

A Complete Model of the *Plasmodium falciparum* Bifunctional Enzyme Dihydrofolate Reductase-Thymidylate Synthase. A Model to Design New Antimalarials

Tanos C. C. França, André L. R. de Medeiros, Edison C. P. dos Santos, Osvaldo A. Santos-Filho and José D. Figueroa-Villar*

Departamento de Química, Instituto Militar de Engenharia, Praça General Tibúrcio 80, Praia Vermelha, 22290-270 Rio de Janeiro - RJ, Brasil

É proposto um modelo teórico para a *pf*DHFR-TS que inclui os 55 aminoácidos que não foram contemplados no modelo cristalográfico. O cálculo do potencial eletrostático sobre a superfície do modelo, revelou uma região contínua de potencial positivo conectando os dois sítios ativos, sugerindo um mecanismo otimizado de transporte de dihidrofolato.

We propose a theoretical model for *pf*DHFR-TS, which includes the 55 aminoacid residues ignored in the crystallographic model. The electrostatic potential calculation on the model surface revealed a continuous positive potential region between the two active sites, suggesting an optimized mechanism for dihydrofolate transport.

Keywords: malaria, homology modeling, DHFR-TS, optimized substrate transport, *Plasmodium falciparum*

Introduction

Malaria is one of the most serious widespread parasitic diseases in humans. Nowadays, at least 300 million people are infected by malaria every year, with something in between 2 and 3 million deaths. To worsen the situation, the population that lives in endemic areas is about 2 billion people or 40% of the world population.¹ It is possible that the rise in Earth temperature will eventually increase the endemic areas to include greater extensions of Europe and North America. In some parts of Africa, 10% of the deaths of children less than five years old are due to the direct effects of malaria. In fact, the actual epidemic of malaria makes it impossible to calculate the contribution of this disease to the mortality rate of children of this age due to other sickness.² The World Health Organization (WHO) estimates that 3,000 children under 5 years old die of malaria in Africa every day,³ a situation that Mankind should not tolerate. In addition and frequently overlooked in the case of malaria caused by *Plasmodium falciparum*, there is considerable morbidity associated with this disease that manifests as chronic severe anemia in children and adults. In Brazil, the most important endemic areas of malaria are situated in Amazonian Region, where this

disease is one of the major hindrances to the progress of the region, leading to 300,000 to 400,000 new reported cases every year.⁴

Malaria is a disease of the poorest regions of the planet, and because of this, the interest of the major pharmaceutical companies in the development of economically feasible chemotherapy is very low or does not exist. In fact, the strategies used against malaria, like chemotherapy and the use of insecticides, are often not accessible to the populations of the endemic areas due to economic and social reasons.

One of the major obstacles to the control of malaria is the rapid spread of parasite resistance to the current available chemotherapy and of the mosquito vector to insecticides. This phenomenon has rendered other times wonderful antimalarials, such as chloroquine, completely useless in several regions.⁵⁻⁷ Also, it should be expected that resistance should eventually rise for any new drug that we develop in the future, thus making the search of new antimalarial chemotherapy a continuous task for the scientific community.

Among the most used antimalarials are the antifolates.⁸ These drugs are divided in two groups: the first are the type I antifolates, which includes compounds that compete with *para*-amino-benzoic acid (PABA) and interrupt the synthesis of tetrahydrofolate; The type II antifolates act as

* e-mail: figueroa@ime.eb.br

antagonists in the enzymes dihydrofolate reductase (DHFR) and thymidylate synthase (TS). Two of the most important type II antifolates that act against DHFR are pyrimethamine and cycloguanil, which are now much less useful to treat malaria in South America.^{9,10} Those drugs were mainly used in combination with type-I antifolates, such as sulfones or sulfonamides. For example, the combination pyrimethamine-sulfadoxime was one of the mixtures of choice to treat malaria up to 1979,¹¹ but lost most of its usefulness due to the rapid appearance of parasite resistance.¹² The scientific community is worried about such a fact and is looking for non-traditional strategies to develop new treatments for malaria. The fact that all protozoa, including *Plasmodium*, have the enzymes DHFR and TS as parts of a bifunctional homodimeric enzyme, in contrast with mammalian, that have these enzymes as monofunctional separate proteins, may be used in the development of new and more selective antimalarials.

