

**GENOMIC, PROTEOMIC AND MOLECULAR MODEL
ANALYSES OF TMP KINASE FROM SOME BLOOD-
PARASITIC PROTOZOA. ONE-STEP FORWARD IN
DEVELOPMENT OF BROAD SPECTRUM
ANTIPROTOZOAL DRUGS**

Mahmoud Kandeel^{1,3,5} and Yukio Kitade^{1,2,3,4,}*

¹Department of Biomolecular Science, Faculty of Engineering, ²Center for Emerging Infectious Diseases, ³Center for Advanced Drug Research, ⁴United Graduate School of Drug Discovery and Medical Information Sciences, Gifu University, Yanagido 1-1, Gifu 501-1193, Japan

⁵Department of Pharmacology, Faculty of Veterinary Medicine, Kafr El-Shikh University, Kafr El-Shikh, Egypt

ABSTRACT

We had introduced TMP kinase as a new hopeful antiprotozoal drug target. We provided theoretical model for the eligibility of TMP kinase from Plasmodium falciparum as a new drug target and we confirmed our assumption by experimental trials based on mutational, catalytic and molecular interference assays. Here we extended our new antimicrobial model to include other blood protozoal infections as anaplasmosis and babesiosis. Anaplasma and babesia`s TMP kinase shows important structural differences which can be targeted to develop highly selective antiprotozoal drugs. Therefore, we estimate that we are able to provide new broad spectrum antiprotozoal capable of binding the more than one parasitic species. Furthermore, we investigated the inhibitory activity of one new compound capable of inhibiting the Plasmodium TMP kinase.

INTRODUCTION

The discovery of new antimicrobial targets is the only way to control microbial infections. Lack of an effective vaccination and the spread of drug resistance necessitate the development of new drug targets. In the development of chemotherapeutics, it may be suitable to select targets from pathways present in the micro-organism but absent in the host. Nevertheless, even if a target was common to both parasite and host, slight structural differences could enhance the optimization of a new drug (1). A difference in only one amino acid residue between human and Plasmodium SAH hydrolases led to the discovery of highly selective inhibitor drugs (2). The rational approach to drug development starts with the identification of new molecular targets critical for the life cycle of the parasite. Enzymes of nucleic acid metabolism seem to be important candidates due to their indispensable biological role. In this respect, we had introduced the plasmodiumTMP kinase (PfTMK) as a possible new antimicrobial target (3,4).

The most remarkable and unique feature of PfTMK is its ability to utilize both purine and pyrimidine nucleosides as substrates. In human TMK, the cavity in which the base binds is very tight to accommodate purine bases, and thus reduces specificity to pyrimidines (5, 6).

The inhibitory studies demonstrated that PfTMK can be inhibited by dG (4). Interestingly, dG showed no inhibitory activities against human TMK at concentrations up to 2 mM. Thus, our findings open a new era in development of antimicrobial drug targets.

In order to check the possibility for developing a broader spectrum antimicrobial we investigated the binding of ligands with the modeled active sites of TMP kinase from *Anaplasma marginale* (amTMK) as well as *Babesia bovis* (bbTMK). This paper provides genomic and proteomic aspects of amTMK and bbTMK with special attention to the structure differences for future drug developments.

MATERIALS

Tris (hydroxy methyl) methane (tris-base), N,N,N',N'-tetramethylethylenediamine (TEMED), Agarose (DNA grade), ampicillin (sodium salt), ethidium bromide (EtBr), isopropyl- β -D-galactopyranoside (IPTG), Nucleosides, nucleotides, nucleoside derivatives and phosphoenolpyruvate were obtained from Sigma (MO, USA). *Escherichia coli* JM109, Sephacryl S200HR were obtained from GE Healthcare (NJ, USA). AccuPrime Pfx polymerase was obtained from Invitrogen (CA, USA). Talon metal affinity resin was obtained from Clontech (Palo Alto, CA). The vector pQE-30 was obtained from Qiagen (CA, USA). Lactate dehydrogenase and pyruvate kinase were obtained from Wako Pure Chemicals (Osaka, Japan). Protein concentrations were determined by using a protein assay reagent (Bio-Rad, USA) with bovine serum albumin as the standard.

Methods:

Genomic data collection:

The genedb (<http://www.genedb.org/>) and the plasmdb (<http://plasmodb.org/plasmo/>) were used to obtain general information on the genes and proteins of interest. Information obtained from these sources included protein and nucleotide sequence information, gene annotation, paralogs and orthologs information, metabolic functions, signal peptides, transmembrane domains, predicted protein export domains, protein

expression profiles, isoelectric point (pI), the number of predicted transmembrane helices, any predicted sequence motifs, and the E value of the closest PDB sequence homolog.

