Bioinformatics of Thymidine metabolism in camels and Trypanosoma evansi: nucleoside deoxyribosyltransferase (NDRT) as a drug target

CURRENT STATUS: UNDER REVIEW

Mahmoud Kandeel
King Faisal University

Corresponding Author
ORCiD: https://orcid.org/0000-0003-3668-5147

Abdulla Al-Taher
King Faisal University

DOI: 10.21203/rs.2.20383/v1

SUBJECT AREAS
Small Animal Medicine

KEYWORDS
Camel, genome, Trypanosoma evansi, thymidine, deoxyribosyltransferase, nucleotide
Abstract

Background *Trypanosoma evansi* (*T. evansi*), the causative agent for surra or camel trypanosomiasis, is characterized by the widest geographic distribution and infects the widest range of hosts among the known trypanosomes. The recent zoonotic importance and increasing reports of drug resistance necessitate the discovery of new drug targets. Drug discovery process entails finding an interesting difference between the host and the parasite.

Results In this study, the thymidine metabolic pathways were compared in camel and *T. evansi*. Metabolic maps, protein sequence comparisons, domain and motifs contents analysis, phylogenetic relations and 3D structure models were used in comparisons. A unique difference in thymidine metabolism was at the level of recycling of thymidine which was performed by thymidine phosphorylase in camel, while this role is *T. evansi* was associated with nucleoside deoxyribosyltransferase (NDRT), which is a unique enzyme for the trypanosome and was absent in camel. Thymidine in *T. evansi* seems to be governed by thymine through NDRT. In contrast to camel, in which thymidine can be produced from thymidylate by the action of 5’-nucleotidase.

Conclusions NDRT can be regarded as a drug target against *T. evansi* for its strict presence in the parasite but not in the host.

Background

By the decoding of camel genome sequences [1], drug discovery studies against camel pathogens are to be accelerated. In investigating the pyrimidine metabolic pathways, finding unique differences in structure, function, sequence or phylogenetic could contribute to the identification of new drug targets.

Thymidine metabolism was traced by the KEGG maps [2–4] to compare the enzymes involved in pathways in thymidine metabolism in camel and *T. evansi*. In this work, comprehensive bioinformatics tools were used to investigate enzyme sequences and its comparisons, domain and motif content, searching similar proteins and its phylogenetic relations.

In this work, the pathway of thymidine metabolism was plotted in both camel and *T. evansi*. Here, we
provide bioinformatics evidence of the presence of unique differences in thymidine metabolism in both camel and *T. evansi*. This could be used for the future drug discovery process.

**Results**

The pyrimidine metabolic maps at KEGG were used to trace the pathways of thymidine [5]. The de novo pathway, thymidine kinase converts thymidine to thymidine monophosphate (TMP). In the catabolic direction, TMP is converted to thymidine by the action of 5′-nucleotidase (Fig. 1). Further, conversion of thymidine to thymine by thymidine/pyrimidine phosphorolase. Whereas, nucleoside deoxyribosyltransferase catalyses the interconversion of thymidine and thymine.

In the camel thymidine pathway, camels were expected to be devoid of nucleoside deoxyribosyltransferase (Fig. 2). Therefore, there is no interconversion between thymine and thymidine in camel pyrimidine metabolism. Thus, the sole source of thymidine in camels is from TMP by the action of 5′-nucleotidase or potentially by the transport of ready formed thymidine from outside the cells.

In *T. evansi*, the enzymes 5′-nucleotidase and thymidine/pyrimidine phosphorolase were not confirmed, suggesting interesting differences in thymidine metabolism, compared with camels. TMP cannot be converted to thymidine in *T. evansi* due to the absence of 5′-nucleotidase. The sole metabolic source of thymidine could be from thymine by the action of nucleoside deoxyribosyltransferase, which is absent in camel (Fig. 3). The list of enzymes of thymidine pathway as well as the predicted content in camel and *T. evansi* are provided in Tables 1-3.

