

## COMPUTATIONAL STUDY OF THE INTERACTION BETWEEN INDENE PYRAZOLE AND CYCLIN DEPENDENT KINASE 2

Juan Enrique Torres, Juan Pablo Toro, Javier Vergara, Rosa Baldiris, Ricardo Vivas Reyes\*

Quantum and Theoretical Chemistry Group, University of Cartagena, Faculty of Exact and Natural Sciences, Campus of San Pablo, 130014 Cartagena, Colombia

CIPTEc, Tecnológico Comfenalco Foundation Group. Faculty of Engineering, Industrial Engineering, Cartagena, Colombia

Recebido em 08/10/2016; aceito em 09/01/2017; publicado na web em 24/02/17

Proteins have been traditionally out of reach of electronic structure methods. But with technological advances in the development of powerful computers and the need to extend the methods of computational chemistry to problems of biological interest, such as the rational design of drugs, new technologies *in silico* have been developed that allow to study condensed systems of phase, which consist of thousands of atoms. Here, there are some techniques that combine two or more methods of calculating in a calculation that allows precise chemical exploration of very large systems. The aim of this work is to find the binding affinity of CDK2 inhibitors calculating their electronic densities and then comparing the similarities of these with the biological activity of ligands developing a QSAR in order to establish correlations between quantum similarity, which is a physical-chemical property and biological activity of said set of molecules that change their properties by varying any of their substituents.

Keywords: CDK2; QSAR; Docking; Quantum Similarity; PLS.

### INTRODUCTION

Kinases generally constitute one of the most important families of targets, representing 20-30% of drug discovery programs of pharmaceutical companies, surpassed only by the G-protein coupled receptors.<sup>1</sup> Due to the participation in regulatory processes of cell homeostasis and deregulation of cyclin dependent kinases (CDKs) in several disorders, CDK inhibitors have a broad spectrum of applications ranging from protozoan infections (such as malaria, leishmania, trypanosomiasis),<sup>2,3</sup> viral infections (HCMV, HSV, HIV, HPV), reproductive disorders,<sup>4</sup> cardiovascular diseases (such as atherosclerosis, restenosis, cardiac hypertrophy),<sup>5</sup> glomerulonephritis<sup>6</sup> and cancers<sup>7,8</sup> to nervous system diseases (such as Alzheimer's disease, seizures, amyotrophic disease and drug abuse).<sup>9</sup> In the investigation of antineoplastic agents, CDK2 is an important drug target (cyclin dependent kinase 2). The inhibition of CDK2 essential activities can cause apoptosis in cancer cells, but only cell cycle reversible arrest in normal cells. This is why CDK2 inhibitors have potential as anticancer agents with good therapeutic window.<sup>10,11</sup> Although numerous inhibitors of CDK2 have been reported including carbazides, staurosporines, flavonoids, indigoids, indurubins, oxindoles, paullone, pyrimidines, purines and pyrrolidones,<sup>12-21</sup> none have progressed to clinically useful drugs. Despite long and intensive experimental investigations, most CDK2 inhibitors and hinge region of the active site of the enzyme has hardly been explored. Indeed, there is no clear answer yet to such an important and momentous question for the detailed origin of the interaction between indene pyrazoles ligands and the CDK2 target enzyme; therefore, it remains as an open problem. A better understanding of this aspect is vital for the design of new inhibitors. Consequently, the theoretical work presented in this research proposal is an alternative approach to study the interaction of CDK2-indene pyrazoles in detail. A methodology used to find compounds that inhibit this enzyme is through a study of molecular coupling or molecular Docking. The Docking approach is a computational method that has as main objective to identify the correct positions of the ligands in the binding site in order to predict the affinity between

the ligand and the protein. In other words, it describes a process by which a molecule interacts with another one to form a stable complex in a three dimensional space. Based on previously reported research, a theoretical study of docking was performed using a number of synthetic derivatives of indene pyrazole ligands that interact in the active site of the enzyme CDK2, in order to determine its affinities and binding modes and based on the results obtained, conducting biological tests.

