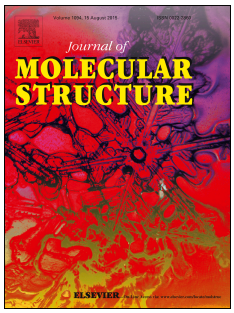
Accepted Manuscript



The electronic density obtained from a QTAIM analysis used as molecular descriptor. A study performed in a new series of DHFR inhibitors

Rodrigo D. Tosso, Marcela Vettorazzi, Sebastian A. Andujar, Lucas J. Gutierrez, Juan C. Garro, Emilio Angelina, Ricaurte Rodríguez, Fernando D. Suvire, Manuel Nogueras, Justo Cobo, Ricardo D. Enriz

PII: S0022-2860(16)31379-5

DOI: 10.1016/j.molstruc.2016.12.060

Reference: MOLSTR 23258

To appear in: *Journal of Molecular Structure*

Received Date: 13 September 2016

Revised Date: 20 December 2016

Accepted Date: 22 December 2016

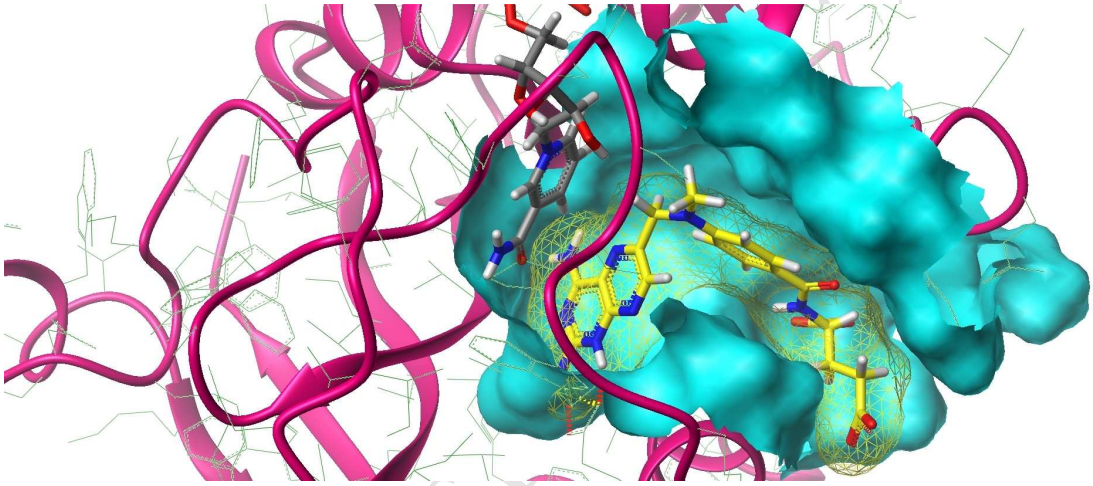
Please cite this article as: R.D. Tosso, M. Vettorazzi, S.A. Andujar, L.J. Gutierrez, J.C. Garro, E. Angelina, R. Rodríguez, F.D. Suvire, M. Nogueras, J. Cobo, R.D. Enriz, The electronic density obtained from a QTAIM analysis used as molecular descriptor. A study performed in a new series of DHFR inhibitors, *Journal of Molecular Structure* (2017), doi: 10.1016/j.molstruc.2016.12.060.

Graphical Abstract

The electronic density obtained from a QTAIM analysis used as molecular descriptor. A study performed in a new series of DHFR inhibitors

Rodrigo D. Tosso, Marcela Vettorazzi, Sebastian A. Andujar, Lucas J. Gutierrez, Juan C. Garro, Emilio Angelina, Ricaurte Rodríguez, Fernando D. Suvire, Manuel Nogueras, Justo Cobo and Ricardo D. Enriz

Departamento de Química, Facultad de Química, Bioquímica y Farmacia, Universidad Nacional de San Luis, Chacabuco 915, 5700 San Luis, Argentina



The electronic density obtained from a QTAIM analysis used as molecular descriptor. A study performed in a new series of DHFR inhibitors

Rodrigo D. Tosso[a,b], Marcela Vettorazzi[a,b], Sebastian A. Andujar[a,b], Lucas J. Gutierrez[a,b], Juan C. Garro[a,b], Emilio Angelina[c], Ricaurte Rodríguez[d,e], Fernando D. Suvire[a,b], Manuel Nogueras[e], Justo Cobo[e] and Ricardo D. Enriz[a,b*]

[a]*Departamento de Química, Facultad de Química, Bioquímica y Farmacia, Universidad Nacional de San Luis, Chacabuco 915, 5700 San Luis, Argentina*

[b]*IMIBIO-CONICET, UNSL, Chacabuco 915, 5700 San Luis, Argentina*

[c]*Laboratorio de Estructura Molecular y Propiedades, Área de Química Física, Departamento de Química, Facultad de Ciencias Exactas y Naturales y Agrimensura, Universidad Nacional del Nordeste, Avda. Libertad 5460, (3400) Corrientes, Argentina.*

[d]*Departamento de Química, Universidad Nacional de Colombia, Ciudad Universitaria, Carrera 30, No. 45-03. Bogotá, Colombia*

[e]*Departamento de Química Inorgánica y Orgánica, Universidad de Jaén, 23071 Jaén, Spain*

---

* To whom correspondence should be addressed.

Phone (54) 266 4423789; e-mail: denriz@unsl.edu.ar

ABSTRACT

The results reported here indicate that the electron density obtained from a QTAIM analysis is an excellent descriptor of molecular interactions that stabilize and destabilize the formation of the ligand-receptor (L-R) complex. The study was conducted on a series of 25 compounds that have inhibitory effects on DHFR. Besides the synthesis and bioassays performed for some of these compounds, various types of molecular calculations were performed. Thus, we performed MD simulations, computations at different levels of theory *(ab initio* and DFT) using reduced models and a QTAIM study on the different complexes.

The resulting model has allowed us to differentiate not only highly active compounds with respect to compounds weakly active, but also among compounds that have similar affinities in this series. The model also showed a high degree of predictability which allows predicting the affinity of non-synthesized compounds. Very important additional information can be obtained through this type of study, it is possible to visualize which amino acids are involved in the interactions determining the different affinities of the ligands.

Keywords: QTAIM; Molecular Descriptor; Molecular Dynamics; Quantum Mechanics

INTRODUCTION

In general, majority of docking algorithms are able to predict the bind correctly, with accuracy of ~ 2 Å root-mean-square deviation (RMSD) to that of observed in the crystal structure (of course depending of the structural characteristics of both the ligand and the binding pocket). The challenge is in scoring these L-R bindings; an ideal scoring function should be able to reproduce binding energy and to rank the ligands according to their binding affinities. However, the majority of scoring functions, bundled with docking packages, often perform a very poor reproduction of the binding affinity; hence, the use of them is limited to screening of databases of a large number of ligands.

In order to predict binding affinity of small molecule inhibitors, a variety of post-docking methods have been established. These methods range from simple consensus scoring to free energy perturbation (FEP) [1-4] among others. Undoubtedly, the post-docking methods can improve significantly the prediction of the energies of L-R binding, however they are still far from being able to predict with a high degree of accuracy the differences in L-R affinities for those ligands having similar binding energies. The situation is even more complex when we are in front of compounds with structural differences. In such cases, most of the times one must accept only if we can differentiate between very active compounds from compounds with low affinity for the receptor (very poor activity). It is clear that any progress or improvement that we can find to enhance these post-docking methods is of paramount importance for the structure based drug design and they are very welcome.

In a recent paper we attempted to find a correlation that would allow us to differentiate between DHFR inhibitors with similar affinities, but we had no success [5]. In fact, we were only able to differentiate between highly active compounds of those who had very poor activity, but we were not able to differentiate between compounds with similar affinities. In that paper we also showed that if one has a good geometry, the QTAIM study provides an important insight into the molecular interactions between ligand and receptor. In this new work, we used the electron density obtained from QTAIM analysis as a descriptor of the molecular interactions of the L-R complex, which has allowed us to discriminate very well between compounds with similar binding affinities.

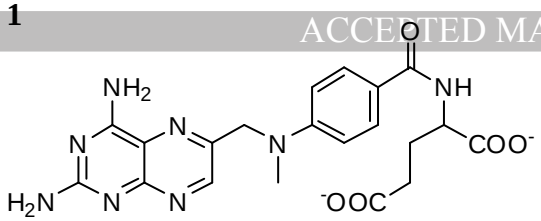
Dihydrofolate reductase (DHFR) is an excellent molecular target for this study because it has been and is currently studied by using different molecular modeling techniques [6-9]. Kerrigan *et al*. have reported an interesting review about recent progress in molecular dynamics simulations of DHFR [10] and they conclude that “molecular mechanics calculations can work well to model the initial binding step of an inhibitor or substrate with

DHFR. However, DHFR continues to be a challenge for free energy estimation methods and caution is recommended when interpreting these results".

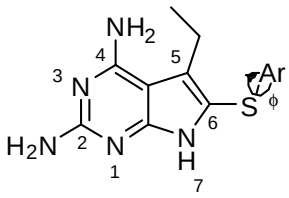
It should be noted that there are very few simulations specifically focused in the molecular interactions involved in the formation of the L-R complexes. Thus, interesting details about the intricacies of molecular interactions of DHFR interacting with its inhibitors remain unknown. Recently, we reported some molecular modeling studies using reduced models for the binding pockets [5, 11-15]. This approach allows performing more accurate quantum mechanical calculations, as well as to obtain a detailed electronic analysis by using the method of Quantum Theory of Atoms in Molecules (QTAIM).

