Bioinformatics of thymidine metabolism in Trypanosoma evansi: exploring nucleoside deoxyribosyltransferase (NDRT) as a drug target

Mahmoud Kandeel ( mkandeel@kfu.edu.sa )
King Faisal University  https://orcid.org/0000-0003-3668-5147

Abdulla Al-Taher
King Faisal University

Research article

Keywords: Camel, Trypanosoma, Protozoa, Drug discovery

DOI: https://doi.org/10.21203/rs.2.20383/v2

License: ☺️  © This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

*Trypanosoma evansi*, the causative agent of surra or camel trypanosomiasis, is characterized by the widest geographic distribution and host range among the known trypanosomes. Its zoonotic importance and increasing evidence of drug resistance necessitate the discovery of new drug targets. The drug discovery process entails finding an exploitable difference between the host and the parasite. In this study, the thymidine metabolic pathways in camel and *T. evansi* were compared by analyzing their metabolic maps, protein sequences, domain and motif contents, phylogenetic relationships, and 3D structure models. The two organisms were revealed to recycle thymidine differently: performed by thymidine phosphorylase in camels (*Camelus* genus), this role in *T. evansi* was associated with nucleoside deoxyribosyltransferase (NDRT), a unique trypanosomal enzyme absent in camels. Thymidine in *T. evansi* seems to be governed by thymine through NDRT, whereas in camels, thymidine can be produced from thymidylate via 5'-nucleotidase. As a result, NDRT may be a promising drug target against *T. evansi*.

Introduction

Decoding camel (*Camelus* genus) genome sequences (Jirimutu et al., 2012) will accelerate drug discovery studies against camel pathogens. One specific line of inquiry is investigating pyrimidine metabolic pathways, and finding unique differences in structure, function, sequence, or phylogeny associated with these pathways could contribute to identifying new drug targets. KEGG maps have been used (Kanehisa et al., 2007; Kanehisa et al., 2016; Ogata et al., 1998) to compare the enzymes involved in thymidine metabolism pathways in camels and *T. evansi*.

In this work, comprehensive bioinformatic tools were used to investigate enzyme sequences in camels and *T. evansi* and compare their domains and motif content, searching for similar proteins and identifying their phylogenetic relationships. We plotted the pathway of thymidine metabolism in the two organisms, providing bioinformatics evidence of unique differences between camels and *T. evansi* that could be exploited for future drug discovery processes.

Materials And Methods

Construction of the metabolic map

The Kyoto Encyclopedia of Genes and Genomes (KEGG) website was used to retrieve the thymidine pathway maps. The map was accessed at the following link (https://www.genome.jp/kegg-bin/show_pathway?map00240). The structure of nucleotides and the pathways were written by ChemDraw software (Fig. 1).

Retrieval of protein sequences
The NCBI protein database was used to obtain the sequences of camel enzymes (http://www.ncbi.nlm.nih.gov/protein), and the *T. evansi* protein sequences were retrieved from the Kinetoplastom genome project (http://tritrypdb.org/tritrypdb/). The sequence files in FASTA format were stored and processed by multiple software programs, including the CLC main workbench, Geneious, and Genedoc. The accession numbers of the retrieved sequences are provided in Figures 2–5.

**Basic local alignment search tool (BLAST)**

A BLAST search was performed using the NCBI search engine (Madden, 2013). The BLAST tool was used to find orthologues with high similarity to the retrieved sequences by searching a database of non-redundant (nr) protein sequences using the default settings. The retrieved camel and *T. evansi* sequences were input into the protein BLAST server at (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome) using the FASTA sequences and searching the nr protein database and blastp algorithm.

**Alignment of multiple and pairwise protein sequences**

Clustal Omega was used to align the protein sequences (Sievers and Higgins, 2014). The obtained alignment file was processed by the CLC main workbench. The aligned sequences were subjected to pairwise comparisons to determine the identity% and the number of amino acid differences.

