• ***Нови подходи*** •
• ***New Approaches***•

TEACHING COMPUTATIONAL CHEMISTRY BY STUDYING MALARIA

João ELIAS, Maycon LOBATO, Antonio FIGUEIREDO, Marcos SANTOS, Márcio FARIAS, Williams MACÊDO, Jardel BARBOSA, José PINHEIRO
Universidade Federal do Pará, BRASIL

Abstract. This paper describes our strategy to introduce Computational Chemistry in the undergraduate courses by studying malaria. Artemisinin and its 8-deoxyartemisinin derivative with significant differences in biological activities against this disease are studied. Complete geometry optimization is done with the Hartree-Fock-Roothaan (HFR) theory with several basis sets and with semiempirical methods. The Principal Component Analysis (PCA) and Hierarquical Cluster Analysis (HCA) methods are used to select the most appropriate method and basis set to model molecules. Maps of molecular electrostatic potential (MEP) and the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) are used to correlate structure-activity. Also the docking is employed to investigate the interaction between artemisinin and heme. By this strategy students can study Computational Chemistry, Theoretical Chemistry, Chemometrics and Medicinal Chemistry.

Keywords: in the classroom, computers in chemistry, theoretical chemistry, chemometrics, interdisciplinary

Introduction

Computers are important in all areas of the human knowledge. Chemistry uses computers to simulate and calculate properties of simple and complex systems. Our group is responsible to implement Computational Chemistry in the undergraduate courses at our university. So we develop methodologies to facilitate the learning by considering interesting themes. In this way, the students' interest in extracting chemical information with the use of methods implemented in computer programs increases. In this paper the theme used as motivation is malaria: a disease responsible for a great number of deaths in the world, mainly in tropical and subtropical areas. We report the use of Computational Chemistry to investigate the structure-activity relationship (SAR) of artemisinin (Fig. 1a) and its 8-deoxyartemisinin derivative (Fig. 1b).

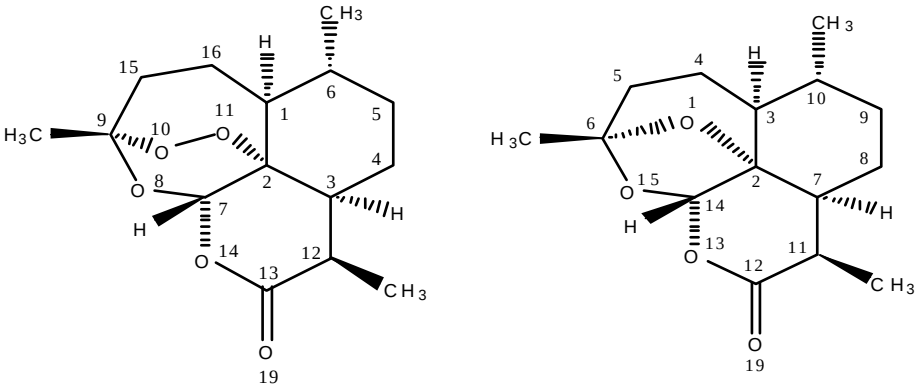

Figure 1. (a) Artemisinin and (b) 8-deoxyartemisinin

Those compounds demonstrate, in experiments, significant differences in biological activities tested *in vitro* against D-6 Indochina clone of *Plasmodium falciparum* [1], the major species of parasite that cause malaria. The artemisinin (IC_{50} = 2.97 ng/mL) is approximately 254 times more potent than 8-deoxyartemisinin (IC_{50} = 761.58 ng/mL), as we can note by their inhibition concentration. According to literature, the severity of the disease results primarily from its ability to modify the surface of infected red blood cells by inserting parasite proteins. The active compounds present a region of negative electrostatic potential of similar shape near the trioxane-ring due to the peroxide linkage. This region is displaced in inactive compounds because they do not present the trioxane-ring. Those artemisinin and its active derivatives' characteristics allow correlating structure-activity in those compounds and with this in mind we can propose new derivatives for future syntheses and biological evaluation [2,3].

For the success of the methodology, the students study previously the topics of the Table 1, which give support to understand the theory and develop the strategy proposed here. This is a way of making connections involving topics from different areas. In general, those topics are studied separate and so students see no relationship among them unfortunately.