Searching for sequence homologues using BLAST and PSI-BLAST:

Protein sequence homologues were searched using the NCBI BLAST (Basic Local Alignment Search Tool) or PSI-BLAST (Position-Specific Iterated-BLAST) servers (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) searching against the non redundant (nr) database, with filtering of low complexity regions. Reverse Position Specific BLAST searches were undertaken against the conserved domain database at the NCBI (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>). BLAST searches of the *Plasmodium falciparum* genome were undertaken using NCBI BLAST from the interface at (<http://www.ncbi.nlm.nih.gov/genome/seq/BlastGen/BlastGen.cgi?taxid=5833>).

Multiple sequence alignment and phylogenetic analysis:

Multiple sequence alignment program for DNA or proteins was carried out using the tools available at (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>). A phylogenic tree was constructed using the amino acid sequences from different prokaryotic and eukaryotic sources. Multiple sequence alignment program (ClustalW2) was used to calculate the best match of the selected sequences. The resultant alignment was used to generate a phylogenic tree, which is visualized by Dendroscope phylogenic tree viewer.

Conserved domain analysis:

Putative domains were searched by the domain prediction program available at (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>). The genomic and molecular information were obtained from Kyoto Encyclopaedia of Genes and Genomes (<http://www.genome.jp/kegg/>).

Genomics and proteomics tools:

ExPASy Proteomics Tools (<http://us.expasy.org/tools/>) and tools available at the website of the European Bioinformatics Institute (<http://www.ebi.ac.uk/Tools/>) were also used for analysis of nucleotide and gene sequences. Protein sequences of target genes were analysed for data such as pI, extinction coefficient and MW for the tagged protein sequence by PROTParam, available from ExPASy (<http://us.expasy.org/cgi-bin/protparam>). PROSITE (<http://us.expasy.org/tools/#pattern>) was used to search for patterns and profiles in the protein sequences of the target genes.

Enzymes three dimensional structure modeling and model analysis:

Requests for protein structure homology-modeling were either submitted to the SWISS-MODEL server (<http://swissmodel.expasy.org/>) or to Protein Homology Analogy Recognition Engine (<http://www.sbg.bio.ic.ac.uk/phyre/html/index.html>). The quality of the produced models were checked with VADAR (Volume, Area, Dihedral Angle Reporter) which is a compilation of more than 15 different algorithms and programs for analyzing and assessing peptide and protein structures from their PDB coordinate data (<http://redpoll.pharmacy.ualberta.ca/vadar/index.html>). The output included Ramachandran plot, fractional accessible surface area, fractional residue volume, stereo/packing quality index, 3D profile quality index, main-chain, side-chain and H-bond statistics. The compatibility of an atomic model (3D) with its own amino acid sequence (1D) was verified at (<http://nihserver.mbi.ucla.edu/Verify>)

_3D/) by using Verify3D program. The MOLprobit server (<http://molprobit.biochem.duke.edu/>) was used to analyze models geometry including rotamer outliers and residues of bad bonds and angles.

Production of recombinant enzyme:

The recombinant *E. coli* cells were grown overnight in an LB medium containing 50 µg/ml ampicillin. The culture was diluted to 1:100 with the same fresh medium and cells were grown at 37°C to mid-log phase ($D_{600}= 0.6$). Induction of expression was carried out by the addition of IPTG to a final concentration of 1 mM and cell growth was continued at 37°C for 4 h. Cells from 2 liters of culture were harvested by centrifugation at $5000 \times g$ for 15 min and then stored at -20°C until use. Cells were lysed in an extraction buffer (25 mM Tris-HCl buffer pH 7.2 containing 150 mM NaCl) and disintegrated by sonication for 40 s (3 cycles with 3 min intervals). The lysate was centrifuged at $16,000 \times g$ for 15 min at 4°C and the precipitate is discarded. The hexahistidine-tagged PfTMK was purified from the soluble cell extract by using talon metal affinity resin. Flow rate was kept constant at 30 ml/min. After binding and washing with the extraction buffer, the protein was eluted using a linear imidazole gradient (from 0 to 150 mM).