The main objective of this work is to find a way that allows us to differentiate between ligands possessing similar affinities for the DHFR. To achieve this goal, different calculations techniques have been used either alone or in combination in order to find a molecular descriptor that allows getting such differentiation. Thus, the present study was carried out at different stages. In the first step, seven new compounds were synthesized and then they were evaluated for their inhibitory activities against human DHFR. These results were added to thirteen compounds reported by Gangjee *et al.* (**1**-**13**) [16], two compounds reported in our earlier work (**14a** and **15a**) [5] and two new compounds (**14e** and **15f**) recently reported in reference [17] in order to have a more complete and representative number of compounds (25 molecules in the complete series (Figure 1 and Table 1)). In the next step, we performed MD simulations, QM calculations (using different levels of theory) and QTAIM analysis with the aim to obtain a correlation which allows the discrimination between compounds possessing similar affinities by the enzyme. The conclusions are presented at the end.

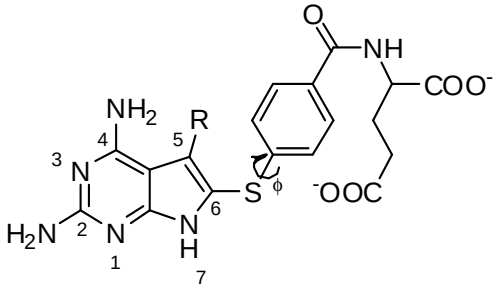
Figure 1



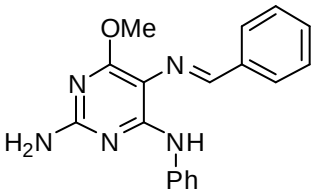
MTX (0.022 μM)



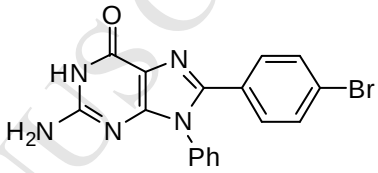
3 Ar = C_6H_5 (18 μM) *
4 Ar = 2,6-diMeC_6H_3 (32 μM) *
5 Ar = 2-naphthyl (6 μM) *
6 Ar = 3,4-diClC_6H_3 (2.8 μM) *
7 Ar = 1-naphthyl (2.9 μM) *
8 Ar = 4-$NO_2C_6H_4$ (5.8 μM) *
9 Ar = 4-BrC_6H_4 (2.7 μM) *
10 Ar = 2,4-diClC_6H_3 (1.4 μM) *
11 Ar = 4-$OMeC_6H_4$ (2.9 μM) *
12 Ar = 2,5-di$OMeC_6H_3$ (5.8 μM) *
13 Ar = 8-quinoline (1.8 μM) *



1 R = Et (0.066 μM) *
2 R = Me (0.21 μM) *



14a (68.01 μM) **

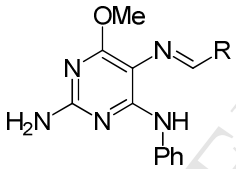
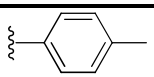
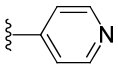
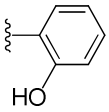
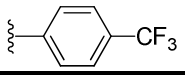
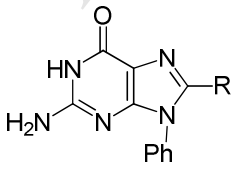
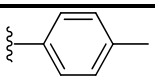
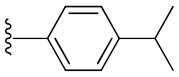
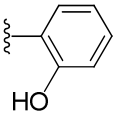
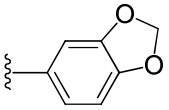
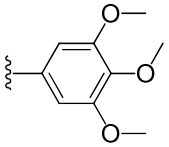


15a (54.45 μM) **

*Compounds reported in reference 16

**Compounds reported in reference 5

Table 1. Structural features and IC_{50} values of compounds type **14** and **15**.

Series	Compound	R	IC_{50} (μM)
 **14**	**14b**		65.98 ± 6.5
	14c		63.05 ± 6.3
	14d		52.89 ± 5.2
	14e*		64.54 ± 6.4
 **15**	**15b**		27.87 ± 2.7
	15c		43.84 ± 4.3
	15d		77.09 ± 7.7
	15e		20.81 ± 2.0
	15f*		49.39 ± 4.9

*Compounds reported in reference 17

METHODS OF CALCULATIONS

The results of this work have been compared with those recently reported in the reference 5; therefore all calculations and molecular simulations have been performed using the same techniques previously used.

MD simulations

The starting structure of human DHFR was obtained from Protein Data Bank of Brookhaven National Laboratory (PDB entry code 2W3M) and the topologies of the ligands were built using the MKTOP program [18]. MD simulations for each L-R complex have been carried out using the GROMACS 4.5.5 simulation package [19, 20]. For these simulations, the OPLS-AA force field [21-26] and the rigid SPC water model [27, 28] in a cubic box with periodic boundary conditions were employed. The box size was 7.437 x 7.437 x 7.437 nm and the total number of water molecules was approximately 14,500 for each simulation. Besides, three Na^+ ions were added to the systems by replacing water molecules in random positions, thus making the whole system neutral.

A steepest-descent algorithm for 1,000 steps was used, in order to minimize the energy of each system. Next, the complexes were equilibrated during 100 ps in NVT and NPT ensembles to stabilize the temperature and the pressure of each system, respectively. Then a 5 ns MD simulation was performed for each complex. The time step was 0.002 ps and the temperature was maintained constant at 310 K with the V-rescale algorithm [29]. Long range interactions were treated by the Particle Mesh Ewald (PME) method [30, 31] with a 1 nm cutoff and a Fourier spacing of 0.12 nm. The compressibility was 4.8 x 10^{-5} bar^{-1}.

The Linear Interaction Energy (LIE) method was used to calculate L-R binding free energy [32, 33]. The following equation was used for such calculations:

$$\Delta G_{LIE}^{bind} = \alpha\left(\langle V_{l-s}^{vdW}\rangle_{bound} - \langle V_{l-s}^{vdW}\rangle_{free}\right) + \beta\left(\langle V_{l-s}^{el}\rangle_{bound} - \langle V_{l-s}^{el}\rangle_{free}\right) \qquad (1)$$

where α and β parameters are dispersion and electrostatic adjustable energy scale factors [34]. Their values (0.181 and 0.5, respectively) were previously reported by Marelius *et al.*, who adjusted them for DHFR from experimental parameters [35]. The $\langle V_{l-s}^{vdW}\rangle$ and $\langle V_{l\text{-}s}^{el}\rangle$ terms denote MD energy averages of the nonbonded van der Waals and electrostatic interactions between ligand and its surrounding environment (subscript l-s), respectively.

MM-GBSA Free Energy Decomposition

The MM-GBSA free energy decomposition was carried out in order to evaluate which amino acids were interacting with the ligand. This calculation allows decomposing the interaction energies to each residue considering molecular mechanics and solvation energies [36-40]. Four energy terms are considered for each inhibitor-residue pair: van der Waals contribution (ΔE_{vdW}), electrostatic contribution (ΔE_{ele}), polar desolvation term (ΔG_{GB}), and nonpolar desolvation term (ΔG_{SA}). These terms can be summarized as the following equation:

$$\Delta G_{inhibitor-residue} = \Delta E_{vdW} + \Delta E_{ele} + \Delta G_{GB} + \Delta G_{SA} \quad (2)$$

For these calculations, the mm_pbsa program in AMBER12 [41] has been employed and snapshots were taken at 10 ps time intervals from the corresponding last 1,000 ps MD trajectories, while the explicit water molecules were removed from the snapshots.

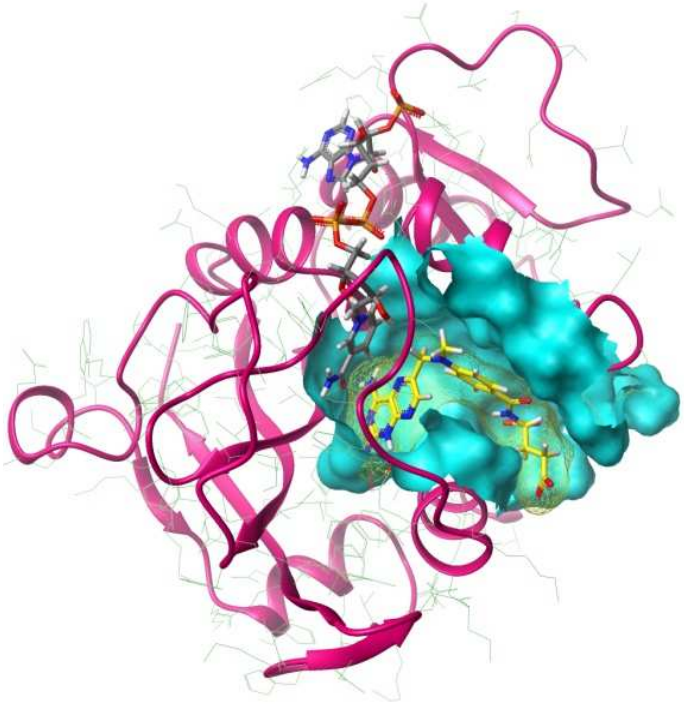
Constructing the reduced models for the binding site

The use of model systems to calculate and simulate molecular interactions (MI) is necessary since the inhibitors interacting at the active site of DHFR constitute a molecular system too large for accurate quantum mechanics molecular orbital calculations and the number of ligands to be screened is large as well.

Since we are interested in comparing the results obtained in this work with those reported in our previous work, we use the same reduced system that was used in that study [5]. This reduced model is composed by 23 amino acids: Ile7, Val8, Ala9, Leu22, Trp24, Glu30, Phe31, Tyr33, Phe34, Gln35, Met52, Thr56, Ser59, Ile60, Pro61, Asn64, Leu67, Lys68, Arg70, Val115, Tyr121, Val135 and Thr136. In addition all water molecules within a 5 Å radius from the ligand were also included in the reduced model.

