Molecular BioSystems

RSC Publishing

PAPER

View Article Online
View Journal | View Issue

Cite this: *Mol. BioSyst.*, 2013, **9**, 1308

Homology modeling of dihydrofolate reductase from *T. gondii* bonded to antagonists: molecular docking and molecular dynamics simulations†

Débora Pacheco Homem,[a] Rogério Flores Jr,[a] Priscilla Tosqui,[b] Thiago de Castro Rozada,[a] Ernani Abicht Basso,[a] Arquimedes Gasparotto Junior[c] and Flavio Augusto Vicente Seixas*[a]

The aim of this work was to solve the structure of the enzyme dihydrofolate reductase from *Toxoplasma gondii* (*Tgondii*DHFR) as a target for drug discovery on account of recent reports of parasite's growing resistance to pyrimethamine (CP6), which is the reference pharmaceutical used to treat toxoplasmosis and malaria. The tertiary structure of the protein bonded to $NADP^+$ and CP6 was solved by homology modeling. The best output model was subjected to conjugate gradient minimization and the comparison with templates shows important replacements at the inhibitor's binding site allowing selective drug design. CP6 redocking in *Tgondii*DHFR shows a $\Delta G_{binding}$ of -8.66 kcal mol^{-1}, higher than those found for templates *Plasmodium vivax* (-9.01) and *P. falciparum* (-8.99). Virtual screening of ligands similar to CP6 was performed using the ZINC database and docking procedures were carried out. The result indicates the substances ZINC14966516, ZINC13685962, ZINC13685929 and ZINC13686062 with a $\Delta G_{binding}$ of -10.57, -10.09, -9.87, and -9.76 kcal mol^{-1}, respectively, as the best choices. NPT molecular dynamics with the complexes indicates that they remained stable along the 10 ns simulation and they dock to *Tgondii*DHFR by salt bridges to the Asp 30 and to nine other residues in the contact region, which makes it more difficult for single mutations to acquire resistance. The contact frequency of protein residues with ligands suggests plausible explanations for site-directed mutagenesis studies regarding CP6 resistance described previously in the literature. All results indicate that the new ligands could be tested as pyrimethamine substitutes in the treatment of toxoplasmosis, in addition to other protozoonosis diseases.

Received 19th November 2012,
Accepted 7th February 2013

DOI: 10.1039/c3mb25530a

www.rsc.org/molecularbiosystems

1. Introduction

Toxoplasmosis is a congenital or acquired disease, caused by the infection of *Toxoplasma gondii*, an obligate intracellular coccidian protozoan that belongs to the phylum Apicomplexa, which also includes other parasites, such as the genera *Plasmodium*, *Leishmania*, *Trypanosoma* and *Cryptosporidium*.[1]

It is a worldwide distributed protozoonosis and it has been shown that at several places up to 95% of the population were infected with *T. gondii*. In the United States it is estimated that 22.5% of the population twelve years and older have been infected with this parasite.[2]

Despite being asymptomatic, immunocompromised patients such as transplanted or in chemotherapy treatment may develop cerebritis, chorioretinitis, pneumonia, myocarditis and maculopapular rash, which may lead to death. In patients with AIDS, cerebral toxoplasmosis is a frequent component. Furthermore, a group at serious risk of toxoplasmosis is pregnant women, who can be infected just before or during pregnancy and can pass the infection on to the unborn infant. Although the disease does not always show the symptoms of the infection, it is possible that the child manifests blindness or mental disability during life.[2] A study conducted with pregnant women in a Brazilian city showed that 70.7% of subjects were

[a] *State University of Maringá, Av. Colombo 5790, Maringá, PR, Brazil. E-mail: favseixas@uem.br; Fax: +55 44 36219326; Tel: +55 44 36219300*
[b] *Regional College of Technology, Av. Francisco das Chagas Oliveira 791, São José do Rio Preto, SP, Brazil. Fax: +55 17 33646598; Tel: +55 17 33646589*
[c] *Paranaense University, Pça. Mascarenhas de Moraes, 4282, Umuarama, PR, Brazil. E-mail: gasparotto@unipar.br; Fax: +55 44 3621-2830; Tel: +55 44 3621-2830*

† Electronic supplementary information (ESI) available: PDB file of *Tgondii*DHFR bonded to pyrimethamine and supporting figures. See DOI: 10.1039/c3mb25530a

seropositive for the disease and the other 29.3% were at risk of infection.[3]

Currently the reference treatment for toxoplasmosis involves the concomitant use of pyrimethamine for inhibition of the enzyme dihydrofolate reductase (DHFR) and sulphadiazine for the inhibition of the enzyme dihydropteroate synthase (DHPS). This treatment has been used for several decades[4] and continuous use eventually led to the inevitable selection of resistant strains of the parasite, a fact that has been described by many studies reporting parasite's resistance to pyrimethamine and other drugs used to treat toxoplasmosis[5–8] and malaria.[9–11]

The role of DHFR is to catalyze the nicotinamide adenine dinucleotide phosphate (NADPH) dependent reduction of dihydrofolate to tetrahydrofolate, an important component in the folic acid pathway. The tetrahydrofolate is methylated to methylene tetrahydrofolate, which is directly involved in thymidine synthesis, by assisting the methylation of deoxyuridine monophosphate to give thymidine monophosphate, and it is also indirectly implicated in the metabolism of amino acids and purine nucleotides.[12] Drugs that inhibit DHFR are known as antifolates thus preventing biosynthesis of DNA, leading to cell death. DHFR is a successful drug target because it is an essential enzyme and also its structure differs subtly from species to species, allowing selective drug design.[13]

The structure of DHFR from *Toxoplasma gondii* (*Tgondii*DHFR) has not yet been solved by experimental methods like crystallography; however, its amino acid sequence has the status confirmed in the UniProt database, presenting notations for the residues of the active site, cofactors and ions. This sequence shares significant identity with DHFR from other Apicomplexan parasites such as *Plasmodium* and *Trypanosoma*, species that already have their homologue crystallographic structures deposited in the Protein Data Bank (PDB). These features allow the homology modeling of *Tgondii*DHFR in order to use it in virtual screening studies aimed at the discovery of new drugs for the treatment of toxoplasmosis.

