J Mol Model (2012) 18:4061–4072
DOI 10.1007/s00894-012-1399-y

ORIGINAL PAPER

# CoMFA/CoMSIA 3D-QSAR of pyrimidine inhibitors of *Pneumocystis carinii* dihydrofolate reductase

**Osvaldo A. Santos-Filho · Delphine Forge · Lucas V. B. Hoelz · Guilherme B. L. de Freitas · Thiago O. Marinho · Jocley Q. Araújo · Magaly G. Albuquerque · Ricardo B. de Alencastro · Nubia Boechat**

Received: 9 January 2012 / Accepted: 6 March 2012 / Published online: 14 April 2012


**Abstract** *Pneumocystis carinii* is typically a non-pathogenic fungus found in the respiratory tract of healthy humans. However, it may cause *P. carinii* pneumonia (PCP) in people with immune deficiency, affecting mainly premature babies, cancer patients and transplant recipients, and people with acquired immunodeficiency syndrome (AIDS). In the latter group, PCP occurs in approximately 80% of patients, a major cause of death. Currently, there are many available therapies to treat PCP patients, including *P. carinii* dihydrofolate reductase (*Pc*DHFR) inhibitors, such as trimetrexate (TMX), piritrexim (PTX), trimethoprim (TMP), and pyrimethamine (PMT). Nevertheless, the high percentage of adverse side effects and the limited therapeutic success of the current drug therapy justify the search for new drugs rationally planned against PCP. This work focuses on the study of pyrimidine inhibitors of *Pc*DHFR, using both CoMFA and CoMSIA 3D-QSAR methods.

**Keywords** 3D-QSAR · CoMFA · CoMSIA · Dihydrofolate reductase · *Pneumocystis carinii*

**Electronic supplementary material** The online version of this article (doi:10.1007/s00894-012-1399-y) contains supplementary material, which is available to authorized users.

O. A. Santos-Filho (✉) · N. Boechat
Laboratório de Modelagem Molecular, Departamento de Síntese Orgânica, Instituto de Tecnologia em Fármacos,
Fundação Oswaldo Cruz, Rua Sizenando Nabuco, 100, Manguinhos,
21041-250 Rio de Janeiro, RJ, Brazil
e-mail: osvaldofilho@far.fiocruz.br

D. Forge
Laboratoire de Chimie Organique, Institut de Chimie,
Université de Mons - UMons,
Place du Parc,
7000 Mons, Belgium

L. V. B. Hoelz · T. O. Marinho · J. Q. Araújo · M. G. Albuquerque · R. B. de Alencastro
Departamento de Química Orgânica, Instituto de Química,
Universidade Federal do Rio de Janeiro,
CT, Bloco A, Lab. 609, Cidade Universitária,
21941-909 Rio de Janeiro, RJ, Brazil

G. B. L. de Freitas
Departamento de Farmácia, Setor de Ciências da Saúde,
Universidade Estadual do Centro-Oeste,
Rua Camargo Varela de Sá, 3, Vila Carli,
85015-430 Guarapuava, PR, Brazil

## Introduction

Acquired immunodeficiency syndrome (AIDS) is a disease of the immune system caused by the human immunodeficiency virus (HIV) [1]. This illness is a major health problem in many parts of the world. In 2009, the World Health Organization (WHO) estimated that there were 33.4 million people worldwide living with HIV/AIDS, with 2.7 million new HIV infections per year, and 2.0 million annual deaths due to AIDS. According to the UNAIDS 2009 report, some 60 million people have been infected worldwide since the start of the pandemic, with some 25 million deaths, and 14 million orphaned children in southern Africa alone.

Because patients infected with HIV are more susceptible to opportunistic infections and the cost of antiretroviral therapy remains high, AIDS has had a major impact on several diseases caused by protozoan pathogens. For example, *Pneumocystis carinii* pneumonia (PCP), caused by *Pneumocystis jirovecii* (formerly *Pneumocystis carinii*), is still the most common AIDS-defining opportunistic infection in HIV-infected patients in the United States and

Europe [2]. PCP was described almost simultaneously by Chagas, in 1909, and Carini, in 1910 [3]. The drugs currently prescribed for the prophylaxis and treatment of PCP in AIDS patients fall into several classes. All of them consist of lipophilic dihydrofolate reductase inhibitors.

Dihydrofolate reductase [DHFR; 5,6,7,8-tetrahydrofolate-$NADP^+$ oxidoreductase (E.C.1.5.1.3)] is an important target for drug development against cancer and a variety of infectious diseases caused by bacteria, protozoa, and fungi [4]. Its importance arises from its function in DNA biosynthesis and cell replication. DHFR catalyzes the reduction of dihydrofolate (DHF) to tetrahydrofolate (THF), an essential cofactor in the thymidylate monophosphate (dTMP) biosynthesis. Inhibition of DHFR leads to a deficiency of dTMP, since DHF cannot be recycled, and thus causes inhibition of cell growth.

Trimethoprim (TMP) and pyrimethamine (PMT) are weak inhibitors of DHFR from *P. carinii* and *T. gondii* [5]. An association with sulfonamides is necessary to increase their potency. However, the use of sulfa drugs induces an increase in the frequency of adverse side effects. In addition, several active site mutations of parasitic DHFR have rendering infections refractive to known DHFR inhibitors.

Trimetrexate (TMX), and piritrexim (PTX), which are 100–10000 times more potent than TMP and PMT against DHFR from *P. carinii* and *T. gondii*, are also potent inhibitors of mammalian DHFR; hence, their use is associated with significant toxicity [6]. As a result, TMX is coadministered with leucovorin, a classical folate. The clinical use of TMX and PTX is therefore limited because of their systemic host toxicity and requirement of expensive cotherapy with the rescue agent leucovorin.

Thus, there is a dire need for new single agents with high potency and selectivity against these organisms. Several research groups have established programs aimed at discovering potent and pathogen selective DHFR inhibitors as agents against AIDS associated opportunistic infections. The overall goal is to improve therapy and to minimize the adverse side effects.

This work is an application of CoMFA/CoMSIA 3D-QSAR modeling to a set of pyrimidine inhibitors of *P. carinii* DHFR (*Pc*DHFR).

## Computational methods

### Molecular data set and biological activity

Table 1 shows the data set of 64 pyrimidine derivatives, including some known drugs, assembled from the literature [4, 6–11]. All compounds have two aromatic systems, a heteroaromatic quinazoline, pyrido[2,3-*d*]pyrimidine, or pteridine ring and a phenyl or naphthyl ring, linked by a spacer group ($X_1$-$X_2$) (Table 1). The corresponding biological activity measures are expressed as the half maximal (50%) inhibitory concentration (IC) against *Pc*DHFR ($IC_{50}$). However, in order to construct the 3D-QSAR models, the $IC_{50}$ values were expressed in negative logarithmic units ($pIC_{50}$, M).

The molecular sampling procedure of this data set from the literature was based on three prerequisites. The first one is the structural similarity (isosteric pharmacophoric groups) between the data set, which makes the alignment process more reliable [12]. These similarities can be noted in Table 1.

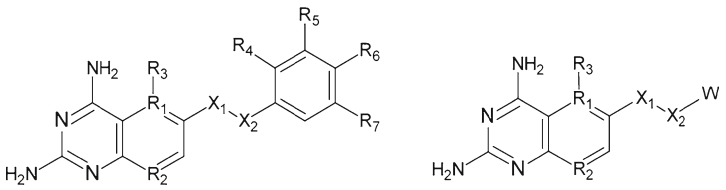
The second prerequisite referred to the pharmacological protocols. Every compound of the data set used in a QSAR study should be tested according to the same pharmacological protocol, which increases robustness and makes the results uniform [13]. A spectrophotometric assay was used for the biological activity determination. It contained sodium phosphate buffer pH 7.4 (40.7 mM), 2-mercaptoethanol (8.9 mM), NADPH (0.117 mM), 1 to 3.7 IU of enzyme (1 IU=0.005 molar absorptivity units per min), and dihydrofolic acid (0.092 mM). The first four reagents were combined in a disposable cuvette and were brought to 37 C. Dilutions of drug were added at this stage. The enzyme was added 30 seconds before the reaction was initiated with dihydrofolic acid. The reaction was followed for 5 min with continuous recording of molar absorptivity at 340 nm [14].