To be sure that the different amino acids that stabilize and destabilize the complex formation have been included in the reduced system, we performed an MM-GBSA free energy decomposition analysis for the different complexes under study. The information obtained from these calculations is very important for quantitative analysis and is highly useful for the understanding of the binding mechanism. Figure 1S in Supplementary Material displays the spectra obtained for the most representative compounds reported here. A surface of the binding site considering the residues of this model is shown in Figure 2.

Figure 2



QM calculations

First, the twenty-three residues included in our reduced model and the ligand were calculated at the PM6 [42] level of theory, using the MOPAC2009 program [43]. The ligand and the torsional angles, bond angles and bond lengths of the side chains of the amino acids were optimized, while the atoms of the backbone were frozen during calculations. For these optimizations, four different starting geometries from MD simulations were taken into account: the global minimum and three local minima, obtained from the potential energy calculations. This step is important because in this way, a more representative sample of the spatial arrangements of the complexes was evaluated. Then, we have performed single point calculations of optimized geometries at RHF/6-31G(d) and PBE1PBE/6-31G(d) levels of theory, using Gaussian 09 program [44].

The binding energy (BE) of each complex L-R was calculated employing the following equation:

$$BE_{QM} = E_{L-DHFR} - (E_{DHFR} + E_L) \quad (3)$$

where BE_{QM} is the binding energy, $E_{\text{L-DHFR}}$ is the complex energy, E_{DHFR} is the energy of the reduced model (binding pocket), and E_L is the energy of the ligand.

Topological Analysis of the Electron Charge Density Distribution

The reduced models constructed for the studied complexes were then submitted to topological analysis of the charge density (ρ(r)). QTAIM calculations were performed over Kohn-Sham DFT wave functions employing the hybrid PBE functional and 6-31G(d) as basis set.

The terminology of QTAIM was extensively reviewed in the standard literature [45]. Next, we briefly summarize some basic concepts that are needed for the discussion of results. From the QTAIM point of view, two interacting atoms share three topological elements related to each other, a point, a line and a surface. The first element is the bond critical point (BCP), namely the critical point in ρ(r) topology that is found between any two interacting nuclei. From each BCP, two unique trajectories of gradient vectors of electronic density, $\nabla\rho$(r), originate at that point and terminate at each of the neighboring nuclei. These trajectories define a line along which ρ(r) is a maximum with respect to any neighboring line. This line, that constitutes the second element, is the bond path, BP. Additionally, the set of trajectories that terminate at a BCP define the interatomic surface (IAS) that separates the atomic basins of the neighboring atoms [45].

The determination of all the intermolecular BCPs and the corresponding BPs were performed with Multiwfn [46] and AIMPAC [47] software. The molecular graphs were depicted with Pymol [48].

In this paper, QTAIM calculations were performed in order to determine the ρ(r) values at the BCPs established between each atom (belonging to the backbone or the side chain) of the amino acids of the receptor and each atom of the ligand. In order to obtain the total ρ(r) value of the interaction between a particular residue and the inhibitor, we performed the sum of the ρ(r)values of each BCPs between that amino acid and the inhibitor according to the following equation:

$$\rho_{Res} = \sum \rho_{Interaction\ Res-L} \tag{4}$$

where ρ_{Res} is the ρ(r) value for the total interaction with a particular residue and $\rho_{Interaction\ Res-L}$ represents the ρ(r) value of each interaction (BCP) between the atoms of the ligand and the corresponding residue.

Multivariable Linear Regression

A data set of 25 compounds with known DHFR inhibition activity was utilized to perform a Multivariable Linear Regression (MLR) analysis with the aim to establish a mathematical relationship between the biological activity and the ρ(r) value calculated from QTAIM study.

In this analysis, we considered each ρ(r) value of each interaction as a molecular descriptor in order to use them as independent variables while the biological activity (expressed as –log IC_{50}) was utilized as the dependent variable. A total of 23 interactions, which belong to the reduced model, were used as molecular descriptors. The values of the biological activity as well as the numbering of the compounds included in the data set are given in Table 1.

The mathematical models were developed using a subset of 22 compounds (calibration set) from the full data set. Then, the optimal model was validated through a test set containing 3 compounds which do not form part of the calibration set. The elements of each set (calibration and test set) were selected in such a way that they share similar structural characteristics and the experimental data of the test set represent of the whole span. Thus, the compounds **7**, **14c** and **15c** were selected as test set.

The optimal model was developed by performing an exhaustive search selecting the best linear regressions (minimum standard deviation (S)) from each combination from 1 to 20 molecular descriptors. The quality of the models is suggested for the calculated coefficients of determination (R^2) and the standard deviations (S). All the MLR calculations were carried out using the MATLAB 7.0 software [49].

EXPERIMENTAL SECTION

General Experimental Section

Solvents and Reagents

Reagents and solvents used were purchased from commercial suppliers and use without further purification procedures.

Chromatography

Thin layer chromatography (TLC) was used to monitorize the reaction progress and product purity, it was performed on Merck Kieselgel 60 F254 aluminium precoated plates, and spots visualized with ultra-violet irradiation.

Melting Points

Melting points were recorded on a *Digital Melting Point Apparatus,* model *IA9300 series, Barnstead Electrothermal* and are uncorrected.

NMR Spectroscopy

1D (^{1}H, ^{13}C, DEPT) and 2D (COSY, HSQC and HMBC) NMR spectra were recorded on *Brucker Advance 400* spectrometer. Chemical shifts (δH) are quoted in parts per million (ppm) downfield of tetramethylsilane, and residual proton of the solvent (δH ($(CH_3)_2SO$) = 2.49 ppm) used as internal reference. Coupling constants (*J*) are given in Hertz (Hz), and multiplicity abbreviated as: d (doublet), t (triplet), dd (double-doublets), m (multiplet). The ^{1}H

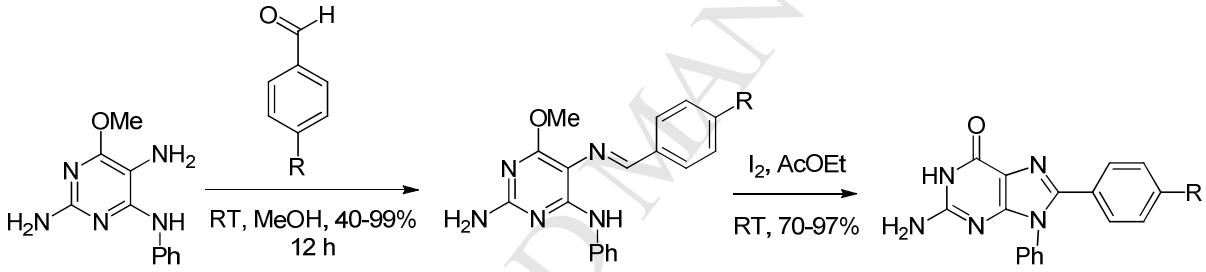
NMR spectra are reported as follows: δ/ppm (multiplicity, number of protons, coupling constants *J*/Hz).

Mass Spectrometry

Low resolution mass spectrometry by electron impact was recorded on a Hewlett Packard HP Engine-5989 spectrometer (equipped with a direct inlet probe) at *70 eV*. High Resolution Mass Spectra by electron impact were recorded on a *Micromass Auto Spec-Ultima*, magnetic sector mass spectrometer at *70 eV*.

General Procedure for the synthesis of imine derivatives (14b-d): To a solution of triaminepyrimidine **1** (1 mmol) in 15 mL of methanol, 1 mmol of appropriate arylbenzaldehyde was added, and this solution was stirred overnight at room temperature. The solid formed was filtered off, washed with fresh methanol and dried at 50 ºC, if necessary it can be recrystallized from MeOH (see Scheme 1).

Scheme 1



6-methoxy-N^5-[(*E*)-(4-methylphenyl)methylidene]-N^4-phenylpyrimidine-2,4,5-triamine (14b). 80% yield; Yellow solid; pf 168-169 ºC; R_f *0.38* (Chloroform); IR (ν cm^{-1}): 3330, 3226, 2923, 1639, 1603, 1560; ^{1}H NMR (DMSO-d_6): δ 2.37 (s, 3H), 3.94 (s, 3H), 6.56 (s, 2H), 6.99 (t, 1H, 8 *Hz*), 7.27 – 7.32 (m, 4H), 7.83 – 7.87 (m, 4H), 8.66 (s, 1H), 9.01 (s, 1H). ^{13}C NMR (DMSO-d_6): δ 21.0, 53.0, 102.7, 120.7, 121.7, 127.6, 128.4, 129.2, 135.2, 139.7, 140.0, 153.7, 158.2, 159.6, 161.2. MS (*m/z*, *%*) (assignation, abundance %): 333 (M^+, 68.2), 332 (M-1, 48.5), 331 (M-2, 100), 302 (M-31, 10.1), 242 (M-91, 89.9), 215 (M-118, 9.4), 77 ($C_6H_5^+$, 24.7). Calculated HRMS for $C_{19}H_{19}N_5O$: 333.1590; found: 333.1578.

6-methoxy-N^4-phenyl-N^5-[(*E*)-pyridin-4-ylmethylidene]pyrimidine-2,4,5-triamine (14c). 85% yield; Yellow solid; pf 208-210 ºC; R_f *0.09* (Chloroform); IR (ν cm^{-1}): 3431, 3335, 3215, 1653, 1607, 1566, 1543; ^{1}H NMR (DMSO-d_6): δ 3.97 (s, 3H), 6.78 (s, 2H), 7.03 (t, 1H, 8 *Hz*), 7.32 (t, 2H, 8 *Hz*), 7.86 – 7.91 (m, 4H), 8.66 (s, 2H), 8.77 (s, 1H), 9.01 (s, 1H). ^{13}C NMR (DMSO-d_6): δ 53.1, 102.6, 120.6, 121.3, 122.1, 128.3, 139.7, 144.5, 149.7, 150.0, 158.8, 160.3, 161.9. MS (*m/z*, *%*) (assignation, abundance %): 320 (M^+, 100), 319 (M-1, 6.1), 242

(M-78, 100), 215 (M-105, 3.6), 77 ($C_6H_5^+$, 6.4). Calculated HRMS for $C_{17}H_{16}N_6O$: 320.1386; found: 320.1385.