New nucleoside analogue development and enzyme inhibition:

For inhibitor studies, the protozoal enzyme was incubated with the natural substrates TMP or dGMP in the presence or absence of an inhibitor and the progress curve was monitored by spectrophotometric. The data was analyzed by non-linear regression analysis. The assay for

inhibitors is complicated by sensitivity of the coupling enzymes. The reaction was carried out in 1-ml final volume of a solution of 50 mM Tris HCl, pH 7.2, 40 mM KCl, 2 mM MgCl₂, 0.5 mM dithiothreitol and variable concentrations of ATP or TMP/dGMP.

RESULTS AND DISCUSSION

Bioinformatics:

We have constructed phylogenetic tree to predict the phylogenetic relationships of amTMK and bbTMK with known TMKs (Fig. 1). The analysis of phylogenetic tree showed two major branches. The first branch represents prokaryotes, while the other represents higher organisms. The plasmodium and babesia TMK were found to be originating at an intermediate branch between the eukaryotic and prokaryotic enzymes. Interestingly, the figure shows that both of the plasmodium and babesia TMKs are forming a common branch. This indicates the high relationship in sequence; and hence, structure and function. This raises the high possibility of applying our previous model on PfTMK on the babesia enzymes. Furthermore, we predict that the inhibition profile of both enzymes could be highly similar and one new compound could be a drug acting on both enzymes.

Fortunately, we had found that anaplasmaTMK is highly divergent from that of higher organisms and especially from the bovines enzyme. Interestingly, the previous finding indicates the abundance of structural differences between the bovines and anaplasma TMKs. Consequently, it is easier to develop new antimicrobials targeting the anaplasma enzyme than those of plasmodium and babesia enzymes.

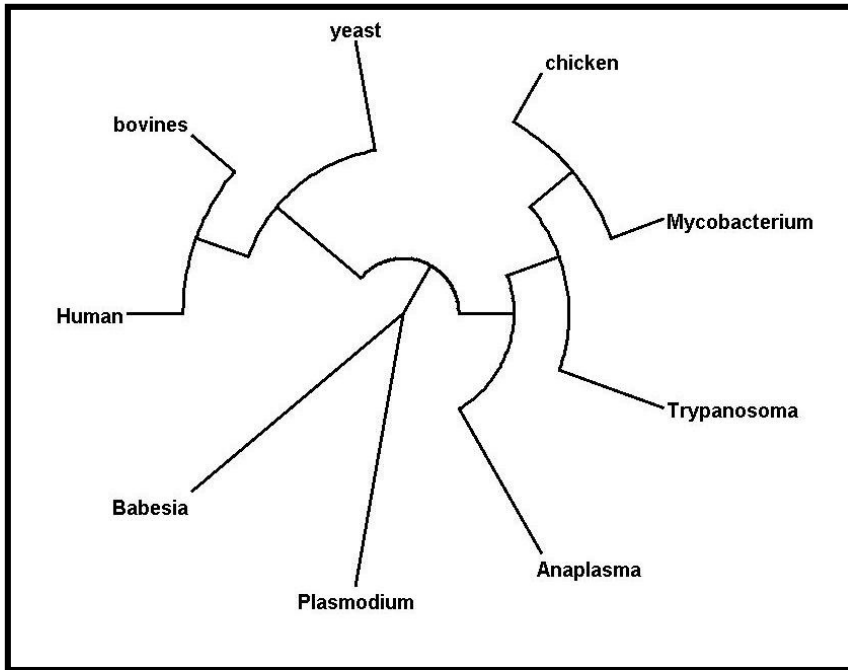


Fig. (1): Circular cladogram showing the basic phylogenetic relationships of *Anaplasma marginale* and *Babesia bovis* TMKs with other prokaryotic and eukaryotic TMKs.

Putative domains:

Babesia TMK showed the classical thymidylate kinase domain, unfortunately, we did not find any other important domain differences from that of the bovines host. Screening of the sequence databases revealed that thymidylate kinase from *Anaplasma* parasites is likely to be functionally different from the bovines enzyme. Analysis of the amino-acid sequence using the domain prediction program revealed that the amTMK possesses a single NMP binding domain with properties different from that of bovine TMK. In addition to the thymidylate kinase, it was predicted that amTMK can possess NTPase and related nucleoside hydrolase e.g. ATPase properties (Fig. 2). The former finding is an attractive finding the field of drug discovery and design.

