African Journal of Biotechnology Vol. 10(36), pp. 6846-6851, 18 July, 2011
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB10.2498
ISSN 1684–5315 © 2011 Academic Journals

Full Length Research Paper

Molecular cloning and characterization of cDNA encoding *Camelus dromedarius* putative glucose 6-phosphate dehydrogenase

Hesham Mahmoud Saeed[1,2]*, Mohammad Saud Alanazi[1] and Mohammad Daoud Bazzi[1]

[1]The Genome Research Chair, Biochemistry Department, College of Science King Saud University, Bld. 5, Lab AA10, P. O. Box: 2454 Riyadh, Kingdom Saudi Arabia.
[2]Department of Biotechnology, Institute of Graduate Studies and Research, Alexandria University, Alexandria, Egypt.

Accepted 3 June, 2011

This study determined the partial coding sequence of glucose-6-phosphate dehydrogenase (G6PD) cDNA from the Arabian camel *Camelus dromedarius* using reverse transcription polymerase chain reaction (RT-PCR). Degenerate PCR primers were designed from the available data base deposited in the GenBank. Applying these primers to cDNA prepared from isolated and purified total RNA from Arabian camel liver led to the generation of two PCR products corresponding to 969 and 477 bp. These PCR products were cloned on pGEM-T-Easy cloning vector separately. Sequencing of the PCR generated cDNA fragments indicated that the 969 bp fragment represents partial coding sequence of camel G6PD and showed 93, 91 and 86% similarity with the *Homo sapiens*, *Bos taurus* and *Rattus norvegicus* G6PD, respectively. The *Camelus dromedaries* G6PD partial coding sequence encodes an open reading frame of a truncated protein consists of 323 amino acid residues that showed the conserved substrate binding site and the dinuclotide-binding fingerprint.

Key words: Arabian camel, degenerate primers, glucose 6-phosphate dehydrogenase, reactive oxygen species (ROS).

INTRODUCTION

The Arabian camel *Camelus dromedarius* is able to live in the harsh desert conditions such as high temperature, direct exposure to sunlight and without drinking water for weeks. Despite being an essential source of food in many parts of the Arabian Peninsula, the one-humped camel (*C. dromedarius*) is the least studied of all domestic species of mammals (Hussein et al., 1992; Mejdoub et al., 1994). The camel genome consists of 74 chromosomes (Al-Khedhairy, 2004; Thomas et al., 1985) and although, there are wide variations in the DNA sequences between camel and other species including human, still there are functional similarities which make the camel protein studies a good target for human applications. For example, the nucleotide sequences of the camel ubiquitin are different from those of humans and mice, the deduced amino acid sequence was found to be fully conserved and 100% similar to that of the polyubiquitin sequence of various species.

The complete conservation of the amino acid sequence of ubiquitin, which was observed in the functional genes that have been sequenced, is in accordance with the fact that ubiquitin is completely conserved through evolution (Al-Khedhairy, 2004). Although, more attention has been given to camel research recently, the camel is still currently a minor species in terms of pharmacological and biomolecular research. Living organisms including camel are continuously exposed to harmful agents from the environment and from the endogenous metabolic processes (Barnes and Lindahl, 2004). Reactive oxygen species (ROS) exert damage to DNA, and therefore pose a major threat to the genetic integrity of cells. While modified proteins and lipids can be degraded and re-synthesized, DNA must be repaired before replication and cell division takes place.

*Corresponding author. E-mail: hesham25166@yahoo.com, hesham25166@hotmail.com, hsaeed1@KSU.EDU.SA. Tel: 00966590236357. Fax: 00966124675791.

In this study, carefully designed highly degenerate PCR primers were used to isolate glucose-6-phosphate dehydrogenase (G6PD) cDNA from one-humped Arabian camel. G6PD is an important cytosolic enzyme involved in carbohydrate metabolism. This enzyme is responsible for the conversion of glucose-6-phosphate to D-ribulose-5-phosphate in two steps mediated by the formation of 6-phosphogluconate with the release of NADPH. The released NADPH consumed in the elimination of reactive oxygen species through many metabolic pathways that involved NADPH as a co-enzyme. Hence, G6PD deficiency impairs the ability of some cells such as erythrocytes to form NADPH resulting in hemolysis and anemia (Beutler, 1994; Abboud and Al-Awaida, 2010). The aim of this study was to undertake the isolation of *C. dromedarius* G6PD cDNA by using degenerate PCR primers. This will enable a study of substrate specificities that may differ from those for related proteins in man, the extent to which these genes are inducible, whether they respond to the same or different inducers and so indicate avenues for the complete characterization of these important proteins in camel.