In this work, the results of two docking programs were compared and drug candidates suggested in common were carefully evaluated by ADMETox criteria and molecular dynamics simulations, as a way to find effective drug candidates.

2. Materials and methods

2.1 Material retrieval and sequence alignment

The sequence of *Tgondii*DHFR (Uniprot: Q07422) was used to find templates for sequence alignment through default BLASTp[14] parameters on PDB and also by structural alignment through fold recognition *via* pGenThreader.[15] The crystal structures of the DHFR domain from *Trypanosoma cruzi* (*Tcruzi*DHFR),[16] *Trypanosoma brucei* bonded to NADP (*Tbrucei*DHFR)[17] and *Plasmodium falciparum* (*Pfalcip*DHFR)[18] with respective PDB ID: 2H2Q, 3RG9 and 3JSU were selected as best templates.

2.2 Model building and evaluation

All templates were firstly put together in the same spatial orientation by secondary structure match (SSM) overlay fit[19] implemented in the Coot software.[20] Therefore, the crystal structures of pyrimethamine (CP6) and NADP bonded to DHFR from *Plasmodium vivax* (*Pvivax*DHFR), PDB ID: 2BL9,[21] were added to templates by geometrical docking. The secondary structure predictions by pGenThreader were used to verify the probable fold of the three segments of the *Tgondii*DHFR sequence without match in alignment, aiming to use it as restrains in the modeling process. One hundred models of *Tgondii*DHFR bonded to CP6 and NADP were generated by Modeller v9.10[22] and the top five models selected by the Modeller DOPE score and the best was selected by Procheck results.[23] The best model was then refined through the generation of 300 new models by loop optimization of the regions around residues placed in disallowed regions of the Ramachandran plot, and the best one was selected as described previously. Subsequently, the model was energetically minimized with 20 000 steps of conjugate gradient to remove stereochemical clashes. Details of the minimization procedure will be discussed ahead. Finally the minimized model was used as an input model for docking and long time molecular dynamics simulations.

2.3 Partial charges calculation

The topologies and parameters for protein and NADP are well described on Charm C35b2-C36a2 force fields.[24] The structure of pyrimethamine as well as the other ligands were downloaded from the ZINC database[25] and had their protonation states adjusted to pH 7.0. Then they were subjected to the SwissParam approach[26] to generate the force fields. The Mülliken charges for each atom of the ligands were calculated at B3LYP/6-31G* level of theory in order to obtain accurate force fields.

2.4 Molecular dynamics simulations

The program package NAMD2[27] was used in the simulations. Firstly, the modeled structures of *Tgondii*DHFR bonded to CP6 and NADP were solvated with TIP3P water in a periodic box with limits at least 10 Å away from the outer surface of protein. An appropriate number of chloride counter ions were added to neutralize the system. Thereafter, 20 000 steps of conjugate gradient minimization were applied and the minimized structure was used for subsequent docking procedures and further molecular dynamics evaluation.

The long time dynamics was carried out in sets. Firstly the minimized protein and ligands were fixed in space and water molecules and ions were subjected to 60 ps of equilibration, and then, the system was minimized again by another 20 000 steps. Finally, the system was equilibrated for 10 ns. All MD simulations were performed at a temperature of 300 K and a pressure of 1 atm. Electrostatic interactions were calculated using the particle-mesh Ewald algorithm with a spacing of 1.0 Å and a cutoff radius for the Van der Waals interactions of 10 Å. The simulations were performed with a time step of 2 fs in the integration of the equations of motion. All analyses were performed on the ensemble of system configurations extracted at 2 ps time intervals from the simulation.

The convergences of the different simulations were analyzed in terms of the root mean square deviation (rmsd) from the

initial model structures. The radius of gyration (R_{gyr}) of the whole protein–ligand complex, the residues in contact with the inhibitors with distances around 4.0 Å, as well as the root mean square fluctuation (rmsf) for Cα atoms of each residue were extracted from the trajectory file.

The simulations took place in a single node of an SGI Altix ICE 8400 LX running in parallel on 2 processors Intel Six Core 5680 of 3.33 GHz (12 cores) and 36 GB RAM, at CENAPAD/Unicamp, Brazil.

2.5 Redocking

Assuming that the minimized structure of *Tgondii*DHFR represents the more stable conformation of the complex and consequently the one with the lowest energy, it may now be used in redocking simulations in order to validate the docking protocols that will be used in virtual screening. To test the efficiency of the protocols, they were applied in the redocking of CP6 on the crystal structure of the template *Pvivax*DHFR (2BL9). The simulations were carried out using two different software programs with different search algorithms.

All protocols restricted the torsion angles of the enzyme and the cofactor, but allowed flexibility for the tested ligands. All water molecules and ions were removed from the structures. The protocol for Autodock 4.2[28] implemented in the Pyrx-0.8 graphical interface uses the Lamarckian GA search algorithm, a grid spacing of 0.375 Å, and grid centers (number of points) of: 55.28 (50.0), 28.08 (50.0) and 38.27 (50.0) at x, y and z respectively. The protocol for Molegro-4.3.0 (trial version)[29] uses the Moldock Score [Grid] as a score function; the Moldock Optimizer as a search algorithm and the center of the radius of the sphere, which defines the docking simulation, was set to 10 Å around the catalytic site cavity. All other options were set as default. Redocking simulations were repeated four times and the results were reproducible.