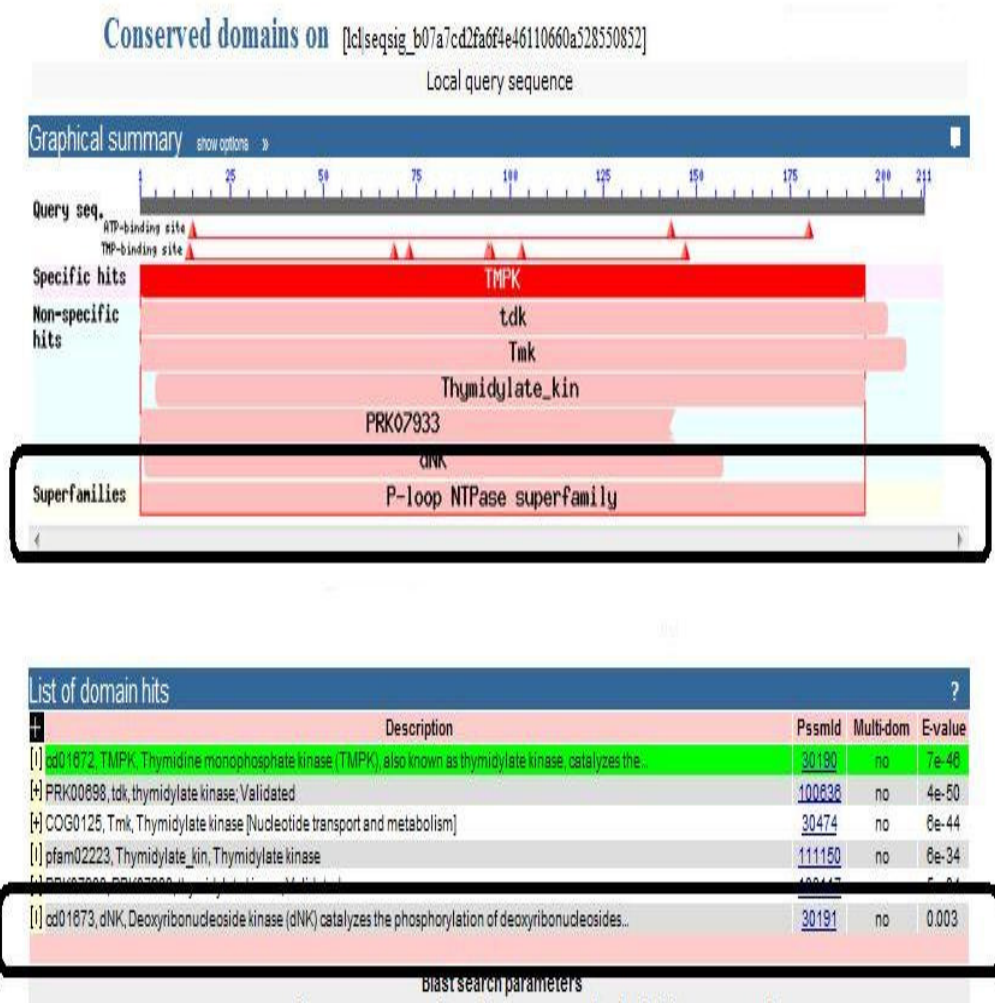


Fig. (2): showing the results of conserved putative domains search of Anaplasma TMK.

Structural differences:

To the best of our knowledge, there is no data available on the structure of bovines, anaplasma and babesia TMKs. So that we developed a three dimensional structure model for these enzymes as described in material and methods.

The model of babesia TMK is highly comparable with that of PfTMK (2,3). This concedes with the phylogentic analysis and contributes to the trials of developing broad spectrum antiprotozoal. The comparison of bovines and amTMKs active sites showed important differences in residual structure (Fig. 3). The p-loop of bovines TMK contain the classical eukaryotic arginine which is completely absent in the anaplasma TMK. Another difference is located at the proximity of active site in which the anaplasma TMK shows the classical prokaryotic glutamine which is replaced by therionine in the bovines TMK. This raises the possibility of development of highly selective drugs.

Enzyme inhibition:

We tested the inhibitory properties of one new compound developed in our lab (Fig. 4). We termed this compound yktk06. The enzyme inhibited PfTMK in the presence of different amount of natural substrate with K_i value of 208 μ M. Interestingly, the compound yktk06 showed more potent inhibitory effect on the deoxyguanylate kinase pathway with K_i of 55 μ M. Therefore, targeting PfTMK not only affects thymidine nucleotide synthesis but may also affect purine nucleotides, and thus it has become an attractive antimicrobial target.

Acknowledgement:

We would like to thank Aya Kato for the synthesis of the compound yktk06. We appreciate the financial support from the Egyptian government to (M.K.). This research was in part supported by Grants-in-Aid for Scientific Research on Priority Area No. 16017239 (to Y.K.).