2-[(*E*)-{[2-amino-4-methoxy-6-(phenylamino)pyrimidin-5-yl]imino}methyl]phenol (14d). 78% yield; Yellow solid; pf 178-180 ºC; R_f *0.29* (Chloroform); IR (ν cm^{-1}): 3462, 3345, 3229, 1632,1601, 1555, 1120, 1082; ^{1}H NMR (DMSO-d_6): δ 3.90 (s, 3H), 6.52 (s, 2H), 6.91 – 6.99 (m, 3H), 7.26 – 7.33 (m, 3H), 7.70 – 7.76 (m, 3H), 8.44 (s, 1H), 9.09 (s, 1H), 11.84 (s, 1H). ^{13}C NMR (DMSO-d_6): δ 52.9, 103.2, 115.9, 118.7, 119.9, 121.4, 121.5, 128.0, 129.5, 131.2, 140.0, 156.7, 157.7, 158.3, 159.2, 160.6. MS (*m/z, %*) (assignation, abundance %): 335 (M^+, 54.8), 334 (M-1, 16.2), 242 (M-93, 100), 215 (M-120, 8.8), 77 ($C_6H_5^+$, 11.9). Calculated HRMS for $C_{18}H_{17}N_5O_2$: 335.1382; found: 335.1380.

Procedure for the one-pot synthesis of guanine derivatives (15b-e): To a solution of triaminepyrimidine **1** (1 mmol) in 15 mL of methanol, 1 mmol of appropriate arylbenzaldehyde was added, and this solution was stirred overnight at room temperature. The solid formed was filtered off, washed with fresh methanol and used directly without further purification with 1 mmol of iodine in 20 mL of AcOEt; the reaction mixture was stirred for 24 h at room temperature. The solid formed is filtered off and washed with a solution of $NaHCO_3$, then with a solution of sodium thiosulfate and, dried in an oven at 100 ºC) (see Scheme 1).

2-amino-8-(4-methylphenyl)-9-phenyl-1,9-dihydro-6*H*-purin-6-one (15b). 60% global yield; Beige solid; pf > 300 °C; R_f *0.79* ($CHCl_3$/CH_3OH, 9:1); IR (ν cm^{-1}): 3420, 3183, 1695, 1647; ^{1}H NMR (DMSO-d_6): δ 2.25 (s, 3H), 6.54 (s, 2H), 7.08 (d, 2H, 8 *Hz*), 7.24 (d, 2H, 8 *Hz*), 7.33 (m, 2H, broad band), 7.48 (m, 3H, broad band), 10.69 (s, 1H). ^{13}C NMR (DMSO-d_6): δ 20.7, 116.3, 127.2, 128.0, 128.2, 128.6, 128.7, 129.3, 135.6, 138.2, 145.2, 153.7, 153.8, 156.8. MS (*m/z,%*) (assignation, abundance %): 317 (M^+, 100), 316 (M-1, 37), 300 (M-17, 6), 275 (M-42, 6), 194 (M-123, 19), 77 ($C_6H_5^+$, 12). Calculated HRMS for $C_{18}H_{15}N_5O$: 317.1277; found: 317.1269.

2-amino-9-phenyl-8-[4-(propan-2-yl)phenyl]-1,9-dihydro-6*H*-purin-6-one (15c). 59% global yield; Beige solid; pf > 300 °C; R_f *0.80* ($CHCl_3$/CH_3OH, 9:1); IR (ν cm^{-1}): 3424, 3308, 3184, 1694, 1645, 1599, 1368; ^{1}H NMR (DMSO-d_6): δ 1.15 (d, 6H, 8 *Hz*), 2.83 (sp, 1H, 8 *Hz*), 6.53 (s, 2H), 7.15 (d, 2H, 8 *Hz*), 7.29 (d, 2H, 8 *Hz*), 7.36 (m, 2H, broad band), 7.50 (m, 3H, broad band), 10.69 (s, 1H). ^{13}C NMR (DMSO-d_6): δ 23.5, 33.0, 116.3, 126.0, 127.5, 127.9, 128.2, 128.7, 129.4, 135.6, 145.1, 148.9, 153.7, 153.8, 156.8. MS (*m/z,%*) (assignation, abundance %): 345 (M^+, 100), 344 (M-1, 20), 330 (M-15, 55), 222 (M-123, 7), 77 ($C_6H_5^+$, 17). Calculated HRMS for $C_{20}H_{19}N_5O$: 345.1590; found: 345.1579.

2-amino-8-(2-hydroxyphenyl)-9-phenyl-1,9-dihydro-6*H*-purin-6-one (15d). 66% global yield; Beige solid; pf > 300 °C; R_f *0.79* ($CHCl_3/CH_3OH$, 9:1); IR (ν cm^{-1}): 3424, 3308, 3161, 1699, 1651, 1211; ^1H NMR (DMSO-d_6): δ 6.58 (t, 1H, 8 *Hz*), 6.79 (d, 1H, 8 *Hz*), 6.84 (s, 2H), 6.90 (d, 1H, 8 *Hz*), 7.17 (t, 1H, 8 *Hz*), 7.43 – 7.45 (m, 3H), 11.45 (s, 1H)12.07 (s, 1H). ^{13}C NMR (DMSO-d_6): δ 113.9, 116.8, 118.1, 127.2, 128.1, 129.1, 129.6, 130.4, 135.7, 144.1, 153.1, 154.6, 156.5, 157.3. MS (*m/z, %*) (assignation, abundance %): 319 (M^+, 100), 318 (M-1, 75), 302 (M-17, 20), 301 (M-18, 10), 276 (M-43, 5), 196 (M-123, 13), 77 ($C_6H_5^+$, 16), 43 (M-278, 6). Calculated HRMS for $C_{17}H_{13}N_5O_2$: 319.1069; found: 319.1073.

2-amino-8-(1,3-benzodioxol-5-yl)-9-phenyl-1,9-dihydro-6*H*-purin-6-one (15e). 52% global yield; Beige solid; pf > 300 °C; R_f *0.82* ($CHCl_3/CH_3OH$, 9:1); IR (ν cm^{-1}): 3422, 3186, 1694, 1645, 1232, 1032; ^1H NMR (DMSO-d_6): δ 6.01 (s, 2H), 6.49 (s, 2H), 6.82 – 6.86 (m, 3H, broad band), 7.35 (s, 1H, broad band), 7.51 (s, 4H, broad band),10.66 (s, 1H). ^{13}C NMR (DMSO-d_6): δ 116.5, 122.1, 128.0, 128.8, 129.1, 129.4, 129.8, 131.1, 135.2, 144.0, 153.8, 153.9, 156.7. MS (*m/z,%*) (assignation, abundance %): 347 (M^+, 100), 346 (M-1, 45), 304 (M-43, 4), 224 (M-123, 13), 77 ($C_6H_5^+$, 5). Calculated HRMS for $C_{18}H_{13}N_5O_3$: 347.1018; found: 347.1015.

Bioassays

To avoid including experimental errors in the correlation, we used exactly the same bioassay as those used in reference [5]. The assay is based on the ability of DHFR to catalyze the NADPH-dependent reduction of dihydrofolic acid to tetrahydrofolic acid. The rate of NADPH consumption in the presence of the compound under investigation is monitored by the decrease in absorbance at 340 nm [50-55] every 15 seconds, during 2.5 minutes. Reactions were performed in a solution containing saturating concentrations of cofactor (80 μM NADPH) and substrate (50 μM dihydrofolate), 15 μL of enzyme solution, 50 mM Tris−HCl, 0.001 M 2-mercaptoethanol, and 0.001 M EDTA at pH 7.4 and 30 °C. The enzyme was purchased from Sigma-Aldrich Co. (St. Louis, MO). In order to determine IC_{50} values, five different concentrations of compounds as inhibitors were tested (each assay was made by triplicate) and percent inhibition graphs were drawn by using statistical packing program on a computer.

RESULTS AND DISCUSSION

We recently reported two new DHFR inhibitors: compounds **14a** and **15a** [5]. In order to have a more extensive series and based on these two structures, we decide to synthesize seven novel derivatives from 2,4,5-triamino-6-methoxypyrimidine with diverse aryl aldehydes to render well compounds **14b-d** (structurally related to **14a**) or compounds **15b-e** (structurally

related to compound **15a**), which were prepared in a two-step one-pot procedure (Scheme 1 and Table 1).

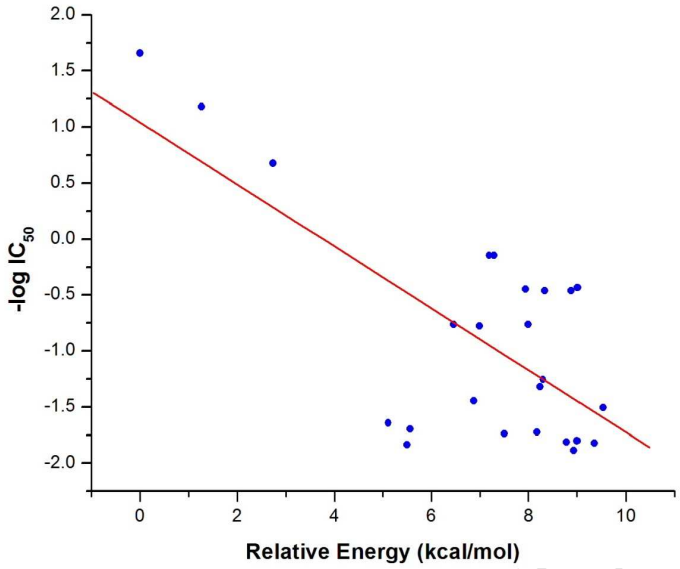
In the next step we tested the inhibitory effect of these compounds; such results are summarized in Table 1. As we can see from Table 1 some of these compounds displayed a relatively significant activity, such as compound **15b** which shows inhibitory effect at concentrations of 27.87μM.

