

# Quantitative Activity Structure Relationship (QSAR) of a Series of Azetidinones Derived from Dapsone by the Method of Density Functional Theory (DFT)

Jean Stéphane N'DRI<sup>1</sup>, Mamadou Guy-Richard KONÉ<sup>1</sup>, Charles Guillaume KODJO<sup>1,2</sup>, Sopi Thomas AFFI<sup>1</sup>, Ahmont Landry Claude KABLAN<sup>3</sup>, Ouanlo OUATTARA<sup>1</sup>, Doh Soroand<sup>1</sup> Nahossé ZIAO<sup>1</sup>

<sup>1</sup>Laboratoire de Thermodynamique et de Physico-Chimie du Milieu, UFR SFA, Université Nangui Abrogoua, 02 BP 801 Abidjan 02, Côte-d'Ivoire,

<sup>2</sup>Laboratoire de Chimie BioOrganique et de Substances Naturelles, Université Nangui Abrogoua, UFR-SFA, 02 B.P. 801 Abidjan 02 Côte-d'Ivoire,

<sup>3</sup>UFR des Sciences Biologiques, Université Péléforo Gon Coulibaly de Korhogo, BP 1328 Korhogo, Côte d'Ivoire

**Type of Review:** Peer Reviewed.

DOI: <http://dx.doi.org/10.21013/jas.v8.n2.p2>

#### How to cite this paper:

N'DRI, J., KONE, M., KODJO, C., AFFI, S., KABLAN, A., OUATTARA, O., Soroand, D., ZIAO, N. (2017). Quantitative Activity Structure Relationship (QSAR) of a Series of Azetidinones Derived from Dapsone by the Method of Density Functional Theory (DFT). *IRA International Journal of Applied Sciences* (ISSN 2455-4499), 8(2), 55-62. doi:<http://dx.doi.org/10.21013/jas.v8.n2.p2>

© Institute of Research Advances.



This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License subject to proper citation to the publication source of the work.

**Disclaimer:** The scholarly papers as reviewed and published by the Institute of Research Advances (IRA) are the views and opinions of their respective authors and are not the views or opinions of the IRA. The IRA disclaims of any harm or loss caused due to the published content to any party.

*Institute of Research Advances is an institutional publisher member of Publishers Inter Linking Association Inc. (PILA-CrossRef), USA. The institute is an institutional signatory to the Budapest Open Access Initiative, Hungary advocating the open access of scientific and scholarly knowledge. The Institute is a registered content provider under Open Access Initiative Protocol for Metadata Harvesting (OAI-PMH).*

*The journal is indexed & included in CAS Source Index of Chemical Abstracts Service of American Chemical Society (USA), WorldCat Discovery Service (USA), CrossRef Metadata Search (USA), WorldCat (USA), OCLC (USA), Open J-Gate (India), EZB (Germany) Scilit (Switzerland), Airiti (China), Bielefeld Academic Search Engine (BASE) of Bielefeld University, Germany, PKP Index of Simon Fraser University, Canada.*