MATERIALS AND METHODS

Chemicals

All chemicals used were of analytical reagent, molecular biology or chromatographic grade as appropriate. Water was de-ionized and distilled.

Animals and tissue preparation

Liver from three adult male Arabian one-humped camels (*C. dromedarius*, 500-800 kg, 6 to 8 years old) were obtained from the local slaughterhouse, after the animals were killed under the observation of a skilled veterinarian.

RNA isolation and cDNA synthesis

Tissue samples for RNA analysis were immediately submerged in RNA later solution (Ambion, Courtabeuf, France) to avoid RNA degradation, stored at 4°C for 24 h and then at -20°C until used for RNA isolation. 30 to 60 mg of tissue specimens were homogenized in RTL lyses buffer (Qiagen) supplemented with 1% 2-mercaptoethanol, using a rotor-stator homogenizer. Total RNA were extracted using the RNeasy Mini Kit (Qiagen), with a DNase digestion step, according to the manufacturer's instruction.

Elution was performed with 50 µl nuclease free water. Concentrations and integrity of RNA samples were assessed using NanoDrop-8000 and formaldehyde agarose gel (1%) electrophoresis in 1X TAE buffer according to Sambrook et al. (1989). Two microgram of the total RNAs were retrotranscribed into single stranded cDNA using ImProm-II Reverse Transcription System (Cat# A3800, Promega Co.).

cDNA isolation and cloning

Complementary DNA was synthesized by reverse transcription as described earlier. G6PD cDNA was isolated by polymerase chain reaction (PCR) using degenerate primers designed from the available data base deposited in the GenBank. The designing of the degenerate oligonucleotide primer sets is based on the highly conserved nucleotide sequences at the 5' and 3' cDNA ends of G6PD cDNA from *Homo sapiens*, *Bos taurus*, *Sus scrofa*, *Capra hircus*, *Ovis aries*, *Oryctolagus cuniculus, Xenopus tropicalis*, *Rattus norvegicus* and *Mus musculus*. The degenerate primers were as follow: G6PD forward and reverse; 5' – TGA GCC GGA CCC AGG TGT G- 3'; 5' – GGG GTT CCC CAC RTA CTG RCC- 3'. PCR was carried out in a final volume of 50 µl as follow: 25 µl of high fidelity PCR master mix, 5 µl of cDNA (200 ng), 3 µl of each forward and reverse primers (30 pmole) for each gene then the final volume was adjusted to 50 µl with nuclease free water. The PCR condition was 1 cycle at 95°C for 5 min followed by 30 cycles at 94°C for 1 min, 52°C for 1 min and 72°C for 1.5 min. Final extension was carried out at 72°C for 5 min. The PCR products were analyzed using agarose gel (1.5%) electrophoresis in 1X TAE buffer. To facilitate cloning of the PCR products onto pGEM-T-Easy vector (Zhou et al., 1995), the PCR products were treated with *Taq* polymerase as follow: 10 µl of each PCR product was taken into a 0.5 ml tube to which 2 µl of *Taq* polymerase buffer and 2 units of *Taq* polymerase and dATP were added. The final volume was adjusted to 20 µl and the tube was incubated at 72°C for 30 min. This treatment facilitates the addition of 5' A to the PCR product to enhance the T/A cloning strategy. To ligate the treated PCR products onto pGEM-T- Easy vector, 2 µl of PCR product was taken in a clean 0.5 ml tube to which 1 µl pGEM-T-Easy vector (50 ng) and 1 µl of 10X ligase buffer were added followed by the addition of 2 units of ligase enzyme. The final volume of the ligation reaction was adjusted to 10 µl by the addition of nuclease free water. The tube was incubated at 16°C for 16 h. Transformation of *Escherichia coli* DH5α competent cells was carried out according to Sambrook et al. (1989).