As it was stated above the main objective of this work is to obtain a correlation between the binding energies of these compounds and their respective IC_{50}. Therefore MD simulations and quantum mechanical calculations were performed for all the L-R complexes. To try to use a sample as representative as possible of different types of ligands, we also include in this analysis the sixteen compounds reported in our previous work [5], thus forming a complete set of 25 compounds (Figure 1).

As we expected the general results obtained from MD simulations were very similar to those previously reported for compounds **14a** and **15a** [5]. A highly conserved glutamic acid (Glu30) is functioning as an anchoring point. In the present study, all the simulated compounds were docked into the receptor with the N1 and 2-amino group near to Glu30. After 5 ns of MD simulations, the ligands moved slightly but in a different form compared with the initial position. However, the strong interaction with Glu30 was maintained for all the complexes. Other important MI stabilizing the different complexes are: π···π stacking interactions with Phe31 and Phe34, and hydrophobic interactions with Ile7 and Val115.

Next, the BE obtained for the different complexes were evaluated. From the BE obtained from our MD simulations, a very good binder can be differentiated from a very weak binder (-11.85 kcal/mol for **MTX** vs. -6.74 and -3.61 kcal/mol for compounds **14b** and **15c**, respectively) but ligands with similar binding affinities cannot be easily differentiated. In function of the IC_{50} values we were expecting exactly the opposite values for compounds **14b** and **15c**. Similar to these unexpected results, there are others that might be observed in Table 1S in Supplementary Material. In addition, it is important to note that the value of R^2 obtained for this correlation is very low (0.49) (Figure 3). This result clearly indicates that by using this approximation is not possible to discriminate between compounds with similar binding affinities; in addition the low correlation obtained indicates a very poor predictive power of the method. This result was not surprising since we obtained similar results in our previous work using this approach [5]. Considering that MD simulations might neglect or poorly approximate terms that are playing determinant roles (such as lone pair directionality in hydrogen bonds, explicit π···π stacking polarization effects, hydrogen bonding networks, induced fit, and conformational entropy), we cannot expect to detect clear differences between compounds possessing relatively similar BE.

Figure 3



In the next step of our study, reduced model systems were optimized using combined semiempirical, *ab initio* and DFT calculations. To perform these calculations, reduced system models were employed whose design is explained in the calculation methods section. PM6 optimizations were performed considering all receptor amino acids that might interact after initial positioning of the ligands against Glu30 residue. Next, RHF/6-31G(d) and DFT (PBE1PBE/6-31G(d)) single point calculations were carried out for each complex optimized from PM6 computations.

It is important to note that the L-R interaction is a dynamic process and therefore in order to have a more accurate description of such situation, four different snapshots for each complex were considered. This resulted in different energy values and such variation can be observed in the error bars which are shown in Figures 4, 5 and 6. Once the BE of the different complexes were obtained from the theoretical calculations, the different correlations between these theoretical calculations and our experimental data (Table 1) were calculated. The Figure 4 shows that semi-empirical calculations (PM6) gave a correlation between the BE and the inhibitory activity with an R^2 value of 0.49, which does not improve those results obtained with the LIE method (Figure 3). However, the results obtained from RHF/6-31G(d) (Figure 5) and PBE1PBE/6-31G(d) (Figure 6) were significantly better with R^2 values of 0.77 and 0.76, respectively. Nevertheless it is important to note that not only the value of R^2 is important in a correlation but also how the distribution of the various points along the line is. Regarding Figures 5 and 6 it is evident that the different points are clustered into two well-defined groups. This would indicate that although these correlations allow us to differentiate between

a very active compound with respect to a compound with low affinity, however there is room for doubt whether it is possible to distinguish between two compounds having similar affinities. To corroborate this assumption we removed the three classical type inhibitors (**MTX**, compounds **1** and **2**) from the series and a new correlation was obtained with this new series. Such as we expected the new correlations gave very low R^2 values (0.64 and 0.61 for *ab initio* and DFT calculations, respectively). These results show the severe limitations of these approaches in order to correctly predict the inhibitory activity between compounds with similar affinities. This result was somewhat disappointing for us so we seek a new parameter or kind of molecular descriptor that might be able to predict the inhibitory activity between compounds with similar BE in this series.

Figure 4

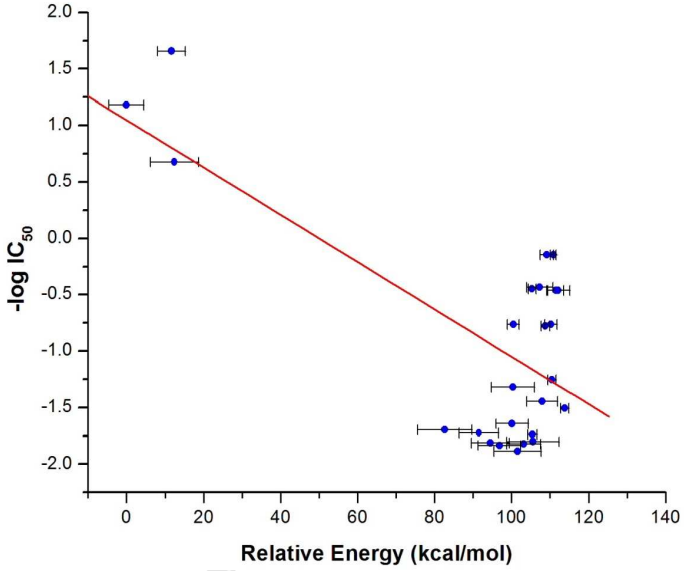


Figure 5

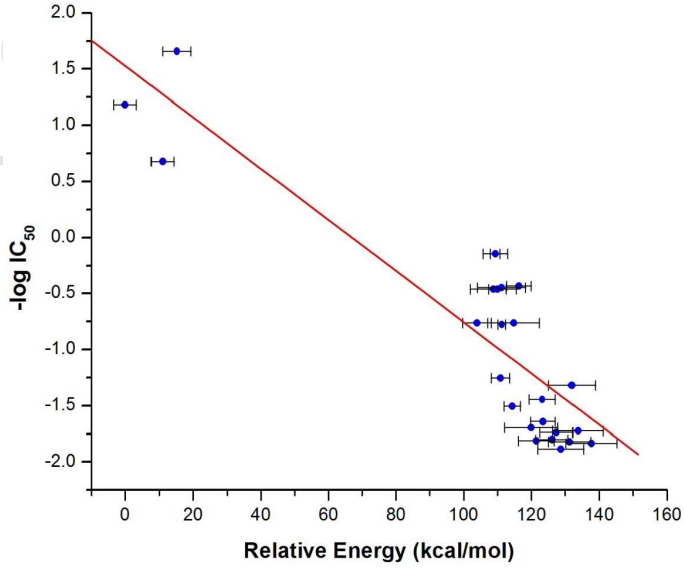
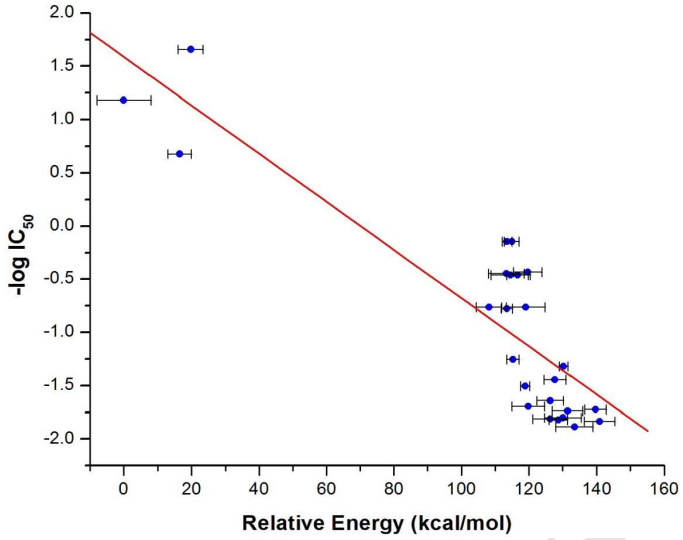


Figure 6



We recently reported that the information about the electronic density obtained from a QTAIM analysis is useful to describe the molecular interactions that stabilize and destabilize the different complexes L-R [13-15]. Specifically in our previous work reported for DHFR inhibitors, the QTAIM analysis gave very interesting results when it was applied to this molecular target [5]. Therefore in the next stage of our study we performed a QTAIM analysis of all the complexes in order to find a new descriptor for this series.

QTAIM Analysis

The Figure 7 shows the sum of the ρ(r) values corresponding to the interactions of the Pyr system and C6 and/or N7 side chains obtained for the different inhibitors analyzed here. The different moieties of each type of inhibitor are showed in Figure 8. The sum of the ρ(r) values for all the interactions of one part of the inhibitor (i.e., the Pyr ring system and the C6-side chain) provides a measure of the anchoring strength of such moiety of the inhibitor to the binding pocket. Figure 7 clearly shows that the classical inhibitors (**MTX, 1** and **2**) bind to the enzyme with similar strength through both parts of the molecule, the Pyr system and the C6-side chain. Non-classical inhibitors reported by Gangjee *et al*. (compounds **3-13**) bind their Pyr portions with similar strength to that of the classical inhibitors, but the anchoring through the C6-side chain is much weaker in these inhibitors with respect to the classical ones. Regarding the inhibitors reported here (compounds **14b-e** and **15b-f**), on average, these compounds are more weakly bonded to the binding pocket than the rest of the inhibitors. Analyzing the anchoring strength of each part of these compounds to the binding pocket, it can be observed that the equivalent Pyr systems are more weakly bonded to the pocket than

the C6- and N7-side chains. It can also be observed that the equivalent Pyr ring system is more weakly H-bonded to the active site in these new compounds than in the rest of the inhibitors. In contrast, the C6- and N7-side chain from compounds **14a-e** and **15a-f** are on average more strongly bonded to the binding pocket than the C6-side chain from the non-classical inhibitors.