The third prerequisite concerns the distribution of the biological activity values (in logarithmic scale) of the compounds along the whole data set. Thus, the $pIC_{50}$ values must correspond to compounds of low, middle and high potency [13]. The graphical distribution of the experimental values for the 64 selected molecules for the 3D-QSAR study, both training and test sets, is shown in Fig. S1 (Electronic supplementary material).

### Training and test sets selection and models refinement

The ultimate models achieved in this study underwent a process of refinement carried out in two steps. In the first step, all 64 compounds (51 as training set) were used to build the CoMFA/CoMSIA [15, 16] non-refined models. In order to improve robustness and predictability of the final QSAR models, ten detected outliers were excluded from the original data set. Outliers are compounds that the residual of fit (difference between experimental and predicted $pIC_{50}$ values) exceed twice the standard deviations from the mean of the model. Thus, the training set encompassed 51 compounds in the non-refined models and 41 in the refined models (Table 1). The test set encompassed 13 compounds exhibiting good structural variation along the data set, which were conserved in both sets of 3D-QSAR models (non-refined and refined).

**Table 1** The data set: pteridine derivatives




| # | $R_1$ | $R_2$ | $R_3$ | $X_1$ | $X_2$ | $R_4$ | $R_5$ | $R_6$ | $R_7$ | W | $pIC_{50}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| M1[b] | C | N | H | $CH_2$ | NH | Cl | H | H | Cl | - | 4.96 |
| M2 | N | N | - | $CH_2$ | S | H | H | H | H | - | 5.02 |
| M3 | C | N | H | $CH_2$ | S | Cl | H | H | Cl | - | 5.23 |
| M4 [b] | C | N | H | $CH_2$ | S | H | Cl | H | Cl | - | 4.96 |
| M5 | C | N | H | $CH_2$ | S | H | $OCH_3$ | $OCH_3$ | H | - | 5.66 |
| M6 | N | CH | - | $CH_2$ | S | H | H | H | H | - | 5.70 |
| M7 | N | CH | - | $CH_2$ | S | - | - | - | - | 1-naphthyl | 6.33 |
| M8 | N | CH | - | $CH_2$ | S | - | - | - | - | 2-naphthyl | 6.42 |
| M9[a] | C | N | H | $CH_2$ | S | - | - | - | - | 1-naphthyl | 6.02 |
| M10[b] | C | N | H | $CH_2$ | S | - | - | - | - | 2-naphthyl | 6.77 |
| M11 | C | N | H | $CH_2$ | NH | - | - | - | - | 1-naphthyl | 7.15 |
| M12[b] | C | N | H | $CH_2$ | NH | - | - | - | - | 2-naphthyl | 5.70 |
| M13[a] | C | CH | H | $NCH_3$ | $CH_2$ | - | - | - | - | 1-naphthyl | 7.77 |
| M14 | C | CH | H | $N(C_2H_5)$ | $CH_2$ | - | - | - | - | 2-naphthyl | 8.14 |
| M15 | C | CH | H | NH | $CH_2$ | - | - | - | - | 1-naphthyl | 6.14 |
| M16 | C | N | H | NH | $CH_2$ | $OCH_3$ | H | H | $OCH_3$ | - | 5.42 |
| M17 | C | N | H | $NCH_3$ | $CH_2$ | $OCH_3$ | H | H | $OCH_3$ | - | 7.08 |
| M18[a] | C | CH | H | NH | $CH_2$ | $OCH_3$ | H | H | $OCH_3$ | - | 5.34 |
| M19 | C | CH | H | $N(C_2H_5)$ | $CH_2$ | H | H | H | H | - | 7.68 |
| M20 | C | CH | H | $N(C_2H_5)$ | $CH_2$ | $OCH_3$ | H | H | $OCH_3$ | - | 8.00 |
| M21 | C | CH | H | $N(C_3H_7)$ | $CH_2$ | $OCH_3$ | H | H | $OCH_3$ | - | 7.42 |
| M22 | C | N | $CH_3$ | $CH_2$ | NH | H | $OCH_3$ | $OCH_3$ | $OCH_3$ | - | 7.07 |
| M23[b] | C | N | $CH_3$ | $CH_2$ | $N(CH_3)$ | H | $OCH_3$ | $OCH_3$ | $OCH_3$ | - | 7.89 |
| M24 | C | N | H | $CH_2$ | NH | H | $OCH_3$ | $OCH_3$ | $OCH_3$ | - | 5.82 |
| M25[b] | C | N | H | $CH_2$ | $N(CH_3)$ | H | $OCH_3$ | $OCH_3$ | $OCH_3$ | - | 6.62 |
| M26 | C | N | H | $CH_2$ | NH | H | H | H | H | - | 5.66 |

**Table 1** (continued)

| # | $R_1$ | $R_2$ | $R_3$ | $X_1$ | $X_2$ | $R_4$ | $R_5$ | $R_6$ | $R_7$ | W | $pIC_{50}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| M27 | C | N | H | $CH_2$ | NH | H | $OCH_3$ | H | H | - | 6.28 |
| M28 | C | N | H | $CH_2$ | NH | $OCH_3$ | H | H | $OCH_3$ | - | 5.21 |
| M29 | C | N | H | $CH_2$ | NH | H | $OCH_3$ | $OCH_3$ | H | - | 6.37 |
| M30 | C | N | H | $CH_2$ | NH | H | Cl | Cl | H | - | 6.60 |
| M31 | C | N | H | $CH_2$ | NH | H | Cl | Cl | Cl | - | 6.46 |
| M32[a] | C | N | H | $CH_2$ | $N(CH_3)$ | H | Cl | Cl | H | - | 6.88 |
| M33[b] | C | N | H | $CH_2$ | $N(CH_3)$ | H | Cl | Cl | Cl | - | 6.71 |
| M34[a] | C | CH | H | NH | $CH_2$ | H | H | H | H | - | 5.06 |
| M35 | C | CH | H | NH | $CH_2$ | H | $OCH_3$ | H | $OCH_3$ | - | 5.66 |
| M36[a] | C | CH | H | NH | $CH_2$ | $OCH_3$ | H | $OCH_3$ | H | - | 5.36 |
| M37 | C | CH | H | NH | $CH_2$ | H | $OCH_3$ | $OCH_3$ | $OCH_3$ | - | 5.17 |
| M38 | C | CH | H | NH | $CH_2$ | $OCH_3$ | $OCH_3$ | $OCH_3$ | H | - | 5.31 |
| M39 | C | CH | H | NH | $CH_2$ | $OCH_3$ | H | $OCH_3$ | $OCH_3$ | - | 5.27 |
| M40[a] | C | CH | H | $NCH_3$ | $CH_2$ | H | $OCH_3$ | H | $OCH_3$ | - | 7.62 |
| M41 | C | CH | H | $NCH_3$ | $CH_2$ | $OCH_3$ | H | $OCH_3$ | H | - | 7.00 |
| M42 | C | CH | H | $NCH_3$ | $CH_2$ | $OCH_3$ | $OCH_3$ | $OCH_3$ | H | - | 7.28 |
| M43 | C | N | $CH_3$ | $CH_2$ | $N(CH_3)$ | H | Cl | Cl | Cl | - | 6.98 |
| M44 | C | N | $CH_3$ | $CH_2$ | NH | H | $OCH_3$ | $OCH_3$ | H | - | 7.36 |
| M45[b] | C | N | $CH_3$ | $CH_2$ | $N(CH_3)$ | H | $OCH_3$ | $OCH_3$ | H | - | 6.49 |
| M46[b] | C | N | $CH_3$ | $CH_2$ | $N(CH_3)$ | $OCH_3$ | H | H | $OCH_3$ | - | 6.67 |
| M47[b] | C | N | $CH_3$ | $CH_2$ | $N(CH_3)$ | H | Cl | Cl | H | - | 7.00 |
| M48[a] | C | N | $CH_3$ | $CH_2$ | NH | H | Cl | Cl | H | - | 6.49 |
| M49 | C | N | $CH_3$ | $CH_2$ | NH | H | H | H | H | - | 7.10 |
| M50 | C | N | $CH_3$ | $CH_2$ | NH | $OCH_3$ | H | H | H | - | 6.93 |
| M51[a] | C | N | $CH_3$ | $CH_2$ | NH | H | $OCH_3$ | H | H | - | 7.16 |
| M52[a] | C | N | $CH_3$ | $CH_2$ | NH | H | H | $OCH_3$ | H | - | 7.02 |
| M53 | C | N | $CH_3$ | $CH_2$ | $N(CH_3)$ | Cl | H | H | H | - | 7.08 |
| M54 | C | N | $CH_3$ | $CH_2$ | $N(CH_3)$ | H | H | Cl | H | - | 7.54 |
| M55 | C | N | $CH_3$ | $CH_2$ | $N(CH_3)$ | H | H | Br | H | - | 7.43 |
| M56[a] | C | N | $CH_3$ | $CH_2$ | $N(CH_3)$ | H | H | $OCH_3$ | H | - | 7.46 |
| M57 | C | N | $CH_3$ | $CH_2$ | $N(CH_3)$ | H | $OCH_3$ | H | H | - | 7.52 |
| M58[a] | C | N | $CH_3$ | $CH_2$ | NH | H | H | Br | H | - | 7.09 |