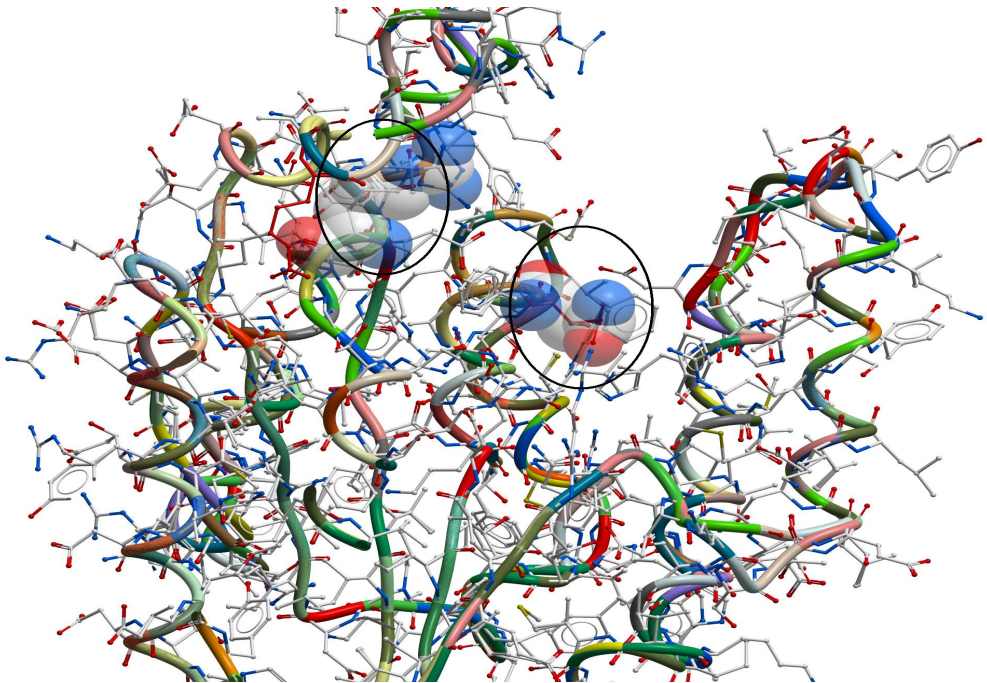


Fig. (3): showing alignment of bovines and anaplasma TMKs.

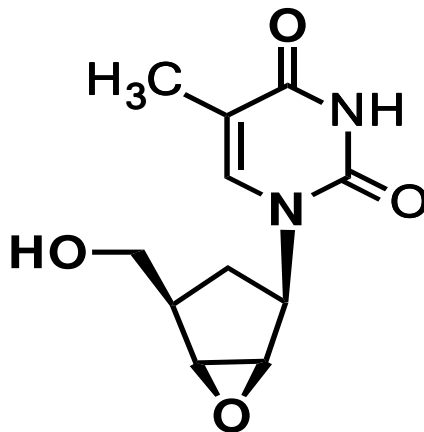


Fig. (4): showing the structure of compound ytk06.

REFERENCES

- **Fidock DA, Rosenthal PJ, Croft SL, Brun R and Nwaka S (2004):** Antimalarial drug discovery: efficacy models for compound screening. *Nat. Rev. Drug Discov.* 3:509-520.
- **Kitade Y, Kojima H, Zulfiqur F, Kim HS, and Wataya Y (2003):** Synthesis of 2-fluoronoraristeromycin and its inhibitory activity against *Plasmodium falciparum* S-adenosyl-L-homocysteine hydrolase. *Bioorg. Med. Chem. Lett.* 13:3963-3975.
- **Kandeel M, Ando T, Kitamura Y, Abdel-Aziz M, Kitade Y. (2009):** Mutational, inhibitory and microcalorimetric analyses of *Plasmodium falciparum* TMP kinase. Implications for drug discovery. *Parasitology.* 136:11-25.
- **Kandeel M, Kitade Y. (2008):** Molecular characterization, heterologous expression and kinetic analysis of recombinant *Plasmodium falciparum* thymidylate kinase. *J Biochem.* 144:245-250.
- **Kenyon GL (1997):** AZT monophosphate knocks thymidylate kinase for a loop. *Nat. Struct. Biol.* 4:595-597.
- **Lavie A, Vetter IR, Konrad M, Goody RS, Reinstein J and Schlichting I (1997):** Structure of thymidylate kinase reveals the cause behind the limiting step in AZT activation. *Nat. Struct. Biol.* 4:601-604.