The first study involving the *pf*DHFR-TS bifunctional enzyme occurred in 1984.¹³ This study demonstrated that the protein has a molecular weight of about 140 kDa, being a dimer composed of two subunits of identical size (70 kDa each). Afterwards, Bzik,¹⁴ making use of cloning and isolation of the gene of the *pf*DHFR-TS, accomplished the sequential analysis of this protein. It was shown that each monomer was composed by 608 residues, with the DHFR domain composed by 228 residues (27 kDa), the TS domain composed by 286 residues (33 kDa) and that both domains were connected through a polypeptide junction chain composed by 94 residues (10 kDa).

Shallom *et al.*¹⁵ related that there are physical interactions between the DHFR domain and the TS domain of the bifunctional protein and that these interactions are necessary to obtain a catalytically active TS site, suggesting, in this manner, that interference with these essential protein-protein interactions could lead to new selective strategies to treat malaria resistant to traditional inhibitors.

Liang and Anderson,¹⁶ in their studies with the bifunctional *Leishmania major* DHFR-TS (*lm*DHFR-TS) protein, provided definitive evidence for some type of optimized substrate transportation, as for example channeling of dihydrofolate, from the TS site to the DHFR site and suggested that the DHFR site is catalytically activated by the TS site and vice versa.

The crystallographic structure of *lm*DHFR-TS proposed by Knighton¹⁷ showed that this enzyme is also a dimer, with each monomer consisting of two domains, one for DHFR and one for TS. No obvious structure was found to support the existence of a tunnel connecting both sites, as it is the case for tryptophan synthase,¹⁸ for example.

However, Knighton noticed that, in his model, the substrate is exposed to the solvent and that it could move by the outside of the enzyme, as he showed that the electrostatic potential in the region between the sites is strongly positive and dihydrofolate carries a net charge of -2 . Based on this evidence, he suggested that the optimized transfer is achieved by interaction of the charged substrate with the positive electrostatic potential surface connecting the DHFR and TS sites of the *lm*DHFR-TS bifunctional enzyme.¹⁹

Clearly, to design compounds that could interfere with the optimized substrate transport in *pf*DHFR-TS, or with the suspected protein-protein interactions in this bifunctional enzyme, it was necessary to obtain first a structural model of the whole protein. Three models for *pf*DHFR obtained by molecular modeling have been reported^{20,21,22} including one from our research group²¹ but none of them included the TS domain region. Before this domain could be properly modeled, Yuvaniyama *et al.*²³ reported the crystallographic structure of *pf*DHFR-TS (PDB code: 1J3I), this representing a very important step towards the rational development of more efficient antimalarials. Unfortunately, the reported structure lacked several aminoacid residues, including 46 residues situated just in the junctional region. Certainly, the consideration of all the aminoacids is important to understand a possible optimized substrate transport in the *P. falciparum* enzyme. Accordingly, we have used the incomplete crystallographic *pf*DHFR-TS structure to obtain a better-adjusted model for the whole protein. This model was used to rationalize some experimental observations concerning the *pf*DHFR-TS, especially the existence of an optimized substrate transport process from one active site to another, as proposed by Liang and Anderson.¹⁶

Experimental

Homology modeling of the complete pfDHFR-TS protein based on the incomplete crystallographic structure

The crystallographic structure of *pf*DHFR-TS bifunctional enzyme, published by Yuvaniyama *et al.*²³ was used as template to obtain our model. As first step in the generation of the initial model of the target, the complete sequence of *pf*DHFR-TS was aligned, in the Swiss-PdbViewer program,²⁴ to the incomplete crystallographic structure.²³ The alignment was fitted manually to match all residues, but the ones corresponding to the missing residues in the crystal structure. The so fitted alignment was submitted to the Optimize mode of the Swiss-Model server^{25,26} that generated the initial model of

the target-enzyme. This model of *pf*DHFR-TS was minimized using the GROMOS force field,²⁷ implemented in the computational package GROMACS 3.1.4.^{28,29} The steepest descent and conjugate gradient minimization algorithms were employed during this minimization procedure. The minimized model was submitted to the BIOTECH server³⁰ for validation.