**Table 1** (continued)

| # | $R_1$ | $R_2$ | $R_3$ | $X_1$ | $X_2$ | $R_4$ | $R_5$ | $R_6$ | $R_7$ | W | $pIC_{50}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| M59 | C | N | $CH_3$ | $CH_2$ | NH | H | H | Cl | H | - | 7.26 |
| M60 | C | N | $CH_3$ | $CH_2$ | NH | H | Cl | H | H | - | 7.64 |
| M61 | C | N | $CH_3$ | $CH_2$ | NH | Cl | H | H | H | - | 7.33 |
| M62 | C | CH | H | $NCH_3$ | $CH_2$ | $OCH_3$ | H | H | $OCH_3$ | - | 7.06 |
| M63 | C | N | $CH_3$ | $CH_2$ | NH | $OCH_3$ | H | H | $OCH_3$ | - | 7.47 |
| M64[a] | C | CH | $CH_3$ | $CH_2$ | NH | H | $OCH_3$ | $OCH_3$ | $OCH_3$ | - | 7.38 |

[a] Test set molecules.
[b] Molecules excluded in refined models construction.

## General procedures

Compound M62 (Table 1) has its 3D-structure solved in the RCSB Protein Data Bank (PDB) [17] in complex with *Pc*DHFR (PDB ID: 1LY3) [18], and it was used as a molecular template in the assembling of the pyrimidine derivatives. The "grow hydrogen atoms" tool of the Spartan'06 software [19] was used to add the missing hydrogen atoms of M62, which often are not detected in the X-ray diffraction experiments used to solve ligand-protein complexes. All structures were built in the Spartan'06 program [19], aligned along the *Pc*DHFR bound conformation of M62, assuming that it represents the most probable "bioactive" conformation of all pyrimidine derivatives at the *Pc*DHFR active site and that all analogs have the same binding mode [13].

Visual molecular dynamics (VMD) [20] was used to generate the molecular electrostatic potential (MEP) map of *Pc*DHFR (Fig. S2) and to perform the superposition of active site residues of human DHFR (hDHFR; PDB ID: 3NZ9) [21] and *Pc*DHFR (PDB ID: 1LY3) [18].

Both CoMFA [15] and CoMSIA [16] 3D-QSAR analyses were carried out as implemented on the SYBYL v.8.0 package [22]. The Gasteiger-Hückel charges [23] were assigned to all molecules.

## Alignment rules

The molecular alignment is the crucial step for any successful application of CoMFA/CoMSIA approaches [15, 16]. The common molecular moiety should be in the same conformation in all molecules, and the other molecular groups should be superimposed as closely as possible [12]. Again, since compound M62 (Fig. 1a) has its 3D-structure solved in the RCSB Protein Data Bank (PDB) [17] in complex with *Pc*DHFR (PDB ID: 1LY3) [18], it was used as the molecular template for the alignment of all compounds, assuming its conformation as the bioactive one (Fig. 1b).

In this study, a two-atom based alignment was carried out using the "fit atoms" tool available in the SYBYL v.8.0 software. Atoms 1, 3, 4', 6, 9, and 10 (Fig. 1c) of M62 were selected for alignment of all structures. The root-mean-square (RMS) value obtained for this particular selection was equal to zero Å since all atoms selected belong to the common substructure.

## CoMFA studies

CoMFA studies were performed with SYBYL v.8.0 molecular modeling package. Both steric and electrostatic CoMFA fields were sampled at each point of regularly spaced grids of 2.0 Å, and were calculated using Lennard-Jones and Coulomb potentials, respectively. The interaction energies were calculated for a $sp^3$ carbon probe atom with a charge of +1 [15]. The cutoff value for both fields was set to 30 $Kcal.mol^{-1}$.

## CoMSIA studies

The same alignment generated for CoMFA (Fig. 1) was used in the CoMSIA [16] analysis. Steric, electrostatic, hydrophobic, and steric/electrostatic CoMSIA fields were built using a Gaussian distance dependent function. Thus, the CoMFA/CoMSIA studies could be compared for similarities and differences, with the aim of selecting the most predictive and robust model. The default value of 0.3 was used as the attenuation factor.

## CoMFA/CoMSIA models statistical analysis

In this 3D-QSAR modeling, CoMFA/CoMSIA [15, 16] descriptors were used as the explanatory (independent) variables, and $pIC_{50}$ values were used as the target (dependent) variable. Moreover, partial least square (PLS) regression analysis was the statistical tool used for constructing the models.

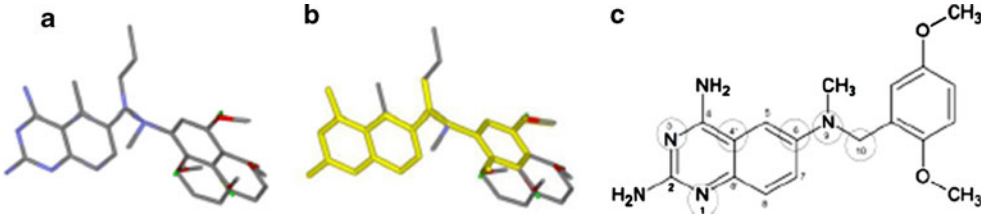



**Fig. 1** **a** 64 aligned structures superimposed; **b** in yellow the template structure M62; **c** atoms from M62 used in the two-atoms based alignment

The optimum number of components (C) was determined by leave-one-out (LOO) cross-validation [24] by using a maximum of 15 principal components in each model.

The statistical $q^2$ parameter (LOO-cross-validated squared correlation coefficient) was used to evaluate robustness of the CoMFA/CoMSIA models. These models are considered robust when $q^2$ is greater than 0.5 and the standard error of prediction (SEP) is lower than 0.5 [13]. However, the external validation (test set), by predicting the potency of compounds not included in the training set, was the main criterion for inferring the predictive strength of each constructed 3D-QSAR model.