Screening of recombinant *E. coli* DH5α

The recombinant *E. coli* DH5α harboring pGEM-T-Easy vector obtained was screened in selective LB/IPTG/X-gal/Ampicillin/agar plates. Moreover, colonies PCR was conducted to screen recombinant bacteria for ligated DNA insert using T7/SP6 multiple cloning site promoter primers that amplify the cloned cDNA inserts. A small part of each bacterial colony was transferred to a clean sterile Eppendorff tube, to which the rest of the PCR reaction components was added as described earlier. The PCR condition was as follow; 1 cycle at 95°C for 5 min followed by 30 cycles at 94°C for 1 min, 50°C for 1 min and 72°C for 2 min. The PCR products were analyzed by agarose gel (1.5%) electrophoresis in 1X TAE buffer. Plasmids were prepared from some positive clones using the PureYield Plasmid Miniprep System (A1222, Promega, USA).

Sequencing of the PCR products

Sequencing of the PCR product cloned onto pGEM-T-Easy vector was carried out according to Sanger et al. (1977) using MegaBACE 1000 DNA Sequencing System (Pharmacia/Amersham Co.). Sequencing primers used were as follows; SP6, 5'-TATTT AGGTGACACTATAG-3'; T7, 5'-TAATACGACTCACTATAGGG-3'. The chain termination sequencing reaction was conducted utilizing the DYEnamic ET terminator kit as an integral part of the MegaBACE 1000 DNA sequencing system. The cDNA and amino acid sequences of the camel G6PD were aligned with those of other mammals using the ClustalW and DNA STAR programs. Neighbor-Joining phylogenetic tree of G6PD from different species was constructed by the use of a program to compute the distance