2.6 Virtual screening

The library for the VS simulations was built by the search for compounds similar to CP6 (Tanimoto index of 90%) in the ZINC database. The protocol established and validated in redocking, with a rmsd of 0.427 ± 0.002 Å (mean of four simulations), was employed in VS. Different orientations of the ligands were searched and ranked based on their energy scores. Compounds with lowest energy were filtered using Lipinski's rule of five[30] and toxicity (ADMETox) was implemented in FAF drugs,[31] thus showing promising features of toxicity and pharmacokinetics. The filtered molecules were then analyzed using the Similarity Ensemble Approach (SEA predictions),[32] a server that correlates the chemical properties of a compound with a database with activity determined molecules.

Finally, the calculation of ligand-binding affinity for protein–ligand complexes (pK_d) for the four best results was also done through a polynomial empirical scoring function implemented in the Polscore program.[33] The comparison of theoretical values of pK_d and inhibition constant (K_i) was made with experimental kinetic parameters for each compound available in the Binding database.[34]

3. Results and discussion

3.1 Sequence alignment and model building

The enzyme DHFR in some protozoan parasites such as *Toxoplasma*, *Leishmania*, *Trypanosoma* and *Plasmodium* is condensed into a bifunctional enzyme, which also features a Thymidylate synthase domain (TS). The alignment of the entire protein (DHFR-TS) shows that the TS domain shares twice the structural identity with mentioned organisms as compared to the DHFR domain. However, the *Tgondii*DHFR domain shares 35%, 36% and 27% identity with *Tcruzi*DHFR, *Tbrucei*DHFR and *Pfalcip*DHFR respectively.

As recent studies have demonstrated, sequence identity higher than 25% between two proteins indicates similarity in three-dimensional structures,[35,36] suggesting that the homology modeling of *Tgondii*DHFR is possible.

The amino acid sequence of the *Tgondii*DHFR model structure does not follow the numbering of the UniProt database sequence, because the initial Met is a start codon which normally does not appear in the final expressed protein. Thus, the first residue of the *Tgondii*DHFR model is the Gln1, which corresponds to Gln2 in the Uniprot's sequence.

The SSM overlay fit of the templates showed that the NADP site and the active binding site are structurally well conserved, in spite of other regions of the molecules that exhibit significant differences such as lack of electron density or the difference in the length of the amino acid sequences. This observation is key to manually adjust the sequence alignment by firstly matching the notations of all binding sites and then aligning the rest of the sequences. Nevertheless, five regions of the *Tgondii*DHFR sequence, ranging residues 62–69, 73, 184, 189–192, 221–222, did not fit the alignment because they are larger than the templates. Secondary structure prediction was used to solve the problem of lack of template information and the result indicates that those regions may be coils. Following this information, no restriction in the topology was set during the modeling process. The final alignment used in modeling and the secondary structure prediction for *Tgondii*DHFR are shown in Fig. 1.

Pyrimethamine is the reference medicine in the treatment of toxoplasmosis. The modeling of this compound bonded to *Tgondii*DHFR was intended to visualize structural details of its interaction with the active site, as well as to serve as a template for proper guidance in the docking studies of other inhibitors. The minimization MD procedures can ensure that the binding site was stereochemically adapted to the presence of the ligand and *vice versa*, in a way to drive the complex to a more stable conformation of low energy, a necessary condition for redocking studies.[37] The Ramachandran plot of this minimized model showed that 99.0% of the residues are placed in allowed regions, indicating that *Tgondii*DHFR is ready for further analysis.

The DHFR tertiary structures of apicomplexan parasites vary significantly among their species. The structures from *Trypanosoma* and *Plasmodium* used as templates in this work differ in size and in the presence of some accessory domains, while the structure from *T. gondii* appears to be a mix of both (see ESI†). However, the folding of the substrate and cofactor binding sites

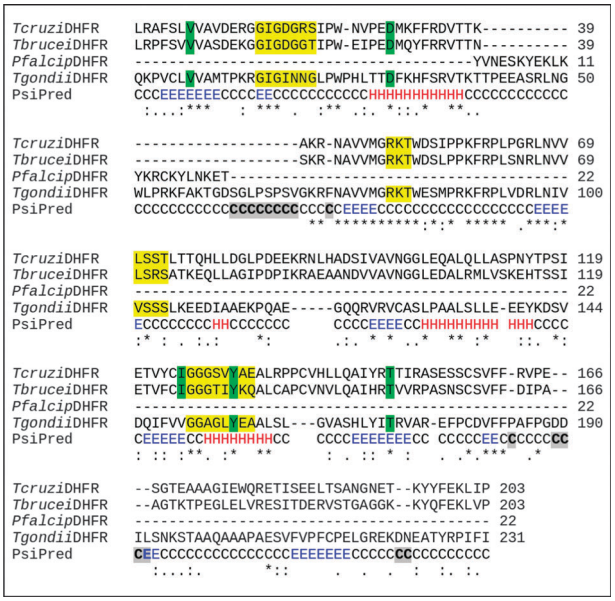

Fig. 1 ClustalW alignment of the DHFR target and templates showing notation for the NADP binding site (yellow boxes) and active site (green boxes). pGenThreader secondary structure prediction for coil (C), strand (E) and helix (H) overall and at sequence gaps (grey boxes). Identical residues (*), conserved substitutions (:), semi-conserved substitution (.).