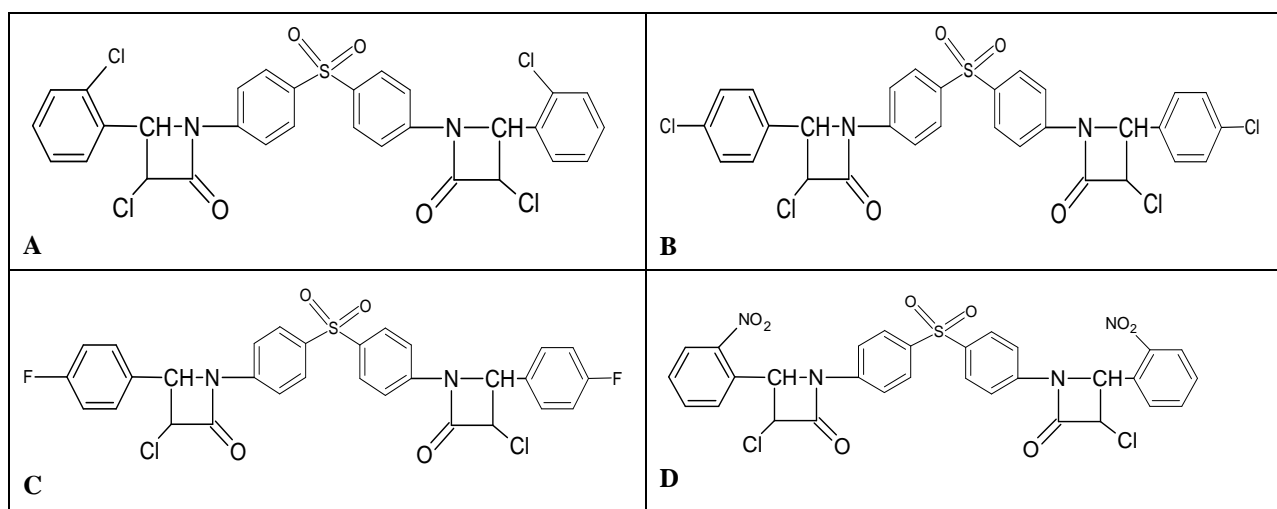
## ABSTRACT

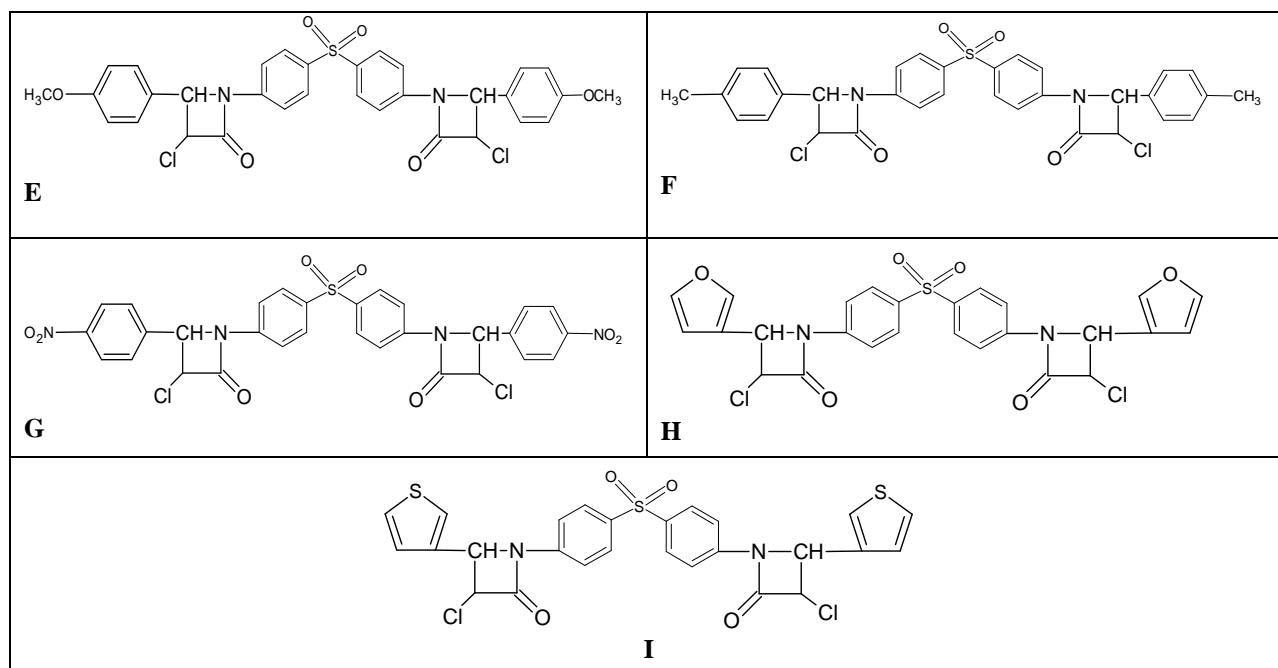
This QSAR study, which involved a series of Azetidinones derived from 4,4'-diaminodiphenylsulfone (dapsone), yielded two models based on molecular descriptors and the antibacterial activities *Escherichia coli* and *Staphylococcus aureus*. The molecular descriptors were obtained by applying the methods of quantum chemistry at the B3LYP/6-31G (d) level. The statistical indicators of the first model which is a function of the *Escherichia coli* activity are: the coefficient of determination  $R^2$  equals 0.992, the standard deviation  $S$  equals 0.342, the Fischer coefficient  $F$  equals 185.088 and the cross-validation coefficient  $Q^2_{cv}$  equals 0.992. Those of the second model showing the activity of *Staphylococcus aureus* are: the regression coefficient  $R^2 = 0.987$ , a standard deviation  $S = 0.193$ , the Fischer coefficient  $F = 114.955$  and the cross-validation coefficient  $Q^2_{cv} = 0.987$ . These models have good statistical performances. The quantum descriptors of dipole moment ( $\mu$ ), global softness ( $\sigma$ ) and electronegativity ( $\chi$ ) are responsible of the antibacterial activity of the Azetidinones derived from dapsone. In addition, the dipole moment is the priority descriptor for the prediction of the antibacterial activity of the studied compounds. The Eriksson *et al.* acceptance criteria used for the test set is verified. The values of the *dtheo/dexp* ratio of the theoretical and experimental activities for the test set tend towards unity.