Figure 7

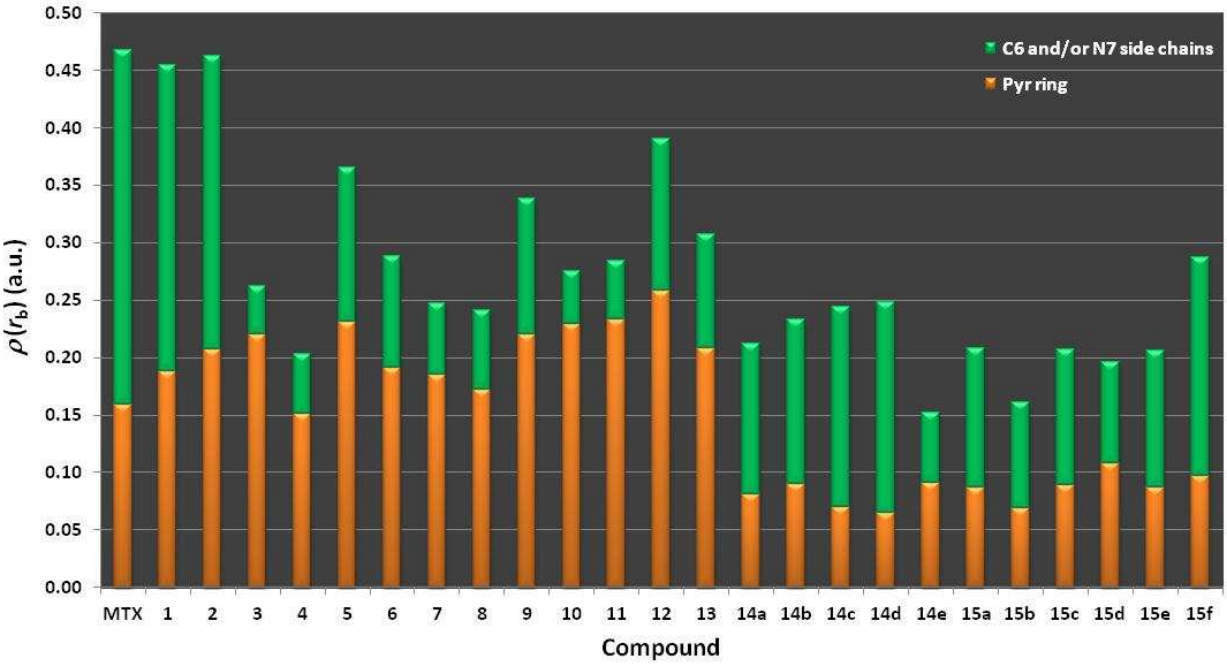
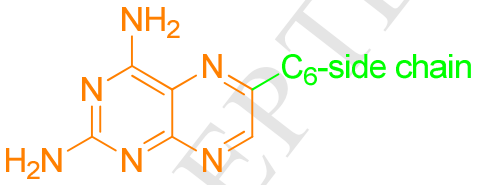
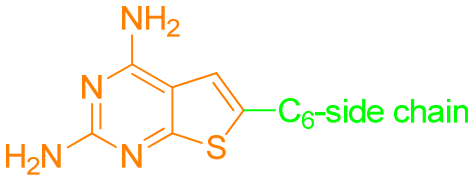


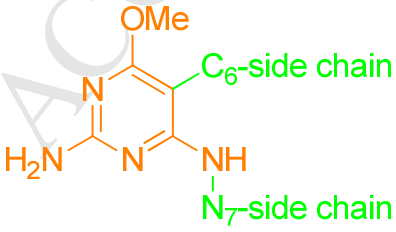
Figure 8



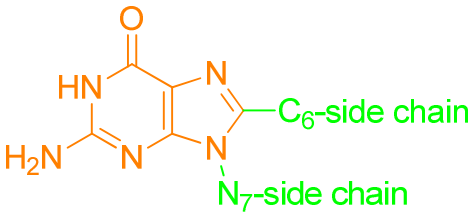
MTX



Gangjee´sinhibitors



Compounds of Series 14



Compounds of Series 15

The electronic density ρ(r) as a molecular descriptor for the L-R affinity

The QTAIM study allows not only analyzing quantitatively each interaction but also getting a correlation using the calculated $\rho(r)$ values and the IC_{50} data obtained in our laboratory. For this purpose, we have considered the $\rho(r)$ value obtained from the QTAIM analysis as a molecular descriptor in order to use them as independent variables in the multivariable lineal regression analysis (for details about this analysis see the method section).

Of the 25 total compounds under study, 22 were used for the construction and calibration of the model, while the remaining three (compounds **7**, **14c** and **15c**) were left to test it. Therefore it should be noted that a model of up to 20 amino acids was obtained. Thus, we obtained the corresponding correlations using different number of amino acids. In the Figure 9, it can be appreciated that the utilization of 7 amino acids is enough to get a very good value of R^2 and increasing the number of them does not significantly improved the correlation. In the Table 2, the main results obtained for models using different numbers of amino acids (from 3 to 8) are summarized. Importantly, all the models showing significant correlations maintain the same four amino acids (Val8, Ala9, Leu67 and Arg70), indicating that they are essential to discriminate the L-R affinity among these compounds. The model possessing only three amino acids maintains only two of these amino acids and has a significantly lower value of R^2 (see the first column of Table 2). Although the model with eight amino acids has a high value of R^2, however its power of prediction is lower with respect to the model of seven amino acids (see the third column in Table 2). These results indicate that the model considering seven amino acids would be the most robust in order to analyze this series maintaining a reazonable number of aminoacids.

Figure 9

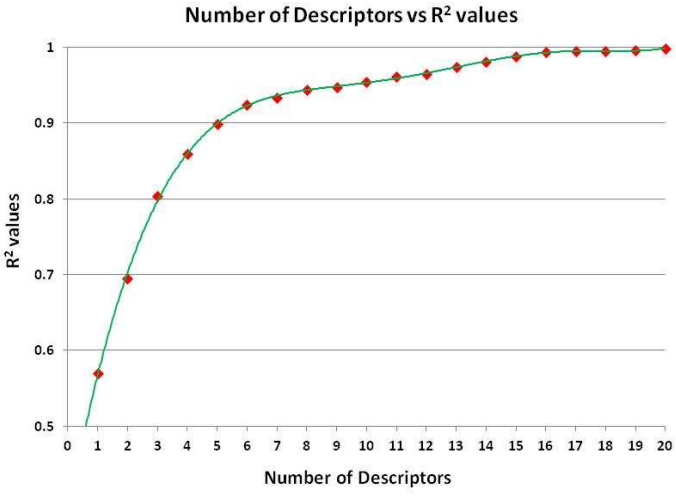


Table 2. R^2, R^2_{val} and S values obtained from the models employing from 3 to 8 amino acids.

Number of residues	Residues	Considering all compounds			Excluding MTX, 1 y 2		
		$R^{2\,(*)}$	$S^{(*)}$	$R^2_{val}{}^{(**)}$	$R^{2\,(*)}$	$S^{(*)}$	$R^2_{val}{}^{(**)}$
3	Leu67, Arg70 and Tyr121	0.80	0.48	0.98	0.66	0.40	0.93
4	Val8, Ala9, Leu67 and Arg70	0.86	0.42	0.99	0.75	0.35	0.97
5	Val8, Ala9, Asn64, Leu67 and Arg70	0.90	0.37	0.99	0.79	0.34	0.97
6	Val8, Ala9, Asn64, Leu67, Arg70 and Thr136	0.92	0.33	0.99	0.88	0.26	0.98
7	Val8, Ala9, Asn64, Leu67, Arg70, Val115 and Thr136	0.93	0.32	0.99	0.90	0.25	0.99
8	Val8, Ala9, Thr56, Asn64, Leu67, Arg70, Val115 and Thr136	0.94	0.30	0.94	0.91	0.25	0.96

(*) Calculated R^2 y S values from the corresponding model.
(**) Calculated R^2 value for the 3 compounds (**7**, **14c** and **15c**) used to validate the model.