**Camel thymidine/pyrimidine phosphorolase (TP)**

Thymidine phosphorolase (EC 2.4.2.4; TP): pyrimidine phosphorolase; thymidine orthophosphate deoxyribosyltransferase; animal growth regulators, blood platelet-derived endothelial cell growth factors; blood platelet-derived endothelial cell growth factor; deoxythymidine phosphorolase; gliostatins; pyrimidine deoxynucleoside phosphorolase; thymidine:phosphate deoxy-D-ribosyltransferase. It is a glycosyltransferase stimulate reversible dephosphorylation of thymidine phosphate supplying thymine and 2-deoxy -D-ribose 1-phosphate. In addition, it catalyzes deoxyribosyltransferase reactions that catalyzed by nucleoside deoxyribosyltransferase in various
tissues. The enzyme can also convert thymidine to thymine, which is not used by the trypanosomes [6]. The properties of the phosphorylases differ significantly from prokaryotes to those of the mammalian animals. *Giardia lamblia* is mainly dependent on salvage synthesis for its pyrimidine requirements. In which, one enzyme of uridine, deoxyuridine and thymidine phosphorylases are responsible for the activities of three enzymes. The enzyme could utilize both uracil and thymine as substrates [7]. Complicated infection with Mycoplasmas reduced the efficiency of anticancer and antiviral nucleoside analogue-based therapies due to the presence of Mycoplasma TP. Impaired activity of TP provoked the elevated uptake and inclusion of deoxyuridine and uracil but thymidine uptake was not affected. Thus, enzymes of Mycoplasma nucleotide synthesis pathway are prospective targets for imminent creation of antibiotics [8]. In *Trypanosoma cruzi*, a phosphorylase activity was detected but was more specific to uridine phosphorylase without specificity to thymidine or purine phosphorylase [9, 10].

The obtained sequence of camel TP was low quality protein and very short amino acids length compared with human TP. The recorded sequence of camel TP was 202 amino acids, bearing 18.44% similarity rate, compared with 482 amino acids length in human TP (Fig. 4). The same finding applies for the TP sequence in the three camel species, dromedary, Bactrian and ferus camels, in which *Camelus ferus* showed the shortest length of 162 amino acids (Fig 5). Comparison of TPs from different prokaryotic and eukaryotic sources (Fig. 6) showed a general low similarity rate from 9.5-41.8% among the tested species. Owing to this difference, the camel TPs were forming a monophyletic group that shared common origin with the prokaryotic TPs and not the vertebrate TPs (Fig. 7). The motif search retrieved zero hits by sing the motif finder, pfam and prosite prediction tools, while a phosphorylase domain was predicted by using the NCBI conserved domain search tool.

**evansi nucleoside deoxyribosyltransferase (NDRT)**

This enzyme was found to be a unique enzyme for *T. evansi* and not present in camel, suggesting its use as a safe drug target. After searching the gene databases, NDRT could be predicted in some protozoa and bacteria e.g. lactobacillus, Leishmania spp., and Trypanosomes. A previous study investigated the crystal structure of *T. brucei* NDRT and found that its structure is highly similar to
NDRT from *Lactobacillus helveticus* [11]. In the previous study, several crystal bound compounds were tested against the blood forms of *T. brucei* and found weak inhibition of parasite growth with IC50 values above 100 µM.

Sequence comparisons between *T. evansi*, *Lactobacillus fecum* and *Enterococcus fecum* NDRTs revealed 17.68-21.21% sequence similarity (Fig. 8). Comparison of *T. evansi* and *T. brucei* NDRTs revealed 100% similarity. This may account for sharing the common features in the published structure, function and inhibition.

The models for *T. evansi* NDRT were built at Swiss-Model server [12]. The model of *T. evansi* was predicted based on the deposited PDB ID 2a0k, which is the NDRT from *T. brucei* (Fig. 9). The modeling statistics comprised 99% coverage, 0.6 sequence similarity and 98% sequence identity.