## Results and discussion

### CoMFA/CoMSIA models statistical analysis

Five CoMFA/CoMSIA 3D-QSAR models were constructed for each training set (51 and 41 compounds) of pyrimidine inhibitors of *Pc*DHFR. Both the CoMFA/CoMSIA non-refined (N=51) and refined (N=41) models have statistical parameters shown in Table 2. Forty-one training set compounds were used for deriving the optimized models, and 13 compounds were used as the test set for external validation purpose. All further analysis will be based on these final refined models.

The statistical significance of the constructed CoMFA/CoMSIA models is shown in Tables 2 and 3, and Fig. 2. The biological predictions of the training set molecules (internal validation) of the CoMFA/CoMSIA models are shown in Table 3. The majority of molecules allowed good QSAR predictions. Training set molecules were considered outliers when residual values exceed twice the standard deviation of the residuals. Thus, according to the CoMFA/CoMSIA steric/electrostatic models, M48 was the only outlier. In the CoMSIA hydrophobic model, four molecules were outliers (M9, M28, M40, and M48).

The statistical significance of the CoMFA models is shown in Tables 2 and 3, and Fig. 2a, which shows high

**Table 2** Statistical parameters of the non-refined (51 training set compounds) and refined (41 training set compounds) CoMFA/CoMSIA models

| | $q^{2b}$ | C[c] | SEP[d] | $r^{2e}$ | SEE[f] | F[g] |
|---|---|---|---|---|---|---|
| Non-refined models (N=51)[a] | | | | | | |
| CoMFA | 0.67 | 2 | 0.50 | 0.78 | 0.42 | 83.68 |
| CoMSIA steric | 0.66 | 6 | 0.53 | 0.81 | 0.40 | 31.67 |
| CoMSIA electrostatic | 0.64 | 10 | 0.58 | 0.85 | 0.37 | 23.16 |
| CoMSIA steric/electrostatic | 0.63 | 9 | 0.58 | 0.86 | 0.36 | 27.92 |
| CoMSIA hydrophobic | 0.60 | 13 | 0.64 | 0.86 | 0.37 | 17.78 |
| Refined models (N=41)[a] | | | | | | |
| CoMFA | 0.83 | 7 | 0.40 | 0.95 | 0.21 | 98.83 |
| CoMSIA steric | 0.80 | 5 | 0.43 | 0.89 | 0.31 | 57.55 |
| CoMSIA electrostatic | 0.83 | 10 | 0.43 | 0.95 | 0.22 | 60.89 |
| CoMSIA steric/electrostatic | 0.82 | 10 | 0.43 | 0.96 | 0.21 | 70.50 |
| CoMSIA hydrophobic | 0.72 | 6 | 0.51 | 0.88 | 0.33 | 42.01 |

[a] Number (N) of training set compounds

[b] Cross-validated squared correlation coefficient ($q^2$)

[c] Optimal number of components (C) used by the model

[d] Standard error of prediction (SEP)

[e] Squared correlation coefficient ($r^2$)

[f] Standard error of estimation (SEE)

[g] F-Test (F)

**Table 3** Experimental and predicted $IC_{50}$ (M) and the corresponding residual ($pIC_{50Exp}$ - $pIC_{50Pred}$) values for the data set molecules. The predictions were based on the CoMFA, CoMSIA steric/electrostatic, and CoMSIA hydrophobic refined models

| # | $pIC_{50Exp}$ | CoMFA $pIC_{50Pred}$ | CoMFA Residual | CoMSIA $_{STE/ELEC}$ $pIC_{50Pred}$ | CoMSIA $_{STE/ELEC}$ Residual | CoMSIA $_{HYDRO}$ $pIC_{50Pred}$ | CoMSIA $_{HYDRO}$ Residual |
|---|---|---|---|---|---|---|---|
| M2 | 5.02 | 5.42 | −0.40 | 5.23 | −0.21 | 5.13 | −0.10 |
| M3 | 5.23 | 5.12 | 0.11 | 5.15 | 0.08 | 5.39 | −0.16 |
| M5 | 5.66 | 5.63 | 0.02 | 5.50 | 0.15 | 6.25 | −0.59 |
| M6 | 5.70 | 5.46 | 0.24 | 5.54 | 0.16 | 5.57 | 0.12 |
| M7 | 6.33 | 6.38 | −0.05 | 6.53 | −0.20 | 6.33 | 0.00 |
| M8 | 6.42 | 6.34 | 0.09 | 6.22 | 0.20 | 6.51 | −0.09 |
| M9[a] | 6.02 | 6.45 | −0.43 | 6.52 | −0.50 | 7.09 | −1.07 [b] |
| M11 | 7.15 | 6.82 | 0.33 | 6.98 | 0.18 | 6.82 | 0.34 |
| M13[a] | 7.77 | 7.90 | −0.13 | 8.13 | −0.36 | 7.67 | 0.10 |
| M14 | 8.14 | 8.47 | −0.32 | 8.27 | −0.13 | 8.40 | −0.25 |
| M15 | 6.14 | 6.21 | −0.06 | 6.07 | 0.07 | 6.33 | −0.18 |
| M16 | 5.42 | 5.43 | 0.00 | 5.24 | 0.18 | 5.12 | 0.30 |
| M17 | 7.08 | 7.16 | −0.08 | 7.00 | 0.07 | 6.47 | 0.60 |
| M18[a] | 5.34 | 5.43 | −0.09 | 5.55 | −0.21 | 5.57 | −0.23 |
| M19 | 7.68 | 7.60 | 0.08 | 7.59 | 0.09 | 7.46 | 0.22 |
| M20 | 8.00 | 7.69 | 0.32 | 7.69 | 0.31 | 7.46 | 0.54 |
| M21 | 7.42 | 7.49 | −0.06 | 7.51 | −0.09 | 7.88 | −0.46 |
| M22 | 7.07 | 7.27 | −0.21 | 7.19 | −0.12 | 7.02 | 0.05 |
| M24 | 5.82 | 5.98 | −0.15 | 5.97 | −0.15 | 5.82 | 0.00 |
| M26 | 5.66 | 5.93 | −0.28 | 5.99 | −0.33 | 6.06 | −0.40 |
| M27 | 6.28 | 6.23 | 0.05 | 6.37 | −0.08 | 6.15 | 0.13 |
| M28 | 5.21 | 5.25 | −0.04 | 5.24 | −0.02 | 6.04 | −0.83 [b] |
| M29 | 6.37 | 6.05 | 0.32 | 6.10 | 0.27 | 5.98 | 0.39 |
| M30 | 6.60 | 6.55 | 0.05 | 6.70 | −0.10 | 6.57 | 0.04 |
| M31 | 6.46 | 6.17 | 0.28 | 6.19 | 0.27 | 5.94 | 0.52 |
| M32[a] | 6.88 | 6.38 | 0.51 | 6.71 | 0.17 | 6.65 | 0.23 |
| M34[a] | 5.06 | 5.35 | −0.29 | 5.44 | −0.38 | 5.56 | −0.50 |
| M35 | 5.66 | 5.62 | 0.04 | 5.62 | 0.04 | 5.50 | 0.16 |
| M36[a] | 5.36 | 5.30 | 0.06 | 5.20 | 0.16 | 5.55 | −0.20 |
| M37 | 5.17 | 5.39 | −0.22 | 5.36 | −0.19 | 5.34 | −0.17 |
| M38 | 5.31 | 5.41 | −0.10 | 5.55 | −0.24 | 5.65 | −0.34 |
| M39 | 5.27 | 5.27 | 0.00 | 5.27 | 0.00 | 5.39 | −0.12 |
| M40[a] | 7.62 | 7.35 | 0.27 | 7.38 | 0.24 | 6.85 | 0.77 [b] |
| M41 | 7.00 | 7.02 | −0.02 | 6.96 | 0.04 | 6.91 | 0.09 |
| M42 | 7.28 | 7.06 | 0.23 | 7.24 | 0.04 | 7.00 | 0.28 |
| M43 | 6.98 | 7.28 | −0.29 | 7.38 | −0.39 | 7.23 | −0.24 |
| M44 | 7.36 | 7.32 | 0.04 | 7.18 | 0.18 | 7.18 | 0.18 |
| M48[a] | 6.49 | 7.83 | −1.33 [b] | 7.78 | −1.29 [b] | 7.77 | −1.27 [b] |
| M49 | 7.10 | 7.23 | −0.14 | 7.20 | −0.11 | 7.26 | −0.16 |
| M50 | 6.93 | 7.28 | −0.35 | 7.31 | −0.38 | 7.42 | −0.49 |
| M51[a] | 7.16 | 7.51 | −0.35 | 7.45 | −0.29 | 7.35 | −0.19 |
| M52[a] | 7.02 | 7.12 | −0.10 | 7.05 | −0.03 | 7.08 | −0.06 |
| M53 | 7.08 | 7.09 | −0.02 | 7.17 | −0.10 | 7.03 | 0.04 |
| M54 | 7.54 | 7.29 | 0.25 | 7.32 | 0.22 | 7.48 | 0.06 |
| M55 | 7.43 | 7.34 | 0.09 | 7.29 | 0.14 | 7.58 | −0.15 |
| M56[a] | 7.46 | 6.99 | 0.47 | 7.05 | 0.40 | 7.17 | 0.29 |
| M57 | 7.52 | 7.34 | 0.18 | 7.42 | 0.10 | 7.44 | 0.08 |