The field of the quantum similarity was introduced by Carbó-Dorca and co-workers;<sup>22-28</sup> they defined the quantum similarity measurements  $Z_{AB}$  between molecules A and B with the electronic density  $\rho_A(r)$  and  $\rho_B(r)$  taking into account the minimizing of the expression for the Euclidean distance as:

$$D_{AB}^2 = \int |\rho_A(r) - \rho_B(r)|^2 dr = \int \rho_A^2(r) dr + \int \rho_B^2(r) dr - 2 \int \rho_A(r) \rho_B(r) dr = Z_{AA} + Z_{BB} - 2Z_{AB} \quad (1)$$

Considering the overlap integral involving the  $Z_{AB}$  between the electronic density of the molecule A and B, where  $Z_{AA}$  and  $Z_{BB}$  represent the Self-similarity.<sup>29</sup> Using the cosine of the angle between the density functions<sup>30</sup> can be expressed mathematically as:

$$I_{AB} = \frac{\int \rho_A(r) \rho_B(r) dr}{\sqrt{\int \rho_A(r) \rho_A(r) dr} \sqrt{\int \rho_B(r) \rho_B(r) dr}} \quad (2)$$

A simple manner using a general operator ( $\Omega$ ), can be expressed as:

$$I_{AB} = \frac{Z_{AB}(\Omega)}{\sqrt{Z_{AA}(\Omega) Z_{BB}(\Omega)}} \quad (3)$$

### COMPUTATIONAL PROCEDURE

#### Dataset, Molecular Docking and optimization

The *in vitro* biological activity data reported as  $IC_{50}$  for inhibition

\*e-mail: rvivasr@unicartagena.edu.co

of CDK2 by the indene pyrazole derivatives was taken from the Singh's article.<sup>31</sup> All the molecules from this study were obtained from sources by the same research group reported at different times. As biological data are normally slanted, the reported  $IC_{50}$  values were turned into the corresponding  $pIC_{50}$  values utilizing the formula:

$$pIC_{50} = -\log IC_{50} \quad (4)$$

The model of the macromolecular target used in this study was the crystallographic structure (PDB Code: 1GZ8), resolved to 1.3 Å with an R value of 0.153 and published by Gibson *et al.*<sup>32</sup> which corresponds to the CDK2 in complex with inhibitor 1 - [(2-amino-6,9-dihydro-1H-purin-6-yl) oxy] -3-methyl-2-butanol (Figure 1). Based on this structure, a system consisting of 5 amino acid residues and the inhibitor will be adopted (hinge region).



**Figure 1.** Structure of co-crystallized ligand (MBP) of the Protein Databank 1GZ8 entry

Amino acid residues included were F<sup>80</sup>, E<sup>81</sup>, F<sup>82</sup>, L<sup>83</sup> and H<sup>84</sup>. With the exception of the catalytic triad K<sup>33</sup>, E<sup>51</sup> and D<sup>14</sup>. Protonation condition of amino acids were determined by the H++ program.<sup>33</sup> For this calculation seven residues were manually rebuilt in the 1GZ8 model (residues 37 to 43). Furthermore, it is assumed that all bond and torsion angles are the same as in the crystallographic structure. Hydrogen atoms missing in the system were added, a model taken from the crystallographic structure. The positions of these hydrogens were optimized by the semi-empirical PM6 method.<sup>34</sup> The structure fully converged was used as an initial model for the remaining calculations. The indene pyrazole ligands were coupled with this converged 1GZ8 structure using the FlexiDock program available in Sybyl-X.<sup>35</sup> FlexiDock parameters were set at 1000 generations for Genetic Algorithms, the CDK2 structure was minimized using Tripos force field with an implicit solvation model. The PRCG method was used with convergence criteria set to Gradient. All indene pyrazole ligands as well as protein residues within 5 Å were allowed to move freely during energy minimization, while residues at a distance between 5 and 15 Å were constrained by a parabolic force constant of 30 kJ Å<sup>-1</sup>. Optimization calculation of indene pyrazoles ligands were performed later in order to use them for the study of quantum similarity applying the *ab initio* DFT/B3LYP method to describe atoms that characterize the electronic population determined by steric and electronic effects in the local atomic shells.<sup>36</sup> Charges and multiplicity of these were also specified. Then it was used the calculation basis 6-31G,<sup>37</sup> to add the Cartesian-Gaussian polarization functions on each of the atoms, followed by an overall minimization of bond distances and angles using an optimizer. For this purpose, the option Opt =QuadMacro was used from the GAUSSIAN09 package.<sup>38</sup>