**Construction of phylogenetic tree**

The phylogenetic tree was built from the obtained alignment files and visualized by the Dendroscope phylogenetic tree viewer (Huson *et al*., 2007) or CLC main workbench (Qiagen software, Hilden, Germany). The tree was constructed by the neighbor-joining method. Bootstrap analysis was performed and set to 100 replicates.

**Putative domain and motif search**

The NCBI domain prediction program available at (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml) (Marchler-Bauer *et al*., 2005) was used to map the domains and motifs of the retrieved proteins. Domains and motifs were also explored using a KEGG motif finder tool.

**Proteomic and genomic tools**

The tools available at ExPASy Proteomics (http://us.expasy.org/tools/) (Gasteiger *et al*., 2003) and (http://www.ebi.ac.uk/Tools/) (Labarga *et al*., 2007) were used to analyze the protein sequences. The tools utilized comprised Clustal Omega for protein alignment, the Swiss Model for protein structure model construction, the Prosite for protein family and domain checking, and BLAST protein for the protein search. All analyses were run using the default parameters.

**Building structure model**
Based on the *T. evansi* sequence, a structure model was built using the Swiss Model server. The sequence of *T. evansi* NDRT was entered in FASTA format in SWISS-MODEL project mode (https://swissmodel.expasy.org/interactive).

**Results**

The pyrimidine metabolic maps at KEGG were used to trace the thymidine pathways (Kanehisa and Goto, 2000). In the de novo pathway, thymidine kinase converts thymidine to thymidine monophosphate (TMP). In catabolism, TMP is converted to thymidine by the action of 5’-nucleotidase (Figure 1A). Further, thymidine is converted to thymine by thymidine/pyrimidine phosphorylase, whereas deoxyribosyltransferase catalyzes the interconversion of thymidine and thymine.

In the camel thymidine pathway, camels were expected to be devoid of NDRT (Figure 1B). 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 contrast, in the *T. evansi* metabolic map, the enzymes 5’-nucleotidase and thymidine/pyrimidine phosphorylase were not confirmed, suggesting notable differences in thymidine metabolism. 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 camels (Figure 1C). The list of enzymes of the thymidine pathway and the predicted content in camels and *T. evansi* are provided in Tables 1–3.

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

Thymidine phosphorylase (EC 2.4.2.4, TP, pyrimidine phosphorylase, thymidine orthophosphate deoxyribosyltransferase, pyrimidine deoxynucleoside phosphorylase, or thymidine:phosphate deoxy-D-ribosyltransferase) is a glycosyltransferase that stimulates reversible dephosphorylation of thymidine phosphate supplying thymine and 2-deoxy-D-ribose 1-phosphate. In addition, it catalyzes deoxyribosyltransferase reactions catalyzed by NDRT in various tissues. The enzyme can also convert thymidine to thymine, which is not used by trypanosomes (Al Chalabi and Gutteridge, 1977; Elamin et al., 2016). The properties of the phosphorylases differ significantly from prokaryotes to those of mammalian animals. For example, in *Giardia lamblia*, mainly dependent on salvage synthesis for its pyrimidine requirements, uridine, deoxyuridine, and thymidine are phosphorylated with one broad-spectrum nucleoside phosphorylase, which can also utilize uracil and thymine as substrates (Lee et al., 1988). Complicated infection of cancer tissues or cells with mycoplasmas reduced the efficiency of anticancer and antiviral nucleoside analog-based therapies due to the presence of mycoplasma TP. In *Mycoplasma pneumoniae* cultures, impaired activity of TP provoked elevated uptake and inclusion of deoxyuridine and uracil, but thymidine uptake was not affected. Thus, the enzymes of the mycoplasma nucleotide synthesis pathway are prospective antibiotic targets (Wang et al., 2014). In *Trypanosoma cruzei*,
phosphorylase activity was detected but was more specific to uridine phosphorylase without specificity to thymidine or purine phosphorylase (Silva et al., 2012; Silva et al., 2011).