Table 1. Topics and Subtopics explained in the Computational Chemistry Study

Topics	Subtopics
Solutions of the Schrödinger Equation for Hydrogenlike Atoms	Hydrogenlike Orbitals; Slater-Type Orbitals (STOs); Gaussian-Type Orbitals (GTOs).
Atomic Basis Sets of Guassian Functions	Minimal Basis Sets; Extended sp Basis Sets; Polarization Basis Sets; Basis Sets Incorporating Diffuse Functions.
Introduction to *Ab Initio* Molecular Orbital Theory	The Hartree-Fock-Roothaan (HFR) Equations; Optimization of Geometry; Potential Surfaces; Molecular Properties; Frontier Orbitals.
Introduction to Semiempirical Molecular Orbital Theory	AM1; PM3; ZINDO.
Molecular Electrostatic Potential (MEP)	Fundamental Relationships; Some General Features of Atomic and Molecular Electrostatic Potentials; Molecular Surfaces; Molecular Electrostatic Potentials and Chemical Reactivity.
Introduction to Multivariate Analysis	Basic Statistics; Visualization of Data; Principal Components Analysis (PCA); Hierarchical Cluster Analysis (HCA).

Results and Discussion

Method and Basis set for the Description of Artemisinin Geometry

Initially the artemisinin is built with aid of GaussView software.[1]) Complete geometry optimization is done with the HFR approach and the 3-21G, STO-3G, 6-31G, 6-31G*, 6-31G**, 6-311G, CEP-31G, and CEP-31G* basis sets and with the AM1, PM3, and ZINDO semiempirical methods as implemented in Gaussian 98 program[2]) (the atomic numbering which we adopt to study the compounds corresponds to the same of

the output of that program). After the theoretical and experimental [4] geometrical parameters for 1,2,4-trioxane ring of the artemisinin are used in PC and HC analyses. The selection of descriptors is done by PCA through the correlation matrix of the variables and those descriptors that show small correlation (< 0.302) are discarded. The best method/basis sets selected by PCA and HCA are used to model the molecular structure of the 8-deoxyartemisinin with the aid of the Gaussview software and to calculate the molecular properties for the two compounds. To correlate quantitatively structure and antimalarial activity, frontier orbitals (HOMO and LUMO) are calculated and MEP maps are computed from the electronic density. The visualization of the MEP maps and frontier orbitals for the studied molecules is done by using the Molekel program.[3]) All molecular calculations are performed with Gaussian 98 code, while PCA and HCA are performed with the Pirouette program.[4]) Autodock 4.0,[5]) an automated docking program, is used for the docking calculations.

For data analysis, four descriptors related to trioxane ring are selected (Table 2). They are responsible for the separation of the samples into two classes: semiempirical methods (as one class) and *ab inito* and experimental methods (as other class). The descriptors are identified by the atoms related to the measures. C9O8 is the interatomic distance between those two atoms, O11O10C9 is the angle connecting those three atoms and O8C7C2O11 and C2O11O10C9 are related to dihedral angles. The correlation between descriptors is always less than 0.82.

Table 2. Descriptors selected for PCA and the correlation matrix

Methods/Basis sets	C9O8	O11O10C9	O8C7C2O11	C2O11O10C9
3-21G	1.435	107.10	-50.85	50.33
STO-3G	1.446	108.01	-55.23	53.89
6-31G	1.435	108.80	-49.41	46.70
6-31G*	1.408	109.45	-50.14	48.68
6-31G**	1.408	109.46	-50.15	48.68
6-311G	1.435	109.20	-49.50	46.88
CEP-31G	1.449	109.40	-49.70	46.60
CEP-31G*	1.418	109.44	-50.26	48.94
AM1	1.427	112.53	-41.77	47.05
PM3	1.428	110.34	-40.51	35.63
ZINDO	1.396	114.31	-46.61	40.13
Exp	1.445	108.10	-51.30	47.80
C9O8		-0.643	-0.302	0.360
O11O10C9			0.655	-0.641
O8C7C2O11				-0.806

The first three PCs explain 95.65% of the total variance in data as follows: PC1= 66.78, PC2 = 21.89, and PC3 = 6.98%. The PC1-PC2 scores are shown in Fig. 2, in which we can see that methods are discriminated into two classes according to PC1. The semiempirical methods are on the left side (AM1, PM3, and ZINDO), while the *ab initio* (3-21G, STO-3G, 6-31G, 6-31G*, 6-31G**, 6-311G, CEP-31G, and CEP-31G*) and experimental methods are on the right side. Moreover we can see that the 3-21G, 6-31G, and CEP-31G basis sets are closer to the experimental method, indicating that any of them could be used in the development of our study.