Validation of the model

The programs PROCHECK³¹ and WHATCHECK³² were used in this step. Results not satisfactory in the validation tests provoked changes in the model. Thus, modified models were minimized as before and submitted to a new validation. Unsatisfactory results, in these new tests, led to new changes in the last models. This iterative process, of constant validation and generation of models, was carried out until the obtainment of the best model. The three-dimensional structure that introduced the most satisfactory result in the evaluation tests constituted the final model of the *pf*DHFR-TS bifunctional enzyme. The hardware resource used in this work was a PC Athlon 2000.

Calculation of the electrostatic potential of *P. falciparum* DHFR-TS bifunctional enzyme

The electrostatic potential of *pf*DHFR-TS bifunctional enzyme was calculated in Swiss-PdbViewer program.²⁴ The method of Poisson-Boltzmann was used, with a dielectric constant of the medium equal to 80 and a dielectric constant for the protein of 4.

Results and Discussion

A few months ago, Yuvaniyama and collaborators published the first X-ray diffraction structure of this enzyme.²³ Examination of this structure from PDB showed that it is incomplete, as there are 55 aminoacid residues missing. Unfortunately from the 55 aminoacids, 46 are part of the junction region. As pointed out by Shallom *et al.*¹⁵ there are important interactions between the DHFR and the TS domains of this enzyme which are necessary for the activation of the TS active site. Since both domains are connected by the junction region, we believe that this region plays an important role on the activity of the whole enzyme. The fact that this region could be important for any possible interactions between both active sites, as well as for the existence of any optimized transfer mechanism of dihydrofolate between both domains, led us to believe that it would be important to obtain a complete model for *pf*DHFR-TS.

One structural aspect that is certainly correct in the experimental structure of *pf*DHFR-TS is the relative position of both domains, this information being very important to define the structure of the junction. Accordingly, we have used the incomplete crystallographic structure of *pf*DHFR-TS to generate a new model for *pf*DHFR-TS. This model is complete, and the only homology modeled parts are the ones relative to the aminoacids missing from the crystallographic structure. The pair alignment between the crystallographic *pf*DHFR-TS structure and its sequence was done in order to match exactly all residues present in the experimental model. The residues absent in the crystal structure fall into two regions. The first region corresponds to the residues located from Val-86 to Ser-95 and the second region to the residues between Leu-234 and Asp-281, corresponding to the 46 missing residues of the junction. The final alignment was submitted to the Optimize mode of SWISS-MODEL server.^{25,26} The analysis of the generated model revealed that the region located between residues Val-86 and Ser-95 was modeled based on the twenty fifth loop of the three-dimensional structure of transketolase from *Sacharomices cerevisiae* (PDB code: 1TRK), while the other region was modeled based on the initial loop of the three-dimensional structure of human carbonic anhydrase II (PDB code: 2CBA) (Figure 1).

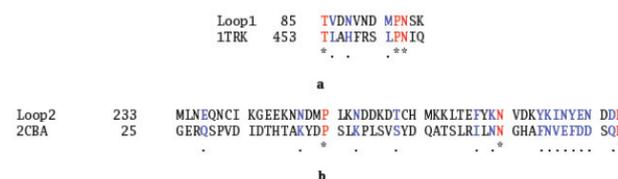


Figure 1. Alignments used to build the *pf*DHFR-TS complete model: a) Alignment of the residues from loop 1 (res Thr85 – Lys96) to the 25th loop of 1TRKA. b) Alignment of the residues from loop2 (Met233 – Asp282) to the 1st loop of 2CBA. Similar residues are shown in blue while the identical ones are shown in red.