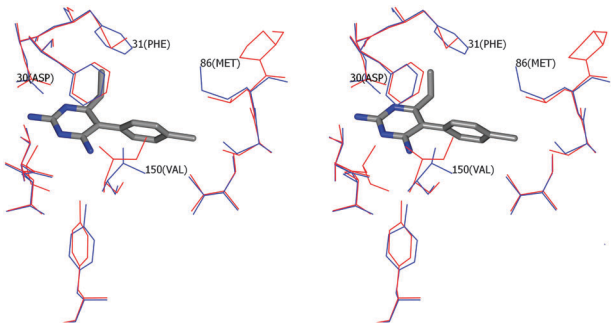

Fig. 2 Stereoview of pyrimethamine bonded to the DHFR site at *T. gondii* (blue) and *T. brucei* (red).

appear to be quite conserved with subtle differences such as the replacement of a few residues (Fig. 2), which allows the design of selective drugs.

3.2 Virtual screening

The library of compounds was built by search for substances having 90% of structural similarity with pyrimethamine in the ZINC database. No purchasable filter was used because, in this level of similarity, we were able to use CP6 as a skeleton for the synthesis of similar compounds.

One hundred and thirty-three candidates were selected and downloaded in the structure-data file format for docking evaluations using the protocols previously validated by redocking studies. As a final result, four compounds were selected among the top 10 presented by both AutoDock and Molegro (Fig. 3) as follows: compound CP7 (ZINC14966516), CP8 (ZINC13685962), CP9 (ZINC13685929) and CP0 (ZINC13686062).

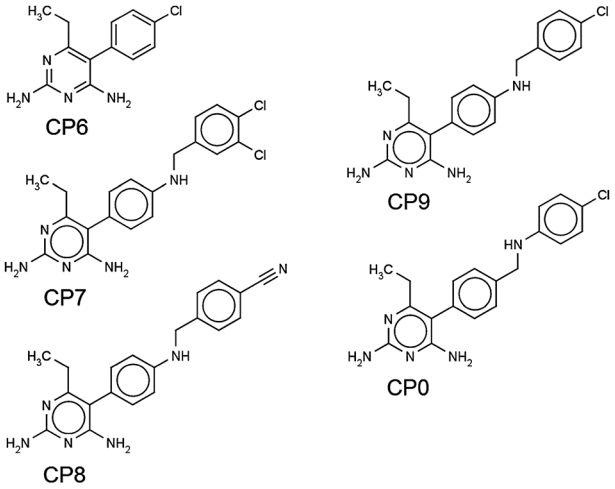

Fig. 3 The 2D structure of pyrimethamine and derived compounds selected as drug candidates against toxoplasmosis in this work.

All those ligands were successfully filtered by Lipinski's rule of five and the ADMETox criterion using FAF-Drugs. The physicochemical properties of CP6 and other ligands, exactly for the structural similarity, are very alike. All of them work as weak bases, being monoprotonated at pH 7.0.

The inhibitory effect of pyrimethamine on DHFR is a well-described event. Thus it is expected that ligands derived from its basic structure (compounds 7, 8, 9 and 0) could be docked by a similar mechanism. Furthermore, the affinity of ligand binding was also assessed by the calculation of the dissociation coefficient (pK_D), as well the theoretical $\Delta G_{binding}$. All parameters indicate that discovered ligands CP7, CP9, CP8 and CP0 have greater affinity to *Tgondii*DHFR than the reference ligand CP6 (Table 1), a key feature for possible new competitive inhibitors.

The main reason for the increased affinity of these compounds is the extra number of interactions made with the enzyme. Actually there is no report of using CP7, CP9, CP8 and CP0 as drug candidates against toxoplasmosis, leading this work to open up perspectives for preclinical studies with these compounds to confirm their *in vivo* effectiveness against toxoplasmosis.

3.3 Reference ligand CP6

The structure of pyrimethamine is composed of two rings (Fig. 3). The data extracted from the trajectory of the simulation show that CP6 docks to *Tgondii*DHFR by two salt bridges made between CP6 atoms, N1 and N2, and atoms Oδ1 and Oδ2 of the Asp30 carbonyl group and by more nine important close contacts with Val7, Val8, Ala9, Phe31, Phe34, Thr82, Met86, Val150 and Thr171. The only hydrogen bond present is made between the CP6 meta-amine group and the atom O20 of the amide group of NADP. This bond stabilizes the anchoring of CP6 to the cofactor.

3.4 Ligand CP7

The ligand CP7 is identified by IUPAC (International Union of Pure and Applied Chemistry) as 5-[4-[(3,4-dichlorophenyl)-methylamino]phenyl]-6-ethylpyrimidine-2,4-diamine (Fig. 3). Similar to pyrimethamine and all other discovered ligands,

Table 1 Docking parameters and literature report for discovered ligands

Compound	XLogP3	SEA prediction[a] (e-value)	Molegro score	Polscore pK_D	Autodock ΔG_{bind} (kJ mol^{-1})	Autodock, K_i (nM)	Protein target in the Binding database	Binding database IC_{50} (nM)
Pyrimethamine (CP6)	2.7	DHFR (3.35×10^{-158})	−78.41	5.34	−8.66	447.0	*Tgondii*DHFR	—
ZINC14966516 (CP7)	4.6	DHFR (3.14×10^{-115})	−99.52	6.94	−10.57	18.0	-	—
		RENIN (3.33×10^{-6})	—	—	—	—	Human renin	27000.0
ZINC13685929 (CP9)	4.2	DHFR (4.32×10^{-122})	−99.44	6.55	−10.09	40.0	H. Ghrelin receptor	310.0
ZINC13685962 (CP8)	3.3	DHFR (9.33×10^{-95})	−101.48	6.52	−9.87	59.0	H. Ghrelin receptor	160.0
ZINC13686062 (CP0)	4.2	DHFR (4.32×10^{-122})	−101.26	6.50	−9.76	70.0	H. Ghrelin receptor	300.0

[a] Prediction for discovered ligands.

it is also a weak base. This fact explains its positive charge at pH 7.0, which is essential to maintain salt bridges with the protein residue Asp30 under physiological conditions.