**Discussion**

The blood-protozoa *T. evansi* is the causative agent for Surra or camel trypanosomiasis, which is a devastating disease in camels and other animals. *T. evansi* is characterized by the widest geographic distribution and large number of hosts including camels, large and small ruminant farm animals, bats, equines, pigs, carnivores, deer, gazelles and elephants [13]. The recent affections of humans with *T. evansi* raised its public health importance [14, 15].

The discovery of new drug targets against *T. evansi* is essential to combat this infection, which is ubiquitous in animals and has zoonotic importance. The ideal target is that is present in the parasite and absent in host. Recently, we have introduced several new antimicrobial targets, which were investigated by bioinformatics tools to elucidate the differences between the host and microbial proteins [16–24]. In this context, pyrimidine enzymes are indispensable for life owing to its rule in the synthesis nucleic acids.

NDRT was found to an important drug target in *T. brucei* [11]. Additionally, owing to its efficient broad-spectrum activity, NDRT had been used as a catalyst in the industrial synthesis of nucleotides [25, 26] and antiviral drugs [27].

In this study, NDRT was confirmed in *T. evansi*. While, by bioinformatics tools, the enzyme was not detected in camels. This renders NDRT as an attractive drug target against *T. evansi*. Future drug
discovery studies targeting T. evansi NDRT is expected to yield safe and specific inhibitory compounds.

Conclusions
In analyzing the thymidine pathways in both camel and T. evansi, we could predict the differences in nucleosides recycling. The recycling of thymidine nucleosides in camels is governed by thymidine phosphorylase, which is absent in T. evansi. On the other hand, the thymidine recycling in T. evansi is performed by nucleoside deoxyribosyltransferase, which a unique enzyme for the protozoan. This signifies the importance of nucleoside deoxyribosyltransferase as a drug target in T. evansi.

Methods

**Construction of metabolic map**
The website of Kyoto Encyclopedia of genes and genomes (KEGG) was used to retrieve the thymidine pathways maps. The structure of nucleotides were written by ChemDraw software.

**Retrieval of protein sequences**
The database of proteins at NCBI was used to obtain the sequences of camel enzymes (http://www.ncbi.nlm.nih.gov/protein). The sequences of T. evansi proteins were retrieved from the Kinetoplastom genome project (http://tritrypdb.org/tritrypdb/). The sequence files in FASTA format were stored and processed by several software including CLC main workbench, Geneious and Genedoc.

**Basic Local Alignment Search Tool (BLAST)**: BLAST search was performed at the NCBI search engine [28]. The BLAST tool was used to find orthologues with high similarity with the retrieved sequences by searching a database of non-redundant (nr) protein sequences.

**Proteins multiple and pairwise sequence alignment** Clustal omega was used to align the protein sequences [29]. The obtained alignment file was processed by CLC main workbench.

**Construction of phylogenetic tree**
The phylogenetic tree was implemented from the obtained alignment files and visualized by Dendroscope phylogenetic tree viewer [30] or CLC main workbench [31].

**Putative domains and motifs search** The NCBI domain prediction program available at
(http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml) [32] was used to map the domains and motifs of the retrieved proteins. Domains and motifs were also explored at KEGG motif finder tool.

**Proteomic and genomic tools** the tools available at ExPASy Proteomics tools (http://us.expasy.org/tools/) [33] and (http://www.ebi.ac.uk/Tools/) [34] were used to analyze the protein sequences.

**Building structure model.** A model of structure, based on *T. evansi* sequence, was built by using the SwissModel server.

**Declarations**

**Acknowledgements**

The authors acknowledge the financial support of this project by King Abdul-Aziz City for Science and Technology (KACST), Basic Research Programs, National Transformation Program, under Research and Development Grants Program for National Research Institutions and Centers (GPURC), Kingdom of Saudi Arabia (Grant No. 2-17-04-004-0001).