This unrefined model for *pf*DHFR-TS was next minimized with the GROMOS force field,²⁷ implemented in the GROMACS 3.1.4 package,^{28,29} using the steepest descent and conjugate gradient algorithms. At this stage, our goal was to relax the modeled enzyme to obtain the best conformations, which would be reflected in good Ramachandran plots.³³ To accomplish that, we conducted several tests varying the convergence criterion. We concluded that 10.00 kcal mol⁻¹ Å⁻¹ gave the best results when using explicit solvent during the minimization.

The final model is shown in Figure 2. Most secondary structure elements of the model are obviously the same as for the crystallographic structure of *pf*DHFR-TS, except for

the two modeled regions. The first region was modeled as a single loop involving residues Val-86 to Ser-95 and the region corresponding to the junction was modeled as 2β sheets, 3 loops and 1 α helix in the following order: β sheet – loop – β sheet – loop – α helix – loop.

In order to check the quality of the model for *pf*DHFR-TS, it was submitted to validation analysis, using the PROCHECK³¹ and WHATCHECK³² programs. In the Ramachandran plot³³ generated in the PROCHECK³¹ program, none of the modeled residues in the loops fall into non allowed regions of the plot, with 95.6% of the amino acid residues falling in favorable regions (73.0% in the most favored plus 22.6% in the additionally favored regions). Regarding the main-chain properties of the modeled enzyme, no considerable bad contacts, C_α tetrahedron distortion nor hydrogen-bond energy problems were found. Moreover, the average G-factor,³¹ the measure of the normality degree of the protein properties, was inside permitted values. Also, no distortions of the side-chain torsion angles were found. Accordingly, the model was considered satisfactory.

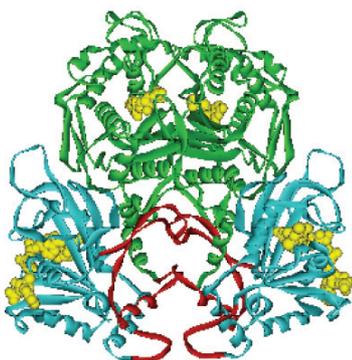


Figure 2. Proposed *P. falciparum* dimeric DHFR-TS homology-model. DHFR domains are shown in blue, TS domains in green, the modeled regions in red and the substrates in the active sites in yellow CPK.

In order to determine if there could exist a mechanism for some kind of optimized substrate transport between both active sites inside the same protein, it was necessary to carefully examine the model. It was evident that there was not any channel present in the model and that the only possible alternative to the channel was some kind of electrostatic path to conduct the substrate from one site to another, as was suggested for *lm*DHFR-TS.¹⁹ Accordingly, the electrostatic contours were calculated for the model and the results are represented for the monomer in Figure 03. Those electrostatic isopotential contours around the model for *pf*DHFR-TS were calculated using the Swiss-PdbViewer program.²⁴ In Figure 3, the blue surface connects

all points having a potential of +0.5 kT and the red surface connects all points having a potential of -0.5 kT, while the TS and DHFR substrates in the active sites are shown in yellow CPK. The analysis of Figure 3a reveals the presence of a continuous region of positive electrostatic potential, of approximately 44 Å of length, connecting the TS and DHFR active sites. This suggests that as TS processes its substrate producing dihydrofolate, this compound should be repelled by the negative electrostatic potential of the TS active site, being conducted to the positive electrostatic potential zone, and from there to the DHFR active site. We also found a narrow region of low negative electrostatic potential decreasing the area of the region of positive electrostatic potential. We believe that this decrease on the positive surface in the model does not hinder the possibility of an electrostatic path being involved in the process of active substrate transport between both sites for two reasons: first in the model there is a positive electrostatic potential continuous region connecting both sites, as shown in Figure 3b, and second, since the negative region is tenuous and strait, the substrate, following the shortest route between both sites, could easily cross over it on its way from one site to the other. It is important to point out that the generation of the electrostatic potential surface on the crystallographic model of *pf*DHFR-TS afforded a discontinuous surface, with at the least two important gaps. This is another indication of the importance of the inclusion of the missing aminoacids from that model, especially those from the junction between both domains. On this regard, it is also important to state that one of the published hypothesis on the putative optimized transport mechanism for dihydrofolate considers that this phenomenon may be related to conformational changes in the protein structure,³⁴ a fact that is certainly related to the junction region.