**Table 3** (continued)

| # | $pIC_{50Exp}$ | CoMFA $pIC_{50Pred}$ | CoMFA Residual | CoMSIA $_{STE/ELEC}$ $pIC_{50Pred}$ | CoMSIA $_{STE/ELEC}$ Residual | CoMSIA $_{HYDRO}$ $pIC_{50Pred}$ | CoMSIA $_{HYDRO}$ Residual |
|---|---|---|---|---|---|---|---|
| M58[a] | 7.09 | 7.52 | −0.42 | 7.31 | −0.21 | 7.48 | −0.38 |
| M59 | 7.26 | 7.47 | −0.21 | 7.35 | −0.09 | 7.39 | −0.13 |
| M60 | 7.64 | 7.66 | −0.02 | 7.64 | 0.00 | 7.64 | 0.00 |
| M61 | 7.33 | 7.18 | 0.15 | 7.20 | 0.13 | 6.95 | 0.38 |
| M62 | 7.06 | 7.08 | −0.02 | 7.24 | −0.18 | 6.92 | 0.14 |
| M64 | 7.47 | 7.28 | 0.19 | 7.30 | 0.17 | 7.27 | 0.20 |
| M64[a] | 7.38 | 7.25 | 0.13 | 7.37 | 0.01 | 7.47 | −0.09 |

[a] Test set molecules

[b] Outliers.

correlation between the experimental and predicted $pIC_{50}$ values for all molecules, but M48, which has an overestimated predicted value. The squared linear correlation coefficient ($R^2$) for the CoMFA model is 0.89 (Fig. 2a) and the $q^2$ value is equal to 0.83 (Table 2).

The statistical significance of the CoMSIA models is shown in Tables 2 and 3, and Fig. 2b-c. In accordance with CoMFA, M48 was outlier in all other models. However, the $CoMSIA_{STE/ELEC}$ model showed higher $R^2$ (Fig. 2b) value than the corresponding CoMFA (Fig. 2a) model. The high

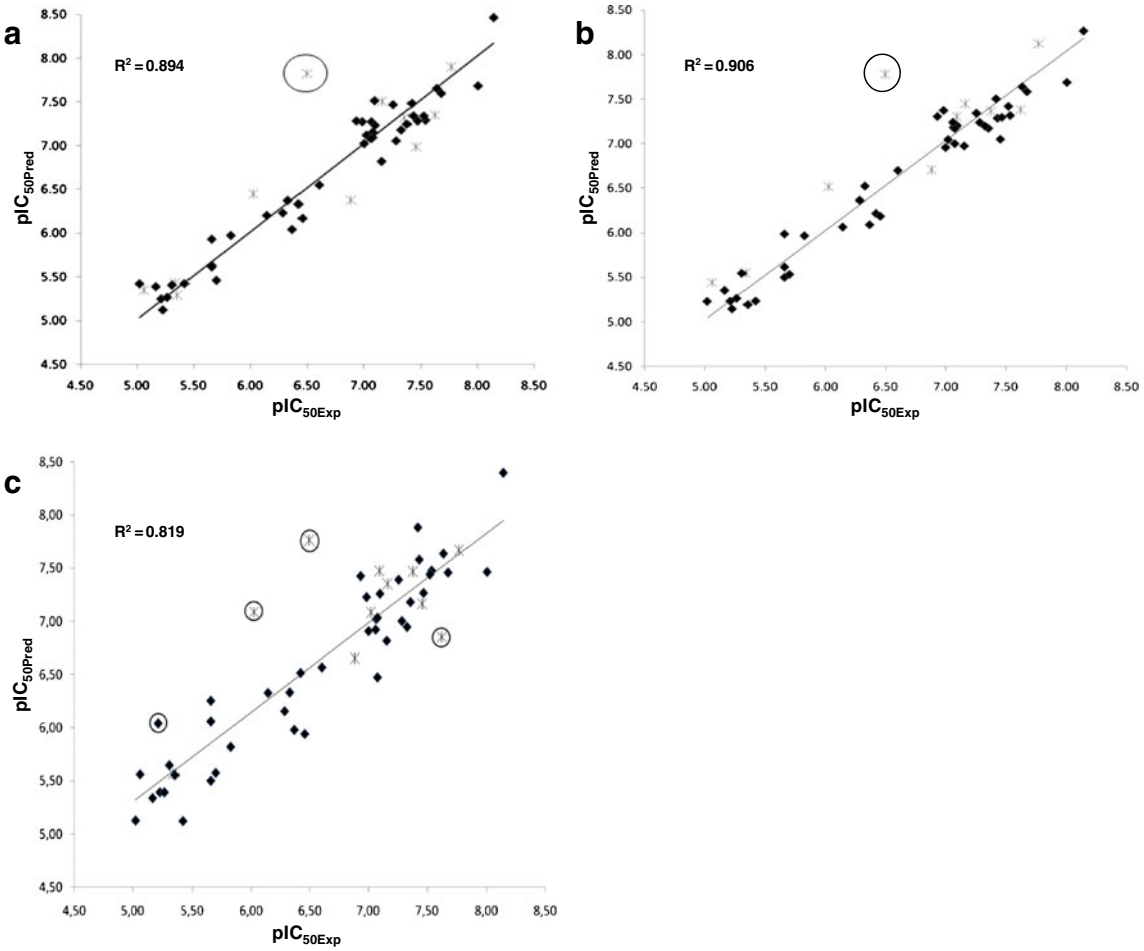



**Fig. 2** Experimental *versus* calculated $pIC_{50}$ values for the training (■) and test (*) set compounds (circled points are outliers) according to: **a** the refined CoMFA model (outlier=M48); **b** the refined CoMSIA steric/electrostatic model (outlier = M48); **c** the refined CoMSIA hydrophobic model (outliers = M9, M28, M40, and M48)

$R^2$ value of the $CoMSIA_{STE/ELEC}$ model (Fig. 2b), as well as its low residual predictions (Table 3), make it a good tool for predicting the potency of *Pc*DHFR inhibitors.

The CoMSIA hydrophobic ($CoMSIA_{HYDRO}$) model has a $R^2$ value equal to 0.82 (Fig. 2c), and $q^2$ value equal to 0.72, confirmed it as a very promising model. However, the standard deviation residual predictions from that model (0.38) are higher than the residuals from CoMFA and $CoMFA_{STE/ELEC}$ (0.29 and 0.27, respectively). Consequently, in the $CoMSIA_{HYDRO}$ model, more molecules were outliers (M9, M40, and M48 test set, and M28 training set), as shown in Fig. 2c.