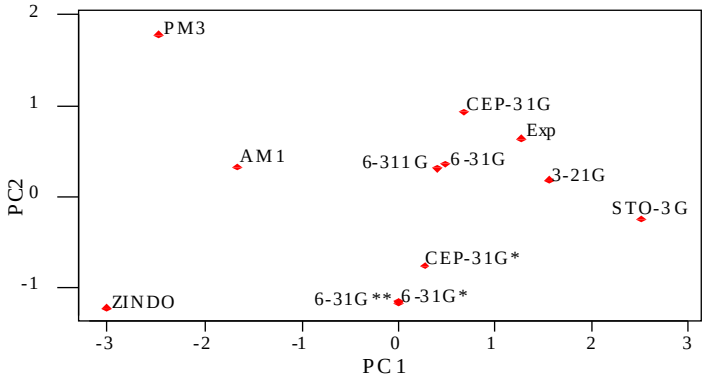

Fig. 2. PC1-PC2 scores plot for the studied methods and basis sets

However, the results of the HCA, displayed in the dendogram in Fig. 3, show the similarity between CEP-31G and the experimental method. So this basis set reveals as the most adequate to be employed in this study. Also we can see that the two classes obtained are the same in PCA analysis (Fig. 2).

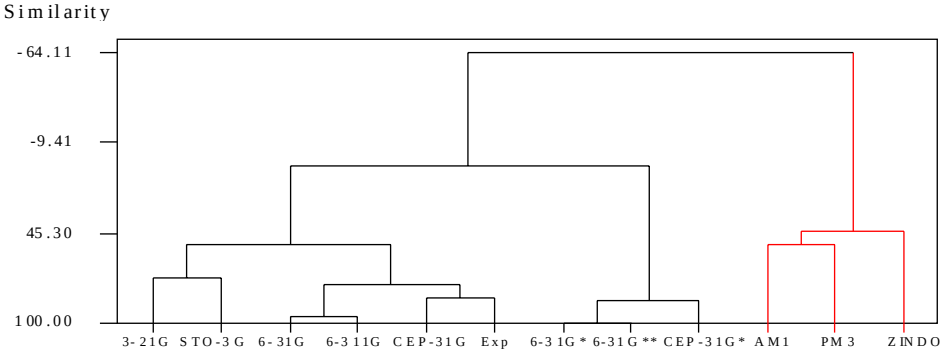

Fig. 3. HCA dendogram for methods/basis sets into two classes (semiempirical methods; ab initio methods)

Structure activity correlation

An important aspect explored in this proposal is to correlate the structure-activity of artemisinin and 8-deoxyartemisinin through the characteristics of the electrostatic potential in the region of the 1,2,4-trioxane-ring in those compounds. MEPs can be useful in predicting the reactive behavior of a wide variety of chemical systems.

Fig. 4 shows the MEP maps of the studied compounds. In this figure, artemisinin (a) has a region of negative potential characteristic shape near the trioxane-ring (red color). However, this region is displaced (yellow color) in the 8-deoxyartemisinin derivative (b). Most negative portions of the surface of MEP map are found near the trioxane ring involved in the complexation of artemisinin with heme.