### Molecular Quantum Similarity

The equation 3 was used in this study to obtain the Molecular Quantum Similarity Measurements (MQSM) and characterizations from the point of view of the atomic shells, described through the

polarizability function in the B3LYP/6-31G. For the implementation of quantum calculation and obtaining similarity matrix, the SIMILARITY<sup>39</sup> program was used, which is part of a computational strategy for QSAR calculations in three dimensions (3D-QSAR). It is currently under development in the group of Theoretical Quantum Chemistry at University of Cartagena.

### Statistical Analysis by PLS

The statistical analysis was performed with 117 ligands derived from indene pyrazoles, of which 7 components were chosen, using a cross-validation method known as "taking two at a time" (k mpteenth with k = 2), which consists of adjusting n times, each time leaving out two of the observations and re-adjusting the model by using the other n - 1 observations. Then the omitted observation is predicted with the model that was excluded. This validation is used to shorten the process in large data sets, as shown in this study. While the meaning of each successive latent variable may be evaluated with partial F tests, the numbers of latent variables to include were chosen from the results of cross-validation. In cross-validation multiple models were developed for which one or more molecules were omitted. Each model was then used to predict the biological activity and potency of molecules omitted. The overall predictability of the model is then expressed in terms of Q<sup>2</sup>, which is formally equivalent to R<sup>2</sup> except that predicted values, rather than setting values are used in the equation:

$$R^2_{\text{pred}} = 1 - \frac{\sum(Y_{\text{pred}} - Y_{\text{test}})^2}{\sum(Y_{\text{test}} - Y_{\text{train}})^2} \quad (5)$$

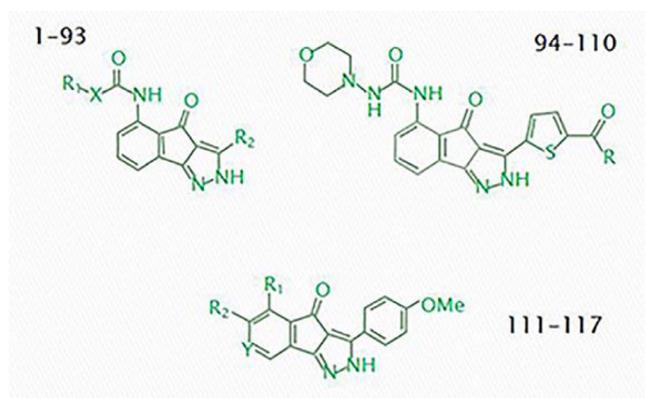
Y<sub>pred</sub> and Y<sub>test</sub> indicate predicted and observed activity values respectively, and Y<sub>train</sub> indicates mean activity value of the training set.

PLS statistical analysis was carried out in the STATGRAPHICS Centurion XVI program, version 16.02.4, which is fairly complete and provides information based on most of the data required in this study as component factor graphics, model and prediction graphics.

### RESULTS AND DISCUSSION

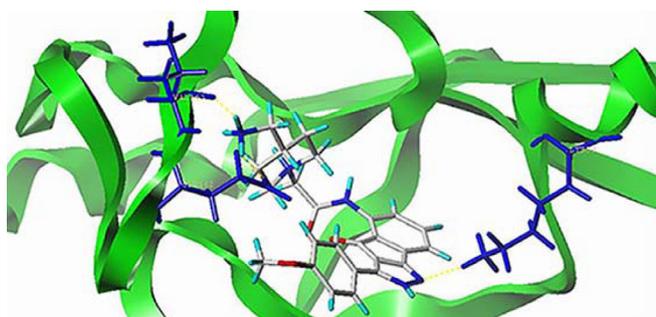
In Table 1S is depicted 117 derived indene pyrazole ligands with their respective  $pIC_{50}$  values that were coupled with CDK2 target and their total coupling scores, a Scoring function can be adapted from force field approaches, estimating the enthalpy of binding via the pair-energy of the complex and also can estimate the entropy of binding, incorporating terms for desolvation and loss of conformational flexibility. The Total Score is the main study descriptor, since it is the resulting sum from the other docking molecular descriptors; this defines more precisely the ligand affinity for the active site of the protein. The ligands in Table 1S are listed according to the different basic structural indene pyrazole cores, as is shown in Figure 2.

