DOI : 10.5958/2277-8934.2020.00022.3 *Vol 27 No 2, p 145-158*

BIOINFORMATICS OF PYRIMIDINE METABOLISM IN CAMELS AND *Trypanosoma evansi*: URIDINE 5′-DIPHOSPHATE (UDP) METABOLIC PATHWAYS AND TARGETING ATP DIPHOSPHATASE

Mahmoud Kandeel[1,3,*], Abdulhafez Dalab[2], Abdulkarem Al-Shabebi[2] and Abdulla Al-Taher[1]

[1]Department of Biomedical Sciences, College of Veterinary Medicine, King Faisal University, Alahsa, 31982, Saudi Arabia
[2]Department of Anatomy, College of Veterinary Medicine, King Faisal University, Alahsa 31982, Saudi Arabia
[3]Department of Pharmacology, Faculty of Veterinary Medicine, Kafrelsheikh University, Kafrelsheikh, 35255, Egypt

ABSTRACT

Comparison of host and parasite genomes can lead to discovery of metabolic or structural differences that can be used in finding new drug targets. In this study, the pyrimidine metabolic pathway uridine 5′-diphosphate (UDP) and deoxyuridine 5-diphosphate (dUDP) paths were researched. In this perspective, several UDP/dUDP metabolising enzymes were investigated by bioinformatics tools. UMP/CMP kinase was not detected in *Trypanosoma evansi* (*T. evansi*) allowing ATP-diphosphatase as an important target for inhibition studies as the interconversion between UMP and UDP are affected. While 100% similarity rate was evident in camel species, there was about 30% differences between the camel and human ATP-diphosphatase. The great sequence differences between the camel and *T. evansi* ATP-diphosphatases (73% differences) suggests for vulnerability of using this target for future studies. However, further investigations are required to establish the biological aspects of the enzyme.

Key words: Camel, CTP synthase, genome, nucelotide, *Trypanosoma evansi*, UDP

The sequence of camel genome was recently published (Jirimutu *et al,* 2012). Resolving the sequence of the genes in various metabolic pathways gives comprehensive insights into the metabolic pathways and the requirements for camel adaptation to its harsh environment. In our group, bioinformatics tools were used for characterisation of new drug target (Alfuwaires *et al,* 2017; Alnazawi *et al,* 2017; Kandeel *et al,* 2018; Mahmoud *et al,* 2019). Recently, we provided some interesting differences in the metabolic pathways of pyrimidines in camel and *T. evansi* (Kandeel and Al-Taher 2020a; 2020b; Kandeel *et al,* 2020).

In this exploration, the pyrimidine metabolic pathways was investigated in both camels and the blood protozoan, *T. evansi*. During this work, the enzymes metabolising UDP/dUDP were reviewed. UDP/dUDP metabolism was traced by the KEGG maps (Kanehisa *et al,* 2007; Kanehisa *et al,* 2016; Ogata *et al,* 1998).

The ultimate goal of this work was to explore the UDP/dUDP metabolic enzymes and compare its sequence in camel, *T. evansi*, human and other prokaryotes and eukaryotes. The enzymes involved in this pathway were listed and searched for its coding genes in camel and *T. evansi* genomes. The expected map of UDP/dUDP metabolism was drawn and compared to find potential new drug targets against the blood protozoa.

Materials and Methods

Retrieval of genomic data

Collection of genomic data was carried out by extracting the information from the gene database (http://www.genedb.org) (Hertz-Fowler *et al,* 2004), Kinetoplastom genome resources (http://tritrypdb.org/tritrypdb/), protein and genome databases (http://www.ncbi.nlm.nih.gov) and the Arabian camel genome project (http://www.camel.kacst.edu.sa). 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,

SEND REPRINT REQUEST TO MAHMOUD KANDEEL email: mkandeel@kfu.edu.sa

any predicted sequence motifs and the E value of the closest PDB sequence homolog.

Protein sequence homologues were searched using the NCBI, BLAST (Basic Local Alignment Search Tool) (Madden, 2013) 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.

Multiple sequence alignment programme for proteins and construction of phylogenetic tree was carried out using the tools available (*https://www.ebi.ac.uk/Tools/msa/clustalo/*) (Sievers and Higgins, 2014). Multiple sequence alignment programme (ClustalW2) was used to calculate the best match of the selected sequences. The resultant alignment was used to generate a phylogenic tree, which is visualised by Dendroscope phylogenic tree viewer (Huson *et al,* 2007) or CLC genomics workbench (Sequencing, 2011).

Putative domains was searched by the domain prediction program available (*http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml*) (Marchler-Bauer *et al,* 2005). The genomic and molecular information was obtained from Kyoto Encyclopedia of Genes and Genomes (*http://www.genome.jp/kegg/*).

Proteomic and genomic tools ExPASy Proteomics tools (http://us.expasy.org/tools/) (Gasteiger *et al,* 2003) and tools available at the website of the European Bioinformatics Institute (http://www.ebi.ac.uk/Tools/) (Labarga *et al,* 2007) were also used for analysis of nucleotide and gene sequences. The protein parameters including the molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity was searched (http://us.expasy.org/tools/protparam.html). Protein sequences of target genes was analysed for data such as pI, extinction coefficient and MW for the tagged protein sequence by PROTParam. PROSITE (http://www.expasy.org/proteomics/families__patterns_and_profiles) was used to search for patterns and profiles in the protein sequences of the target genes. The protein domains, families and functional sites was searched (http://prosite.expasy.org/).

Results and Discussion

Uridine 5′- diphosphate (UDP) production process is proposed by 3 production routes, derived from the metabolic pathways of the KEGG (Kyoto Encyclopedia of Genes and Genomes) (Kanehisa *et al,* 2016). In the synthetic direction, uridine-5′-monophosphate (UMP) is phosphorylated by the kinase activity of UMP/CMP kinase and/or UMP kinases. The phosphorylation of uridine 5′-triphosphate (UTP) by nucleoside diphosphate phosphatase, thymidine triphosphatase and/or ATP-diphosphatase can yield uridine 5′-diphosphate (UDP) as shown in (Fig 1). In the catabolic pathway, UDP is phosphorylated by nucleoside diphosphatase or ATP diphosphatase to generate uridine 5′-monophosphate (UMP) (Fig 1 and Table 1). After bioinformatics tools application, the revised metabolic pathways of camels and *T. evansi* was provided in Fig 2 and 3, respectively.

The generation of UDP in camels was brought by the action of NDK, ATP diphosphatase and UMP/CMP kinase (Fig 2 and Table 2). In contrast, *T. evansi* was similar to camel profile, but devoid of UMP/CMP kinase (Fig 3, Table 3). Therefore, inhibition of ATP diphosphatase might be deleterious for the parasite life as it is the only predicted enzyme for catabolising UTP and UDP to UMP.

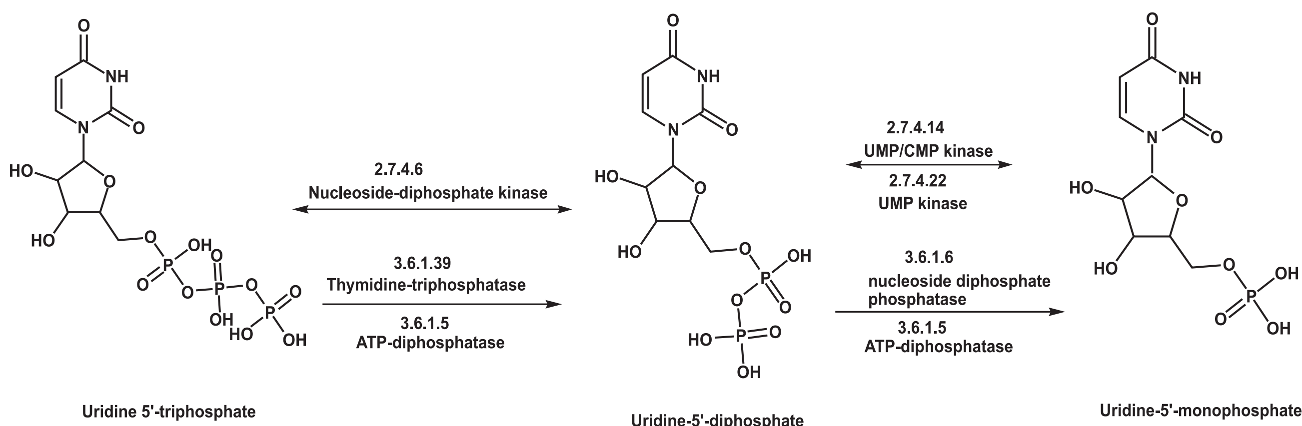

Fig 1. The proposed metabolic pathways of uridine 5′-diphosphate (UDP).