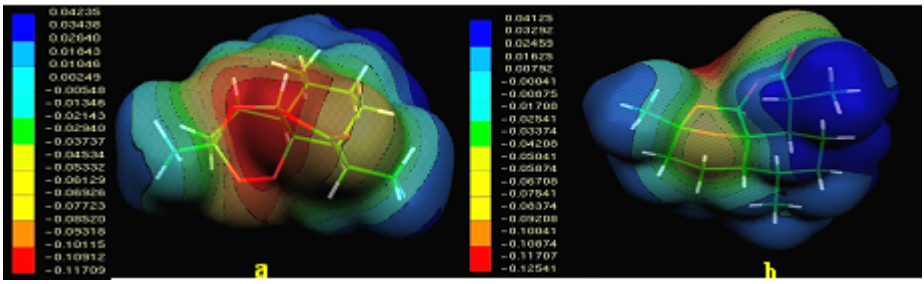

Fig. 4. MEP (in au) maps to (a) artemisinin and (b) 8-deoxyartemisinin

In Fig. 5, when HOMO and LUMO orbitals are compared, there is a clear distinction between the active and the inactive molecule. It is verified that HOMO orbital lobes in artemisinin (active compound) are positioned on the atoms C_2-C_3, C_2-C_7, C_7-O_8, C_7-O_{14}, O_{14}-C_{13}, C_{12}-C_{13}, C_3-C_{12}, C_{13}-O_{19}, C_2-O_{11}, O_{10}-C_{11} with contribution in the [+0.16 $2p_z$ (C_2)] - [+ 0.12 $2p_z$ (C_3)], [+0.12 $2p_x$ (C_2)] - [+0.15 2s (C_7)], [-0.19 $2p_z$ (C_7)] - [-0.16 $2p_z$ (O_8)], [-0.19 $2p_z$ (C_7)] - [-0.47 $2p_z$ (O_{14})], [-0.19 $2p_z$ (C_{13})] - [-0.47 $2p_z$ (O_{14}), [+0.19 $2p_y$ (C_{12})] - [+0.11 2s (C_{13})], [+0.22 $2p_z$ (C_{12})] - [+0.17 $2p_z$ (C_3)], [-0.17 $2p_z$ (C_{13})] - [-0.14 $2p_z$ (O_{19})], [+0.16 $2p_z$ (C_2)] - [+0.13 $2p_z$ (O_{11})], [+0.14 $2p_z$ (O_{10})] - [+0.13 $2p_z$ (O_{11})] bonds. For the 8-deoxyartemisinin (inactive compound), the HOMO orbital lobes are positioned on all the atoms with contribution of all the bonds of the framework. Artemisinin presents LUMO orbital lobes positioned on the atoms C_{13}-O_{14}, C_{13}-O_{19}, C_{13}-C_{12} with contribution in the [-0.45 $2p_z$ (C_{13})] - [-0.23 $2p_z$ (O_{14})], [-0.45 $2p_z$ (C_{13})] - [-0.46 $2p_z$ (O_{19})], [+1.0 $2p_y$ (C_{13})] - [+0.70 $2p_y$ (C_{12}) bonds. 8-deoxyartemisinin also reveals the LUMO orbital lobes positioned on the atoms C_{12}-O_{13}, C_{12}-C_{11}, C_{12}-O_{19} with contribution in the [-0.68 2s (C_{12})] - [+0.40 $2p_y$ (O_{13})], [-0.21 $2p_z$ (C_{12})] - [-0.26 $2p_z$ (C_{11})], [+1.1 $2p_y$ (C_{12})] - [+0.79 $2p_y$ (O_{19})] bonds.

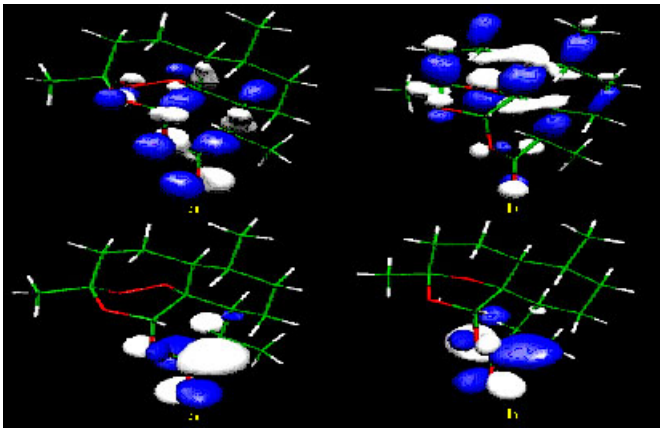

Figure 5. HOMO (upper) and LUMO (lower) orbitals for (a) artemisinin and (b) 8-deoxyartemisinin

Molecular docking

The docking showed (Fig. 6) that artemisinin approaches heme by pointing O11 at the endoperoxide linkage toward the iron center, a mechanism that is controlled by steric hindrance. Although the mechanism of antimalarial activity of artemisinin is still in doubt, there is general agreement that endoperoxide group plays an important role to the antimalarial activity since the 8-deoxyartemisinin compound, which is much less active, does not present the endoperoxide moiety.