**Funding**

The authors acknowledge the financial support of this project by King Abdul-Aziz City for Science and Technology (KACST), Basic Research Programs, National Transformation Program, under Research and Development Grants Program for National Research Institutions and Centers (GPURC), Kingdom of Saudi Arabia (Grant No. 2-17-04-004-0001).

**Author’s contribution**

MK and AA designed the study, performed research, analyzed data, contributed new methods or models and wrote the paper.

**Availability of data and materials**

All data are within the manuscript.

**Compliance with Ethical Standards**

**Disclosure of potential competing interest**
The authors declare no conflict of interest.

**Research involving Human Participants and/or Animals**

No animals / humans were used for the studies of this research.

**Informed consent**

Not applicable in this research.

**References**

1. Jirimutu, Wang Z, Ding G, Chen G, Sun Y, Sun Z, Zhang H, Wang L, Hasi S, Zhang Y et al: *Genome sequences of wild and domestic bactrian camels*. Nature communications 2012, 3:1202.

2. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K: *KEGG: new perspectives on genomes, pathways, diseases and drugs*. Nucleic acids research 2016, 45(D1):D353-D361.

3. Ogata H, Goto S, Fujibuchi W, Kanehisa M: *Computation with the KEGG pathway database*. Biosystems 1998, 47(1-2):119-128.

4. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T: *KEGG for linking genomes to life and the environment*. Nucleic acids research 2007, 36(suppl_1):D480-D484.

5. Kanehisa M, Goto S: *KEGG: kyoto encyclopedia of genes and genomes*. Nucleic acids research 2000, 28(1):27-30.

6. Al Chalabi K, Gutteridge WE: *Catabolism of deoxythymidylate in some trypanosomatids*. Parasitology 1977, 74(3):299-312.

7. Lee CS, Jimenez BM, O'Sullivan WJ: *Purification and characterization of uridine (thymidine) phosphorylase from Giardia lamblia*. Molecular and biochemical parasitology 1988, 30(3):271-277.

8. Wang L, Schmidl SR, Stulke J: *Mycoplasma pneumoniae thymidine phosphorylase*. Nucleosides, nucleotides & nucleic acids 2014, 33(4-6):296-304.
9. Silva RG, Vetticatt MJ, Merino EF, Cassera MB, Schramm VL: Transition-state analysis of Trypanosoma cruzi uridine phosphorylase-catalyzed arsenolysis of uridine. *Journal of the American Chemical Society* 2011, 133(25):9923-9931.

10. Silva RG, Kipp DR, Schramm VL: Constrained bonding environment in the Michaelis complex of Trypanosoma cruzi uridine phosphorylase. *Biochemistry* 2012, 51(34):6715-6717.

11. Bosch J, Robien MA, Mehlin C, Boni E, Riechers A, Buckner FS, Van Voorhis WC, Myler PJ, Worthey EA, DeTitta G: Using fragment cocktail crystallography to assist inhibitor design of Trypanosoma brucei nucleoside 2-deoxyribosyltransferase. *Journal of medicinal chemistry* 2006, 49(20):5939-5946.

12. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L et al: *SWISS-MODEL*: homology modelling of protein structures and complexes. *Nucleic Acids Res* 2018, 46(W1):W296-w303.

13. Desquesnes M, Holzmuller P, Lai DH, Dargantes A, Lun ZR, Jittapalapong S: *Trypanosoma evansi and surra*: a review and perspectives on origin, history, distribution, taxonomy, morphology, hosts, and pathogenic effects. *Biomed Res Int* 2013, 2013:194176.

14. Joshi PP, Shegokar VR, Powar RM, Herder S, Katti R, Salkar HR, Dani VS, Bhargava A, Jannin J, Truc P: Human trypanosomiasis caused by Trypanosoma evansi in India: the first case report. *Am J Trop Med Hyg* 2005, 73(3):491-495.