CP7 makes van der Waals contacts with other eleven residues: Val7, Val8, Ala9, Leu22, His26, Thr82, Ser85, Met86, Pro87, Val150 and Tyr156. This ligand makes two hydrogen bonds with the cofactor: one as previously described for pyrimethamine and a second H-bond between atom N5 of the CP7 amine group and atom O10 of the NADP ribose hydroxyl subgroup. This additional stabilization is reflected in the theoretical values of pK_D, K_i and $\Delta G_{binding}$ for CP7, which showed the best parameters among all ligands (Table 1).

The analysis of CP7 by the target similarity ensemble approach related proteins using SEA predictions, suggesting DHFR and renin as target candidates with low and high probability of false positive (SEA e-value) results respectively (Table 1). However, a search in the Binding database for experimental data using CP7 as a drug candidate indicates that it was already tested as an antihypertensive using human renin as a target.[38,39]

3.5 Ligand CP8

The ligand CP8 is identified by IUPAC as 4-[[4-(2,4-diamino-6-ethylpyrimidin-5-yl)anilino]methyl]benzonitrile (Fig. 3). It is positively charged under physiological conditions, and also binds to the *Tgondii*DHFR making salt bridges with Asp30. The additional interaction with the protein occurs by the close contacts with residues Val7, Val8, Ala9, Gli21, Leu22, His26, Phe34, Thr82, Ser85, Met86, Pro87 and Val150. The interaction of CP8 with the cofactor occurs through two hydrogen bonds, one more than the reference ligand has, thus contributing to additional stabilization of the cofactor. It is clear that there is a more number of closer contacts holding CP8 bonded to protein than CP6 and this fact explains the better binding parameters found in the docking studies (Table 1) for CP8. The ligand CP8 has already been tested as a growth hormone secretagogue receptor (GHS-R) antagonist,[40] however, there is no report of using it as an antifolate drug. The SEA predictions suggest DHFR as a target for CP8 with a low probability of false positive (Table 1) and there was no indication of GHS-R as a target.

3.6 Ligand CP9

The ligand CP9 differs from CP7 by having one less chlorine, which increases its solubility. The CP9 is identified by IUPAC as 5-[4-[(4-chlorophenyl)methylamino]phenyl]-6-ethylpyrimidine-2,4-diamine (Fig. 3), it is also positively charged under physiological conditions, and also binds to *Tgondii*DHFR by two salt bridges with Asp30. Additional stabilization is made by close contact with ten residues of the binding site: Val7, Val8, Ala9, Gly21, Phe34, Thr82, Ser85, Met86, Pro87 and Val150. The interaction with cofactor NADP is also made by two hydrogen bonds. The CP9 has already been tested as a growth hormone secretagogue receptor (GHS-R) antagonist in the same paper that described CP8,[40] however, there is no report of using it as an antifolate drug. The SEA predictions suggest DHFR as a target for CP9 (Table 1) and there is no indication for GHS-R as a target.

3.7 Ligand CP0

The ligand CP0 is an isomeric form of CP9 and is identified by IUPAC as 5-[4-[(4-chloroanilino)methyl]phenyl]-6-ethylpyrimidine-2,4-diamine (Fig. 3). As described for all other ligands, it is also positively charged under physiological conditions, and also binds to *Tgondii*DHFR by two salt bridges with Asp30. Additional stabilization is made by close contact with eleven residues of the protein binding site: Val7, Asn20, Gly21, Leu22, His26, Phe31, Phe34, Ser85, Met86, Pro87 and Val150. The CP0 also interacts with the cofactor by two hydrogen bonds. This compound is less stable in molecular dynamics studies but still docks to protein by more contacts than CP6. The CP0 has already been tested as a growth hormone secretagogue receptor (GHS-R) antagonist in the same paper that described CP8 and CP9,[40] however, there is no report of using it as an antifolate drug. The SEA predictions suggest DHFR as a target for CP0 (Table 1) however, no prediction for GHS-R as a target was found.

3.8 MD simulations and the overall stability of structure

To explore the stability of the modeled *Tgondii*DHFR domain as well as to obtain information about the contribution of each residue of the protein binding site to the stabilization of inhibitors, we performed 10 ns MD simulations of protein at Apo and bonded forms without any harmonic restraints. In these studies, the final model of *Tgondii*DHFR docked with the best pose of each ligand was minimized by 20 000 steps and further equilibrated as previously described in methodology.

The equilibrated state of the complexes was considered achieved when the RMSD taken from the protein backbone over the MD trajectory, stabilized in a plateau for at least 5 ns.

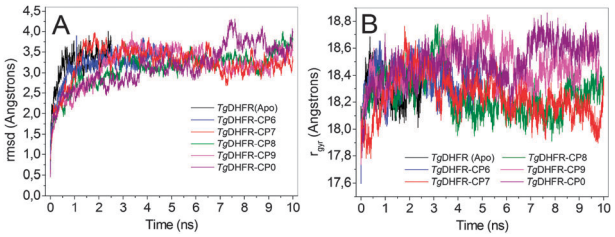

Fig. 4 Simulation behaviour of *Tgondii*DHFR complexes as a function of time. (A) Root main square deviation. (B) Radius of gyration.

These results shown in Fig. 4A indicate overall stabilization of RMSD after 3 ns simulation with values ranging between 3.0 to 3.5 Å. This result also indicates that 10 ns of simulation is a sufficient time to get an overall picture of protein dynamics in the equilibrated state.