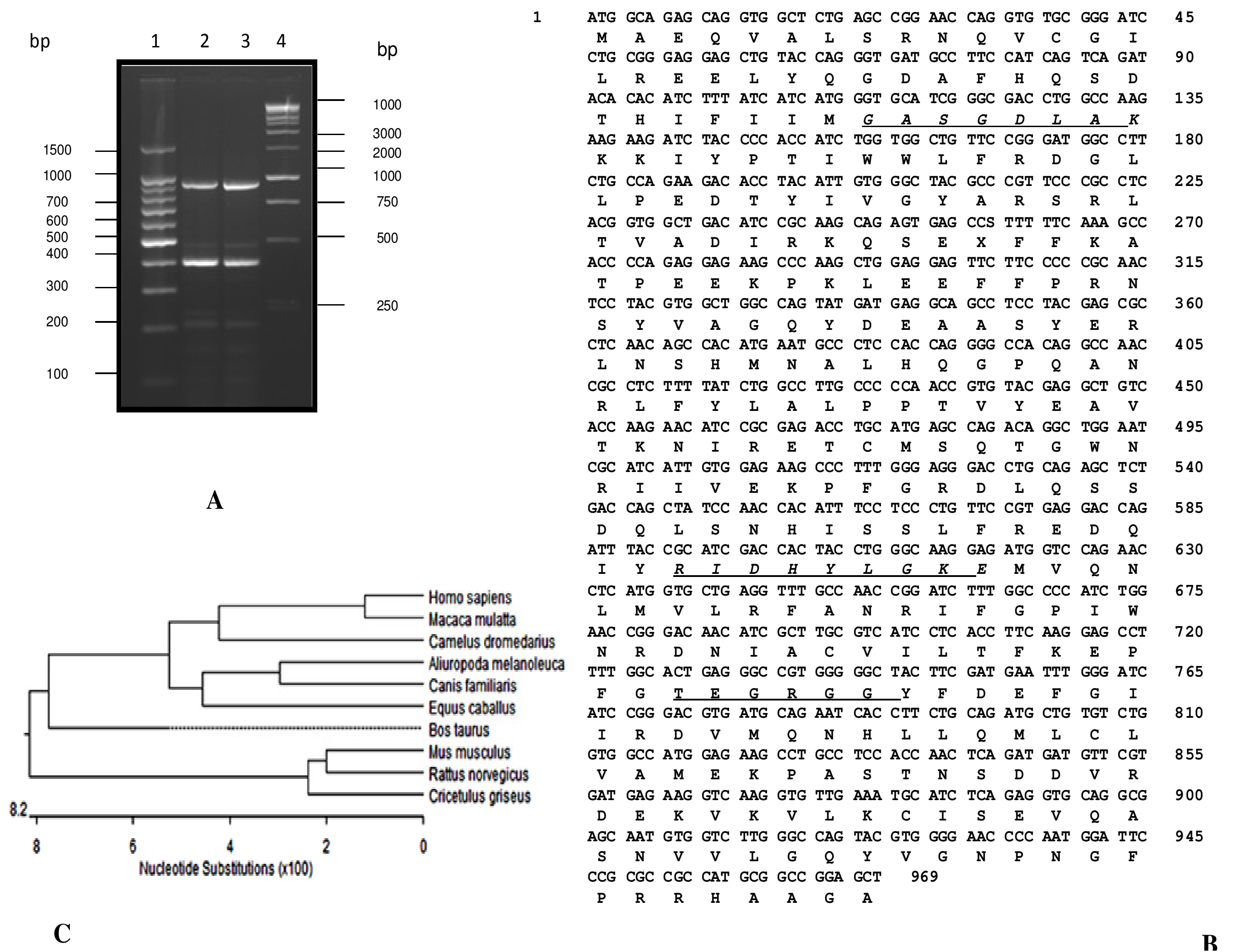

Figure 1. (A) Agarose gel (1.5%) electrophoresis of PCR products for *C. dromedarius* G6PD at 50 and 52°C annealing temperatures (Lanes 2 and 3). Lanes 1 and 4 represent 100 bp and 1 kbp Promega DNA molecular weight markers. (B) The deduce amino acid sequences of the *C. dromedarius* G6PD partial coding cDNA. The conserved substrate binding site and the dinucleotide-binding fingerprint ($NADP^+$) are in italic underline. (C) Phylogenetic tree constructed by the neighbour-joining method based on the partial coding sequence of *C. dromedarius* G6PD and G6PD from other species deposited in the GenBank. Bootstrap values are indicated when ≥50% as a percentage obtained from 1,000 resamplings of the data. The scale bar represents a genetic distance of 0.05 substitutions per site.

matrix DNADIST version3.5c to show the evolutionary relationships.

RESULTS

PCR and G6PD cDNA isolation

In an attempt to isolate *C. dromedarius* G6PD, PCR based technique was used. G6PD gene specific degenerate primers were designed from the available data in the GenBank. These primers specify some conserved sequences in the G6PD gene covering the NADP binding site and part of the conserved C terminal end. Figure 1A shows the result of the PCR at 50 and 52°C annealing temperatures. Two different specific PCR products were obtained corresponding to 969 and 471 bp, respectively (Figure 1A). Both of these PCR products were found to be reproducible at different annealing temperatures and under identical amplification conditions indicating that these generated DNA fragments or at least one of them were likely to encode part of the *C. dromedarius* G6PD cDNA.

Cloning and sequencing of the PCR products

To facilitate the sequencing of these amplified products, T/A cloning strategy was applied. Purified PCR products

were ligated into pGEM-T-Easy cloning vector as previously mentioned in the materials and methods. The presence of a cloned cDNA insert was tested and verified by PCR amplification of the plasmid with T7/SP6 primers, which anneal to the vector regions flanking the cDNA insert. Several clones containing the cDNA insert were then sequenced to ensure their uniqueness. Sequencing of several clones revealed that most of the amplified 969 bp cDNA inserts were uniform and corresponded to G6PD. Nucleotide sequences corresponding to 969 bp were then compared with the nucleotide sequences deposited in the GenBank database using the BLASTn program on the NCBI BLAST server. Only those clones that have statistically significant similarity scores to numerous G6PD and no other genes were considered to be G6PD. The partial cDNA sequence of *C. dromedarius* G6PD was deposited in the GenBank under the accession number GU810157.

Figure 1B shows the nucleotide sequence and the partial coding sequence of the 969 bp PCR generated fragment. The deduced amino acid sequence was found to consist of an open reading frame of 323 amino acid residues. The multiple sequence alignment includes some of the best characterized representative features of G6PD such as glucose 6-phosphate substrate binding site and the dinucleotide-binding fingerprint, $NADP^+$ (Figure 2).

Based on the amino acids multiple alignment results, it was observed that higher similarity was obtained with the *H. sapiens* G6PD, accession number X03674 (93%), followed by *B. taurus,* accession number XM_002699717 (91%) and *Canis familiaris*, accession number ABW-89488 (90%) as shown in Figure 2 and Table 1. The phyolgenetic analysis of the *C. dromedarius* G6PD cDNA indicated that the camel G6PD was located beside *H. sapiens*, *Macaca mulatta* and *Aliuropoda melanoleuca* as shown in Figure 1C.

