GAFLEX: A Hybrid GA Engine for Mapping the Configuration Hyper-Space of Organic Compounds in Solution

Carlos A. Del Carpio
carlos@translell.eco.tut.ac.jp
Laboratory for Informatics & AI in Moleclular and Biological Sciences
Department of Ecological Engineering, Toyohashi University of Technology
Tempaku, Toyohashi 441-8580, Japan

1 Introduction

Biological activity of drugs and other organic compounds is determined by their solubilities in water or any other solvent or solution. Several methodologies to compute the solubilities of organic compounds have been devised, however most of them are bound to aqueous solutions, and many times the methodology works for a limited number of compounds. Furthermore, those methodologies state very little or nothing on the structure that the compound adopts in solution, and to the date comparisons of the structural characteristics of a molecule in solution as referred to those in the gas phase have been rarely reported.

In the present work we present a new and general approach for the prediction of the most stable three dimensional structure of an organic compound in water which may be extended to any polar or apolar solvent. From this structure (in fact, several structures), thermodynamical properties of the organic molecules as well as solubilities and bio-activities in solution can be derived.

2 Methodology

The methodology proposed here to predict the conformation of organic compounds in solution consists in: (a) a heuristic algorithm to search the conformation space of the solute, and (b) the development of the solvent effect module. The conformational search is performed by a genetic algorithm (GA) engine [1], suitable for finding minima in complicated multidimensional surfaces. The evolutionary engine operates on rotable bonds of the molecule as well as on ring vertices and edges, flipping and flapping them. These perturbations are performed as part of the GA mutation operation.

The intramolecular force field on which the GA operates consists of a bond and angle stretching function, a torsional energy function, van der Waals interactions, and electrostatic interactions among the atoms constituting the solute. The GA here is hybridized with a Newton-Raphson optimization procedure to refine the energy minimization.

On the other hand, the solvent effect module expresses the constraints that the solvent imposes to the solute in solution. The module takes into account the energy due to the cavity formation of the solute in solution [2], the van der Waals interaction between the solute atoms and the surrounding solvent molecules, and the solvent-solute electrostatic interaction. The solvent is treated as a continuum, and the scale particle theory is used to express the thermodynamics of cavity formation. Electrostatic interactions are expressed by means of the generalized Born equation [3], while van der Waals interactions are computed by an empirical model based on the calculation of the solvent accessible surface area. GAFLEX (the acronym for Genetic Algorithms for molecular FLEXibility analysis) incorporates all these functions, and comparisons of geometrical characteristics in gas and solution can be carried out for any organic molecule.

3 Results and Discussion

GAFLEX was applied to the prediction of the configuration in solution of several types of organic compounds. Fig. 1, shows the relation between experimental and the GAFLEX computed values of solvation energies for different kinds of compounds. The correlation is very high, although still some parameterization has to performed in order to handle several other kinds of small organic molecules. Correlation of the solvation energies with any of the different solubilities scales may also allow the prediction of the solubility of a compound. Here we have introduced a new force field based on SASA to compute the 3D structures of small organic compounds in solution. The GA operating on the variable elements of the structure allows the mapping of the configuration hyper-space of the molecule, making possible the prediction of the most stable conformation in solution. Further development of the system may allow prediction of the 3D of molecules in several other solvents.

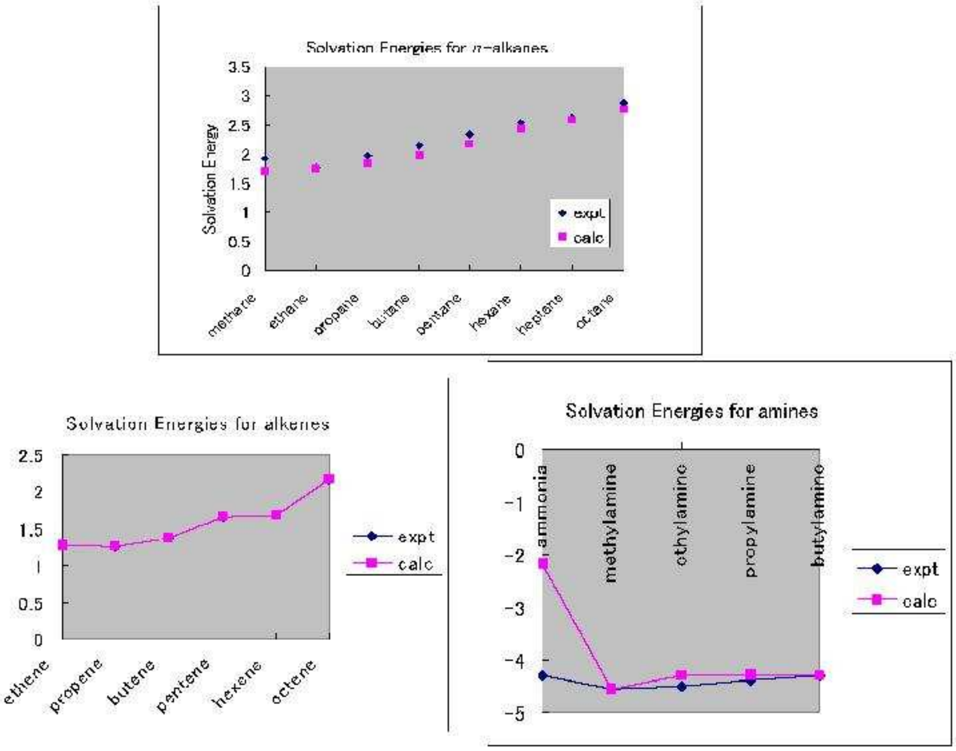

Figure 1: Calculated and experimental solvation energies for several types of organic compounds.

References

[1] Del Carpio, C.A. and Gogonea, V., Modelling Proteins Conformation in Solution. Part I: A Parallel GA Engine for Protein Conformational Space Mapping, *Genome Informatics 1996*, Universal Academy Press, 108–109, 1996.

[2] Ben Naim, A. and Marcus, Y., Solvation Thermodynamics of Nonionic Solutes, *J. Them. Phys.*, 81:2016, 1984.

[3] Still W.C., Tempczyk, A. and Hawley, R.C., and Hendrickson, T., Semianalytical Treatment of Solvation for Molecular Mechanics and Dynamics, *J. Am. Chem. Soc.*, 112:6127–6129, 1990.