The radius of gyration (r_{gyr}) is defined as the root mean square distance of the collection of atoms from their common centre of gravity. As shown in Fig. 4B, all complexes displayed a common r_{gyr} lying between 18.0 and 18.5 Å after 3 ns equilibration. This result indicates that both Apo and liganded forms of *Tgondii*DHFR kept the fold of their original design, and the binding of the inhibitors does not let the protein to unfold. However, the protein is not static *in vivo* and certain regions of the structure are more flexible than others. In this work, we accessed the regions of greater flexibility in *Tgondii*DHFR by analyzing the root mean square fluctuation (rmsf) of each residue. The rmsf was calculated from Cα atoms in the last 2 ns of simulation during which the protein is at the equilibrated stage. As expected, the major fluctuations occurred between residues 50 to 75 and 188 to 210, indicated by R1 and R2 regions in Fig. 5. Those regions lie away from the active site causing no influence on ligand binding. It is interesting to note that R1 and R2 regions still remain flexible in the structures of the templates as demonstrated by the high values of the B-factor and the lack of electron density found in those analogous regions of templates.

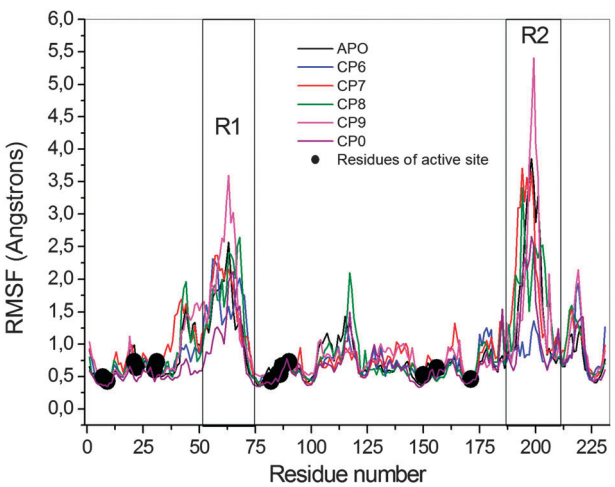

Fig. 5 Root main square fluctuation of Cα atoms from *Tgondii*DHFR complexes at last 2 ps simulation.

3.9 Active site

In this work, the active site was delimited by the residues making any contact with the ligand in a distance range of 4.0 Å. In order to obtain a more accurate picture about the contribution of each residue to the building of the active site, the contact frequency of each residue with the ligand was taken in each one of the 5000 frames of the molecular dynamics trajectory. A complete list of contact frequencies from *Tgondii*DHFR to the discovered ligands is shown in Table 2.

The analysis of Table 2 shows that the high contact frequencies of Val7, Val8, Ala9, Asp30, Phe34, Thr82, Met86, Val150 and Thr151 make them important in the stabilization of all ligands. However, the residues Leu22, Ser85 and Pro87 play an important role only in the extra stabilization of the discovered ligands beyond those found for pyrimethamine, since the side chain of these three residues has a high frequency of contact with ligands. The Ser85 has a polar side chain and it is placed in the interface between the ligand and the cofactor contributing to the anchoring of both. These results may suggest that any replacement in the position of those three residues in *Tgondii*DHFR could result in change in ligand affinity, which could let the parasite to develop a possible resistance to those discovered inhibitors. This observation may be important in the future molecular studies aiming to identify strains of *T. gondii* resistant to those agents.

The comparison of *Tgondii*DHFR and *Tbrucei*DHFR active sites by overlay fit shows important stereochemical differences (Fig. 2). The most significant changes are due to the replacement of the nonpolar residues Met30, Leu65 and Ile135 in *Tbrucei*DHFR by Phe31, Met86 and Val150 in *Tgondii*DHFR. The position of those three residues places them in close contact with pyrimethamine, a fact that may affect and even justify less affinity of this inhibitor for *Tgondii*DHFR than *Pfalcip*DHFR and *Tbrucei*DHFR.[41–43]

Site-directed mutagenesis studies of *Tgondii*DHFR showed that the replacement of Trp24Arg did not alter the value of IC_{50}

Table 2 Contact frequency of the *Tgondii*DHFR residues with ligands

Residue	CP6	CP7	CP8	CP9	CP0
Val-7	1.00	1.00	0.99	0.98	0.98
Val-8	0.96	0.89	0.89	0.82	0.59
Ala-9	0.94	0.97	0.80	0.87	0.68
Asn-20	0.01	0.06	0.66	0.37	0.82
Gly-21	0.11	0.20	0.70	0.90	0.98
Leu-22	0.43	0.91	0.97	0.66	0.92
Trp-24	0.02	0.11	8×10^{-4}	4×10^{-4}	—
His-26	0.25	0.71	0.76	0.36	0.81
Leu-27	—	0.11	—	—	—
Asp-30	1.00	1.00	1.00	1.00	1.00
Phe-31	0.81	0.35	0.28	0.69	0.71
His-33	2×10^{-4}	—	0.17	0.23	0.02
Phe-34	0.90	0.02	0.74	0.94	0.97
Met-78	0.10	4×10^{-4}	0.04	2×10^{-4}	2×10^{-4}
Thr-82	0.78	0.94	0.96	0.94	0.49
Ser-85	0.34	0.98	0.93	0.88	0.99
Met-86	0.80	0.97	0.93	0.94	0.95
Pro-87	0.00	0.88	0.80	0.91	0.75
Phe-90	6×10^{-4}	0.37	0.08	0.20	0.09
Val-150	0.93	0.96	0.96	0.91	0.72
Tyr-156	0.47	0.76	0.38	0.14	0.02
Thr-171	0.88	0.62	0.52	0.46	0.47
Tyr-226	—	0.11	2×10^{-4}	2×10^{-4}	0.27

for CP6 when compared to the wild type.[7] Our results in Table 2 may now explain that this experimental observation may be due to the low contact frequency of all ligands with this residue and, as a consequence, replacement at this position should not have significant effects in *Tgondii*DHFR.