**Keywords:** Azetidinone, Quantum chemistry, QSAR, Quantum Descriptors, Antibacterial activity

## INTRODUCTION

Leprosy is an infectious, transmissible disease caused by *Mycobacterium leprae* (M. leprae) or Hansen's bacillus (1873), preferentially reaching the skin, mucous membranes, peripheral nervous system, eyes. These disease realise different clinical forms according to the cellular immunity of the infected subject [1]. One of the major aim for the control of leprosy in the world is the treatment of patients. Dapsone or diaminodiphenylsulfone is a sulfone active against many bacteria, but it is mainly used for its action against *Mycobacterium leprae*, the agent responsible for leprosy. Dapsone is also used as an alternative to cotrimoxazole in the treatment and prophylaxis of *Pneumocystis pneumoniae* and, with pyrimethamine, for the prophylaxis of malaria. It is used in the treatment of herpetic dermatitis and other dermatoses; It has also been used in the prophylaxis of toxoplasmosis and for the treatment of cutaneous leishmaniasis and actinomycetomas [2]. In this QSAR study, a series of nine (9) Dapsone-derived azetidinones were used, among them six (6) were used in the training set and the three others remained in the external test set (Fig.1). These compounds have been synthesized and tested by Wadhera *et al.* [3] and Mehta *et al.* [4] for their biological activities.





**Figure 1:** Molecular structures of test sets and validation of Azetidiones used for QSAR models

These molecules were synthesized by condensing dapsone-derived Schiff bases with chloroacetylchloride in the presence of triethylamine as catalyst to obtain the Azetidiones (A-I). These compounds have been developed to control *Staphylococcus aureus* and *Escherichia coli* bacteria. The general objective of this work is to make a descriptive and predictive study of antibacterial activity. The models obtained will allow to identify the best analogues called «leads» which will be able to fight more effectively against the bacterial strains.

## MATERIALS AND METHODS

### 2. Material and Method of Calculation

The six molecules of the test set and the three other molecules in the validation set used in this study have inhibition diameters ranging from 14.00 to 20.40 millimeters. The diameter of the zones of inhibition observed around the disks makes it possible to deduce the minimum inhibitory concentrations (MIC) of the different bacteria for the strain tested and the clinical categorization (Sensitive, Intermediate, Resistant).

#### 2.1. Calculation Level

The link between the biological activity values of the molecules studied and their molecular structures was demonstrated by quantum chemistry calculations carried out using Gaussian software [5]. DFT methods are generally known to generate a variety of molecular properties [6-13] in QSAR studies that increase predictability, reduce computation time, and involve costs in the design of new drugs [14,15]. The B3LYP / 6-31G (d) theory level was used to determine the molecular descriptors. The modeling was done using the multilinear regression method implemented in Excel spreadsheets [16] and XLSTAT [17].