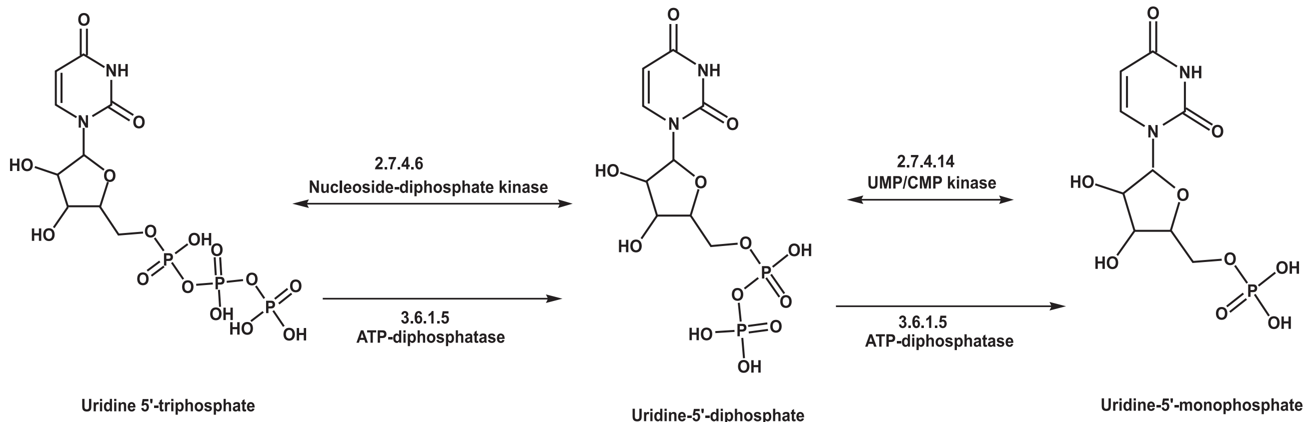

Fig 2. The proposed metabolic pathways of uridine 5′-diphosphate (UDP) in camels.

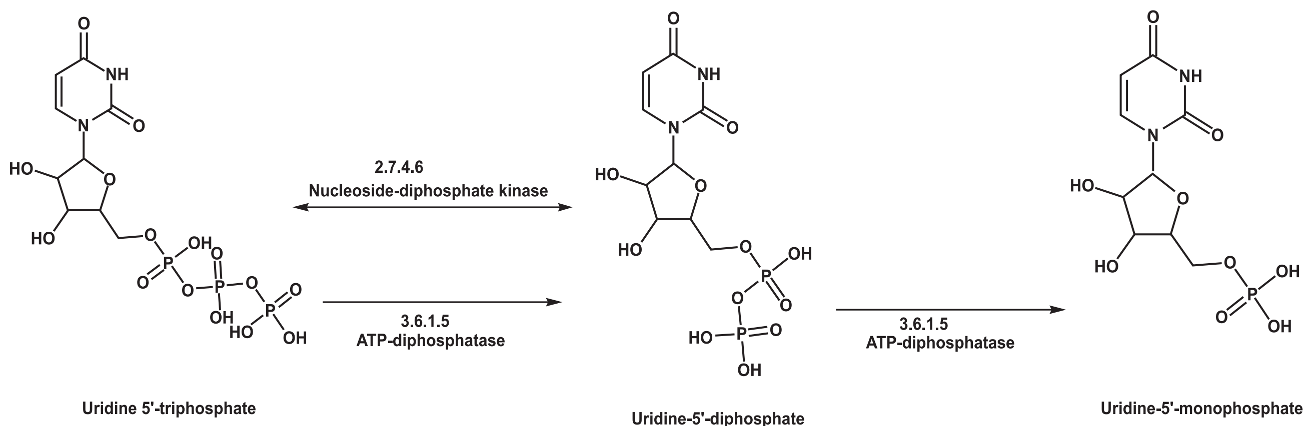

Fig 3. The proposed metabolic pathways of uridine 5′-diphosphate (UDP) in *Trypanosoma evansi*.

Table 1. Enzymes involved in metabolic pathways of UDP.

ID (E.C. number)	Definition (Enzyme name)
2.7.4.6	nucleoside-diphosphate kinase
3.6.1.39	thymidine-triphosphatase, thymidine triphosphate nucleotidohydrolase
3.6.1.5	nucleoside triphosphate phosphohydrolases (nucleoside monophosphoate-forming) or ATP-diphosphatase, apyrase, ectonucleoside triphosphate diphosphohydrolase
2.7.4.14	UMP-CMP kinase, cytidylate kinase; deoxycytidylate kinase; CTP: CMP phosphotransferase
2.7.4.22	UMP kinase; uridylate kinase; UMPK
3.1.3.6	3′-nucleotidase; 3′-mononucleotidase

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

ID (E.C. number)	Definition (Enzyme name)
2.7.4.6	nucleoside-diphosphate kinase
3.6.1.5	nucleoside triphosphate phosphohydrolases (nucleoside monophosphoate-forming) or ATP-diphosphatase, apyrase, ectonucleoside triphosphate diphosphohydrolase
2.7.4.14	UMP-CMP kinase, cytidylate kinase; deoxycytidylate kinase; CTP: CMP phosphotransferase

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

ID (E.C. number)	Definition (Enzyme name)
2.7.4.6	nucleoside-diphosphate kinase
3.6.1.5	nucleoside triphosphate phosphohydrolases (nucleoside monophosphoate-forming) or ATP-diphosphatase, apyrase, ectonucleoside triphosphate diphosphohydrolase

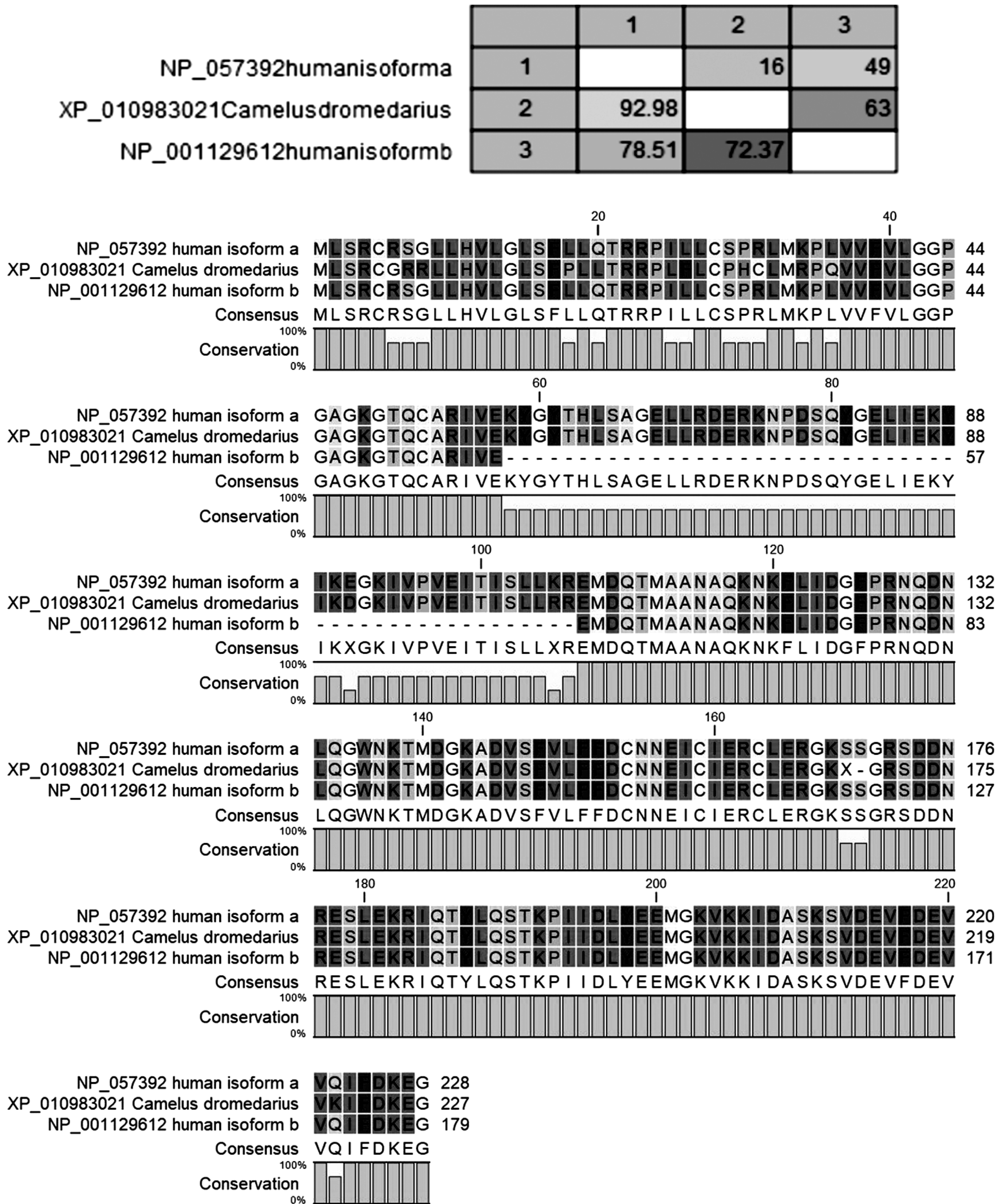

Fig 4. Multiple sequence alignment of dromedary camel and human bifunctional UMP/CMP kinase. The upper panel represents sequence comparison statistics. The upper right diagonal region explain the number of differences between 2 sequences, while the lower left diagonal region explain the per cent of identity between 2 sequences.