In another work regarding site-directed mutagenesis experiments, the replacement of Ser35Arg did not let the parasite acquire resistance to CP6.[44] Our results may explain that Ser35 in this case does not make any contact with the ligand. A plausible explanation has to do with the fact that this residue is placed out of the active site and its contribution may be due to a steric effect leading to displacement of its neighbor Phe34, which is very important in the stabilization of the ligand. However, the same work shows that the mutation Thr82Asn led *Tgondii*DHFR to acquire CP6 resistance. Thr82 has a high contact frequency with all ligands evaluated in this work, so mutations in these residues could significantly affect the binding of inhibitors.

The authors also show that the double mutation Ser35Arg and Thr82Asn provided at least 10 times more resistance to CP6 than Thr82Asn alone. The two residues involved in this mutation are placed on opposite sides of the active site and their Cα atoms are separated by 15 Å. Since those residues are relatively larger than the original ones, this double mutation was simulated in the structure of the native protein and the resulting residues were sterically oriented toward each other, leading to fill part of the active site cavity and let the shortest distance between them to be approximately 8 Å, large enough to prevent any direct interaction. An MD simulation was carried out with this double mutant protein bonded to CP6 for 10 ns. The results showed that both RMSD and radius of gyration were significantly affected (see ESI†), which suggests that the mutant protein complexed to CP6 is not stable.

Our simulation results did not provide direct evidence that might explain why the double mutation Ser35Arg and Thr82Asn provided increased resistance to CP6 than Thr82Asn alone, which makes any attempt to explain this phenomenon purely speculative. However, it is safe to conclude that Thr82Asn mutation decreased the volume of the active site and the mutation Ser35Arg conferred an additional positive charge at the entrance of the active site. In this scenario, the mutations potentiate the effect of each other leading to an unfavorable environment for the interaction of the protein with pyrimethamine, which also has a positive charge at pH 7.0.

Molecular dynamics studies regarding drug–target interaction become an interesting way to investigate the contribution of each residue of the active site in the docking process, thus providing a more elucidative view of the phenomena compared to a single static image from crystallography experiments.

Finally, the parasitic diseases of Apicomplexa are no longer a problem of tropical countries and have become a worldwide health problem, due to the great migration flow of human populations. Studies aiming to seek for new pharmacological agents against toxoplasmosis are just one of the actions that may be taken against pyrimethamine resistant *Toxoplasma* strains.[5–8]

4. Conclusions

Our approach of comparative docking using different software for screening the same drug library resulted in the identification of four potential drug candidates. Every stage of this study was carefully validated before proceeding to the next, which ensures greater reliability of the results. The theoretical affinity indicators led to common results that ended in the discovery of ZINC14966516, ZINC13685962, ZINC13685929 and ZINC13686062 as drug candidates against toxoplasmosis.

Although these hits have been discovered using *Tgondii*DHFR as a target, it is also possible that they may be tested against *Trypanosoma* and/or *Plasmodium* infections, since the drug hits discovered here are derived from pyrimethamine, a common agent used against Apicomplexan parasites.

Acknowledgements

This work was supported by financial aid from Fundação Araucária and CAPES. We also acknowledge CENAPAD for providing computational facilities.

References

1 K. Becker, *Apicomplexan parasites: Molecular Approaches toward Targeted Drug Development*, Wiley, Weinheim, 2011.
2 CDC, Centers for Disease Control and Prevention, 2012.
3 P. R. Leão, J. M. Filho and S. F. Medeiros, *Rev. Bras. Ginecol. Obstet.*, 2004, **26**, 627–632.
4 P. Borst and M. Ouellette, *Annu. Rev. Microbiol.*, 1995, **49**, 427–460.
5 P. Caramello, T. Brancale, B. Forno, A. Lucchini, A. M. Pollono, A. Ullio, P. Gioannini, I. Viano and E. Tonso, *Antimicrob. Agents Chemother.*, 1995, **39**, 2371–2372.
6 H. E. Kaufman, J. Remington, M. L. Melton and S. L. Jacob, *Arch. Ophthalmol.*, 1959, **62**, 611–615.
7 M. G. Reynolds, J. Oh and D. S. Roos, *Antimicrob. Agents Chemother.*, 2001, **45**, 1271–1277.
8 M. G. Reynolds and D. S. Roos, *J. Biol. Chem.*, 1998, **273**, 3461–3469.
9 B. Faye, M. Ndiaye, J. L. Ndiaye, A. Annie, R. C. Tine, A. C. Lo, D. Sow, A. De Sousa and O. Gaye, *Parasitol. Res.*, 2011, **109**, 133–138.
10 J. E. Hyde, *Pharmacol. Ther.*, 1990, **48**, 45–59.
11 S. Zakeri, M. S. Farahani, M. Afsharpad, M. Salehi, A. Raeisi and N. D. Djadid, *Int. J. Infect. Dis.*, 2010, **14**(Suppl 3), e123–e128.
12 F. Zuccotto, M. Zvelebil, R. Brun, S. F. Chowdhury, R. Di Lucrezia, I. Leal, L. Maes, L. M. Ruiz-Perez, D. Gonzalez Pacanowska and I. H. Gilbert, *Eur. J. Med. Chem.*, 2001, **36**, 395–405.
13 E. F. da Cunha, T. C. Ramalho, D. T. Mancini, E. M. Fonseca and A. A. Oliveira, *J. Braz. Chem. Soc.*, 2010, **21**, 1786–1806.
14 S. F. Altschul, W. Gish, W. Miller, E. W. Myers and D. J. Lipman, *J. Mol. Biol.*, 1990, **215**, 403–410.
15 A. Lobley, M. I. Sadowski and D. T. Jones, *Bioinformatics*, 2009, **25**, 1761–1767.