## CoMFA/CoMSIA maps

Fig. 3a-b shows both CoMFA and CoMSIA contour maps of steric molecular field. One notices the presence of a favorable (green) field on the substituent group located around $X_1$-$X_2$ and $R_3$ positions (Table 1). This is an indication that large substituents on those positions could increase the ligand-enzyme interaction. However, it is important to note that near this favorable field on the $X_1$-$X_2$/$R_3$ positions, there are unfavorable (yellow) fields. In 3D-QSAR studies, when analyzing steric properties, it is always expected that molecular "regions" described as favorable also be able to determine its limits.

Thus, the presence of a favorable steric field on $X_1$-$X_2$/$R_3$ positions surrounded by unfavorable steric fields in both CoMFA (Fig. 3a) and CoMSIA (Fig. 3b) models provide information about the "ideal" volume of the substituent at the $X_1$-$X_2$/$R_3$ positions. The same feature was also noticed in the 3D-QSAR steric field models by Gangjee and Lin [25]. However, probably due to the molecular diversity of the used training set, the statistical validation of their models is lower than that of our models.

In order to figure out which steric volume would be desired for the substituents at the $X_1$-$X_2$/$R_3$ positions, a correlation between the chemical structure and the biological activity was carried out for molecules that show

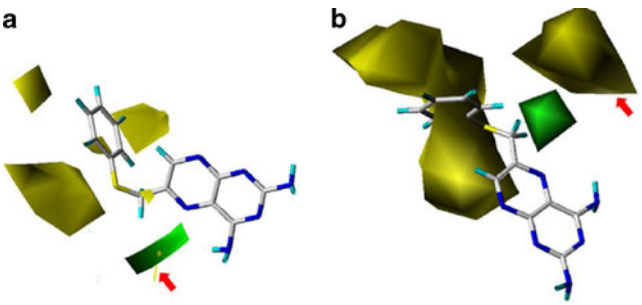



**Fig. 3** The CoMFA (**a**) and CoMSIA (**b**) contour maps of steric molecular field. The red arrow indicates the unfavorable (yellow) field close to the favorable (green) region

structural variation on the $X_1$-$X_2$ spacer group: M16, M17, M18, M20, M21 and M62. A closer look on M16 ($X_1$ = NH; $pIC_{50}$=5.42) and M17 ($X_1$ = $NCH_3$; $pIC_{50}$=7.08) shows that the presence of the methyl group, so often considered chemically, improves the potency. In fact, the presence of methyl groups can have a huge impact on pharmacological properties of a molecule such as selectivity and increased potency [26]. Both M16 and M18 have a secondary amine on $X_1$ position that, when changed by methyl (M62), ethyl (M20), or propyl (M21) groups show an increasing biological potency when compared with the correspondent secondary amines. However, by changing an ethyl (M20) by a propyl (M21) group shows a decrease in potency of 0.58 in logarithmic scale. This suggests that the substituent on $X_1$ position should not exceed an ethyl group.

Some of the data set molecules have a naphthyl group as substituent on W position (M7 to M15). The comparison between M15 ($X_1$ = NH and W=1-naphthyl, $pIC_{50}$=6.14) and M13 ($X_1$ = $NCH_3$ and W=1-naphthyl; $pIC_{50}$=7.77) shows that, as above, the presence of a methyl group improves the biological potency.

The same steric molecular field identified on spacer $X_1$ is present on spacer $X_2$. This finding reinforces the remarks mentioned above regarding the steric-activity relationship on spacer $X_1$. However, on spacer $X_2$ there is no diminishing activity molecular field. It is important to mention that due to the proximity to the $X_1$ spacer, the size of the substituent on $X_2$ spacer is already modulated by its steric field.

In order to hypothesize a drug design strategy based on the constructed QSAR models, structure-activity relationships were analyzed for some molecules. The comparison of M61 to M53 did not provide a useful hypothesis for lead design because the only difference between those molecules is in the $X_2$ position. Despite the smaller $X_2$ substituent of M61, its activity ($pIC_{50}$=7.33) is slightly higher than M53 ($pIC_{50}$=7.08). Since none of those molecules were outliers, probably they have a different binding mode into the active site of the enzyme. It is worthy to note that all molecules that have a sulfur atom as the spacer group $X_2$ (M2 to M10) have low to moderate potency ($4.96 \leq pIC_{50} \leq 6.77$). This activity range can be attributed to the absence of large carbon chains located on $X_1$/$X_2$ spacers.

At a closer look on the training set, it was realized that there is just two possibilities for substituent moieties: carbon or nitrogen on $R_1$, and methyl or hydrogen on $R_3$. Thus, the evaluation of those positions was jointly carried out. Since just four molecules have nitrogen atoms on the $R_1$ position (M2, M6, M7, and M8), it was not possible to propose a structure-activity hypothesis based on that position only.

By comparing the biological activity values of molecules: M24 ($R_3$ = H; $pIC_{50}$=5.82) and M22 ($R_3$ = $CH_3$; $pIC_{50}$=

7.07), M26 ($R_3$ = H; $pIC_{50}$=5.66) and M49 ($R_3$=$CH_3$; $pIC_{50}$=7.10), M25 ($R_3$ = H; $pIC_{50}$=6.62) and M23 ($R_3$ = $CH_3$; $pIC_{50}$=7.89), one noticed that the methyl group on the $R_3$ position seems to contribute for the increasing of biological value. This finding is in accordance with the work by Gangjee and Lin [25], which shows a methyl group in $R_3$ position located closer to a sterically unfavorable yellow region.

Our hypothesis is that this unfavorable region (pointed by the red arrows in Fig. 3) is probably related to the NADPH co-factor, since in the 1LY3 X-ray structure, the NADPH nicotinamide moiety is bound close to $X_1$/$R_3$ positions of M62, and delimited by residues Ile10, Val11, Ala12, Leu13, Ile19, Leu25, Trp27, Thr61, Ile123, Gly124, and Tyr129. Therefore, there is a need of available space for the binding of the nicotinamide ring of NADPH that could not be occupied by large $X_1$/$R_3$ substituents of the ligand (e.g., propyl).

The $R_1$ and $R_2$ positions are located on the heteroaromatic system, a well-known pharmacophoric group of DHFR inhibitors. In order to ascertain any change caused by substitution on that position, we compared M2 and M6, as well as M22 and M64 to each other. $R_2$ substituents for M2 and M6 are, N and CH, and the corresponding $pIC_{50}$ values are 5.02 and 5.70, respectively. The $R_2$ substituents for M22 and M66 are also N and CH, and the corresponding $pIC_{50}$ values are 7.07 and 7.38, respectively, so it is clear that when an N atom is changed to a CH group the biological activity increases. An investigation of the effect of molecular substituents on the heteroaromatic system cannot be satisfactory without considering the substituent effect on spacers $X_1$, $X_2$, and $R_3$, as well. It was noticed that ethylamine at $X_1$, methylamine at $X_2$, and methyl at $R_3$ also increases the biological activity, in comparison to smaller fragments.

The electrostatic field contour map analysis shows that substituents of low electronic density on the $X_1$ and $X_2$ positions may increase the inhibitory activity (Fig. 4a) as will be discussed latter.

Three maps were constructed: one (steric/electrostatic) CoMFA map and two (electrostatic and steric/electrostatic) CoMSIA maps. One notices that the CoMFA model shows good statistics. Moreover, it explains more adequately the structure-activity relationship (Fig. 4a). Thus, our discussion will limit itself to the QSAR model only.

Gangjee and Lin's CoMFA model [25] shows an electrostatic field on the $R_3$ position, which is not present in our model. That specific field is responsible for increasing the activity, when the methyl group is linked to an aromatic carbon atom. It is noteworthy to mention that this attribute is actually related to the atom linked to the methyl group, instead of to the methyl group itself. In other words, compounds with a carbon atom at $R_1$ are more potent than analogs with a nitrogen atom on that position.

