

Characterization of the Nucleotide Sequence of a Polyubiquitin Gene (PUBC1) from Arabian Camel, *Camelus dromedarius*

Abdulaziz Ali A. Al-Khedhairi*

Department of Zoology, College of Sciences, P. O. Box 2455, King Saud University, Riyadh 11451, Saudi Arabia

Received 5 February 2003, Accepted 18 June 2003

Molecular amplification and sequencing of genomic DNA that encodes camel polyubiquitin (PUBC1) was performed by a polymerase chain reaction (PCR) using various sets of primers. The amplification generated a number of DNA fragments, which were sequenced and compared with the polyubiquitin coding sequences of various species. One DNA fragment that conformed to 325 bp was found to be 95 and 88% homologous to the sequences of human polyubiquitin B and C, respectively. The DNA translated into 108 amino acids that corresponded to two fused units of ubiquitin with no intervening sequence, which indicates that it is a polyubiquitin and contains at least two units of ubiquitin. Although, variations were found in the nucleotide sequence when compared to those of other species, the amino acid sequence was 100% homologous to the polyubiquitin sequences of humans, mice, and rats. This is the first report of the polyubiquitin DNA coding sequence and its corresponding amino acid sequence from camels, amplified using direct genomic DNA preparations.

Keywords: Arabian camel, Genomic DNA sequence, Polyubiquitin

Introduction

Ubiquitin is a 76-amino acid highly-conserved protein that appears to be present in all eukaryotes (Hershko and Ciechanover, 1986; Jentsch *et al.*, 1991; Neno *et al.*, 1994). It is both a cytoplasmic and nuclear protein, encoded by a multigene family. Functional ubiquitin is produced from two different types of ubiquitin genes, named polyubiquitin and ubiquitin extension genes (Schlesinger and Bond, 1987; Callis

and Vierstra, 1989). The polyubiquitin genes contain tandem head-to-tail repeats of 228 bp, encoding the ubiquitin protein. These genes normally have no introns in the coding region and end with a small extension of several amino acids between the last ubiquitin unit and stop codon. They also contain in their 5' prime region a heat-shock promoter element (Bond and Schlesinger, 1985). The ubiquitin extension genes contain a single ubiquitin repeat that is followed in frame by sequences encoding one of two ribosomal proteins.

Ubiquitin is functionally a very important protein. It has a fundamental role in mediating intracellular ATP-dependent protein degradation by 26S proteasome (Hershko, 1991; Rechsteiner, 1991; Varshavsky, 1992; Hochstrasser, 1996; Viestra, 1996; Callis, 1997). In animals and fungi, the ubiquitin pathway is involved in response to biotic and abiotic stresses, DNA repair, gene regulation, cell cycle regulation, signal transduction, and cell recognition (reviewed by Ciechanover and Schwartz, 1994; Wilkinson, 1994). It has also been proposed that ubiquitin conjugation plays an essential role in spermatogenesis (Mezquita *et al.*, 1997). Ubiquitin is also found in neurofibrillary tangles that are typical of the Alzheimer syndrome (Mori *et al.*, 1987; Perry *et al.*, 1987).

The conservation of ubiquitin throughout the phylogenetic tree is remarkable. Its amino acid sequence is identical, from arthropod to mammals, and is 96% similar with that of yeast, *Dictyostelium discoideum*, protozoa, and plants. Presently, it is known as one of the most conserved proteins. The individual unit (within a number of poly-ubiquitin genes) is significantly more similar to each other than would be expected if each unit evolved independently (Perelygin *et al.*, 2002). Although the ubiquitin protein sequence is highly conserved, the nucleotide sequences are quite divergent in different organisms. In contrast to the coding regions, the 5' and 3' non-coding sequences of some members of the gene family are entirely different from one another (Xia and Mahon, 1998). In this study, the molecular amplification and sequencing of genomic DNA, encoding a polyubiquitin gene (PUBC1) from camels, was performed by polymerase chain reaction (PCR) using various sets of primers.

*To whom correspondence should be addressed.

Tel: 966-1-4675765; Fax: 966-1-4678514

E-mail: kedhairi@ksu.edu.sa

GenBank accession: AY225193

Table 1. Set of primers

Set	Forward primer	Backward primer
1	5'-ctg acc agc aga ggy tga tct t-3'	5'-gtc ttg cca gtg agt gtc ttc a-3'
2	5'-tga aga ccc tgt ctg gta aga c-3'	5'-tgg act ctt tct gga tgt tgt ag-
3	5'-aag atg gac gca ccc tgt ctg act aca aca t-3'	5'-ctt cct tat ctt gga tct ttg cct tga cat t-3'
4	5'-cct gtc tga cta caa cat cca gaa aga gtc gac-3'	5'-atc ttc cag ctg ttt ccc agc aaa gat caa cct-3'

Materials and Methods

DNA was extracted from camel blood with standard procedures that utilized proteinase K-phenol-chloroform extraction. The primers were designed on the basis of the sequence data for human polyubiquitin C (Wiborg *et al.*, 1985). A polymerase chain reaction (PCR) was performed using PuRe Taq Ready-To-Go PCR Beads (Amersham, Biosciences, Piscataway, USA) with different primers sets (Table 1).

