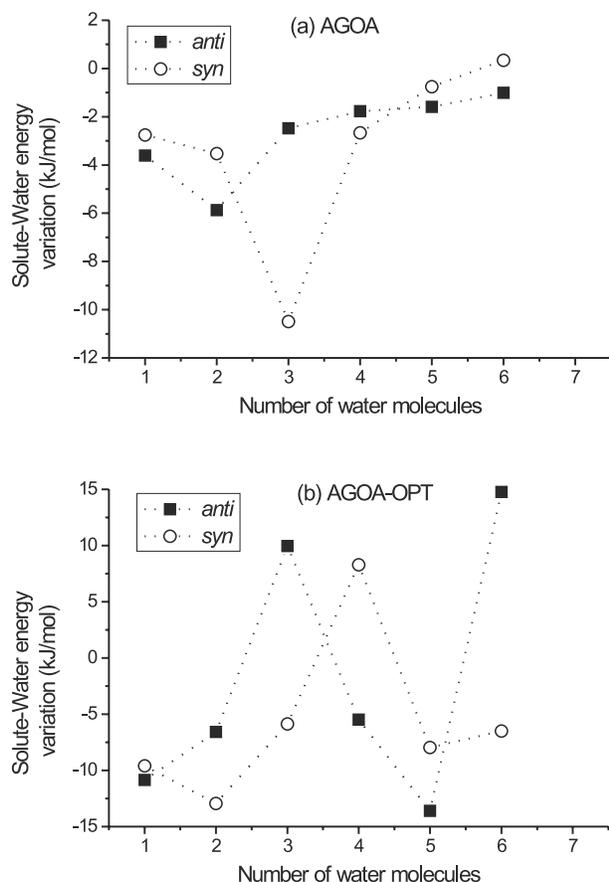


Figure 7. The solute-water energy variation as a function of the number of water molecules.

Consequently, it thus seem very difficult to correctly represent the solute in solution using these optimized cluster models in vacuum. However, the AGOA procedure, without the geometry optimization, provides realistic results, including the interaction with the π -electrons of the indole ring system. More specifically, these interactions involve the water molecules labeled 1 and 2 in the *anti* and *syn* conformers, respectively. The importance of these results is related to the fact that it was necessary to employ high level *ab initio* methods (MP2/DZP) in order to describe these water-indole π -electrons interactions,¹⁴ which are promptly provided by the AGOA procedure using the MEP from the AM1 method.

Conclusions

The AGOA procedure for determining hydration structures has proven to be quite efficient, even for stringent tests such as the hydration of the 1-phenyl- β -carboline molecule. This test can be considered stringent, but not

general, since this molecule contains an indolic ring system for which the AGOA procedure provided the proper description for the interaction of the solvent (water) with the π -electrons of the indole ring system.

The AGOA procedure is computationally efficient, with the bottleneck being the wavefunction calculation, which can be either semiempirical or *ab initio*. The AGOA program can be coupled to a visualization program, such as, RasMol²¹ or Chime,²² which can then provide qualitative tools for understanding the hydration of polar molecules.

The Fortran code has been written in such a way that it can be easily generalized for other polar solvents, such as, methanol, dimethylether, etc. and/or for other quantum chemical programs.

Further improvements of the AGOA procedure include an automated choice of the 3D-grid parameter, which would make the procedure independent of the cutoff radius as well as the implementation of a cutoff radius for the solvent molecules, thus avoiding the generation of highly correlated hydration structures at each AGOA run.

References

- Born, M. Z.; *Physik* **1920**, *1*, 45.
- Cramer, C. J.; Truhlar, D. G.; *Rev. Comput. Chem.* **1995**, *6*, 72.
- Dillet, V.; Rinaldi, D.; Rivail, J.-L.; *J. Phys. Chem.* **1994**, *98*, 5034.
- Cappelli, C.; Mennucci, B.; da Silva, C. O.; Tomasi, J.; *J. Chem. Phys.* **2000**, *112*, 5382.
- Tomasi, J.; Cammi, R.; Mennucci, B.; *Int. J. Quantum Chem.* **1999**, *75*, 783.
- Tapia, O.; Goscinski, O.; *Mol. Phys.* **1975**, *29*, 1653.
- Allen, M. P.; Tildesley, D. J. In *Computer Simulation of Liquids*; Clarendon Press; Oxford, 1987.
- Freitas, L. C. G.; Longo, R. L.; Simas, A. M.; *J. Chem. Soc., Faraday Trans.* **1992**, 189.
- Longo, R. L.; Nunes, R. L.; Bieber, L. W.; *J. Braz. Chem. Soc.* **2001**, *12*, 52.
- Politzer, P.; Murray, J. S.; *Rev. Comput. Chem.* **1991**, *2*, 273.
- Bonaccorsi, R.; Scroco, E.; Tomasi, J.; *J. Am. Chem. Soc.* **1977**, *99*, 4546.
- Bonaccorsi, R.; Scroco, E.; Tomasi, J.; *J. Am. Chem. Soc.* **1976**, *98*, 4049.
- Andreani, A.; Locatelli, A.; Rambaldi, M.; Leoni, A.; Bossa, R.; Fraccari, A.; Galatulas, I.; *Anticancer Res.* **1996**, *16*, 3585.
- van Mourik, T.; Price, S. L.; Clary, D. C.; *Chem. Phys. Lett.* **2000**, *331*, 253.
- Carney, J. R.; Zwier, T. S.; *J. Phys. Chem. A* **1999**, *103*, 9943.
- Mons, M.; Dimicoli, I.; Tardivel, B.; Piuze, F.; Brenner, V.; Millie, P.; *J. Phys. Chem. A* **1999**, *103*, 9958.

17. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T. A.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; . Foresman, J. B.; Peng, C. Y.; Ayala, P. Y.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. In *Gaussian 94* (Revision D.1); Gaussian, Inc.; Pittsburgh PA, 1995.
18. Jorgensen, W. L.; Madura, J. D.; *Mol. Phys.* **1985**, *56*, 1381.
19. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P.; *J. Am. Chem. Soc.* **1985**, *107*, 3902.
20. Bernal-Uruchurtu, M. I.; Ruiz-López, M. F.; *Chem. Phys. Lett.* **2000**, *330*, 118.
21. Roger Sayle, <http://www.umass.edu/microbio/rasmol>, accessed in June/2001.
22. Eric Martz, <http://www.umass.edu/microbio/chime>, accessed in June/2001.

Received: March 19, 2001

Published on the web: September 18, 2001