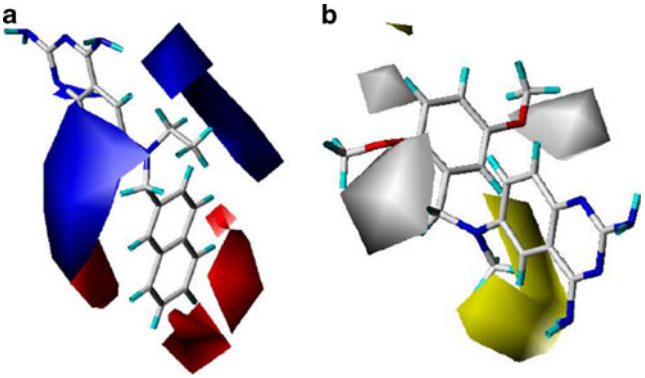



**Fig. 4** **a** The CoMFA electrostatic field contour map showing the most activity molecule (M14) in the data set: negative and positive charge favoring areas are represented by red and blue contours, respectively. **b** CoMSIA hydrophobic map showing the template structure M62: favorable and unfavorable hydrophobic areas are represented by yellow and gray contours, respectively

There are also electrostatic fields around $R_4$-$R_7$ positions, all of them showing favorable contributions on the activity. It can be confirmed by analyzing analogs M50 and M61; M54 and both M55 and M56; and M60 and M51.

The substituent, on $R_4$-$R_7$ positions, with the highest electronic density in our data set is chlorine. Among the high potent analogs containing chlorine on their heteroaromatic systems are the monosubstituted compounds. Any additional chlorine substituent decreases the activity regardless of its position. Monosubstituted analogs on $R_5$ (M60) and $R_6$ (M54) characterize that position as being responsible for the increased activity, even in the absence of hindering groups on the spacer region, as it is the case of M60. A possible explanation of such electronic feature is due to residues present on that active site position, which form a region around $R_4$-$R_7$ position where groups with high electronic density tend to interact through electrostatic interactions (Fig. S2).

The understanding of both steric and electrostatic properties is fundamental to predict the biological activity of a molecule. Although a third parameter, the substituent hydrophobicity [27], is considered by Hansch as important for the structure-activity relationship, this parameter can be represented, partially, in the 3D-QSAR approach by hydrophobic contour maps.

The hydrophobic contour map was constructed through CoMSIA (Fig. 4b). The regions where groups or atoms with hydrophobic features contribute to the activity are in yellow. The gray color represents unfavorable regions. Hydrophobic contour maps are exclusively located around two ligand positions: the spacer group ($X_1$-$X_2$), and the substituent of the heteroaromatic system ($R_4$-$R_7$).

Around $X_1$-$X_2$ there is a yellow contour map, justified by previous discussion. This result is probably due to hydrophobic

interactions of these substituents with hydrophobic enzyme residue Ile123, which is located on the active site (Fig. 5).

The second region considered is located around the heteroaromatic system, on $R_4$-$R_7$ positions. Figure 4b shows gray maps (unfavorable for the presence of hydrophobic groups) surrounding this region, clearly attributable to the chlorine substituent which decreases the activity.

## Comparing the CoMFA/CoMSIA models with the hDHFR and *Pc*DHFR binding sites

When possible, it is always interesting to consider selectivity features in 3D-QSAR modeling. Figure 5 shows the structural superimposition of both human (PDB ID: 3NZ9) [21] and *P. carinii* (PDB ID: 1LY3) [18] DHFRs. The similarity between these two enzymes is high (33.33% of identity; 69 identical positions and 72 similar positions). However, some differences are worthy of comment.

Residue Asn64 from the human enzyme (*h*DHFR) is superimposed to Phe69 from *Pc*DHFR (Fig. 5). This confers to the *Pc*DHFR a slightly more hydrophobic "micro-environment" for ligand-enzyme interaction (67% and 58% of the active site residues are hydrophobic, respectively, for the *P. carinii* and human enzymes). However, Phe is more bulky than Asn, increasing steric hindrance. Thus, both hydrophobic and steric effects of the substituent at $R_7$ position should be considered for designing new lead inhibitors for *Pc*DHFR. According to the CoMSIA hydrophobic map (Fig. 4b), the substituent at $R_7$ position should not be highly hydrophobic, and according to the CoMFA steric map (Fig. 3a), the R7 substituent should not be highly bulky.

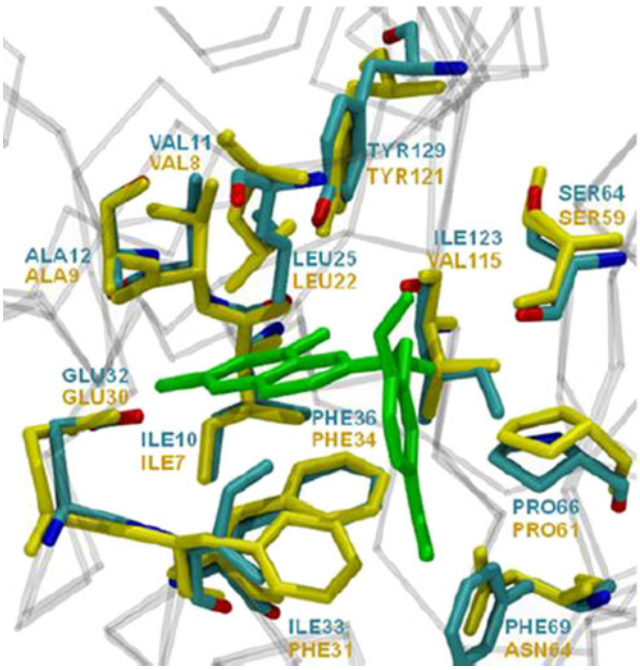



**Fig. 5** Active site residues of *Pc*DHFR (blue) and hDHFR (yellow) around 7 Å of compound M62, which is show in green

Another structural difference between *h*DHFR and *Pc*DHFR regards their active sites steric features. *Pc*DHFR might accommodate larger ligands than *h*DHFR, since its biding site has the Ile33 residue, instead of Phe31 (Fig. 5). Ile33 is not only smaller but also presents more conformational freedom than Phe31. Thus, the larger *Pc*DHFR active site cavity probably has a better chance of "adapting" itself to larger substituents on $R_2$ through induced fit process. Interestingly, we could not correlate any regions of the CoMFA/CoMSIA maps because the data set does not contain compounds with substituents other than N or CH at $R_2$.

Residue Ile123 from *Pc*DHFR is superimposed to Val115 from the human enzyme (*h*DHFR) (Fig. 5). Both residues are similar in terms of hydrophobic and steric properties, and as discussed previously, this residue (Ile123) could be correlated with regions of increasing potency in the CoMSIA hydrophobic (Fig. 4b) and CoMFA/CoMSIA steric (Fig. 3) maps that are correlated to $X_1$-$X_2$/$R_3$ positions.

## Conclusions

A series of 64 pyrimidine derivatives (51 compounds in the training set and 13 in the test set) with *Pc*DHFR inhibitory activity was subjected to a 3D-QSAR study. Five 3D-QSAR models were built, using two different QSAR methods, CoMFA and CoMSIA. All models have shown good predictabilities and statistical validation.

The resulting steric/electrostatic and hydrophobic contour maps have provided useful insights in active-structure relationship, allowing a discussion in terms of drug design. The methyl group has shown a relevant role on the activity, by increasing it. Substituents such as ethylamine on position $X_1$, methylamine on position $X_2$, and methyl on position $R_3$ enhance the inhibitory activity, in comparison to other fragments.