Thymidine-triphosphatase

Thymidine-triphosphatase (dTTP; 3.6.1.39) an enzyme also named thymidine triphosphate monophosphohydrolase (dTTPase); thymidine triphosphate nucleotidohydrolase; dTTPase; deoxythymidine-5′-triphosphatase. This is a hydrolase enzyme, induces the dephosphorylation of deoxythymidine-triphosphate (dTTP) to the equivalent deoxythymidine-diphosphate (dTDP). Likewise, it acts on deoxyuridine-triphosphate (dUTP) and uridine-triphosphate (UTP) but very slowly. Therefore, its molecular weight is 48 500.

		1	2	3
XP_010983021Camelusdromedarius	1		54	71
EQB77421camelusferus	2	76.32		66
XP_010949526Camelusbactrianus	3	70.17	70.93	

Fig 5. Multiple sequence alignment of dromedary, bactrian and feral camels bifunctional UMP/CMP kinase. 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 per cent of identity between two sequences.

The dTTPase is formed of 2 equivalent subunits. Moreover, dTTPase displays optimal activity at pH 7 - 9 and the activation energy is valued to be 7.1 kcal/mol at pH 7.8. Meanwhile, the existence or absence of divalent cations does not affect dTTPase activity and the ethylenediaminetetracetic acid (EDTA) but not phenanthroline which can inhibit the enzyme activity. The Mn^{2+} reverses the inhibition

		1	2	3	4	5	6	7	8	9	10	11	12
NP_001301182Caprahircus	1		4	7	9	13	17	16	23	21	35	55	70
NP_001039509Bostaurus	2	98.25		8	10	14	18	15	23	21	35	53	70
XP_010983021Camelusdromedarius	3	96.93	96.49		10	12	16	16	23	21	34	56	68
XP_023114431Feliscatus	4	96.05	95.61	95.61		11	15	14	21	19	37	55	72
XP_023486088Equuscaballus	5	94.30	93.86	94.74	95.18		15	17	23	21	36	52	73
NP_057392human	6	92.54	92.11	92.98	93.42	93.42		14	23	21	37	51	75
XP_002715160Oryctolaguscuniculus	7	92.98	93.42	92.98	93.86	92.54	93.86		21	18	34	50	72
EDL90325Rattusnorvegicus	8	89.91	89.91	89.91	90.79	89.91	89.91	90.79		9	40	53	76
NP_079923Musmusculus	9	90.79	90.79	90.79	91.67	90.79	90.79	92.11	96.04		37	52	75
BAA06130swine	10	84.65	84.65	85.09	83.77	84.21	83.77	84.89	82.38	83.70		23	72
NP_001026735Gallusgallus	11	75.88	76.75	75.44	75.88	77.19	77.63	77.78	76.65	77.09	88.27		94
XP_010949526Camelusbactrianus	12	70.71	70.71	71.43	69.87	69.46	68.62	69.87	68.20	68.62	69.36	60.00	

```
                                                     20                  40
NP_001301182 Capra hircus         MLSRCGRRLLHVLGLSFPLLTGRPLFLLPHRFMKPQVVFVLGGP 44
NP_001039509 Bos taurus           MLNRCGRRLLHVLGLSFPLLTGRPLFLFPHRFMKPQVVFVLGGP 44
XP_010983021 Camelus dromedarius  MLSRCGRRLLHVLGLSFPLLTRRPLFLCPHCLMRPQVVFVLGGP 44
XP_023114431 Felis catus          MLSRCSRRLLHVLGLSFPLLTCRPLFLWPHRLMKPLVVFVLGGP 44
XP_023486088 Equus caballus       MRSLCARRLLHVLGLRFPLPTRRPLFLCPRRLMKPLVVFVLGGP 44
NP_057392 human                   MLSRCRSGLLHVLGLSFLLQTRRPILLCSPRLMKPLVVFVLGGP 44
XP_002715160 Oryctolagus cuniculus MLNRCRRRLLHVLGLSFPLLTP---LLSPPHLMKPLVVFVLGGP 41
EDL90325 Rattus norvegicus        MLSSCRRWLLHVLVPSFPPLT-RGLRFFPLQLMKPLVVFVLGGP 43
NP_079923 Mus musculus            MLSSCRRLLPHVLVPSCPPLT-RTLRLFPLHLMKPLVVFVLGGP 43
BAA06130 swine                    M-------------------------------RPKVVFVLGGP 12
NP_001026735 Gallus gallus        M--------------------------------KPVVVFVLGGP 12
XP_010949526 Camelus bactrianus   MEGTAWGRF--VPAPSE--LTFWGPSSSPERAGR--------GI 32
Consensus                         MLSRCRRRLLHVLGLSFPLLT-RPLFLXPHRLMKPLVVFVLGGP
Conservation

                                              60                  80
NP_001301182 Capra hircus         GAGKGTQCARIVEKYGYTHLSAGELLRDERKNPDSQYGELIGKY 88
NP_001039509 Bos taurus           GAGKGTQCARIVEKYGYTHLSAGELLRDERKNPDSQYGELIEKY 88
XP_010983021 Camelus dromedarius  GAGKGTQCARIVEKYGYTHLSAGELLRDERKNPDSQYGELIEKY 88
XP_023114431 Felis catus          GAGKGTQCARIVEKYGYTHLSAGELLRDERKNPDSQYGELIEKY 88
XP_023486088 Equus caballus       GAGKGTQCARIVEKYGYTHLSAGELLRDERKNPDSQYGELIEKY 88
NP_057392 human                   GAGKGTQCARIVEKYGYTHLSAGELLRDERKNPDSQYGELIEKY 88
XP_002715160 Oryctolagus cuniculus GAGKGTQCARIVEKYGYTHLSAGELLRDERKNPDSQYGELIEKY 85
EDL90325 Rattus norvegicus        GAGKGTQCARIVEKYGYTHLSAGELLRDERKNPDSQYGELIEKY 87
NP_079923 Mus musculus            GAGKGTQCARIVEKYGYTHLSAGELLRDERKNPDSQYGELIEKY 87
BAA06130 swine                    GAGKGTQCARIVEKYGYTHLSAGELLRDERKNPDSQYGELIEKY 56
NP_001026735 Gallus gallus        GAGKGTQCARIVEKYGYTHLSAGDLLRDERKRPGSQYGELIENY 56
XP_010949526 Camelus bactrianus   AGGSASRDPELAKKYGYTHLSAGELLRDERKNPDSQYGELIEKY 76
Consensus                         GAGKGTQCARIVEKYGYTHLSAGELLRDERKNPDSQYGELIEKY
Conservation

                                            100                 120
NP_001301182 Capra hircus         IKDGKIVPVEITISLLRREMDQTMAANAQKNKFLIDGFPRNQDN 132
NP_001039509 Bos taurus           IKDGKIVPVEITISLLRREMDQTMAANAQKNKFLIDGFPRNQDN 132
XP_010983021 Camelus dromedarius  IKDGKIVPVEITISLLRREMDQTMAANAQKNKFLIDGFPRNQDN 132
XP_023114431 Felis catus          IKDGKIVPVEITISLLKREMDQTMAANAQKNKFLIDGFPRNQDN 132
XP_023486088 Equus caballus       IKDGKIVPVEITISLLKREMDQTMAANAQKNKFLIDGFPRNEDN 132
NP_057392 human                   IKEGKIVPVEITISLLKREMDQTMAANAQKNKFLIDGFPRNQDN 132
XP_002715160 Oryctolagus cuniculus IKEGKIVPVEITISLLKREMDQTMAANAQKNKFLIDGFPRNQDN 129
EDL90325 Rattus norvegicus        IKEGKIVPVEITISLLKREMDQTMAANAQKNKFLIDGFPRNQDN 131
NP_079923 Mus musculus            IKEGKIVPVEITISLLKREMDQTMAANAQKNKFLIDGFPRNQDN 131
BAA06130 swine                    IKDGKIVPVEITISLLRREMDQTMAANAQKNKFLIDGFPRNQDN 100
NP_001026735 Gallus gallus        IKEGEIVPVEITISLLKRAMDQTMAANSQKNKFLIDGFPRNEDN 100
XP_010949526 Camelus bactrianus   IKDGKIVPVEITISLLRREMDQTMAANAQKNKFLIDGFPRNQDN 120
Consensus                         IKDGKIVPVEITISLLKREMDQTMAANAQKNKFLIDGFPRNQDN
Conservation

                                        140                 160
NP_001301182 Capra hircus         LQGWNKTMDGKADVSFVLFFDCNNE-----------ICIERCLE 165
NP_001039509 Bos taurus           LQGWNKTMDGKADVSFVLFFDCNNE-----------ICIERCLE 165
XP_010983021 Camelus dromedarius  LQGWNKTMDGKADVSFVLFFDCNNE-----------ICIERCLE 165
XP_023114431 Felis catus          LQGWNNTMDGKAEVSFVLFFDCNNE-----------ICIERCLE 165
XP_023486088 Equus caballus       LQGWNKTMDGKADVSFVLFFDCNNE-----------ICIERCLE 165
NP_057392 human                   LQGWNKTMDGKADVSFVLFFDCNNE-----------ICIERCLE 165
XP_002715160 Oryctolagus cuniculus LQGWNKTMDGKADVSFVLFFDCNDE-----------ICIERCLE 162
EDL90325 Rattus norvegicus        LQGWNKTMDGKADVSFVLFFDCNNE-----------ICIDRCLE 164
NP_079923 Mus musculus            LQGWNKTMDGKADVSFVLFFDCNNE-----------ICIERCLE 164
BAA06130 swine                    LQGWNKTMDGKADVSFVLFFDCNNE-----------ICIERCLE 133
NP_001026735 Gallus gallus        LQGWNKTMDGKADVSFVLFFDCDNE-----------ICIGRCLE 133
XP_010949526 Camelus bactrianus   LQGWNKTMDGKADVSFVLFFDCNNEVMKIFICLHGLICFEGCLE 164
Consensus                         LQGWNKTMDGKADVSFVLFFDCNNE-----------ICIERCLE
Conservation

                                  180                 200                 220
NP_001301182 Capra hircus         RGKSSGRSDDNRESLEKRIQTYLQSTKPIIDLYEEMGKVKKIDA 209
NP_001039509 Bos taurus           RGKSSGRSDDNRESLEKRIQTYLQSTKPIIDLYEEMGKVRKIDA 209
XP_010983021 Camelus dromedarius  RGKX-GRSDDNRESLEKRIQTYLQSTKPIIDLYEEMGKVKKIDA 208
XP_023114431 Felis catus          RGKSSGRSDDNRESLEKRIQTYLQSTKPIIDLYEEMGKVKKIDA 209
XP_023486088 Equus caballus       RGKSSGRSDDNRESLEKRIQTYLQSTKPIIDLYEEMGKVKKIDA 209
NP_057392 human                   RGKSSGRSDDNRESLEKRIQTYLQSTKPIIDLYEEMGKVKKIDA 209
XP_002715160 Oryctolagus cuniculus RGKSSGRSDDNRESLEKRIQTYLQSTKPIIDLYEEMGKVKKIDA 206
EDL90325 Rattus norvegicus        RGKSSGRSDDNRESLEKRIQTYLESTKPIIDLYEEMGKVKKIDA 208
NP_079923 Mus musculus            RGKSSGRSDDNRESLEKRIQTYLESTKPIIDLYEEMGKVKKIDA 208
BAA06130 swine                    RGKSSGRSDDNRESLEKRIQTYLQSTKPIIDLYEEMGKVKKIDA 177
NP_001026735 Gallus gallus        RGKSSGRSDDNRESLEKRIHTYLQSTRPIIDLYERMGKVRRVDA 177
XP_010949526 Camelus bactrianus   RG----------ESLEKRIQTYLQSTKPIIDLYEEMGKVKKIDA 198
Consensus                         RGKSSGRSDDNRESLEKRIQTYLQSTKPIIDLYEEMGKVKKIDA
Conservation

NP_001301182 Capra hircus         SKSVDEVFDEVVKIFDKEG 228
NP_001039509 Bos taurus           SKSVDEVFDEVVKIFDKEG 228
XP_010983021 Camelus dromedarius  SKSVDEVFDEVVKIFDKEG 227
XP_023114431 Felis catus          SKSVDEVFDEVVKIFDKEG 228
XP_023486088 Equus caballus       SKSVDEVFDEVVKIFDKEG 228
NP_057392 human                   SKSVDEVFDEVVQIFDKEG 228
XP_002715160 Oryctolagus cuniculus SKSVDEVFDEVVKIFDKEG 225
EDL90325 Rattus norvegicus        SKSVDEVFGDVMKIFDKEG 227
NP_079923 Mus musculus            SKSVDEVFGEVVKIFDKEG 227
BAA06130 swine                    SKSVDEVFDEVVKIFDKEG 196
NP_001026735 Gallus gallus        SKSVDEVFEKVVQIFDKEG 196
XP_010949526 Camelus bactrianus   SKSVDEVFDEVVKIFDKEG 217
Consensus                         SKSVDEVFDEVVKIFDKEG
Conservation
```