DISCUSSION

Several enzymes are responsible for NADPH production, including cytosolic NADP isocitrate dehydrogenase, cytosolic NADP malic enzyme and the first two enzymes of pentose phosphate shunt, glucose 6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase. The last two are recognized as the main suppliers in ruminants providing the 50 to 80% of the required NADPH for fatty acid synthesis as well as for the detoxification reactions (Belk et al., 1993; Laliotis et al., 2007; Vernon, 1981). G6PD has been extensively studied in recent years especially in human and rodents and complete characterized gene sequences has already been elucidated for the aforementioned species. In human, G6PD maps to Xq28 and it consists of 13 exons with a coding region of 1545 bp. Moreover, more than 100 deficient variant have been reported. In sheep, although, G6PD is one of the main suppliers of NADPH for fatty acid synthesis, the study of the gene has been mainly conducted on enzymatic level.

The design of degenerate PCR primers is of great important for the identification and characterization of novel members of a class of enzymes. Optimally, highly conserved functional domains or sequences are used for degenerate primers design as it has done before for the detection of halocarboxilic acid dehydrogenase (Hill et al., 1999); ring hydroxylating dehydrogenases and alkane hydroxylases (Smits et al., 1999; Yates et al., 2000). In this study, we showed amplification of the *C. dromedarius* G6PD cDNA partial coding sequence using a degenerate primer set spanning the substrate (G6P) recognition and the $NADP^+$ binding sites. A cDNA fragment with the anticipated size of 969 bp (Figure 1) was obtained and its sequence matched several G6PD sequences in GenBank. *C. dromedarius* G6PD partial cDNA sequence consisted of 323 amino acid residues (Figure 1B). Several observations from the primary structure (Figure 1B) and from the multiple sequence alignment (Figure 2) merit discussion.

First, the primary sequence homology between *C. dromedarius* G6PD and other species was greater than 85% (Table 1 and Figure 2). This confirms the specificity of the degenerate primer set. Second, the primary structure (Figure 1B and 2) showed two conserved sequences motifs. The conserved eight-residue peptide RIDHYLGK (residues from 198 to 205), corresponding to the substrate binding site (Shannon et al., 2000) and the dinucleotide-binding fingerprint ($NADP^+$) GXXGDLX (residues from 38 to 44) that represents a part of the N-terminus of most mammalian G6PD, where R is arginine, I is isoleucine, D is aspartic, H is histidine, Y is tyrosine, L is leucine, G is glycine and X is any amino acid. Bautista et al. (1995) demonstrated that site-directed mutagenesis of Lys^{205} to Thr or Arg (K205T and K205R) resulted in a substantial loss of catalytic activity and these results confirms that Lys^{205} is essential for catalytic activity because it is in or near to the G6P binding site (Bautista et al., 1995). Other features such as the conserved site among most of the mammals F^{372} – N^{413} (where F is phenylalanine and N is asparagine) proposed by Hirono and Beutler, (1989) to be the second coenzyme $NADP^+$ binding domain is not shown in the *C. dromedarius* partial cDNA sequence.

In conclusion, the degenerate primers designed in this study allowed the amplification of a partial sequence of G6PD gene from the *C. dromedaries*; however, more sequencing will be required to get the full length G6PD gene in an attempt to express this gene and to characterize the enzyme in term of kinetic properties. Moreover, the data provided here will also allow the further study of camel G6PD gene, such as the *in vitro* expression of camel G6PD or analysis of the 5′ regulatory region as a tool to examine the mechanism of lipogenesis in camel. To this end, it would be useful to understand