A complete analysis for all Total Scores was carried out in Table 1S in order to find ligands with higher and lower activity reported out of the range, which was chosen with a score convention of 4.0 to 7.0. Here, some ligands with low inhibition were found such as 001, 002, 111 and 113 with a total score of 3.96, 3.50, 3.51 and 2.76 respectively, while we can see ligands with potent inhibition with a total score of 7.03, 7.38 and 7.14, such as 017, 062 and 108 being these ligands the ones with more internal energy of interaction with the active site of CDK2. After analyzing the results of molecular coupling, it was observed that the major interactions between CDK2 protein and all ligands are presented in the catalytic site by means of hydrogen bonds between amino acids of polar nature (Glu 81 and His 84) and the oxygens atoms present in this group. Furthermore, joints between the ligand and the active site of the protein through



**Figure 2.** Structural bases of indene pyrazol ligands used in the study

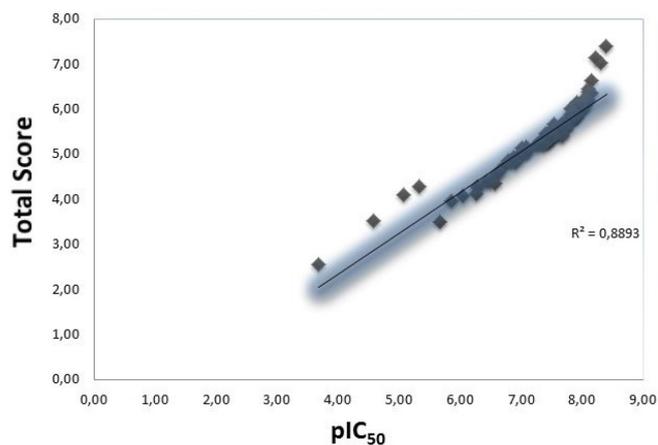
the oxygens atoms of the carboxyl group present in the side chain of the residue Lys 33 and nitrogens atoms present in the indene pyrazol ligands. On the other hand, the nonpolar stabilizing interactions occur between residues Phe 80, Phe 82, Leu 83 and the adjacent aromatic ring on the indene pyrazolic base of ligands. All interactions detailed above are essential for the binding between the enzyme and the ligands and are present in all the molecules studied. Making a more detailed approach of the different basic structural cores of the indene pyrazol derivatives, it can be said that the highest total score of coupling and highest affinity for the target CDK2 were those ones of the first structural basis (1-93), this is due to the majority of the ligands of this group possess strong hydrogen bonding interactions that are formed between the NH groups and thiols with amino acid residues of the active site of CDK2, while with a relatively medium and significant score, ligands of the second structural base could be found (94-110). In the last ligands a few hydrogen interactions and van der Waals interactions were found with less affinity value than the ones in the first base because there are more lipophilic interactions between both piperidine and pyrrolidine rings as well as phenylalanine residues of the hinge region. Finally we find the third structural basis (111-117) as the lowest average score of all coupling. Latter ligands derived from indene pyrazoles have fairly simple radicals and lower electron densities; therefore, few interactions with amino acid residues of the active site as Van der Waals and torsion can be created, which do not contribute so much to the total score as the hydrophilic and lipophilic. In Figure 3 are shown the Van der Waals interactions, lipophilic, hydrogen bonds, among others, that have the amino acid residues of the active site with the most suitable ligand, 062.



**Figure 3.** Ligand interactions of indene pyrazole 062 with the active site of the protein

In order to find the correlation between biological activity and the Total Score of indene pyrazole ligands a scatter plot was performed. According to the results of the regression it can be said that there is a good correlation between the biological activity measured in half

maximal inhibitory concentration ( $IC_{50}$ ) and the total score of the coupling motor (Total Score) with a determination coefficient  $R^2$  of 0.89, and as is shown in Figure 4 these values increase linearly in proportion. The above analysis was carried out based on two specific variables, while the PLS analysis is designed to build a statistical model that links an independent variable X (in this case, the descriptors of molecular quantum similarity) with a dependent variable Y (biological activity of indene pyrazol ligands). All ligands in Table 1S were optimized as was described above to be used in the calculation of quantum similarity rates, which in turn shows a matrix of similarity index. In the matrix, each column represents the similarity between two ligands. Similarity values for 117 compounds were obtained by using the overlap operator, combined with the Carbo index. In Tables 3S and 4S, the values of the multiple linear regression coefficients which were obtained with the PLS technique are given.