#### 2.2. Quantum Descriptors

For the development of QSAR models, some theoretical descriptors related to the conceptual DFT were determined such as the energy of the Lowest Unoccupied Molecular Orbital ( $E_{LUMO}$ ), the energy of the Highest Occupied Molecular Orbital ( $E_{HOMO}$ ), the electronegativity ( $\chi$ ), the softness ( $\sigma$ ) and the dipole moment ( $\mu$ ). These descriptors are all determined from the optimized molecules. It should be noted that the descriptors related to the boundary molecular orbitals have been calculated in a very simple way within the framework of the Koopmans approximation [18]. LUMO energy characterizes the sensitivity of the molecule to nucleophilic attack, and HOMO energy characterizes the susceptibility of a molecule to an electrophilic attack. The electronegativity ( $\chi$ ) is the parameter that translates the ability of a molecule not to let its electrons escape. The global softness ( $\sigma$ ) expresses the resistance of a system to the change of its number of electrons. The dipole moment ( $\mu$ ) indicates the stability of a molecule in water. Thus, a high dipole moment will result in poor

solubility in organic solvents and high solubility in water. These different parameters, except for the dipole moment ( $\mu$ ), are determined from equations (1):

$$\begin{aligned}\chi &= -1/2 (\epsilon_{LUMO} + \epsilon_{HOMO}) \\ \eta &= (E_{LUMO} + E_{HOMO})/2 \\ \sigma &= 1/\eta\end{aligned}\quad (1)$$

For all the descriptors studied, the analysis of the bivariate data, that is to say the calculation of the partial correlation coefficient between each of the pairs of the set of descriptors is less than 0.70 ( $a_{ij} < 0.70$ ), which means that these different descriptors are independent of each other [19].

### 2.3. Estimation of the Predictive Ability of a QSAR Model

The quality of a model is determined on the basis of various statistical analysis criteria including the coefficient of determination  $R^2$ , the standard deviation  $S$ , the correlation coefficients of cross validation  $Q_{cv}^2$  and Fischer coefficient  $F$ . The statistical indicators  $R^2$ ,  $S$  and  $F$  relate to the adjustment of the calculated and experimental values. They describe the predictive capacity within the model limits and allow to estimate the precision of the values calculated on the training set [20,21]. As for the cross-validation coefficient  $Q_{cv}^2$ , it provides information on the predictive power of the model. This predictive power is called "internal" because it is calculated from the structures used to build this model. The determination's coefficient  $R^2$  gives an evaluation of the dispersion of theoretical values around the experimental ones. The quality of the modeling is better when the points are close to the fitting line [22]. The adjustment of the points to this line can be evaluated by the coefficient of determination.

$$R^2 = 1 - \frac{\sum(y_{i,exp} - \hat{y}_{i,theo})^2}{\sum(y_{i,exp} - \bar{y}_{i,exp})^2} \quad (2)$$

Where

$Y_i, exp$ : Experimental value of antibacterial activity

$Y_i, theo$ : Theoretical value of the antibacterial activity

$\bar{Y}_i, exp$ : The mean value of the experimental values of the antibacterial activity

The closer the value of  $R^2$  to 1, the more the theoretical and experimental values are correlated

Moreover, the variance  $\sigma^2$  is determined by the relation 3:

$$\sigma^2 = s^2 = \frac{\sum(y_{i,exp} - y_{i,theo})^2}{n - k - 1} \quad (3)$$

Where  $k$  is the number of independent variables (descriptors),  $n$  is the number of molecules in the test or learning set and  $n-k-1$  is the degree of freedom. The standard deviation  $S$  is another statistical indicator used. It evaluates the reliability and precision of a model:

$$s = \sqrt{\frac{\sum(y_{i,exp} - y_{i,theo})^2}{n - k - 1}} \quad (4)$$

The Fischer coefficient  $F$  is also used to measure the level of statistical significance of the model, ie the quality of the choice of the descriptors constituting the model.

$$F = \frac{\sum(y_{i,theo} - y_{i,exp})^2}{\sum(y_{i,exp} - y_{i,theo})^2} * \frac{n - k - 1}{k} \quad (5)$$