The mathematical model of 7 amino acids was tested using the three compounds which do not form part of the initial statistical analysis. As can be seen in Table 3, the predicted activities of these compounds do not differ from the experimental results, showing the high predictive power of the developed model. The residues included in this model are: Val8, Ala9, Asn64, Leu67, Arg70, Val115 and Thr136. The first two ones interact with the main ring of the ligand (Pyr ring) through hydrogen bonds with the atoms of their backbones. Asn64 presents polar interactions with different groups located on the phenyl ring. In the case of Leu67, this residue displays hydrophobic interactions with the phenyl ring and different non polar groups located on it. It is important to note that the interaction of Arg70 with the ligands allows the differentiation between classical and non-classical compounds. This residue is able to form a salting bridge with the α-carboxilate group of classical inhibitors, which is a very strong interaction. In a minor degree, Arg70 presents MI with polar groups of the non-classical ligands. Val115 and Thr136 can establish hydrogen bond interactions with amino groups located in the Pyr ring. Some residues like Glu30, Phe31 and Phe34, which produce important interactions, have not been considered in this model. This is because all the interactions

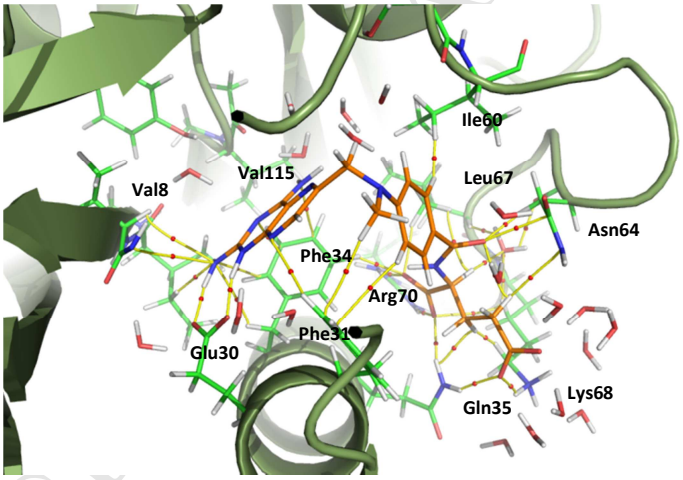
observed in these residues are present in a similar way in all complexes and, therefore, they do not allow us to differentiate the ligands in terms of their activities within the series. Most of the interactions previously described might be well appreciated in Figure10 in which the molecular graph obtained for **MTX** is shown. Molecular graphs were analyzed for the different complexes studied here. We only show that obtained for **MTX** as an example in which the main interactions might be appreciated (Figure 10).

Table 3. Activity values (theoretical and experimental) obtained for compounds **7**, **14c** y **15c**

Compound	-log IC_{50}		Error (%)
	Theoretical	Experimental	
7 (*)	-0.50	-0.46	8.79
14c	-1.46	-1.80	18.88
15c	-1.57	-1.64	4.08

(*) Compound reported in reference 16

Figure 10



The next step was to determine whether this model using the electron density as molecular descriptor can only discriminate between compounds with different affinities, or by otherwise it is also possible to distinguish between compounds that have similar affinities. To test the model, the three classical inhibitors were removed from the test group, resulting in just a very slightly diminution of the R^2 (0.90). It should be noted that the R^2 testing the model was 0.99 showing that the model is able to maintain not only a high correlation value but also an excellent predictive power. This result is extremely important as it shows that this model using the $\rho(r)$ as molecular descriptor can be used to predict the activity of compounds with similar affinities in this series. This result is highly satisfactory as it allows us to achieve the main objective of our work; this is we are able to differentiate between compounds that have

similar affinities and not only compounds with different affinities. An important aspect to note when we use this type of approach is that just one molecular descriptor has been used and therefore we can easily understand the physical-chemical behavior of the different L-R complexes. In this sense it is possible to determine which are those amino acids that might perform stabilizing or destabilizing interactions on the various complexes and what is more important is that such information can be obtained quantitatively.

CONCLUSIONS

There are many techniques and approaches to study the postdocking problem. While many of these techniques are able to distinguish between compounds that have high affinity with respect to those with low affinity for the receptor, unfortunately most of these techniques have failed when they have to distinguish between compounds with similar affinities.

In this paper we have succeeded in what we fail in our previous report [5]. In this case we have been able to get an excellent correlation between the electronic densities obtained from a QTAIM study with the experimental data. It is important to note that through this study is possible to differentiate the L-R affinity even though the compounds possess similar affinities. Very important additional information obtained through this type of study is that it is possible to visualize which amino acids are involved in the interactions determining the different affinities of the ligands. Such information is crucial when we are interested in the design of new specific ligands. Some additional benefits that may be mentioned for this approach are: the technique is relatively simple, easy to interpret and not too demanding about the computing time. However, a somewhat limiting aspect of this type of study is that the QTAIM study is highly dependent on the optimized geometry and therefore the conformational variability can be a serious problem. In the particular case of DHFR this is not too problematic due to the structural characteristics of the active site of the enzyme. It is well known that the DHFR binding pocket is relatively narrow, with little space for the ligand and therefore does not lead to large conformational changes at least in comparison with others more flexible binding sites. These features of the active site have allowed that with care in the calculations (four conformations for each complex were considered), it is possible to obtain highly satisfactory results. Clearly, if the characteristics of the active site are different it is necessary to check whether this technique is effective exactly with this procedure or if it is necessary to consider more carefully the issue of conformational variability of the various complexes.

In summary, in this paper we have shown that the electron density obtained from a QTAIM analysis is an excellent descriptor of molecular interactions that stabilize and destabilize the formation of the L-R complex. If this information is used properly, it is possible the

preparation of models which might allow to differentiate the affinity between ligands even when they show similar affinities. The possibility of using this postdocking technique on other binding sites with different structural characteristics is currently being studied in our laboratory.

ACKNOWLEDGMENTS

Grants from Universidad Nacional de San Luis (UNSL), partially supported this work. This research was also supported by the Spanish "Ministerio de Educación y Ciencia" grant SAF 2007-63142. Rodrigo Tosso thanks a doctoral fellowship of CONICET-Argentina. R. D. Enriz and S. A. Andujar are members of the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET-Argentina) staff. We especially acknowledge MSc. Daniel O. Zamo for technical assistance.

REFERENCES

[1] P. Kollman, Free energy calculations: applications to chemical and biochemical phenomena, Chem Rev 93(7) (1993) 2395-2417.
[2] H.A. Carlson, W.L. Jorgensen, An extended linear response method for determining free energies of hydration, J Phys Chem 99(26) (1995) 10667-73.
[3] P.A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D.A. Case, T.E. Cheatham, 3rd, Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models, Acc Chem Res 33(12) (2000) 889-97.
[4] R. Teramoto, H. Fukunishi, Supervised consensus scoring for docking and virtual screening, J Chem Inf Model 47(2) (2007) 526-34.
[5] R.D. Tosso, S.A. Andujar, L. Gutierrez, E. Angelina, R. Rodriguez, M. Nogueras, H. Baldoni, F.D. Suvire, J. Cobo, R.D. Enriz, Molecular modeling study of dihydrofolate reductase inhibitors. Molecular dynamics simulations, quantum mechanical calculations, and experimental corroboration, J Chem Inf Model 53(8) (2013) 2018-32.
[6] M. Graffner-Nordberg, K. Kolmodin, J. Aqvist, S.F. Queener, A. Hallberg, Design, synthesis, computational prediction, and biological evaluation of ester soft drugs as inhibitors of dihydrofolate reductase from Pneumocystis carinii, J Med Chem 44(15) (2001) 2391-402.
[7] M. Graffner-Nordberg, K. Kolmodin, J. Aqvist, S.F. Queener, A. Hallberg, Design, synthesis, and computational affinity prediction of ester soft drugs as inhibitors of dihydrofolate reductase from Pneumocystis carinii, Eur J Pharm Sci 22(1) (2004) 43-54.
[8] C. Pitts, J. Yin, D. Bowen, C.J. Maxwell, W.M. Southerland, Interaction energy analysis of nonclassical antifolates with Pneumocystis carinii dihydrofolate reductase, Int J Mol Sci 3 (2002) 1188-1202.
[9] S. Bag, N.R. Tawari, M.S. Degani, S.F. Queener, Design, synthesis, biological evaluation and computational investigation of novel inhibitors of dihydrofolate reductase of opportunistic pathogens, Bioorg Med Chem 18(9) (2010) 3187-97.
[10] J.E. Kerrigan, E.E. Abali, J.R. Bertino, Recent Progress in Molecular Dynamics Simulations of Dihydrofolate Reductase, Curr Enzyme Inhib 8 (2012) 140-9.
[11] S.A. Andujar, B.M. de Angel, J.E. Charris, A. Israel, H. Suarez-Roca, S.E. Lopez, M.R. Garrido, E.V. Cabrera, G. Visbal, C. Rosales, F.D. Suvire, R.D. Enriz, J.E. Angel-Guio, Synthesis, dopaminergic profile, and molecular dynamics calculations of N-aralkyl substituted 2-aminoindans, Bioorg Med Chem 16(6) (2008) 3233-44.