As a template, 200-300 ng of camel genomic DNA was used in a 25 μ l reaction. Genomic DNA was amplified for 40 cycles. Each cycle consisted of 94°C for 30 s, 52°C for 30 s, and 72°C for 1 min.

The obtained PCR products were separated by electrophoresis on a 1.5% agarose gel in a Tris-Actate-EDTA (TAE) buffer and visualized by ethidium bromide fluorescence. Fragments with the expected size were cut from the gel, then purified using a GFX PCR DNA Gel band purification kit (Amersham, Bioscience).

The purified DNA segments were sequenced using both forward and backward primers. Subsequently, the results were analyzed and compared with the sequences of the polyubiquitin genes in humans, rats, and mice using blast-2 software (NCBI). The DNA sequence was also translated into an amino acid sequence and compared with that of polyubiquitin from humans, rats, and mice. Homology was then determined.

Results and Discussion

Genomic DNA was amplified by PCR to yield the camel polyubiquitin nucleotide sequence. The amplification generated various DNA fragments (Fig. 1), which were then sequenced. One of these fragments yielded 325 bp nucleotide. The sequence data showed that it is a polyubiquitin since it contains more than one unit of ubiquitin (228 bp). However, the number of units of ubiquitin that are present in this polyubiquitin of camel could not be determined. A comparison of the camel PUBCI nucleotide sequence with its counterparts of polyubiquitin from humans, mice, and rats (located at various chromosomes) indicated a high sequence homology. The nucleotide sequence was found to be 95 and 88% similar to human ubiquitin B and C, which are located at chromosomes 17 and 12, respectively. Homology was also noticed with UBBP1 and UBBP4 human pseudogenes, located at chromosomes 2 and 17 (Baker and Board, 1987; Cowland *et al.*, 1988). It was 94% homologous to the mouse ubiquitin B (UBB) gene that is located at chromosome 11, as well as to the polyubiquitin 10 of rats. It was 88% similar to

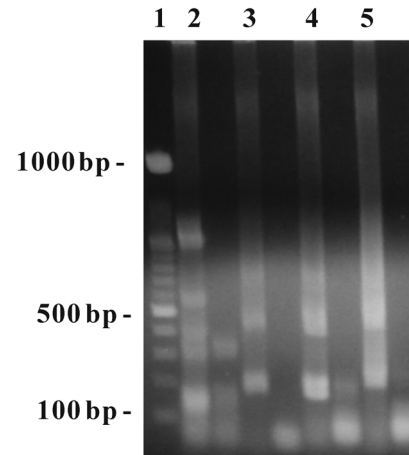


Fig. 1. Amplification of camel polyubiquitin (PUBC1) from genomic DNA. Lane 1 shows DNA size marker, 100 bp ladder. Lane 2 shows the amplified DNA fragment using primers set 1. Lane 3 shows the amplified DNA fragment using primers set 2. Lane 4 shows the amplified DNA fragment using primers set 3. Lane 5 shows the amplified DNA fragment using primers set 4. Lane following each number shows the respective negative control.

the polyubiquitin from mice that are located at chromosome 5. Moreover, it was also 89% homologous to putative polyubiquitin (UBQ10) of *Arabidopsis thaliana* (Callis *et al.*, 1995).

Figure 2 shows some nucleotide differences between camel PUBC1 and polyubiquitin B of humans and mice. The camel PUBC1 nucleotide was translated into 108 amino acids (Fig. 3), which corresponded to two incomplete units of ubiquitin protein. This indicates that PUBC1 contains at least two units of ubiquitin, which are present in the camel genome. The repeat number of ubiquitin coding units in the polyubiquitin genes varies between species, ranging from 3 units in humans and chickens to about 30 units in *Trypanosoma brucei* (Baker and Board, 1987; Wong and Campbell, 1989; Mezquita and Mezquita, 1991).