The coefficient of determination of the cross-validation  $Q_{cv}^2$ , makes it possible to evaluate the accuracy of the prediction on the test set and it is calculated by using the following equation:

$$Q_{cv}^2 = \frac{\sum(y_{i,theo} - \bar{y}_{i,exp})^2 - \sum(y_{i,theo} - y_{i,exp})^2}{\sum(y_{i,theo} - \bar{y}_{i,exp})^2} \quad (6)$$

The performance of a model, according to Eriksson *et al.* is characterized by a value of  $Q_{cv}^2 > 0.5$  for a satisfactory model and for the excellent model  $Q_{cv}^2$  is higher than 0.9[23]. For a given training set, a model will be performing if the acceptance criterion  $R^2 - Q_{cv}^2 < 0.3$  is respected. Moreover, the predictive power of a model can be obtained from the value of the ratio  $d_{theo}/d_{exp}$  for the test set. The model is acceptable when the value of the ratio of theoretical and experimental activity tends towards unity.

## RESULTS AND DISCUSSION

The training set of the six (6) molecules and the three (3) molecules of the Azetidinone test set derived from dapsone are shown in **Table 1**. Thereafter, the values of the partial correlation coefficients  $a_{ij}$  of the descriptors are also presented in **Table 2**.

**Table 1:** Training and Test set of a series of azetidinones derived from Dapsone

Code	$\chi$ (eV)	$\sigma$ (eV)	$\mu$ (Debye)	<i>Staphylococcus Aureus</i> inhibition Diameters(mm)	<i>Escherichia coli</i> inhibition Diameters (mm)
A	4.200	0.212	3.466	16.90	17.30
B	4.241	0.212	1.991	20.30	20.40
C	4.193	0.212	2.011	20.10	20.40
D	4.666	0.282	5.923	18.40	20.10
E	3.979	0.217	2.601	17.90	19.80
F	4.066	0.213	2.822	20.00	18.00
G	4.858	0.254	3.533	20.00	17.00
H	4.171	0.211	7.782	17.00	15.00
I	4.145	0.212	8.062	16.00	14.00

**Table 2:** Values of the partial correlation coefficients of the descriptors.

	$\mu$	$\chi$	$\sigma$
$\mu$	<b>1.000</b>		
$\chi$	0.149	<b>1.000</b>	
$\sigma$	0.172	-	<b>1.000</b>

The correlation matrix is used to see the mutual correlation among the parameters used in the model. This matrix shows that descriptors have low intercorrelation values. The partial correlation coefficients  $a_{ij}$  between the descriptor pairs ( $\mu$ ,  $\chi$ ) and ( $\mu$ ,  $\sigma$ ) are less than 0.70 ( $a_{ij} < 0.70$ ). This demonstrates the independence of the descriptors used to develop the models.

### 3.2. QSAR Model Validation

It should be noted that the negative or positive sign of the descriptor's coefficient of the model reflects the proportionality's effect between the evolution of the biological activity and this parameter of the regression equation. Thus, the negative sign indicates that when the value of the descriptor is high, the biological activity decreases while the positive sign translates the opposite effect. The best QSAR models obtained for the different antibacterial activities of *Escherichia coli* and *Staphylococcus aureus* as well as the statistical indicators are given in **Table 3**. It should be emphasized that these models were established using the descriptors of the training test and test set in **Table 1**.

**Table 3:** Statistical analysis report of the antibacterial activities on the bacterial cells that are *Escherichia coli* and *Staphylococcus aureus*.