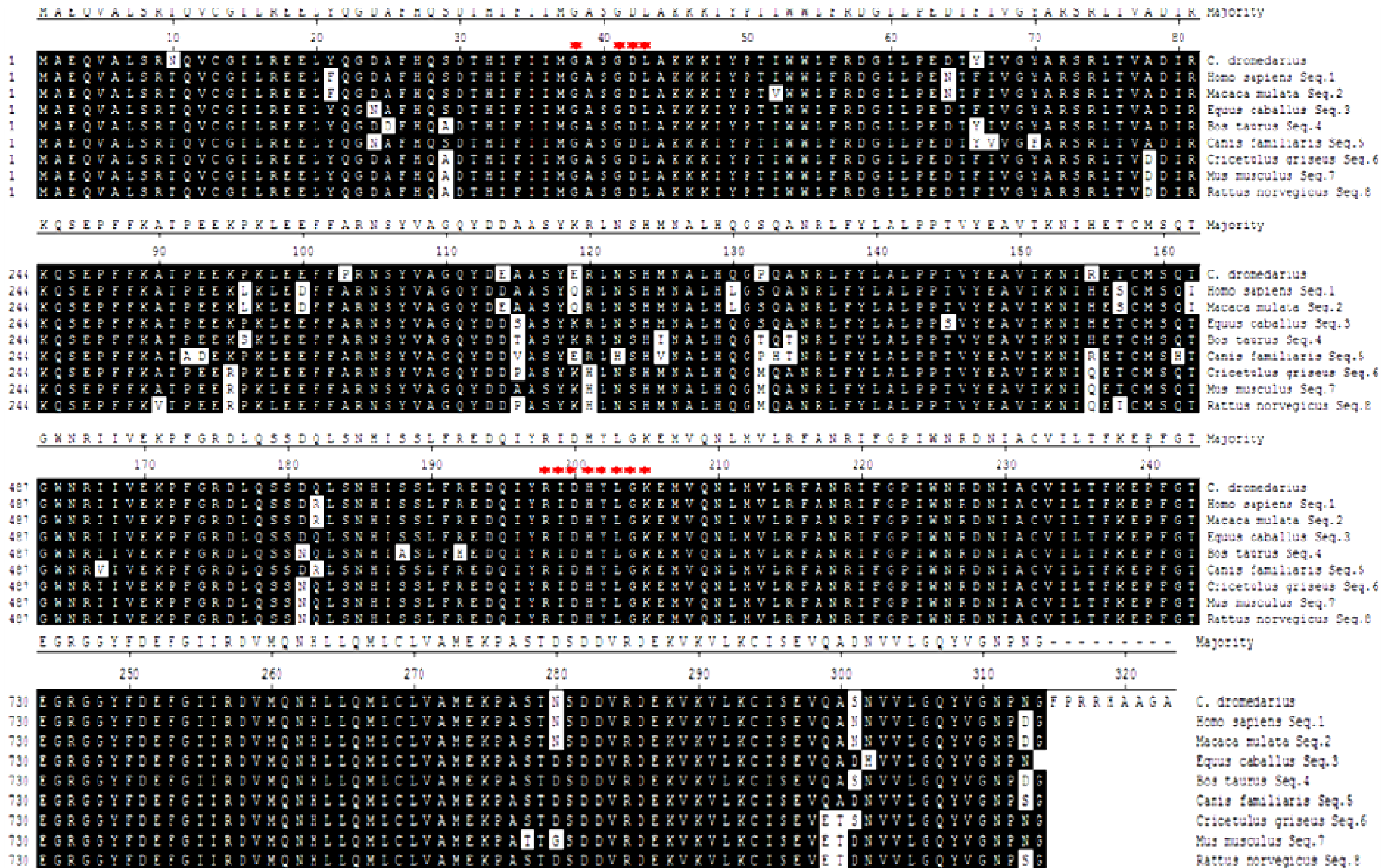

Figure 2. Multiple alignment of the deduced amino acid sequences of *C. dromedarius* G6PD and other organisms (*Homo sapiens*: accession number X03674; *Macaca mulatta*: accession number XM-002806445; *Bos taurus*: accession number XM-002699717; *Equus caballus*: accession number XM-001492232; *Canis familiaris*: accession number ABW89488; *Cricetulus griseus*: accession number AF044676; *Rattus norvegicus*: accession number BC081820 and *Mus musculus*: accession number BC075663). Dark shaded boxes indicate the region conserved among all aligned sequences. Asterisks indicate the conserved substrate and $NADP^+$ binding sites

Table 1. Homology of the *C. dromedaries* deduced amino acid sequence with other G6PD genes.