15. Vanhollebeke B, Truc P, Poelvoorde P, Pays A, Joshi PP, Katti R, Jannin JG, Pays E: Human Trypanosoma evansi infection linked to a lack of apolipoprotein L-I. *N Engl J Med* 2006, 355(26):2752-2756.

16. Kandeel M, Nakanishi M, Ando T, El-Shazly K, Yosef T, Ueno Y, Kitade Y: Molecular cloning, expression, characterization and mutation of Plasmodium
falciparum guanylate kinase. *Molecular and biochemical parasitology* 2008, 159(2):130-133.

17. Kandeel M, Kitade Y: **Molecular characterization, heterologous expression and kinetic analysis of recombinant Plasmodium falciparum thymidylate kinase.** *J Biochem* 2008, 144(2):245-250.

18. Kandeel M, Ando T, Kitamura Y, Abdel-Aziz M, Kitade Y: **Mutational, inhibitory and microcalorimetric analyses of Plasmodium falciparum TMP kinase. Implications for drug discovery.** *Parasitology* 2009, 136(1):11-25.

19. Kandeel M, Miyamoto T, Kitade Y: **Bioinformatics, enzymologic properties, and comprehensive tracking of Plasmodium falciparum nucleoside diphosphate kinase.** *Biol Pharm Bull* 2009, 32(8):1321-1327.

20. Kandeel M, Kitade Y: **Substrate specificity and nucleotides binding properties of NM23H2/nucleoside diphosphate kinase homolog from Plasmodium falciparum.** *J Bioenerg Biomembr* 2010, 42(5):361-369.

21. Kandeel M, Kitade Y: **The substrate binding preferences of Plasmodium thymidylate kinase.** *Biol Pharm Bull* 2011, 34(1):173-176.

22. Kandeel M, Kitade Y: **Binding dynamics and energetic insight into the molecular forces driving nucleotide binding by guanylate kinase.** *J Mol Recognit* 2011, 24(2):322-332.

23. Kato A, Yasuda Y, Kitamura Y, Kandeel M, Kitade Y: **Carbocyclic thymidine derivatives efficiently inhibit Plasmodium falciparum thymidylate kinase (PfTMK).** *Parasitol Int* 2012, 61(3):501-503.

24. Noguchi Y, Yasuda Y, Tashiro M, Kataoka T, Kitamura Y, Kandeel M, Kitade Y: **Synthesis of carbocyclic pyrimidine nucleosides and their inhibitory activities against Plasmodium falciparum thymidylate kinase.** *Parasitol Int*
25. Crespo N, Sánchez-Murcia PA, Gago F, Cejudo-Sanches J, Galmes MA, Fernández-Lucas J, Mancheño JM: 2’-Deoxyribosyltransferase from Leishmania mexicana, an efficient biocatalyst for one-pot, one-step synthesis of nucleosides from poorly soluble purine bases. *Applied microbiology and biotechnology* 2017, 101(19):7187-7200.

26. Fresco-Taboada A, de la Mata I, Arroyo M, Fernández-Lucas J: New insights on nucleoside 2’-deoxyribosyltransferases: a versatile biocatalyst for one-pot one-step synthesis of nucleoside analogs. *Applied microbiology and biotechnology* 2013, 97(9):3773-3785.

27. Hanrahan JR, Hutchinson DW: The enzymatic synthesis of antiviral agents. *Journal of biotechnology* 1992, 23(2):193-210.

28. Madden T: The BLAST sequence analysis tool. In: *The NCBI Handbook [Internet]* 2nd edition. edn.: National Center for Biotechnology Information (US); 2013.

29. Sievers F, Higgins DG: Clustal Omega, accurate alignment of very large numbers of sequences. In: *Multiple sequence alignment methods*. edn.: Springer; 2014: 105-116.

30. Huson DH, Richter DC, Rausch C, Dezulian T, Franz M, Rupp R: Dendroscope: An interactive viewer for large phylogenetic trees. *BMC bioinformatics* 2007, 8(1):460.