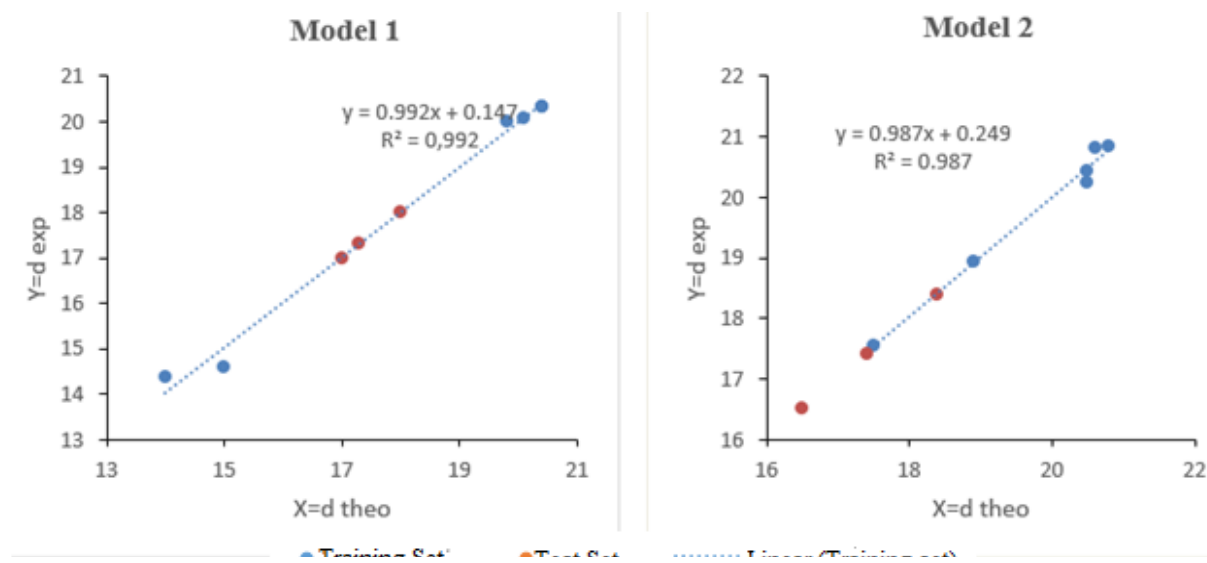
Cell line	Regression equations	R <sup>2</sup>	Q <sup>2</sup> <sub>cv</sub>	S	F	R <sup>2</sup> - Q <sup>2</sup> <sub>cv</sub>
<i>E. Coli</i> (Model 1)	d = 11.372-0.982*μ+51.499*σ	0.992	0.992	0.342	185.088	0.00
<i>S. Aureus</i> (Model 2)	d = 18.415-0.565*μ+0.721*χ	0.987	0.987	0.193	114.955	0.00

The negative signs of the dipole moment coefficient indicate that the antibacterial activities will be improved for a low value of the dipole moment. In contrast, the positive signs of the coefficients of electronegativity and global softness require high values of these descriptors to improve antibacterial activity. As for the significance of the models, it is translated by high value of the coefficient of Fisher F for model 1: 185.088 and for model 2: 114.955. Moreover, the cross-validation correlation coefficient Q<sup>2</sup><sub>cv</sub> for model 1 is 0.992 and for model 2 is 0.987. These different models are all acceptable because all the values of R<sup>2</sup>-Q<sup>2</sup><sub>cv</sub> are less than 0.3. External validation of models 1 and 2 was carried out with Azetidinones (A, F, G) and (A, E, I) respectively. The values of the d<sub>theo</sub>/d<sub>exp</sub> ratio of the external validation set are shown in **Table 4**.

**Table 4:** The values of the diameter ratio of the theoretical and experimental inhibition zone of the external validation set of each model

<b>Model 1</b>			
Compounds	Inhibition Diameter(mm)		
	d <sub>exp</sub>	d <sub>theo</sub>	d <sub>theo</sub> /d <sub>exp</sub>
A	17.30	17.30	1.00
F	18.00	18.00	1.00
G	17.00	17.00	1.00
<b>Model 2</b>			
A	16.90	16.90	1.00
E	17.90	17.90	1.00
I	16.00	16.00	1.00