Animal species	Gene	Accession number	Homology (%)
Homo sapiens	G6PD	X03674	93
Macaca mulana	G6PD	XM-002806445	93
Bos taurus	G6PD	XM-002699717	91
Equus caballus	G6PD	XM-001492232	91
Aliuropoda melanoleuca	G6PD	XM-002929683	90
Canis familiaris	G6PD	ABW89488	90
Cricetulus griseus	G6PD	AF044676	89
Rattus norvegicus	G6PD	BC081820	86
Mus musculus	G6PD	BC075663	86

more about the regulation of camel G6PD gene expression in order to consider up or down regulation of the appropriate promoter element. Furthermore, it could be very significant to analyze the substrate specificity for this protein in greater detail in order to achieve an appropriate functional characterization.

ACKNOWLEDGEMENT

The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for funding the work through the research group project number RGP-VPP-081.

REFERENCES

Al-Khedhairy A (2004). Characterization of the Nucleotide Sequence of a Polyubiquitin Gene (PUBC1) from Arabian Camel, *Camelus dromedaries*. J. Biochem. Mol. Biol. 37: 144-147.

Abboud M, Al-Awaida W (2010). Synchrony of G6PD activity and RBC fragility under oxidative stress exerted at normal and G6PD deficiency. Clin. Biochem. 43: 455-460.

Barnes DE, Lindahl T (2004). Repair and genetic consequences of endogenous DNA base damage in mammalian cells. Annu. Rev. Genet. 38: 445-476.

Bautista JM, Mason PJ, Luzzatto L (1995). Human glucose-6-phosphate dehydrogenase Lysine 205 is dispensable for substrate binding but essential for catalysis. FEBS Lett. 366: 61-64

Belk KE, Savell JW, Davis SK, Taylor JF, Womack JE, Smith SB (1993). Tissue specific activity of pentose cycle oxidative enzymes during feeder lamp development. J. Anim. Sci. 71: 1796-1804.

Beutler E (1994). G6PD deficiency. Blood, 84: 3613-3636.

Hill KE, Marchesi JR, Weightman AJ (1999). Investigation of two evolutionarily unrelated halocarboxylic acid dehydrogenase gene families. J. Bacteriol. 181: 2535-2547.

Hirono A, Beutler E (1989). Alternative splicing of human glucose 6-phosphate dehydrogenase messenger RNA in different tissues. J. Clin. Invest. 83: 343-346.

Hussein MF, Al-Moen AK, Gader AMA (1992). Haemostatic parameter in the camel (*Camelus dromedarius*) comparison with human. Comp. Haematol. Int. 2: 92-96.

Laliotis G, Argyrokastritis PA, Bizelis I, Rogdakis E (2007). Cloning and characterization of an alternative transcript of ovine glucose 6-phosphate dehydrogenase gene:comparative approach between ruminant and non-ruminant species. Gene, 388: 93-101.

Mejdoub H, Reinbolt J, Gargouri K (1994). Dromedary pancreatic lipase: purification and structural properties. Biochim. Biophys. Acta. 1213: 119-126.

Sambrook J, Frisch EF, Maniatis T (eds) (1989). Molecular cloning: A Laboratory Manual, 2nd ed. Cold spring Harbor laboratory press, New York.

Sanger F, Nicklen S, Coulson AR (1977). DNA sequencing with chain terminating inhibitor. Proc. Natl. Acad. Sci. USA. 74: 5463-5467.

Shannon A, Gover S, Lam V, Adams MJ (2000). Human glucose-6-phosphate dehydrogenase:the crystal structure reveals a structural NAD+ molecule and provides insight into enzyme deficiency. Structure, 8(3): 293-303.

Smits TH, Rothlisberger M, Witholt B, van Beilen JB (1999). Molecular screening for alkane hydroxylase genes in Gram negative and Gram positive strains. Environ. Microbiol. 1: 307-317.

Thomas D, Bunch Warren C, Foote A M (1985). Chromosome banding pattern homologies and NORs for the Bactrian camel, guanaco, and llama. J. Heredity, 76: 115-118.

Vernon RG (1981). Lipid metabolism in the adipose tissue of ruminant animals. In: Christie WW (Ed), Lipid Metabolism in Ruminant Animal. Pergamon, Oxford, New York, pp. 279-362.

Yates C, Holmes AJ, Gilling MR (2000). Novel forms of ring-hydroxylating dioxygenases are widespread in pristine and contaminated soils. Environ. Microbiol. 2: 644-653.

Zhou MY, Clark SE, Gomez-Sanchez CE (1995). Universal Cloning Method by TA Strategy. Biotechniques, 19: 34-35.