31. Sequencing H: CLC Genomics Workbench. In.: Workbench; 2011.

32. Marchler-Bauer A, Anderson JB, Cherukuri PF, DeWeese-Scott C, Geer LY, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z: CDD: a Conserved Domain Database for protein classification. *Nucleic acids research* 2005, 33(suppl_1):D192-D196.

33. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A: ExPASy: the
proteomics server for in-depth protein knowledge and analysis. *Nucleic acids research* 2003, 31(13):3784-3788.

34. Labarga A, Valentin F, Anderson M, Lopez R: Web services at the European bioinformatics institute. *Nucleic acids research* 2007, 35(suppl_2):W6-W11.

Tables

**Table 1. Enzymes involved in metabolic pathways of thymidine**

| ID (E.C. number) | Definition (Enzyme name) | comment |
|------------------|---------------------------|---------|
| 19               | 2.7.1.21 Thymidine kinase | in camel and Trypanosoma |
| 54               | 3.1.3.89 5'-deoxynucleotidase; 2'-deoxyribonucleoside 5'-monophosphate phosphohydrolase | *E. coli* only |
| 30               | 3.1.3.5 5'-nucleotidase; uridine 5'-nucleotidase | Present in camel but not in Trypanosoma |
| 17               | 2.4.2.6 Nucleoside deoxyribosyltransferase | This is unique enzyme for Trypanosoma not present in camel |
| 16               | 2.4.2.4 Thymidine phosphorylase; pyrimidine phosphorylase | Present in camel but not in Trypanosoma |

**Table 2. The expected enzymes involved in metabolic pathways of thymidine in camels**

| ID (E.C. number) | Definition (Enzyme name) | comment |
|------------------|---------------------------|---------|
| 19               | 2.7.1.21 Thymidine kinase | in camel and Trypanosoma |
| 30               | 3.1.3.5 5'-nucleotidase; uridine 5'-nucleotidase | Present in camel but not in Trypanosoma |
| 16               | 2.4.2.4 Thymidine phosphorylase; pyrimidine phosphorylase | Present in camel but not in Trypanosoma |

**Table 3. The expected enzymes involved in metabolic pathways of thymidine in Trypanosoma evansi**

| ID (E.C. number) | Definition (Enzyme name) | comment |
|------------------|---------------------------|---------|
| 19               | 2.7.1.21 Thymidine kinase | in camel and Trypanosoma |
| 17               | 2.4.2.6 Nucleoside deoxyribosyltransferase | This is unique enzyme for Trypanosoma not present in camel |

Figures
Figure 1

The proposed metabolic pathways of thymidine.

Figure 2

The proposed metabolic pathways of thymidine in camels.

Figure 3

The proposed metabolic pathways of thymidine in Trypanosoma evansi.
Figure 4

Multiple sequence alignment of dromedary camel and human thymidine/pyrimidine phosphorylase. The upper panel represents sequence comparison statistics. The upper right diagonal region is explains the number of differences between two sequences, while the lower left diagonal region explains the % of identity between two sequences.
Figure 5

Multiple sequence alignment of dromedary and Bactrian camels thymidine/pyrimidine phosphorylase.
Figure 6
The sequence comparison statistics of thymidine/pyrimidine phosphorylase. The upper right diagonal region is explains the number of differences between two sequences, while the lower left diagonal region explains the % of identity between two sequences.
Figure 7

Phylogram of camel thymidine/pyrimidine phosphorylase in relation to a set of prokaryotic and eukaryotic organisms.
Multiple sequence alignment of *Trypanosoma evansi*, *Lactobacillus fecum* and *Enterococcus fecum* nucleoside deoxyribosyltransferase. The upper panel represents sequence comparison statistics. The upper right diagonal region is explains the number of differences between two sequences, while the lower left diagonal region explains the % of identity between two sequences.
Molecular model of T. evansi NDRT. The model was built by using SwissModel.