Fig 6. Multiple sequence alignment of dromedary camel and other eukaryotes bifunctional UMP/CMP kinase. 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 per cent of identity between two sequences.

		1	2	3	4
XP_657222Entamoebahistolytica	1		140	294	255
XP_001610286Babesiabovis	2	32.04		298	260
KNG78296Plasmodiumfalciparum	3	20.33	19.46		349
PWV08006Trypanosomacruzei	4	11.46	10.65	7.67	

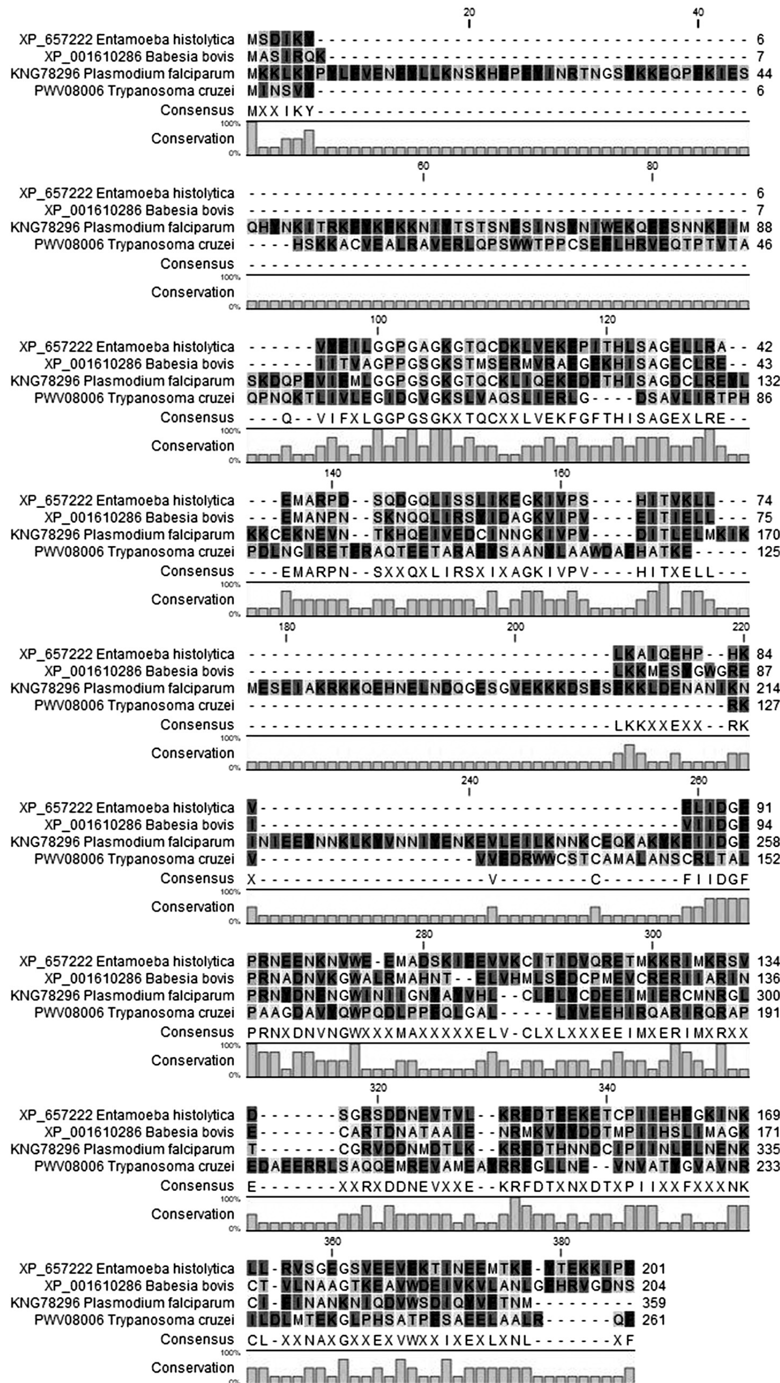

Fig 7. Multiple sequence alignment of *Trypanosoma cruzei* and some protozoal bifunctional UMP/CMP kinase. The upper panel represents sequence comparison statistics. The upper right diagonal region is explains the number of differences between two sequences, while the lower left diagonal region explains the per cent of identity between two sequences.

caused by EDTA and Zn^{2+} triggers the absolute inhibition of dTTPase activity. The breakdown of dUTP and UTP are by about 50 and 20% of the rate of dTTP hydrolysis, respectively. Other deoxyribonucleosides or ribonucleoside triphosphates do not act as substrates for the dTTPase. This enzyme could play a significant role in the regulation of the cellular dTTP levels (Dahlmann, 1982). Thymidine-triphosphatase (thymidine triphosphate nucleotidohydrolase, dTTPase, EC no. 3.6.1.39) has been previously purified from human serum (Dahlmann, 1982). However, genetic and sequence data for other organisms were not available and the sequence of human gene was not present.