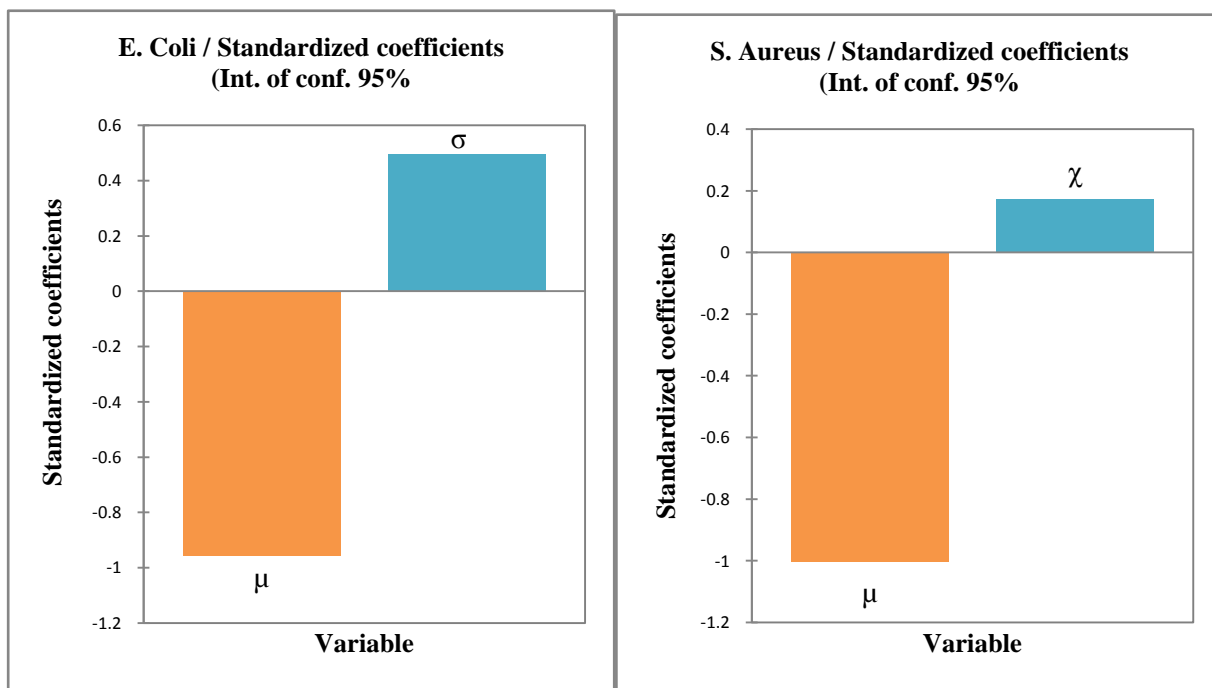
All values of the ratio d<sub>theo</sub>/d<sub>exp</sub> tend to 1. This indicates the good correlation between the theoretical and experimental inhibition diameters of the studied azetidinones. These models are therefore acceptable for the prediction of the antibacterial activities (*Escherichia coli* and *Staphylococcus aureus*) of the series of Azetidinones derived from dapson. The different regression lines between the experimental and theoretical antibacterial activities of the training set (blue dots) and the test set (red dots) for *Escherichia coli* (model 1) and *Staphylococcus aureus* (model 2) are illustrated in **Fig. 2**.



**Figure 2:** Graph between actual and predicted biological activity for training and test set of the different models

**3.3. Analysis of the Contribution of Descriptors in Models**

The study of the relative contribution of the descriptors in the prediction of the antibacterial activity of the compounds was carried out for the bacteria *Staphylococcus aureus* and *Escherichia coli*. The various contributions are presented in **Fig. 3**.



**Figure 3:**Contribution chart of various descriptors in biological activity in different models

The dipole moment presents a large contribution compared to electronegativity or to globalsoftness. Thus, the dipole moment proves to be the priority descriptor in the prediction of the antibacterial activities (*Escherichia coli* and *Staphylococcus aureus*) of the Azetidiones derivatives studied.

## CONCLUSION

The dipole moment ( $\mu$ ), the electro-negativity ( $\chi$ ) and the global softness ( $\sigma$ ) allowed us to predict the behavior of the azetidiones studied in the presence of the bacteria *Escherichia coli* and *Staphylococcus aureus*. This study revealed strong correlations between the calculated and experimental values of the diameter of the inhibition zone. The QSAR models obtained allow us to predict the activity of the best analogues called "lead". These proposed models permit to identify the dipole moment as a priority descriptor for improving the antibacterial activity. Moreover, the negative signs of the dipole moment coefficient indicate that the antibacterial activities will be improved for a low value of the dipole moment. This work constitutes a compass for the design of new molecules more active against the bacteria *Escherichia coli* and *Staphylococcus aureus*. The significance of these models was verified using a test set comprising three molecules. Finally, it is concluded that the work presented here will play an important role in understanding the relationship of physicochemical parameters with structure and biological activity. By studying the QSAR model, one can select the suitable substituent and design new compounds with improved biological activity.