The obtained sequence of camel TP was very short and of a lower-quality protein than that of human TP. The recorded sequence of camel TP was 202 amino acids, only 18.44% similar to the 482 amino acid human TP. Similarly, of the TP sequences in the three camel species (dromedary, Bactrian, and wild camels), Camelus ferus had the shortest length, 162 amino acids (Figure 2). Comparing TPs from different prokaryotic and eukaryotic sources (Figure 3) showed a general low similarity rate from 9.5–41.8% among the tested species. Due to this difference, the camel TPs formed a monophyletic group that shared a common origin with the prokaryotic TPs and not the vertebrate TPs (Figure 4). The motif search retrieved zero hits using the motif finder and the pfam and prosite prediction tools, while a phosphorylase domain was predicted by the NCBI conserved domain search tool.

**Nucleoside deoxyribosyltransferase (NDRT) in *T. evansi***

This enzyme was found to be unique for *T. evansi* and not present in camels, 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* (Bosch et al., 2006). In the previous study, several crystal-bound compounds tested against the blood forms of *T. brucei* demonstrated 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 (Figure 5). Comparison of *T. evansi* and *T. brucei* NDRTs revealed 100% similarity. This finding may account for the observed similarity in the common features in the published structure, function, and inhibition.

The models for *T. evansi* NDRT were built using the Swiss Model server (Waterhouse et al., 2018). The model of *T. evansi* was predicted based on the deposited PDB ID 2a0k, which is the NDRT from *T. brucei* (Figure 6). The *T. evansi* NDRT modeled in dimer form with a monomer is 151 amino acids in length and composed of three beta sheets and five alpha helices. The dimer composition and interface are similar to the previously reported *T. brucei* structure (Bosch et al., 2006). 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, a devastating disease in camels and other animals. *T. evansi* is characterized by the widest trypanosome geographic distribution and a large number of hosts, including camels, large and small ruminant farm animals, bats, equines, pigs, carnivores, deer, gazelles, and elephants (Desquesnes et al., 2013). The recent human
infections of humans with *T. evansi* has increased its public health importance (Joshi *et al*., 2005; Vanhollebeke *et al*., 2006).

The discovery of new drug targets against *T. evansi* is essential to combat this ubiquitous, zoonotic infection. The ideal target is present in the parasite and absent in the host. Recently, we have introduced several new antimicrobial targets, which were investigated using bioinformatics tools to elucidate the differences between the host and microbial proteins (Kandeel and Al-Taher, 2020a, b, 2021; Kandeel *et al*., 2020; Kandeel and Kitade, 2011a, b; Kato *et al*., 2012; Noguchi *et al*., 2013). In this context, pyrimidine enzymes are indispensable for life due to their role in nucleic acid synthesis.

NDRT was found to an important drug target in *T. brucei* (Bosch *et al*., 2006). Additionally, due to its efficient broad-spectrum activity, NDRT has been used as a catalyst in the industrial synthesis of nucleotides (Crespo *et al*., 2017; Fresco-Taboada *et al*., 2013) and antiviral drugs (Hanrahan and Hutchinson, 1992). In this study, NDRT was confirmed in *T. evansi*, and by bioinformatics tools, it was not detected in camels. This finding reveals NDRT as an attractive drug target against *T. evansi*. Future drug discovery studies targeting *T. evansi* NDRT are expected to yield safe and specific inhibitory compounds.

**Conclusions**

By analyzing the thymidine pathways in both camels and *T. evansi*, we could predict the differences in nucleoside recycling. The recycling of thymidine nucleosides in camels is governed by thymidine phosphorylase, absent in *T. evansi*. Similarly, the thymidine recycling in *T. evansi* is performed by NDRT, an enzyme unique to this protozoan, indicating its importance as a drug target in *T. evansi*.

**Declarations**

**ACKNOWLEDGMENTS**

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).