UMP-CMP kinase is a bifunctional enzyme present in eukaryotes that catalyses the phosphorylation of CMP and UMP. Prokaryotes has 2 different monofunctional enzymes EC 2.7.4.25, which act as CMP kinase and EC 2.7.4.22, which act as UMP kinase.

UMP/CMP kinase

UMP kinase (uridylate kinase; UMPK, 2.7.4.22) has restricted kinase activity for UMP only and specific for prokaryotes. It is replaced by the bifunctional UMP-CMP kinase.

Pyrimidine nucleoside monophosphate kinase (UMP/CMP kinase; 2.7.4.14); cytidylate kinase; deoxycytidylate kinase; CTP: CMP phosphotransferase; dCMP kinase; deoxycytidine monophosphokinase; UMP-CMP kinase; ATP:UMP-CMP phosphotransferase; pyrimidine nucleoside monophosphate kinase; uridine monophosphate-cytidine monophosphate phosphotransferase. This is eukaryotic bifunctional transferase enzyme (UMP/CMP kinase) with dual-specificity, transferring phosphorus-containing groups catalyses the phosphorylation of both substrates CMP and UMP

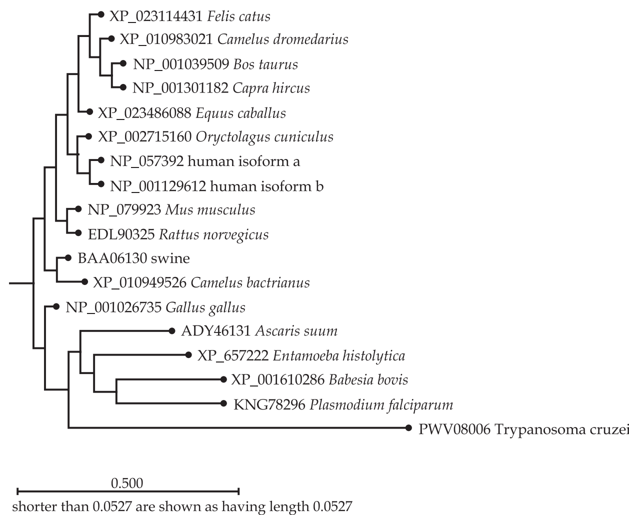

Fig 8. Phylogram of camel and *Trypanosoma cruzei* bifunctional UMP/CMP kinase in relation to a set of eukaryotic organisms.

with analogous efficacy forming UDP, CDP and dCDP, which are required for cellular nucleic acid synthesis. Also, dCMP may possibly act as acceptor. This eukaryotic enzyme dissimilar from the monofunctional prokaryotic enzymes EC 2.7.4.25, CMP kinase and EC 2.7.4.22, UMP kinase. Moreover, several cytidine and deoxycytidine analogues are important anticancer and antiviral drugs. These drugs exert their therapeutic effects *via* phosphorylation of their triphosphate structures. All of the nucleoside triphosphates used by UMP/CMPK as phosphate donors, the best donors are ATP and dATP while CTP being the poorest. Likewise, UMP/CMPK also able to phosphorylate all of the deoxycytidine analogue monophosphates. Strongly, DTT, 2-mercaptoethanol and thioredoxin, are reducing agents which could activate this enzyme, suggesting that its activity might be regulated by redox potential *in vivo*. The localisation of UMP/CMPK chiefly in the cytoplasm (Liou *et al,* 2002; Pasti *et al,* 2003).

Uridine monophosphate kinase (UMP kinase; 2.7.1.48); uridylate kinase; UMPK; PyrH; UMP-kinase; SmbA. UMP kinase is an enzyme transferring phosphate groups, precisely specific for UMP in prokaryotes and used for the *de novo* synthesis of pyrimidines nucleotides. The feedback control of this process *via* repression of carbamoyl phosphate synthetase gene. Moreover, UMPK activated by GTP and inhibited by UTP. Conversely, UMP/CMP kinase is bifunctional for the same purpose in eukaryotes. The structure of UMPK differ from animals nucleoside monophosphate kinases, comprising the UMP/CMP kinase therefore, UMPK

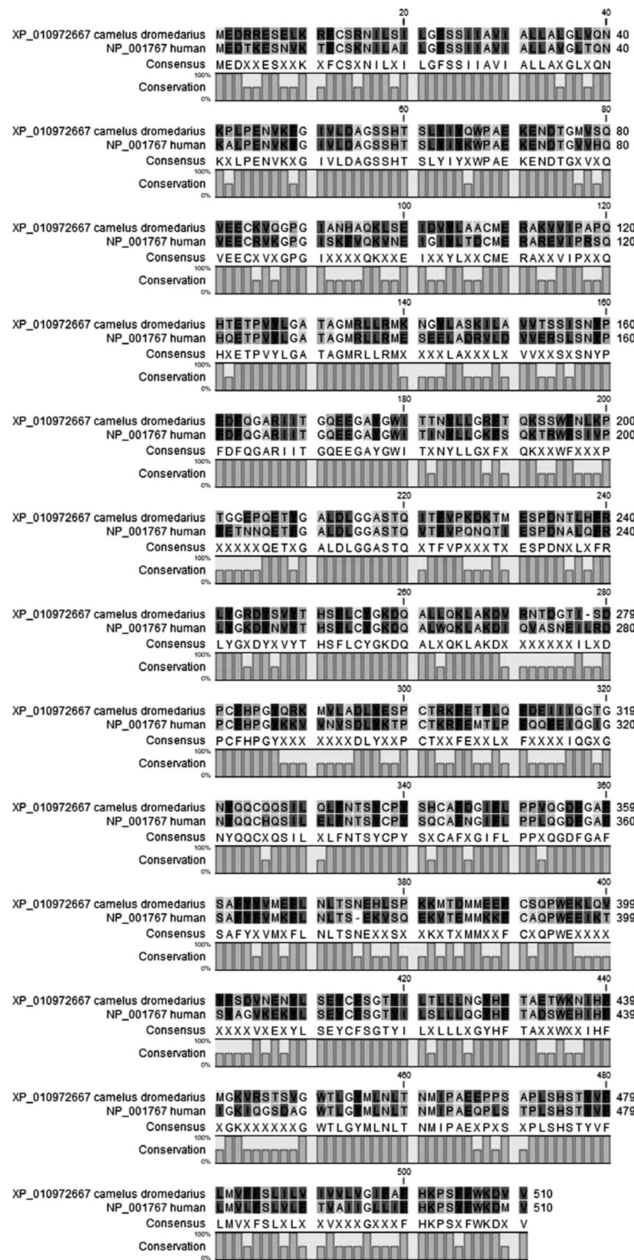

Fig 9. Multiple sequence alignment of dromedary camel and human ectonucleoside triphosphate diphosphohydrolase. The upper panel represents sequence comparison statistics. The upper right diagonal region is explains the number of differences between two sequences, while the lower left diagonal region explains the per cent of identity between two sequences.

may act as antimicrobial target. This enzyme has a homohexameric assembly centred around a hollow nucleus and is organised as a trimer of dimers. The polypeptide of UMPK displays the amino acid kinase family fold, which described in carbamate kinase and acetylglutamate kinase. The substrates of acetylglutamate kinase bind within each subunit at identical, sufficiently adjusted sites. While, the

		1	2	3
XP_010972667camelusdromedarius	1		0	0
XP_014421250Camelusferus	2	100.00		0
XP_010969434bactrianus	3	100.00	100.00	