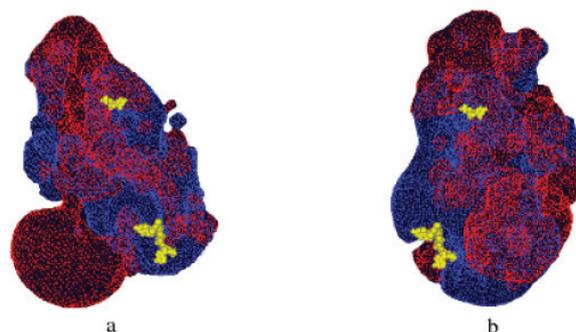


Figure 3. Electrostatic isopotential contours around *P. falciparum* DHFR-TS model. a) front view of the monomeric structure; b) back view of the monomeric structure. In blue are the regions with positive potential and in red the ones with negative potential. The substrates in the active sites are shown in yellow CPK.

Conclusion

We have used the recently reported incomplete crystallographic model of *pf*DHFR-TS²³ to propose a complete virtual model for this protein. This model, which was very useful to propose the existence of an optimized substrate transport mechanism between the TS and the DHFR active sites of this bifunctional enzyme, could be useful for the design of new potential antimalarials.

The fact that the protozoan enzyme is bifunctional while the mammal enzymes are different proteins may serve as a starting point to design new selective antifolates as antimalarials drugs. We have shown the existence of an electrostatic potential path that may serve to conduct dihydrofolate from the TS active site to the DHFR active site, in agreement with similar findings in *lm*DHFR-TS¹⁷ and other literature reports.³⁴ This optimized substrate transport mechanism may be a potential target for new antimalarial drugs, as it may be important for the constant maintenance of thymidylate production by the *Plasmodium* cell. With the missing regions of crystallographic *pf*DHFR-TS modeled and validated, we reached a starting point to proceed on to study this possibility by molecular dynamics simulations of the optimized substrate mechanism transfer between the sites. These studies are now in course in our laboratory. We are also using this model to design new type II antifolates as potential and selective antimalarial drugs that may be active against the forms of the parasite that are resistant to chemotherapy with pyrimethamine and other 2,4-diaminopyrimidines.

Acknowledgments

We thank the Brazilian agencies CNPq, FAPERJ and CAPES for scholarships, studentships and financial support.