**CONFLICT OF INTEREST**

The authors declare that they have no conflict of interests

**CONTRIBUTION OF EACH AUTHOR**

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

**References**


https://doi.org/10.1093/nar/gky427.

## Tables

### Table 1. All enzymes known to be involved in metabolic pathways of thymidine

<table>
<thead>
<tr>
<th>ID (E.C. no.)</th>
<th>Enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.7.1.21</td>
<td>Thymidine kinase</td>
</tr>
<tr>
<td>3.1.3.89</td>
<td>5'-deoxynucleotidase; 2'-deoxyribonucleoside 5'-monophosphate phosphohydrolase</td>
</tr>
<tr>
<td>3.1.3.5</td>
<td>5'-nucleotidase; uridine 5'-nucleotidase</td>
</tr>
<tr>
<td>2.4.2.6</td>
<td>nucleoside deoxyribosyltransferase</td>
</tr>
<tr>
<td>2.4.2.4</td>
<td>thymidine phosphorylase; pyrimidine phosphorylase</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>ID (E.C. no.)</th>
<th>Enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.7.1.21</td>
<td>Thymidine kinase</td>
</tr>
<tr>
<td>3.1.3.5</td>
<td>5'-nucleotidase; uridine 5'-nucleotidase</td>
</tr>
<tr>
<td>2.4.2.4</td>
<td>thymidine phosphorylase; pyrimidine phosphorylase</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>ID (E.C. no.)</th>
<th>Enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.7.1.21</td>
<td>Thymidine kinase</td>
</tr>
<tr>
<td>2.4.2.6</td>
<td>nucleoside deoxyribosyltransferase</td>
</tr>
</tbody>
</table>

## Figures
Figure 1

**Figure 2**

Multiple sequence alignment of dromedary and Bactrian camels thymidine/pyrimidine phosphorylase.

**Figure 3**

The sequence comparison statistics of thymidine/pyrimidine phosphorylase. The upper right diagonal region explains the number of differences between two sequences, while the lower left diagonal region explains the % of identity between two sequences.
Figure 4

Phylogram of camel thymidine/pyrimidine phosphorylase in relation to a set of prokaryotic and eukaryotic organisms.
Figure 5

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 explains the number of differences between two sequences, while the lower-left diagonal region explains the % of identity between two sequences.

*Trypanosoma evansi* NDRT

<table>
<thead>
<tr>
<th>Lactobacillus fecum VEF34489</th>
<th>Enterococcus fecum VEF46003</th>
<th>Trypanosoma evansi nucleoside_2-deoxyribosyltransferase</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Lactobacillus fecum VEF34489</td>
<td>1</td>
<td>127</td>
</tr>
<tr>
<td>Enterococcus fecum VEF46003</td>
<td>21.12</td>
<td>135</td>
</tr>
<tr>
<td>Trypanosoma evansi nucleoside_2-deoxyribosyltransferase</td>
<td>17.71</td>
<td>17.68</td>
</tr>
</tbody>
</table>

Consensus:

- **Lactobacillus fecum VEF34489**: MXXK --- YIXAGP --- FFEXQXXXAAAAALXXNXXTXXL
- **Enterococcus fecum VEF46003**: DQ --- VXAGB --- GEXQXXXAAAAALXXNXXTXXL
- **Trypanosoma evansi nucleoside_2-deoxyribosyltransferase**: MXXK --- YIXAGP --- FFEXQXXXAAAAALXXNXXTXXL

Conservation:

- **Lactobacillus fecum VEF34489**: MXXK --- YIXAGP --- FFEXQXXXAAAAALXXNXXTXXL
- **Enterococcus fecum VEF46003**: DQ --- VXAGB --- GEXQXXXAAAAALXXNXXTXXL
- **Trypanosoma evansi nucleoside_2-deoxyribosyltransferase**: MXXK --- YIXAGP --- FFEXQXXXAAAAALXXNXXTXXL
Figure 6

Molecular model of T. evansi NDRT. The model was built by using SwissModel.