## References

- [1] B.A. Gaüzère, *La lèpre ou Maladie de Hansen*, 2016, 1.
- [2] S.C. Sweetman, Martindale, *The complete Drug Reference*, 34<sup>ème</sup> Ed, Pharmaceutical Press, Londres, 2005, 202-204.
- [3] S. J. Wadher, M. P. Puranik, N. A. Karande and P. G. Yeole, *Synthesis and Biological Evaluation of Schiff base of Dapsone and their derivative as Antimicrobial agents*, 2009, (1), 22-33.
- [4] P. D. Mehta, N. P. S. Sengar, E. V. S. Subrahmanyam1 and D. Satyanarayana, *Synthesis and Biological Activity Studies of Some Thiazolidinones and Azetidiones*, *Indian Journal of Pharmaceutical Sciences*, 2006.
- [5] Gaussian 09, Revision A.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.
- [6] P. K. Chattaraj, A. Cedillo, and R. G. Parr, *J. Phys. Chem.*, 103: 7645, (1991).
- [7] P. W. Ayers, and R. G. Parr, (2000), *J. Am Chem. Soc.*, 122: (2010), 2000.
- [8] F. De Proft, J. M. L. Martin, and P. Geerlings, *Chem. Phys Let.*, 250: 393, (1996).
- [9] P. Geerlings, F. De Proft, and J. M. L. Martin, *In Theoretical and Computational Chemistry*; Seminario, J., Ed.; Elsevier; Amsterdam. Vol-4 (*Recent Developments in Density Functional Theory*), p773, (1996).
- [10] F. De Proft, J. M. L. Martin, and P. Geerlings, *Chem. Phys Let.*, 256: 400, (1996).
- [11] F. De Proft, and P. Geerlings, *J Chem Phys*, 106: 3270, (1997).
- [12] P. Geerlings, F. De Proft, and W. Langenaeker, *Adv. Quantum Chem.*33: 303, (1996).
- [13] R. G. Parr, R. A. Donnelly, M. Levy, and W. E. Palke, *J. Chem. Phys.*, 68: 3801, (1978).
- [14] C. Hansch, P. G. Sammes, and J. B. Taylor, *Computers and the medicinal chemist; in: Comprehensive Medicinal Chemistry*, vol. 4, Eds. Pergamon Press, Oxford, pp. 33-58, (1990)
- [15] R. Franke, *Theoretical Drug Design Methods*, Elsevier, Amsterdam, (1984).
- [16] Microsoft © Excel © 2013 (15.0.4420.1017) MSO (15.0.4420.1017) 64 Bits (2013) Partie de Microsoft Office Professionnel Plus.
- [17] XLSTAT Version 2014.5.03 Copyright Addinsoft 1995-2014 (2014) XLSTAT and Addinsoft are Registered Trademarks of Addinsoft. <https://www.xlstat.com>
- [18] T. Koopmans, *Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den einzelnen Elektronen eines Atoms*, *Physica*, 1: 104-113, (1934).
- [19] A. Vessereau, *Méthodes statistiques en biologie et en agronomie*. Lavoisier (Tec & Doc). Paris, (1988); 538.
- [20] G. W. Snedecor, W. G. Cochran, *Statistical Methods*; Oxford and IBH: New Delhi, India; p. 381, (1967).
- [21] M. V. Diudea, *QSPR/QSAR Studies for Molecular Descriptors*; Nova Science: Huntington, New York, USA, (2000).
- [22] E. X. Esposito, A. J. Hopfinger, J. D. Madura, *Methods in Molecular Biology*, 275, 131- 213, (2004).
- [23] L. Eriksson, J. Jaworska, A. Worth, M.T. D. Cronin, R. M. Mc Dowell, P. Gramatica, *Methods for Reliability and Uncertainty Assessment and for Applicability Evaluations of Classification- and Regression-Based QSARs*, *Environmental Health Perspectives*, 111(10):1361-1375, (2003).