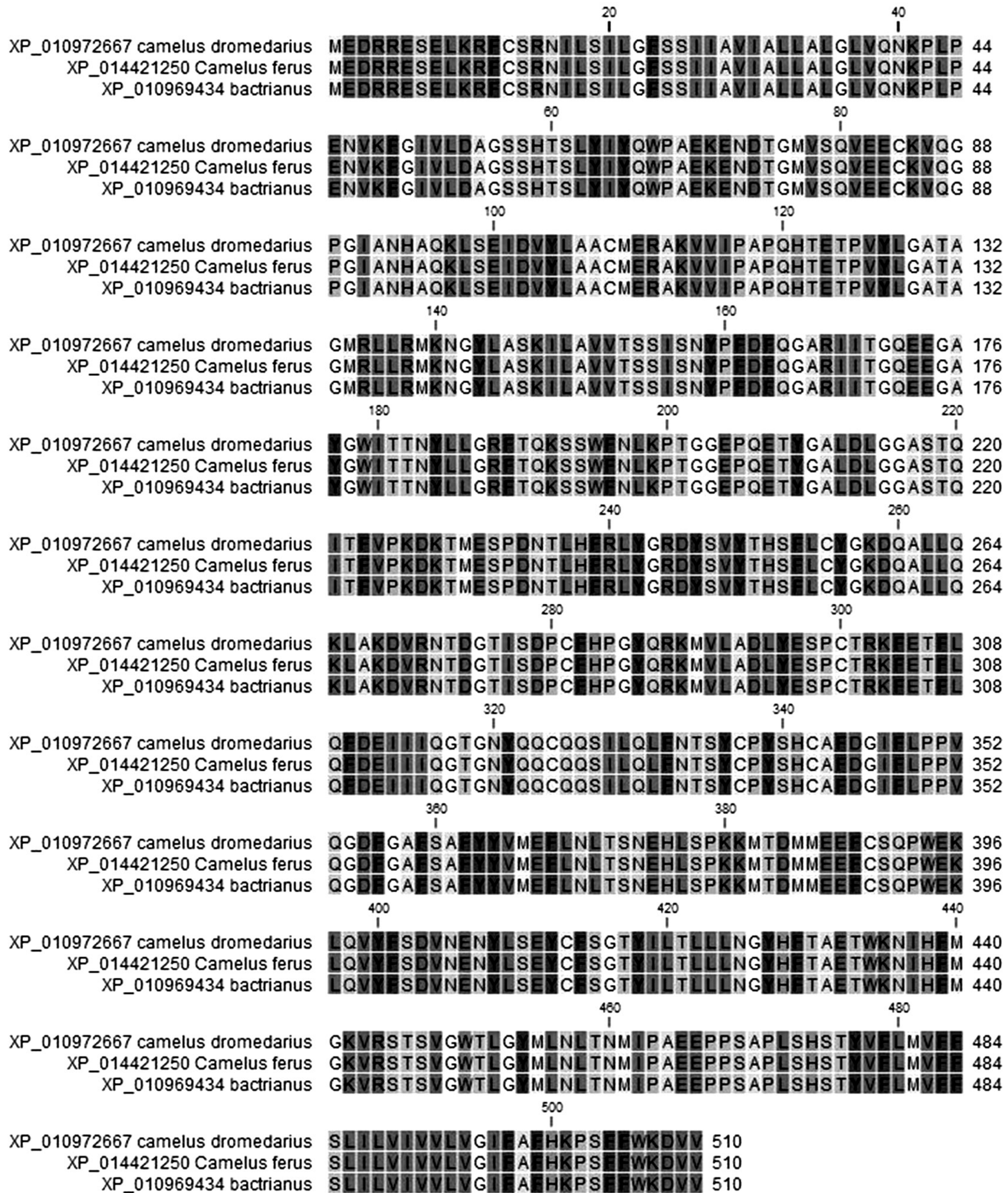

Fig 10. Multiple sequence alignment of dromedary, Bactrian and feral camels ectonucleoside triphosphate diphosphohydrolase. The upper panel represents sequence comparison statistics. The upper right diagonal region explains the number of differences between two sequences, while the lower left diagonal region explains the per cent of identity between two sequences.

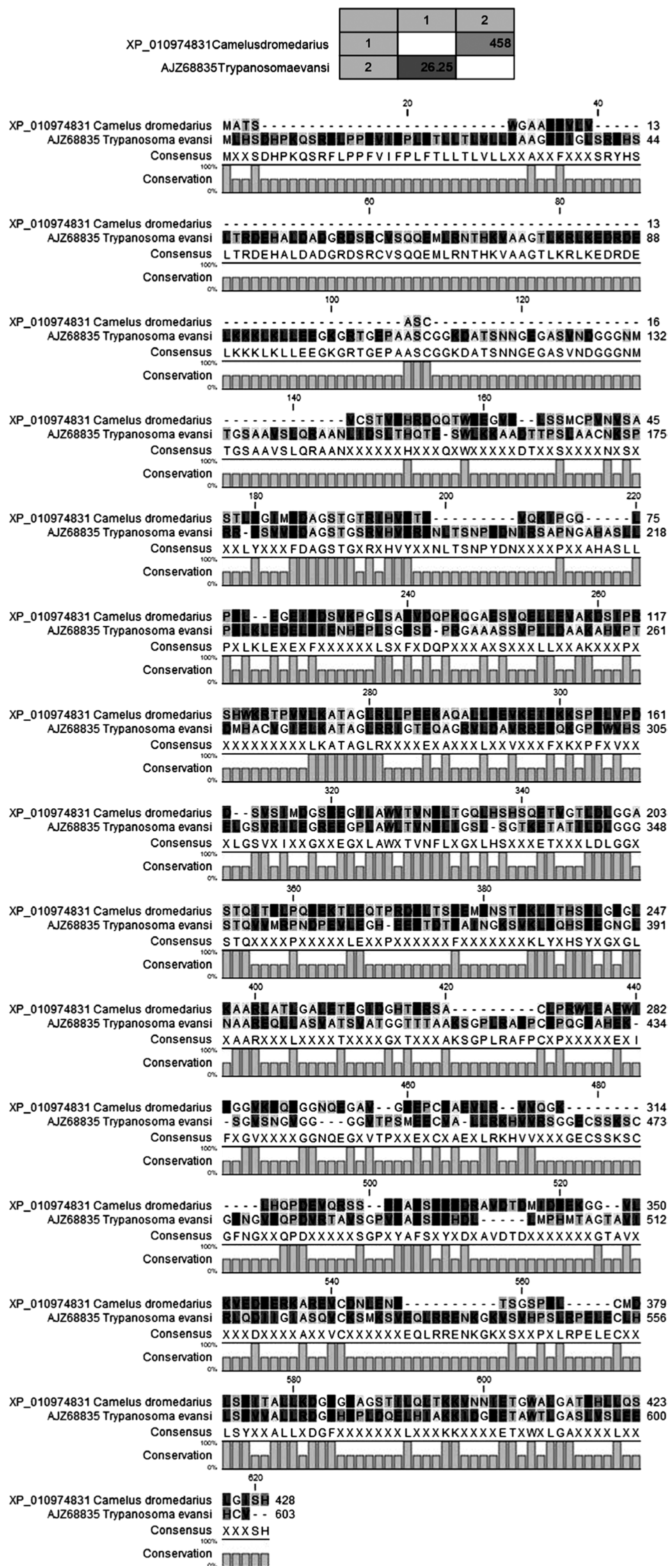

Fig 11. Pairwise sequence alignment of *T. evansi* and dromedary camel (isoform 5) ectonucleoside triphosphate diphosphohydrolase. The upper panel represents sequence comparison statistics. The upper right diagonal region is explains the number of differences between two sequences, while the lower left diagonal region explains the per cent of identity between two sequences.

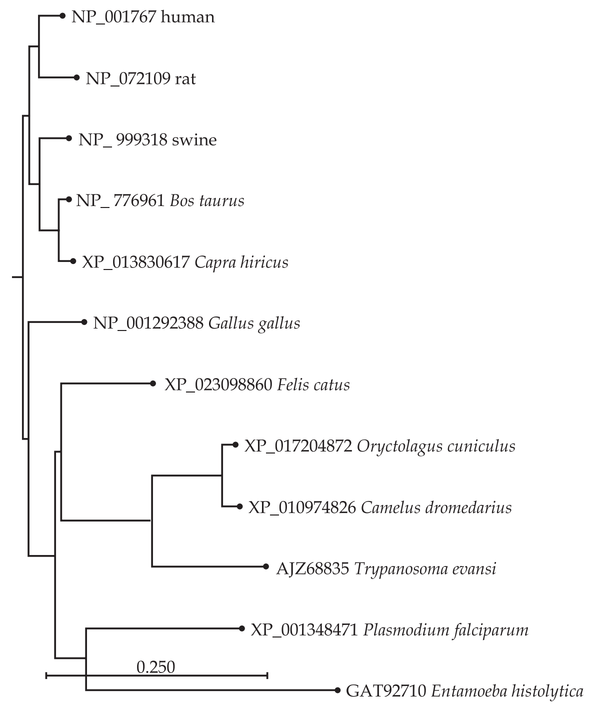

Fig 12. Phylogram of camel ectonucleoside triphosphate diphosphohydrolase in relation to a set of eukaryotic organisms.

structure of UMPK comprises 2 bound Mg^{2+} ions, one ion assists the maintenance of the transition status, hence having the same catalytic role as one lysine residue found in acetylglutamate kinase. Corresponding to carbamate kinase and acetylglutamate kinase, UMPK presents a radically dissimilar dimer architecture, lacking the characteristic 16-stranded beta-sheet backbone (Marco-Marin *et al,* 2005; Serina *et al,* 1995).

The sequence of human UMP/CMP kinase isoform A was aligned to its counterpart in dromedary camels (Fig 4). Among 228 AA, there were only 15 mismatches (6.5%) between human and dromedary camel UMP/CMP kinase. In addition to a missing residue in dromedary camel and insertion in human, it is worth noting that 12 mismatches (80%) are in the first 40 residues of the protein sequence. However, it is yet to determine whether variations at the beginning of the sequence have a considerable effect on the protein function.