**Figure 4.** Correlation between  $pIC_{50}$  VS Total Score of indene pyrazol ligands

In Table 3S, we can see relatively high values for  $R^2$ , reaching a value of 100.0% in the seventh component, so that the correlation between proposed variables with seven components used is notoriously demonstrated. The PRESS value (Prediction Error Sum of Squares) was calculated through the cross-validation test group. These statistics are comparable to the mean square of the residues in Table 2S except that the former is calculated from predictions for observations when these are not used to adjust the model. The lower the values of the PRESS mean square the less error prediction and therefore, the model with the number of components to extract is better. We can see that in the seventh component, PRESS reaches its minimum value (0.63) so this means a model with 7 has the lowest prediction error and is the best model. Table 4S presents percentages of the total variation in variables X and Y explained as the number of components increase. The last column shows the  $R$  Squared of average Prediction through all dependent variables (biological activity). Although the value for 10 components could explain the correlation and model prediction ( $R^2 = 84.78$ ) the average reaches a peak of 7 components, with a value of  $R^2 = 89.37\%$ . A 2D factors graph was also done (Figure 5), in which two factors are chosen, a pair for each axis, and the points representing the rows in the corresponding columns were plotted. In situations where the factors are interpreted, this graph shows the value of each of the samples for these factors.

To analyze these outcomes, it is important to keep in mind that each component chosen in PLS analysis is a linear combination of the similarity matrix columns, and it can generate new descriptors uncorrelated with each other and increase in turn the correlation with biological activities. The difference between the traditional analysis of main components and PLS analysis used in this study, is in the implementation of the correlation between the independent variables

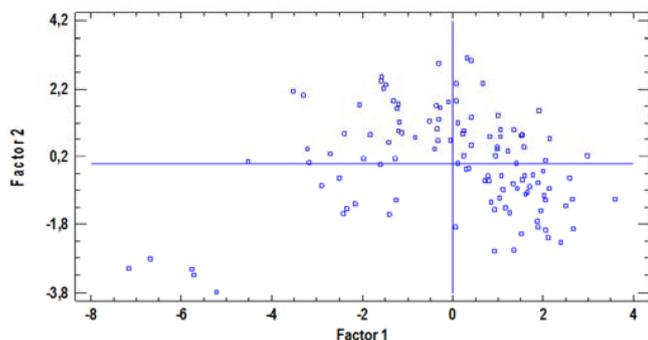


Figure 5. Graph of factor values for PLS in 2D

when generating components. Finally, in Figure 6 are shown the values of the predicted biological activity vs the experimentally measured biological activity, in which the predicted biological activity was calculated with the values of the coefficients obtained by PLS for 7 components.

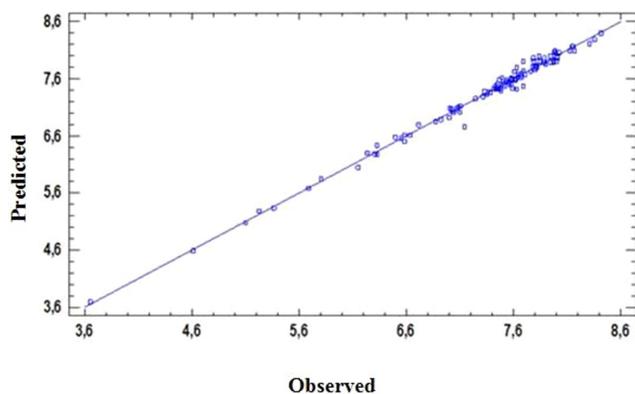


Figure 6. Biological activity chart (current  $pIC_{50}$ ) observed vs predicted

In this study there is a good adjusting model, and is seen that points are aligned along the diagonal line, also the  $R^2$  of prediction for 7 components is 89.37, suggesting that a model with seven components is an excellent choice and quite accurately describes the good correlation between the model and molecular interactions.