[12] S. Andujar, F. Suvire, I. Berenguer, N. Cabedo, P. Marin, L. Moreno, M. Dolores Ivorra, D. Cortes, R.D. Enriz, Tetrahydroisoquinolines acting as dopaminergic ligands. A molecular modeling study using MD simulations and QM calculations, J Mol Model 18(2) (2012) 419-31.
[13] S.A. Andujar, R.D. Tosso, F.D. Suvire, E. Angelina, N. Peruchena, N. Cabedo, D. Cortes, R.D. Enriz, Searching the "biologically relevant"conformation of dopamine: a computational approach, J Chem Inf Model 52(1) (2012) 99-112.
[14] E. Angelina, S. Andujar, R.D. Tosso, R.D. Enriz, N. Peruchena, Non-covalent interactions in receptor–ligand complexes. A study based on the electron charge density, J. Phys. Org. Chem. 27 (2014) 128-134.
[15] J. Parraga, N. Cabedo, S. Andujar, L. Piqueras, L. Moreno, A. Galan, E. Angelina, R.D. Enriz, M.D. Ivorra, M.J. Sanz, D. Cortes, 2,3,9- and 2,3,11-trisubstituted tetrahydroprotoberberines as D2 dopaminergic ligands, Eur J Med Chem 68 (2013) 150-66.
[16] A. Gangjee, Y. Zeng, T. Talreja, J.J. McGuire, R.L. Kisliuk, S.F. Queener, Design and synthesis of classical and nonclassical 6-arylthio-2,4-diamino-5-ethylpyrrolo[2,3-d]pyrimidines as antifolates, J Med Chem 50(13) (2007) 3046-53.
[17] M.F. Andrada, E.G. Vega-Hissi, F.M. Garibotto, M.R. Estrada, J. Cobo, R. Enriz, J.C. Garro Martinez, Application of 2D and 3D-QSAR models for the design the novel Dihydrofolate reductase inhibitors. Synthesis and biological evaluation, Chem Biol Drug Des (In Press).
[18] A. Ribeiro, B. Horta, R.B. de Alencastro, MKTOP: a program for automatic construction of molecular topologies, J Braz Chem Soc 19 (2008) 1433-35.
[19] H.H. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: a message-passing parallel molecular dynamics implementations, Comput Phys Commun 91 (1995) 43-56.
[20] S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M.R. Shirts, J.C. Smith, P.M. Kasson, D. van der Spoel, B. Hess, E. Lindahl, GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit, Bioinformatics 29(7) (2013) 845-54.
[21] A. van Buuren, S. Marrink, H.H. Berendsen, A molecular dynamics study of the decane/water interface, J Phys Chem 36 (1993) 9206-12.
[22] A. Mark, S. van Helden, P. Smith, L. Janssen, W. van Gunsteren, Convergence properties of free energy calculations. A-cyclodextrin complexes as a case study., J Am Chem Soc 116 (1994) 6293-6302.
[23] W. Jorgensen, J. Chandrasekhar, J. Madura, R. Impey, M. Klein, Comparison of simple potential functions for simulating liquid water, J Chem Phys 79 (1983) 926-35.
[24] A.R. van Buuren, H.J. Berendsen, Molecular dynamics simulation of the stability of a 22-residue alpha-helix in water and 30% trifluoroethanol, Biopolymers 33(8) (1993) 1159-66.
[25] H. Liu, F. Muller-Plathe, W. van Gunsteren, A Force Field for Liquid Dimethyl Sulfoxide and Physical Properties of Liquid Dimethyl Sulfoxide Calculated Using Molecular Dynamics Simulation, J. Am. Chem. Soc. 117 (1995) 4363-4366.
[26] S. Miyamoto, P. Kollman, SETTLE-an analytical version of the SHAKE and RATTLE algorithm for rigid water models, J. Comput. Chem. 13 (1992) 952-962.
[27] H.J. Berendsen, H. Postma, W. van Gunsteren, W. Hermans, Interaction models for water in relation to protein hydration, in: B. Pullman (Ed.), Intermolecular Forces, Reidel: Dordrecht, The Netherlands, 1981, pp. 331-342.
[28] T. Darden, D. York, L. Pedersen, Particle mesh Ewald - an N.log(n) method for Ewald sums in large systems., J. Chem. Phys. 98 (1993) 10089-10092.
[29] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, J Chem Phys 126(1) (2007) 014101.
[30] U. Essmann, L. Perera, M. Berkowitz, T. Darden, H. Lee, L. Pedersen, A smooth particle mesh Ewald method, J. Chem. Phys. 103 (1995) 8577-8593.
[31] B. Luty, I. Tironi, W. van Gunsteren, Lattice - sum methods for calculating electrostatic interactions in molecular simulations, J. Chem. Phys. 103 (1995) 3014-3021.

[32] J. Aqvist, C. Medina, J.E. Samuelsson, A new method for predicting binding affinity in computer-aided drug design, Protein Eng 7(3) (1994) 385-91.
[33] H. Gutierrez-de-Teran, J. Aqvist, Linear interaction energy: method and applications in drug design, Methods Mol Biol 819 (2012) 305-23.
[34] T. Hansson, J. Marelius, J. Aqvist, Ligand binding affinity prediction by linear interaction energy methods, J Comput Aided Mol Des 12(1) (1998) 27-35.
[35] J. Marelius, M. Graffner-Nordberg, T. Hansson, A. Hallberg, J. Aqvist, Computation of affinity and selectivity: binding of 2,4-diaminopteridine and 2,4-diaminoquinazoline inhibitors to dihydrofolate reductases, J Comput Aided Mol Des 12(2) (1998) 119-31.
[36] T. Hou, N. Li, Y. Li, W. Wang, Characterization of domain-peptide interaction interface: prediction of SH3 domain-mediated protein-protein interaction network in yeast by generic structure-based models, J Proteome Res 11(5) (2012) 2982-95.
[37] H. Gohlke, C. Kiel, D.A. Case, Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes, J Mol Biol 330(4) (2003) 891-913.
[38] T. Hou, W. Zhang, D.A. Case, W. Wang, Characterization of domain-peptide interaction interface: a case study on the amphiphysin-1 SH3 domain, J Mol Biol 376(4) (2008) 1201-14.
[39] T. Hou, Z. Xu, W. Zhang, W.A. McLaughlin, D.A. Case, Y. Xu, W. Wang, Characterization of domain-peptide interaction interface: a generic structure-based model to decipher the binding specificity of SH3 domains, Mol Cell Proteomics 8(4) (2009) 639-49.
[40] T. Hou, Y. Li, W. Wang, Prediction of peptides binding to the PKA RIIalpha subunit using a hierarchical strategy, Bioinformatics 27(13) (2011) 1814-21.
[41] D.A. Case, T.A. Darden, T.E. Cheatham III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Goetz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, P.A. Kollman, AMBER12, University of California, San Francisco, 2012.
[42] J.J. Stewart, Optimization of parameters for semiempirical methods V: modification of NDDO approximations and application to 70 elements, J Mol Model 13(12) (2007) 1173-213.
[43] J.J. Stewart, MOPAC2009, in: S.C. Chemistry (Ed.) Colorado Springs, CO, USA, 2008.
[44] 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, Ö. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, Gaussian 09, Gaussian Inc., Wallingford, CT, 2009.
[45] R.F.W. Bader, Atoms in Molecules: A Quantum Theory, Oxford, U.K., 1994.
[46] T. Lu, F. Chen, Multiwfn: a multifunctional wavefunction analyzer, J Comput Chem 33(5) (2012) 580-92.
[47] http://www.chemistry.mcmaster.ca/aimpac.
[48] The PyMOL Molecular Graphics System, Schrödinger, LLC., 2010.
[49] MATLAB, MathWorks Inc., Natick, MA, 2004.
[50] J.E. Gready, Dihydrofolate reductase: binding of substrates and inhibitors and catalytic mechanism, Adv Pharmacol Chemother 17 (1980) 37-102.
[51] R.L. Blakley, Eukaryotic dihydrofolate reductase, Adv Enzymol Relat Areas Mol Biol 70 (1995) 23-102.

[52] M.P. Costi, S. Ferrari, Update on antifolate drugs targets, Curr Drug Targets 2(2) (2001) 135-66.
[53] B.I. Schweitzer, A.P. Dicker, J.R. Bertino, Dihydrofolate reductase as a therapeutic target, Faseb J 4(8) (1990) 2441-52.
[54] C.K. Mathews, K.G. Scrimgeour, F.M. Huennekens, Dihydrofolate reductase, Methods Enzymol 6 (1963) 364-68.
[55] B.L. Hillcoat, P.F. Nixon, R.L. Blakley, Effect of substrate decomposition on the spectrophotometric assay of dihydrofolate reductase, Anal Biochem 21(2) (1967) 178-89.

Figure Captions

Figure 1. Structural features of the compounds reported by Gangjee *et al.* (**1-13**) and Tosso *et al.* (**14a** and **15a**). Each IC_{50} value is indicated in parenthesis.

Figure 2. Surface of the active site of DHFR (light blue). **MTX** (yellow sticks), **$NADP^+$** (grey sticks) and the protein (magenta ribbon) are also represented. Leu22 and Trp24 were not considered in the surface for clarify.

Figure 3. Correlation obtained from LIE calculations. The *x-axis* denotes relative energies ($\Delta\Delta G$), and the *y-axis* denotes the inhibitory activity value of each compound (-log IC_{50}).

Figure 4. Correlation obtained from semi-empirical calculations. The *x-axis* denotes PM6 relative energies, and the *y-axis* denotes the inhibitory activity value of each compound (-log IC_{50}).

Figure 5. Correlation obtained from *ab initio* calculations. The *x-axis* denotes *ab initio* (RHF/6-31G(d)) relative energies, and the *y-axis* denotes the inhibitory activity value of each compound (-log IC_{50}).

Figure 6. Correlation obtained from DFT calculations. The *x-axis* denotes DFT (PBE1PBE/6-31G(d)) relative energies, and the *y-axis* denotes the inhibitory activity value of each compound (-log IC_{50}).

Figure 7. Charge density values for the total interactions of the Pyr ring system (orange stacked bars) and the C6 and/or N7 side chains (green stacked bars) for the inhibitors at the binding pocket. The repulsive short C-H···H-C contacts were not included.

Figure 8. Schematic representation of each moiety of the different inhibitors. The main ring (Pyr ring) of the compounds is indicated in orange.

Figure 9. R^2 values of each obtained correlation employing different number of residues.

Figure 10. MTX (in orange sticks) at the DHFR binding site (in green sticks). Also the elements of the topology of the electron density are shown: yellow sticks represent the bond

paths connecting the nuclei and the red circles on them are the BCPs. Due to the complexity of the structure most of the interactions of water molecules are not shown.

Scheme 1. Synthesis of imine and guanine derivatives.

Highlights

The electronic density obtained from a QTAIM analysis used as molecular descriptor. A study performed in a new series of DHFR inhibitors

Rodrigo D. Tosso, Marcela Vettorazzi, Sebastian A. Andujar, Lucas J. Gutierrez, Juan C. Garro, Emilio Angelina, Ricaurte Rodríguez, Fernando D. Suvire, Manuel Nogueras, Justo Cobo and Ricardo D. Enriz

Departamento de Química, Facultad de Química, Bioquímica y Farmacia, Universidad Nacional de San Luis, Chacabuco 915, 5700 San Luis, Argentina

- A new series of guanine and imine derivatives has been synthesized.
- The activities of the new compounds were evaluated against human DHFR.
- MD simulations, QM calculations and QTAIM analysis were performed.
- An excellent correlation between theoretical and experimental data was obtained.