16 N. Schormann, O. Senkovich, K. Walker, D. L. Wright, A. C. Anderson, A. Rosowsky, S. Ananthan, B. Shinkre, S. Velu and D. Chattopadhyay, *Proteins*, 2008, **73**, 889–901.
17 J. Vanichtanankul, S. Taweechai, J. Yuvaniyama, T. Vilaivan, P. Chitnumsub, S. Kamchonwongpaisan and Y. Yuthavong, *ACS Chem. Biol.*, 2011, **6**, 905–911.
18 A. Nzila, M. Rottmann, P. Chitnumsub, S. M. Kiara, S. Kamchonwongpaisan, C. Maneeruttanarungroj, S. Taweechai, B. K. Yeung, A. Goh, S. B. Lakshminarayana, B. Zou, J. Wong, N. L. Ma, M. Weaver, T. H. Keller, V. Dartois, S. Wittlin, R. Brun, Y. Yuthavong and T. T. Diagana, *Antimicrob. Agents Chemother.*, 2010, **54**, 2603–2610.
19 E. Krissinel and K. Henrick, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2004, **60**, 2256–2268.
20 P. Emsley and K. Cowtan, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2004, **60**, 2126–2132.
21 P. Kongsaeree, P. Khongsuk, U. Leartsakulpanich, P. Chitnumsub, B. Tarnchompoo, M. D. Walkinshaw and Y. Yuthavong, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 13046–13051.
22 N. Eswar, B. Webb, M. A. Marti-Renom, M. S. Madhusudhan, D. Eramian, M. Y. Shen, U. Pieper and A. Sali, *Current Protocols in Bioinformatics*, 2006, ch. 5, Unit 5.6.
23 Collaborative Computational Project Number 4, The *CCP*4 suite: programs for protein crystallography, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1994, **50**, 760–763.
24 A. D. Mackerell, Jr., M. Feig and C. L. Brooks, 3rd, *J. Comput. Chem.*, 2004, **25**, 1400–1415.
25 J. J. Irwin, T. Sterling, M. M. Mysinger, E. S. Bolstad and R. G. Coleman, *J. Chem. Inf. Model.*, 2012, **52**, 1757–1768.
26 V. Zoete, M. A. Cuendet, A. Grosdidier and O. Michielin, *J. Comput. Chem.*, 2011, **32**, 2359–2368.
27 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781–1802.
28 G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J. Comput. Chem.*, 2009, **30**, 2785–2791.
29 R. Thomsen and M. H. Christensen, *J. Med. Chem.*, 2006, **49**, 3315–3321.
30 C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Adv. Drug Delivery Rev.*, 2001, **46**, 3–26.
31 M. A. Miteva, S. Violas, M. Montes, D. Gomez, P. Tuffery and B. O. Villoutreix, *Nucleic Acids Res.*, 2006, **34**, W738–W744.
32 M. J. Keiser, B. L. Roth, B. N. Armbruster, P. Ernsberger, J. J. Irwin and B. K. Shoichet, *Nat. Biotechnol.*, 2007, **25**, 197–206.
33 W. F. de Azevedo, Jr. and R. Dias, *Bioorg. Med. Chem.*, 2008, **16**, 9378–9382.
34 T. Liu, Y. Lin, X. Wen, R. N. Jorissen and M. K. Gilson, *Nucleic Acids Res.*, 2007, **35**, D198–D201.
35 A. S. Yang and B. Honig, *J. Mol. Biol.*, 2000, **301**, 691–711.
36 B. Rost, *Protein Eng.*, 1999, **12**, 85–94.
37 R. Dias, L. F. Timmers, R. A. Caceres and W. F. de Azevedo, Jr., *Curr. Drug Targets*, 2008, **9**, 1062–1070.
38 D. D. Holsworth, M. Jalaie, T. Belliotti, C. Cai, W. Collard, S. Ferreira, N. A. Powell, M. Stier, E. Zhang, P. McConnell, I. Mochalkin, M. J. Ryan, J. Bryant, T. Li, A. Kasani, R. Subedi, S. N. Maiti and J. J. Edmunds, *Bioorg. Med. Chem. Lett.*, 2007, **17**, 3575–3580.
39 R. L. Webb, N. Schiering, R. Sedrani and J. Maibaum, *J. Med. Chem.*, 2010, **53**, 7490–7520.
40 M. D. Serby, H. Zhao, B. G. Szczepankiewicz, C. Kosogof, Z. Xin, B. Liu, M. Liu, L. T. Nelson, W. Kaszubska, H. D. Falls, V. Schaefer, E. N. Bush, R. Shapiro, B. A. Droz, V. E. Knourek-Segel, T. A. Fey, M. E. Brune, D. W. Beno, T. M. Turner, C. A. Collins, P. B. Jacobson, H. L. Sham and G. Liu, *J. Med. Chem.*, 2006, **49**, 2568–2578.
41 A. Gangjee, H. D. Jain, S. F. Queener and R. L. Kisliuk, *J. Med. Chem.*, 2008, **51**, 4589–4600.
42 S. Kamchonwongpaisan, J. Vanichtanankul, S. Taweechai, P. Chitnumsub and Y. Yuthavong, *Int. J. Parasitol.*, 2007, **37**, 787–793.
43 F. Gamarro, P. L. Yu, J. Zhao, U. Edman, P. J. Greene and D. Santi, *Mol. Biochem. Parasitol.*, 1995, **72**, 11–22.
44 R. G. Donald and D. S. Roos, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 11703–11707.