Protein sequences of UMP/CMP kinase of the 3 camel species were aligned and shown in figure 5. The alignment showed the same phenomenon observed in figure 4 where the variations were concentrated in the first 60 AA. Only UMP/CMP kinase of *Camelus bactrianus* has

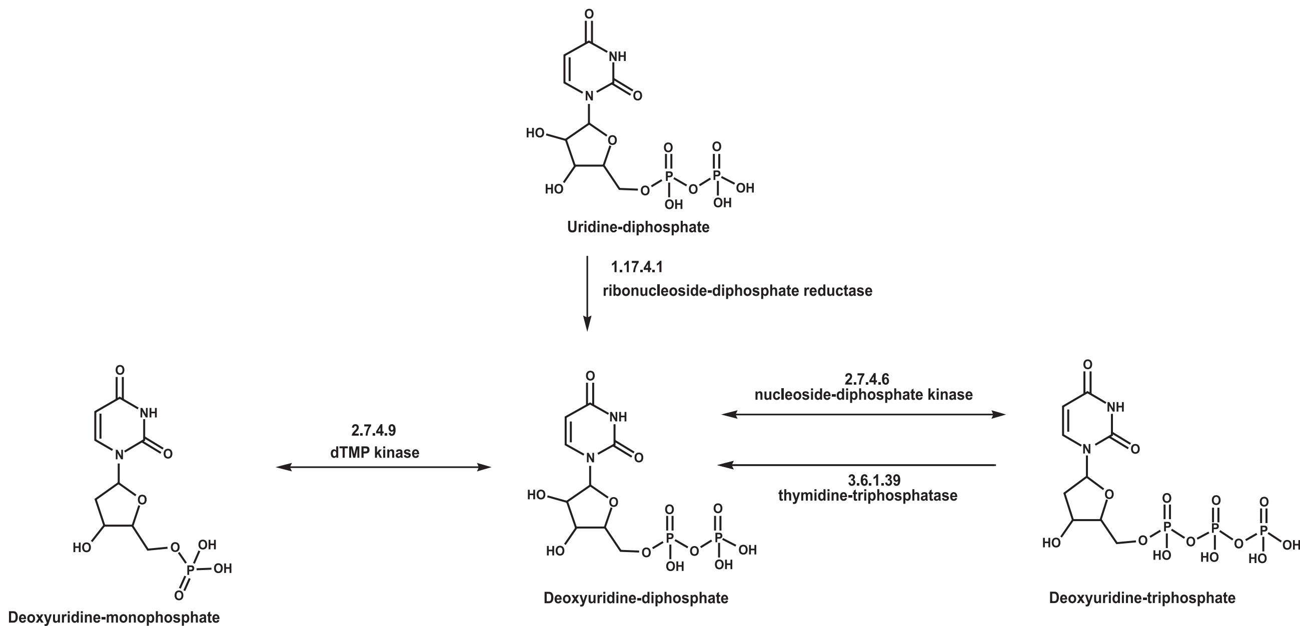

Fig 13. The proposed metabolic pathways of deoxyuridine-diphosphate (dUDP).

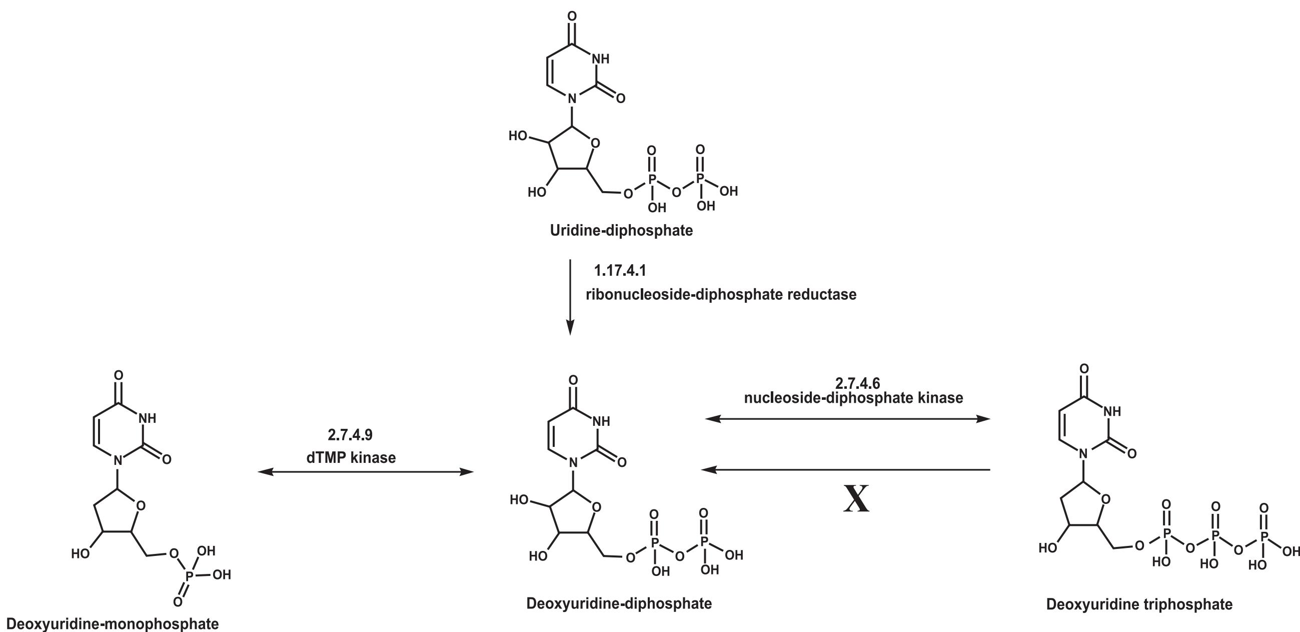

Fig 14. The proposed metabolic pathways of deoxyuridine-diphosphate (dUDP) in camels and *Trypanosoma evansi*.

insertion and deletion of 11 and 8 AA, respectively. This observation suggests that the positions of insertion/deletion may be not crucial for the protein function.

The sequence of UMP/CMP kinase of dromedary camel and other 11 species were aligned to get a larger view of variations across different species (Fig 6). The variations concentrated within the first 40 AA while, the rest showed a relative conservativeness. Only UMP/CMP kinase of *Camelus bactrianus* has insertion and deletion as described previously. In contrast to sequence of UMP/CMP kinase of the studied eukaryotic species, sequence UMP/CMP kinase of *Trypanosoma cruzei* and some protozoal is highly variable (Fig 7). We tried several BLAST searches of *T. brucei* and *T. evansi* genome using several UMP/CMP kinases, however, there outcome was zero hits.

Phylogenetic tree were reconstructed to visualise the evolutionary distance between camel

and *T. cruzei* bifunctional UMP/CMP kinase in relation to UMP/CMP kinase of a set of other eukaryotic organism (Fig 8). The protozoal enzyme was highly related to its protozoal orthologs and highly distant form the camel enzyme.

ATP - diphosphatase/ectonucleoside triphosphate diphosphohydrolase

ATP- diphosphatase (apyrases (ATP diphosphohydrolases; EC 3.6.1.5; the nucleoside triphosphate phosphohydrolases (nucleoside monophosphoate-forming); ATP- diphosphatase) act correspondingly on di- and triphosphate nucleotides (NDPs and NTPs) and dephosphorylate NTPs to nucleotide monophosphates (NMPs) by 2 separate succeeding phosphate-releasing steps, in which NDPs are intermediates. Unlike, apyrases differ from ATPases that distinctively hydrolyse ATP, by dephosphorylating both ATP and ADP. The eukaryotic enzymes necessitates Ca^{2+} and Mg^{2+} can alternative. The bulk of the ecto-ATPases found on be the cell surface and hydrolyse extracellular nucleotides that are target substrates for this enzyme family (Komoszynski and Wojtczak, 1996). The cells of *Trypanosoma brucei* can hydrolyse extracellular ATP. The absence of any divalent metal reduced the hydrolysis level of ATP in this protozoon parasite. Moreover, $MgCl_2$ promote ATP hydrolysis. Similarly, when $MnCl_2$ replacing $MgCl_2$, the acceleration of activity was also detected. Correspondingly, $CaCl_2$ and $ZnCl_2$ could speed up the ATPase activity, while less than $MgCl_2$. This ecto-ATPase activity was not sensitive to inhibitors than other ATPase and phosphatase activities. Living cells successively hydrolysed the ATP molecule creating ADP, AMP and adenosine and administration of ATP and adenosine was able to support the multiplication of *T. brucei*. Furthermore, the availability of purines controlled the activity of *T. brucei* ATP- diphosphatase (de Souza Leite *et al,* 2007).