The electrostatic map of the CoMFA model presented the best statistical validation and, thus, it was used to explain some of the structure-activity relationships. Substituents with high electronic density in $R_4$-$R_7$ positions have shown putative favorable interactions. Thus, compounds monosubstituted with a chlorine atom in this region have shown values of $pIC_{50}$ all above 7.0. This effect is supported by electrostatic potential map of *Pc*DHFR active site, as depicted in Electronic supplementary material.

Hydrophobic contour map around the spacer group ($X_1$-$X_2$) has revealed that hydrophobic interactions enhance activity, justifying the previous discussion about methyl group. Overall, the results of the CoMFA/CoMSIA contour maps analyses may serve as theoretical support for the design of new putative inhibitors for *Pc*DHFR with enhanced potency and selectivity.

**Acknowledgments** This study was supported by the following Brazilian agencies: "Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro" (FAPERJ), "Conselho Nacional de Desenvolvimento Científico e Tecnológico" (CNPq), and "Coordenação de Aperfeiçoamento de Pessoal de Nível Superior" (CAPES)

## References

1. Gonzalez-Scarano F, Martin-Garcia J (2005) The neuropathogenesis of AIDS. Nat Rev Immunol 5(1):69–81. doi:nri1527,10.1038/nri1527
2. Huang L, Crothers K, Atzori C, Benfield T, Miller R, Rabodonirina M, Helweg-Larsen J (2004) Dihydropteroate synthase gene mutations in Pneumocystis and sulfa resistance. Emerg Infect Dis 10(10):1721–1728
3. Rodriguez M, Fishman JA (2004) Prevention of infection due to Pneumocystis spp. in human immunodeficiency virus-negative immunocompromised patients. Clin Microbiol Rev 17(4):770–782. doi:17/4/770,10.1128/CMR.17.4.770-782.2004
4. Gangjee A, Kurup S, Namjoshi O (2007) Dihydrofolate reductase as a target for chemotherapy in parasites. Curr Pharm Des 13(6):609–639
5. Gangjee A, Guo X, Queener SF, Cody V, Galitsky N, Luft JR, Pangborn W (1998) Selective Pneumocystis carinii dihydrofolate reductase inhibitors: design, synthesis, and biological evaluation of new 2,4-diamino-5-substituted-furo[2,3-d]pyrimidines. J Med Chem 41(8):1263–1271. doi:10.1021/jm970537w,jm970537w
6. Graffner-Nordberg M, Kolmodin K, Aqvist J, Queener SF, Hallberg A (2001) Design, synthesis, computational prediction, and biological evaluation of ester soft drugs as inhibitors of dihydrofolate reductase from Pneumocystis carinii. J Med Chem 44(15):2391–2402. doi:jm010856u
7. Gangjee A, Adair O, Queener SF (2001) Synthesis of 2,4-diamino-6-(thioarylmethyl)pyrido[2,3-d]pyrimidines as dihydrofolate reductase inhibitors. Bioorg Med Chem 9(11):2929–2935. doi:S0968089601002231
8. Gangjee A, Adair OO, Pagley M, Queener SF (2008) N9-substituted 2,4-diaminoquinazolines: synthesis and biological evaluation of lipophilic inhibitors of pneumocystis carinii and toxoplasma gondii dihydrofolate reductase. J Med Chem 51(19):6195–6200. doi:10.1021/jm800694g
9. Gangjee A, Adair OO, Queener SF (2003) Synthesis and biological evaluation of 2,4-diamino-6-(arylaminomethyl)pyrido[2,3-d]pyrimidines as inhibitors of Pneumocystis carinii and Toxoplasma gondii dihydrofolate reductase and as antiopportunistic infection and antitumor agents. J Med Chem 46(23):5074–5082. doi:10.1021/jm030312n
10. Gangjee A, Vidwans AP, Vasudevan A, Queener SF, Kisliuk RL, Cody V, Li RM, Galitsky N, Luft JR, Pangborn S (1998) Structure-based design and synthesis of lipophilic 2,4-diamino-6-substituted quinazolines and their evaluation as inhibitors of dihydrofolate reductases and potential antitumor agents. J Med Chem 41(18):3426–3434. doi:10.1021/jm980081y
11. Gangjee A, Adair O, Queener SF (1999) Pneumocystis carinii and Toxoplasma gondii dihydrofolate reductase inhibitors and antitumor agents: synthesis and biological activities of 2,4-diamino-5-methyl-6-[(monosubstituted anilino)methyl] pyrido[2,3-d]pyrimidines. J Med Chem 42(13):2447–2455. doi:10.1021/jm990079m,jm990079m
12. De Freitas GB, da Silva LL, Romeiro NC, Fraga CA (2009) Development of CoMFA and CoMSIA models of affinity and selectivity for indole ligands of cannabinoid CB1 and CB2 receptors. Eur J Med Chem 44(6):2482–2496. doi:S0223-5234(09)00020-8,10.1016/j.ejmech.2009.01.026
13. Kubinyi H, Folkers G, Martin YC (1998) 3D QSAR in drug design: Recent advances, vol 3. Three-dimensional quantitative structure activitity relationships. Kluwer, Dordrecht
14. Bartlett MS, Shaw M, Navaran P, Smith JW, Queener SF (1995) Evaluation of potent inhibitors of dihydrofolate reductase in a culture model for growth of Pneumocystis carinii. Antimicrob Agents Chemother 39(11):2436–2441
15. Cramer RD, Patterson DE, Bunce JD (1988) Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. J Am Chem Soc 110(18):5959–5967. doi:10.1021/ja00226a005
16. Klebe G, Abraham U, Mietzner T (1994) Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. J Med Chem 37(24):4130–4146
17. Cody V, Pace J, Chisum K, Rosowsky A (2006) New insights into DHFR interactions: analysis of Pneumocystis carinii and mouse DHFR complexes with NADPH and two highly potent 5-(omega-carboxy(alkyloxy) trimethoprim derivatives reveals conformational correlations with activity and novel parallel ring stacking interactions. Proteins 65(4):959–969. doi:10.1002/prot.21131
18. Cody V, Galitsky N, Luft JR, Pangborn W, Queener SF, Gangjee A (2002) Analysis of quinazoline and pyrido[2,3-d]pyrimidine N9-C10 reversed-bridge antifolates in complex with NADP+and Pneumocystis carinii dihydrofolate reductase. Acta Crystallogr D Biol Crystallogr 58(Pt 9):1393–1399. doi:10.1107/S0907444902010442,S0907444902010442
19. Wavefunction Inc (2006) Spartan'06, Wavefunction Inc., 18401 Von Karman Avenue, Suite 370 Irvine, CA 92612, USA
20. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph 14(1):33–38. doi:0263785596000185,27–28
21. Cody V, Pace J (2011) Structural analysis of Pneumocystis carinii and human DHFR complexes with NADPH and a series of five potent 6-[5'-(omega-carboxyalkoxy)benzyl]pyrido[2,3-d]pyrimidine derivatives. Acta Crystallogr D Biol Crystallogr 67(Pt 1):1–7. doi:10.1107/S0907444910041004
22. Tripos International (2008) SYBYL v.8.0, Tripos International, 1699 South Hanley Rd., St. Louis, MO
23. Gasteiger J, Marsili M (1980) Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. Tetrahedron 36(22):3219–3228
24. Efroymson MA (1960) Mathematical methods for digital computers. Multiple regression analysis. Wiley, New York
25. Gangjee A, Lin X (2005) CoMFA and CoMSIA analyses of Pneumocystis carinii dihydrofolate reductase, Toxoplasma gondii dihydrofolate reductase, and rat liver dihydrofolate reductase. J Med Chem 48(5):1448–1469. doi:10.1021/jm040153n
26. Barreiro EJ, Kummerle AE, Fraga CA (2011) The methylation effect in medicinal chemistry. Chem Rev 111(9):5215–5246. doi:10.1021/cr200060g
27. Thomas G (2003) Fundamental of medicinal chemsitry. Wiley, Chichester