References

- Nájera, J. A.; Hempel, J.; *The Burden of Malaria*, World Health Organization: Geneva, 1996.
- Rang, H. P.; Dale, M. M.; Ritter, J. M.; *Farmacologia*, 4th ed., Guanabara Koogan: Rio de Janeiro, 2001.
- The World Health Organization Report*, WHO Publications: Geneva, 1997.
- <http://funasa.gov.br>; accessed in July 2003.
- Winstanley, P. A.; *Parasitol. Today* **2000**, *16*, 146.
- Urdaneta, L.; Plowe, C.; Goldman, I.; Lal, A. A.; *Am. J. Trop. Med. Hyg.* **1999**, *61*, 457.
- Toteja, R.; Nair, L.; Bhasin, V. K.; *Mem. Inst. Oswaldo Cruz* **2001**, *96*, 427.
- Delfino, R.T.; Santos-Filho, O. A.; Figueroa-Villar, J.D.; *J. Braz. Chem. Soc.* **2002**, *13*, 727.
- Hyde, J.E.; *Pharmacol. Ther.* **1990**, *48*, 45.
- Sibley, C. H.; Hyde, J. E.; Sims, P. F. G.; Plowe, C. V.; Kublin, J. G.; Mberu, E. K.; Cowman, A. F.; Winstanley, P. A.; Watkins, W. M.; Nzila, A. M.; *Trends Parasitol.* **2001**, *17*, 582.
- Winstanley, P. A.; *Parasitol. Today* **2000**, *16*, 146.
- Foote, S. J.; Cowman, A.F.; *Acta Tropica* **1994**, *56*, 157.
- Garrett, C. E.; Coderre, J. A.; Meek, T. D.; Garvey, E. P.; Claman, D. M.; Beverley, S. M.; Santi, D. V.; *Mol. Biochem. Parasitol.* **1984**, *11*, 257.
- Bzik, D. J.; Li, W.; Horii, T.; Inselburg, J.; *Proc. Natl. Acad. Sci.* **1987**, *84*, 8360.
- Shallom, S.; Zhang, K.; Jiang, L.; Rathod, P. K.; *J. Biol. Chem.* **1999**, *274*, 37781.
- Liang, P.; Anderson, K. S.; *Biochemistry* **1998**, *37*, 12195.
- Knighton, D. R.; Kan, C. C.; Janson, C. A.; Hostomska, Z.; Welsh, K. M.; Matthews, D. A.; *Nat. Struct. Biol.* **1994**, *1*, 186.
- Hyde, C. C.; Ahmed, S. A.; Padlan, E. A.; Miles, E. W.; Davies, D. R.; *J. Biol. Chem.* **1988**, *263*, 17857.
- Elcock, A. H.; Potter, M. J.; Matthews, D.A.; Knighton, D.R.; McCammon, J. A.; *J. Mol. Biol.* **1996**, *262*, 370.
- Lemck, T.; Christensen, I. T.; Jørgensen, F. S.; *Bioorg. Med. Chem.* **1998**, *7*, 1003.
- Santos-Filho, O. A.; Alencastro, R. B.; Figueroa-Villar, J. D.; *Bioph. Chem.* **2001**, *91*, 305.
- Delfino, R. T.; Santos-Filho, O. A.; Figueroa-Villar, J. D.; *Bioph. Chem.* **2002**, *98*, 287.
- Yuvaniyama, J.; Chitnumsub, P.; Kamchonwongpaisan, S.; Vanichtanankul, J.; Sirawaraporn, W.; Taylor, P.; Walkinshaw, M. D.; Yuthavong, Y.; *Nat. Struct. Biol.* **2003**, *10*, 357.
- Guex, N.; Peitsch, M.C.; *Electrophoresis* **1997**, *18*, 2714.
- Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C.; *Nucleic Acids Res.* **2003**, *31*, 3381.
- Peitsch, M. C.; *BioTechnology* **1995**, *13*, 658.
- Jorgenson, W. L.; Tirado-Reeves, J.; *J. Am. Chem. Soc.* **1988**, *110*, 1657.
- Lindal, E.; Hess, B.; Van der Spoel, D.; *J Mol Mod.* **2001**, *7*, 306.
- Berendsen, H. J. C.; Van der Spoel, D.; Van Drunen, R.; *Comp. Physic. Comm.* **1995**, *91*, 43.
- <http://biotech.embl-heidelberg.de:8400/>; accessed in May 2003.
- Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; Thornton, J.M.; *J. Appl. Cryst.* **1993**, *26*, 283.
- Hoof, R. W. W.; Vriend, G.; Sander, C.; Abola, E. E.; *Nature* **1996**, 381.
- Ramachandran, G. N.; Sasisekharan, V.; *Adv. Prot. Chem.* **1968**, *23*, 283.
- Ivanetich, K. M.; Santi, D. V.; *FASEB J.* **1990**, *4*, 1591.

Received: August 6, 2003

Published on the web: May 10, 2004