## CONCLUSIONS

From this study it can be concluded that there is a great affinity between indene pyrazol ligands used and the target CDK2 enzyme. This could be verified through a significant correlation between the inhibitory activity of indene pyrazol ligands and the total score of the complex docking, with a coefficient of determination of  $R^2 = 0.89$  and a very good prediction correlation between molecular quantum similarity and biological activity of ligands, with a squared coefficient of prediction  $R^2 = 89.37$  for seven components, which means that a successful QSAR model has been developed that clearly explains the interaction of the inhibitors with the amino acids of the hinge region in the active site of CDK2, being this one more robust than recurrent models of simple division.

## REFERENCES

- Cohen, P.; *Nat. Rev. Drug Discovery* **2002**, *1*, 309.
- Osorio, E. J.; Montoya, G. L.; Arango, G. J.; *Vitae* **2006**, *13*, 61.
- Tanowitz, H. B.; Machado, F. S.; Jelicks, L. A.; Shirani, J.; Campos de Carvalho, A. C.; Spray, D. C.; Factor, S. M.; Kirchoff, L. V.; Weiss, L. M.; *Prog. Cardiovasc. Dis.* **2009**, *51*, 524.
- Laman, H.; Coverley, D.; Krude, T.; Laskey, R.; Jones, N.; *Mol. Cell. Biol.* **2001**, *21*, 624.
- Tempfer, C.B.; Simoni, M.; Destenaves, B.; Fauser, B. C. J. M.; *Hum. Reprod. Update* **2009**, *15*, 97.
- Suzuki, J.; Isobe, M.; Morishita, R.; Aoki, M.; Horie, S.; Okubo, Y.; Kaneda, Y.; Sawa, Y.; Matsuda, H.; Ogiwara, T.; Sekiguchi, M.; *Nat. Med. (N. Y., NY, U. S.)* **1997**, *3*, 834.
- Zoja, C.; Casiraghi, F.; Conti, S.; Corna, D.; Rottoli, D.; Cavinato, R. A.; Remuzzi, G.; Benigni, A.; *Arthritis Rheum.* **2007**, *56*, 1629.
- Feldmann G, Mishra A.; Hong, S. M.; Bisht, S.; Strock, C. J.; Ball, D. W.; Goggins, M.; Maitra, A.; Nelkin, B. D.; *Cancer Res.* **2010**, *70*, 4460.
- Korzeniewski, N.; Wheeler, S.; Chatterjee, P.; Duensing, A.; Duensing, S.; *Mol. Cancer* **2010**, *9*, 153.
- Nguyen, M. D.; Boudreau, M.; Kriz, J.; Couillard-Després, S.; Kaplan, D. R.; Julien, J. P.; *J. Neurosci.* **2003**, *23*, 2131.
- Shapiro, G. I. J.; *Clin. Oncol.* **2006**, *24*, 1770.
- Thomas, M. P.; McInnes, C.; *IDrugs* **2006**, *9*, 273.
- Lan, P.; Chen, W.N.; Xiao, G. K.; Sun, P. H.; Chen, W. M.; *Bioorg. Med. Chem. Lett.* **2010**, *20*, 6764.
- McGahren-Murray, M.; Terry, N. H. A.; Keyomarsi, K.; *Cancer Res.* **2006**, *66*, 9744.
- Kim, H.; Lee, E.; Kim, J.; Jung, B.; Chong, Y.; Ahn, J. H.; Lim, Y.; *Bioorg. Med. Chem. Lett.* **2008**, *18*, 661.
- Wu, Z. L.; Aryal, P.; Lozach, O.; Meijer, L.; Guengerich, F. P.; *Chem. Biodiversity* **2005**, *2*, 51.
- Moon, M. J.; Lee, S. K.; Lee, J. W.; Song, W. K.; Kim, S. W.; Kim, J. I.; Cho, C.; Choi, S. J.; Kim, I. C.; *Bioorg. Med. Chem.* **2006**, *14*, 237.