Alignment of dromedary camel and human ectonucleoside triphosphate diphosphohydrolase showed variations among the length of the two protein sequences (Fig 9). The longest consecutive conserved sequence did not exceed 25 AA showing 149 differences and homology about 70.84%. However, sequence of ectonucleoside triphosphate diphosphohydrolase of the 3 camel species is identical and have 100% similarity (Fig 10). BLAST search of the obtained *T. evansi* protein against camel database revealed close similarity to ectonucleoside triphosphate diphosphohydrolase 5 isoform X2 (*Camelus dromedarius*). Fig 11 shows comparison of camel and *T. evansi* enzymes. There was low similarity rate of 26.25%.

The evolutionary distance between camel ectonucleoside triphosphate diphosphohydrolase and a set of eukaryotic organisms were shown by constructing a phylogenetic tree (Fig 12). The camel enzyme was distant from other vertebrates and highly related to the protozoal enzymes.

Motif and domain search by using prosite, NCBI domain or motif finder revealed conservative one domain content of nucleoside phophatease family in both of camel and *T. evansi* ATP-diphosphatase.

Nucleoside triphosphate phosphohydrolase

Nucleoside triphosphate phosphohydrolases (nucleoside monophosphoate-forming, ATP-diphosphatase, apyrase, ectonucleoside triphosphate diphosphohydrolase, EC no. 3.6.1.5) present in vertebrates including camel and some protozoa. The enzyme was previously found by experimental procedures (Weiss *et al,* 2015), yet it was not found in the gene sequence databases after searching by accession number or name of enzyme.

Deoxyuridine 5-diphosphate (dUDP)

The production of dUDP was proposed by three production routes. In the synthetic direction, deoxyuridine-5′-monophosphate (dUMP) is phosphorylated by the kinase activity of dTMP kinase. The degeneration of deoxyuridine 5′-triphosphate (dUTP) by nucleoside diphosphate phosphatase, thymidine triphosphatase can yield dUDP as shown in (Fig 13). The reduction of UDP will give dUDP by the action of ribonucleoside diphosphate reductase. After bioinformatics tools application, the revised metabolic pathways of camels and *T. evansi* is provided in Fig 14. Camel and *T. evansi* shareed was common features of dUDP pathways comprising the lack of thymidine triphosphatase.

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).

References

Alfuwaires M, Altaher A and Kandeel M (2017). Molecular dynamic studies of interferon and innate immunity resistance in MERS CoV non-structural protein 3. Biological and Pharmaceutical Bulletin 40:345-351.

Alnazawi M, Altaher A and Kandeel M (2017). Comparative genomic analysis MERS CoV isolated from humans and camels with special reference to virus encoded helicase. Biological and Pharmaceutical Bulletin 40:1289-1298.

Dahlmann N (1982). Human serum thymidine triphosphate nucleotidohydrolase: purification and properties of a new enzyme. Biochemistry 21:6634-6639.

de Souza Leite M, Thomaz R, Fonseca FV, Panizzutti R, Vercesi AE and Meyer-Fernandes JR (2007). *Trypanosoma brucei*: biochemical characterisation of ecto-nucleoside triphosphate diphosphohydrolase activities. Experimental Parasitology 115:315-323.

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

Hertz-Fowler C, Peacock CS, Wood V, Aslett M, Kerhornou A, Mooney P, Tivey A, Berriman M, Hall N and Rutherford K (2004). GeneDB: a resource for prokaryotic and eukaryotic organisms. Nucleic Acids Research 32:D339-D343.

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

Jirimutu, Wang Z, Ding G, Chen G, Sun Y, Sun Z, Zhang H, Wang L, Hasi S, Zhang Y, Li J, Shi Y, Xu Z, He C, Yu S, Li S, Zhang W, Batmunkh M, Ts B, Narenbatu, Unierhu, Bat-Ireedui S, Gao H, Baysgalan B, Li Q, Jia Z, Turigenbayila, Subudenggerile, Narenmanduhu, Wang Z, Wang J, Pan L, Chen Y, Ganerdene Y, Dabxilt, Erdemt, Altansha, Altansukh, Liu T, Cao M, Aruuntsever, Bayart, Hosblig, He F, Zha-ti A, Zheng G, Qiu F, Sun Z, Zhao L, Zhao W, Liu B, Li C, Chen Y, Tang X, Guo C, Liu W, Ming L, Temuulen, Cui A, Li Y, Gao J, Li J, Wurentaodi, Niu S, Sun T, Zhai Z, Zhang M, Chen C, Baldan T, Bayaer T, Li Y and Meng H (2012). Genome sequences of wild and domestic bactrian camels. Nature Communications 3:1202.

Kandeel M and Al-Taher A (2020a). Bioinformatics of uridine/deoxyuridine paths in *Trypanosoma evansi* revealed targeting uridine phosphorylase and cytidine deaminase. Journal of Camel Practice and Research 27:111-119.

Kandeel M and Al-Taher A (2020b). Uridine 5'-monophosphate (UMP) metabolising enzymes uracil phosphoribosyltransferase and orotidine-5'-phosphate decarboxylase/ump synthase in camels and *Trypanosoma evansi*. Journal of Camel Practice and Research 27:81-88.

Kandeel M, Al-Taher A, Li H, Schwingenschlogl U and Al-Nazawi M (2018). Molecular dynamics of Middle East Respiratory Syndrome Coronavirus (MERS CoV) fusion heptad repeat trimers. Computational Biology and Chemistry 75:205-212.

Kandeel M, Dalab A, Al-Shabebi A, Al-Taher A, Altaher Y and Abozahra M (2020). Uridine 5-triphosphate (UTP) metabolising enzymes nucleoside diphosphate kinase and cytidine triphosphate (CTP) synthase in camels and *Trypanosoma evansi*. Journal of Camel Practice and Research 27:61-68.

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

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

Komoszynski M and Wojtczak A (1996). Apyrases (ATP diphosphohydrolases, EC 3.6.1.5): function and relationship to ATPases. Biochimica et Biophysica Acta 1310:233-241.

Labarga A, Valentin F anderson M and Lopez R (2007). Web services at the European Bioinformatics Institute. Nucleic Acids Research 35:W6-W11.

Liou JY, Dutschman GE, Lam W, Jiang Z and Cheng YC (2002). Characterisation of human UMP/CMP kinase and its phosphorylation of D- and L-form deoxycytidine analogue monophosphates. Cancer Research 62:1624-1631.

Madden T (2013). The BLAST sequence analysis tool. In: The NCBI Handbook (Internet) 2nd ed. National Centre for Biotechnology Information (US).

Mahmoud K, Kamal E, Wael E, Mahmoud F and Ibrahim G (2019). Species specificity and host affinity rather than tissue tropism controls codon usage pattern in respiratory mycoplasmosis. Journal of Camel Practice and Research 26:1-12.

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

Marco-Marin C, Gil-Ortiz F and Rubio V (2005). The crystal structure of *Pyrococcus furiosus* UMP kinase provides insight into catalysis and regulation in microbial pyrimidine nucleotide biosynthesis. Journal of Molecular Biology 352:438-454.

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

Pasti C, Gallois-Montbrun S, Munier-Lehmann H, Veron M, Gilles AM and Deville-Bonne D (2003). Reaction of human UMP-CMP kinase with natural and analog substrates. European Journal of Biochemistry 270:1784-1790.

Sequencing H (2011). CLC Genomics Workbench. Workbench.

Serina L, Blondin C, Krin E, Sismeiro O, Danchin A, Sakamoto H, Gilles AM and Barzu O (1995). *Escherichia coli* UMP-kinase, a member of the aspartokinase family, is a hexamer regulated by guanine nucleotides and UTP. Biochemistry 34:5066-5074.

Sievers F and Higgins DG (2014). Clustal Omega, accurate alignment of very large numbers of sequences. In: Multiple Sequence Alignment Methods. Springer. pp 105-116.

Weiss PH, Batista F, Wagner G, Magalhaes Mde L and Miletti LC (2015). Kinetic and biochemical characterisation of *Trypanosoma evansi* nucleoside triphosphate diphosphohydrolase. Experimental Parasitology 153:98-104.

News

LOOK CAMEL FOR MILK AND BEYOND MILK

The astonishing adaptations of the camel aren't restricted to its milk. Here are a few more that have their roots in the animal's genes.

1. The urea cycle conserves nitrogen, used to make protein in the face of limited food.
2. The heart makes a lot of the protein α-actinin, enabling it to beat steadily through fluctuating drought and wetness.
3. Hiked enzyme levels in the brain and liver keep up energy supplies.
4. Vimentin protein in cells in the hump mobilizes fat, releasing energy.
5. A camel's streamlined nanobodies arose from a mutation that removed the hinges that connect the Y-shaped arms of more conventional antibodies. Sometimes a mutation is a good thing!
6. Further infection protection comes from the milk protein lactoferrin, which fights hepatitis C.
7. Tolerating High Blood Sugar by camel-herding people in India, the Raika, drink camel milk and don't get diabetes.

(Camel Milk and Autism: Connecting the Genetic Dots, November 21, 2019 Ricki Lewis, PLOS BLOGS DNA Science)