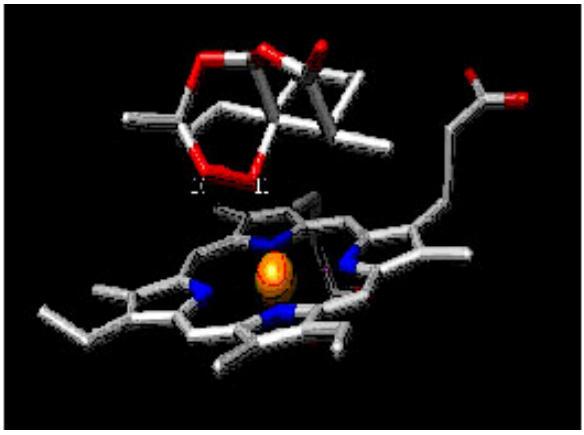

Figure 6. The final docking configuration of artemisinin-heme

Conclusions

By this proposal university teachers are encouraged to develop with their students the application of Computational Chemistry in a daily theme such as malaria. Certainly it is an interesting and useful way of teaching/learning important aspects not only of Chemistry but also related areas since this methodology allows interdisciplinary discussions. Finally students can see the useful side of science, that is, when science is applied to make human life better.

Acknowledgments. We gratefully acknowledge the Brazilian Agencies CNPq and CAPES for financial support. We also thank the Instituto de Química-Araraquara and the Swiss Center for Scientific Computing for the use of GaussView and Molekel programs, respectively and the Centro Nacional de Processamento de Alto Desempenho-UNICAMP and the Laboratorio de Química Teórica e Computacional-UFPA for the computational support.

NOTES

1. Gauss View 1.0, Gaussian Inc., Pittsburg, 1997.
2. Gaussian 98, version A.11, Gaussian Inc., Pittsburg, 2000-2001.
3. Molekel, version 4.1. Swiss Center for Scientific Computing, Mano, 2001.
4. Porouette 3.01. Infometrix Inc., Woodinville, 2001.
5. Auto-Doc 4.0. The Scrips Research Institute http://autodock.scripps.edu/

REFERENCES

1. Brossi, A., B. Venugopalan, L.D. Gerpe, H.J.C. Yeh, J.L. Flippen-Anderson, P. Buchs, X. De-Luo, W. Milhous, W. Peters. Arteether, A New Antimalarial Drug: Synthesis and Antimalarial Properties. *J. Med. Chem.* **31**, 645-650 (1988).

2. Cardoso, F.B., R.B. Costa, A.F. Figueiredo, J.P. Barbarosa, I.N. Nava, Jr., J.C. Pinheiro, O.A.S. Romero. Modeling Artemisinin Derivatives with Potent Activity against *P. falciparum* Malaria with *ab initio* and PLS Methods. *Int. Elect. J. Mol.* Des. **6**, 122-134 (2007).

3. Cardoso, F.B., A.F. Figueiredo, M.S. Lobato, R.M. Miranda, R.C.O. Almeida, J.C. Pinheiro, A Study on Antimalarial Artemisinin Derivatives Using MEP Maps and Multivariate QSAR. *J. Mol. Model*. **14**, 39-48 (2008).

4. Lisgarten, J.N., B.S. Potter, C. Bantuzenko, R.A. Palmer. Structure, Absolute Configuration, and Conformation of the Antimalarial Compound, Artemisinin. *J. Chem. Cryst*. **28**, 539-543 (1998).

ПРЕПОДАВАНЕ НА ИЗЧИСЛИТЕЛНА ХИМИЯ ВЪРХУ ПРИМЕРА НА МАЛАРИЯТА

Резюме. Чрез изучаване на маларията може да се въведе базисен курс по изчислителна химия. Лекарството, което е обект на изследването, е артемисинин и неговото 8-деоксиартемисин производно. Геометричната оптимизация е направена чрез теорията на Hartree-Fock-Roothaan (HFR). Други използвани методи са: Principal Component Analysis (PCA) и Hierarquical Cluster Analysis (HCA) заедно с полуемпирични методи. Карти на молекулния електростатичен потенциал (MEP) и на най-високата заета молекулна орбитала (HOMO) и най-ниската незаета молекулна орбитала (LUMO) са използвани за корелацията структура — активност. Такава учебна стратегия може да се използва при изучаването на изчислителна химия, теоретична химия, хемометрия и медицинска химия.

✉ **Joao Elias** (corresponding author)
Laboratório de Química Teórica e Computacional, Faculdade de Química,
Instituto de Ciências Exatas e Naturais, Universidade Federal do Pará, CP 101101,
CEP 66075-110, Belém, PA, Amazônia, Brasil
E-Mails: joao.elias@yahoo.com.br
antonioflorencio2000@yahoo.com.br
jardelquantun@yahoo.com.br
ciriaco@ufpa.br