- Bramson, H. N.; Corona, J.; Davis, S. T.; Dickerson, S. H.; Edelstein, M.; Frye, S. V.; Gampe, R. T.; Harris, P. A.; Hassell, A.; Holmes, W. D.; Hunter, R. N.; Lackey, K. E.; Lovejoy, B.; Luzzio, M. J.; Montana, V.; Rocque, W. J.; Rusnak, D.; Shewchuk, L.; Veal, J. M.; Walker, D. H.; Kuyper, L. F.; *J. Med. Chem.* **2001**, *44*, 4339.
- Zaharevitz, D.; Gussio, R.; Leost, M.; Senderowicz, A.; Lahusen, T.; Kunick, C.; Meijer, L.; Sausville, E. A.; *Cancer Res.* **1999**, *59*, 2566.
- Elgazwy, S. H.; Ismail, N. S. M.; Elzhabhi, H. S. A.; *Bioorg. Med. Chem.* **2010**, *18*, 7321.
- Nesi, M.; Borghi, D.; Brasca, M. G.; Fiorentini, F.; Pevarello, P.; *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3205.
- Carbó-Dorca, R.; Mercado, L. D.; *J. Com. Chem.* **2010**, *31*, 2195.
- Carbó-Dorca, R.; Besalú, E.; Mercado, L. D.; *J. Com. Chem.* **2011**, *32*, 582.
- Gironés, X.; Carbó-Dorca, R.; *QSAR. Comb. Sci.* **2006**, *25*, 579.
- Carbó-Dorca, R.; Gironés, X.; *Int. J. Quantum Chem.* **2005**, *101*, 8.
- Amat, L.; Carbó-Dorca, R.; *Int. J. Quantum Chem.* **2002**, *87*, 59.
- Carbó-Dorca, R.; Besalú, E.; *J. Comp. Chem.* **2010**, *31*, 2452.
- Carbó, R.; Arnau, M.; Leyda, L.; *J. Quant. Chem.* **1980**, *17*, 1185.
- Bultinck, P.; Carbó-Dorca, R.; *J. Chem. Sci.* **2005**, *117*, 425.
- Parr, W.; Yang, W.; *Density-functional Theory of Atoms and Molecules*, Oxford University Press: New York, **1989**; Ayers, P. W.; Anderson, J. S.; Bartolotti, L. J.; *Int. J. Quantum Chem.* **2005**, *101*, 520; Gazquez, J. L.; *J. Mex. Chem. Soc.* **2008**, *52*, 3.
- Singh, S. K.; Dessalewb, N.; Bharatamb P. V.; *Eur. J. Med. Chem.* **2006**, *41*, 1310.
- Gibson, A. E.; Arris, C. E.; Bentley, J.; Boyle, F. T.; Curtin, N. J.; Davies, T. G.; Endicott, J. A.; Golding, B. T.; Grant, S.; Griffin, R. J.; Jewsbury, P.; Johnson, L. N.; Mesguiche, V.; Newell, D. R.; Noble, M. E.; Tucker, J. A.; Whitfield, H. J.; *J. Med. Chem.* **2002**, *45*, 3381.
- Gordon, J.; Myers, J.; Folta, T.; Shoja, V.; Aguilar, B.; Back, G.; Ruscio, J. Z.; *H++ 1.0: Web-based computational prediction of protonation states*; Virginia Tech, USA, 2005.

34. Stewart, J. P.; *J. Mol. Model.* **2007**, *13*, 1173.
35. Tripos International; SYBYL X; *Molecular Modeling from Sequence Through Lead Optimization*; South Hanley Rd., St. Louis, Missouri, USA, 1997.
36. Becke, A. D. J.; *Chem. Phys.* **1993**, *98*, 5648; Lee, C.; Yang, W.; Parr, R. G.; *Phys. Rev. B.* **1988**, *37*, 785; Davidson, E.; Feller, D.; *Chem. Rev.* **1986**, *86*, 681.
37. Hehre, W. J.; Radom, L.; Schleyer, P. V.; Pople, J. A.; *Ab Initio Molecular Orbital Theory*, Wiley: New York, 1986.
38. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian 09; Gaussian Inc, Wallingford CT, 2004.
39. <http://iqc.udg.es/cat/similarity/ASA/>, accessed on January 2017.