However, the amino acid sequence was 100% homologous to the polyubiquitin sequence of humans, mice, and rats. A comparison of the amino acid sequence of PUBC1 with those of other species indicated that it was well conserved in various species. Mammalian genomes, including that of humans, reportedly contain an ubiquitin multigene family with varied DNA sequence organization. Some of the genes in this family

```

Camel: 1      gaccagcagaggctgatctttgcaaggcaagcagctggaagatggccgcactctttctgat 60
                |||
Human:      gaccagcagaggctcatctttgaggcaagcagctggaagatggccgtactctttctgac
                |||
Mouse:      gaccagcagaggctgatctttgcccggcaagcagctggaagatggccgcaccctctctgat

Camel: 61    tacaacatccagaaggagtcgacctgcacctggtcctgctgaggggtggatgcatg 120
                |||
Human:      tacaacatccagaaggagtcgacctgcacctggtcctgctgaggggtggatgcatg
                |||
Mouse:      tacaacatccagaaggagtcgacctgcacctggtcctcctgctgaggggtggatgcatg

Camel: 121   atcttcgtgaagaccctgactggcaagaccatcacctggaggtggagccttagcgacacc 180
                |||
Human:      atcttcgtgaagaccctgaccggcaagaccatcacctggaagtggagcccagtgacacc
                |||
Mouse:      atcttcgtgaagaccctgaccggcaagaccatcacctggaggtggagcccagtgacacc

Camel: 181   atcgagaacgtgaaggccaagatccaggataaaggaaggcatccctccctgaccagcagagg 240
                |||
Human:      atcgaaaatgtgaaggccaagatccaggataaagaaggcatccctcccgaccagcagagg
                |||
Mouse:      atcgagaacgtgaaggccaagatccaggataaagagggcatccccctgaccagcagagg

Camel: 241   ttgatcttcgaggcaagcagctggaagatggccgcactcttctctgattacaacatccag 300
                |||
Human:      ctcatctttgaggcaagcagctggaagatggccgcactcttctctgactacaacatccag
                |||
Mouse:      ctgatctttgcccggcaagcagctggaagatggccgcaccctctctgattacaacatccag

Camel: 301   aaggagtcgacctgcacctggtcc 325
                |||
Human:      aaggagtcgacctgcacctggtcc
                |||
Mouse:      aaagagtcaaccctgcacctggtcc

```

Fig. 2. Comparison of genomic nucleotide sequence of camel polyubiquitin (PUBC1) with mouse and human ubiquitin B (UBB).

```

Human:  MQIFVKLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQ
Camel :                               DQQ

Human:  RLIFAGKQLEDGRTLSDYNIQKESTLHLVLRRLGGMQIFV
Camel :  RLIFAGKQLEDGRTLSDYNIQKESTLHLVLRRLGGMQIFV

Human:  KLTGKTITLEVEPSDTEINVKAKIQDKEGIPPDQQLIF
Camel :  KLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQLIF

Human:  AGKQLEDGRTLSDYNIQKESTLHLVLRRLGG
Camel :  AGKQLEDGRTLSDYNIQKESTLHLV

```

Fig. 3. Comparison of amino acids sequences of camel polyubiquitin (PUBC1) with 2 units of human ubiquitin B.

encode for polyubiquitin that contains three, four, or nine repeats of ubiquitin, while others encode for only one ubiquitin sequence (Wiborg *et al.*, 1985). In present study, two incomplete units of ubiquitin were amplified from the camel genomic DNA using specific primers. However, these two repeats were not separated by sequence but were immediately adjacent to one another, as expected in a polyubiquitin gene sequence.

In spite of several differences in the nucleotide sequences of the camel PUBC1 and ubiquitin B of humans and mice (Fig. 2), the deduced amino acid sequence was found to be fully

conserved and 100% similar to that of the polyubiquitin sequence of various species. Similarly, Okubo *et al.* (2002) reported that in rainbow trout, although the DNA sequence is not conserved, the encoded amino acid sequence is fully conserved. 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.

In this study, genomic DNA was used to amplify ubiquitin, though most of the previous studies were based on cDNA. This is the first report on the ubiquitin gene from camels. It indicates that it is a polyubiquitin gene, since at least two units of ubiquitin are present in the camel genome.

References

- Baker, R. T. and Board, P. G. (1987) The human ubiquitin gene family: Structure of a gene and pseudogene from UbB subfamily. *Nucleic Acids Res.* **15**, 443-463.
- Bond, U. and Schlesinger, M. J. (1985) Ubiquitin is a heat shock protein in chicken embryo fibroblast. *Mol. Cell Biol.* **5**, 949-956.
- Callis, J. (1997) Regulation of protein degradation in plants; in *Genetic Engineering; Principles and Methods*, Setlow, J. K. (ed.), p. 324, Plenum Press, New York, USA

- Callis, J. and Vierstra, R. D. (1989) Ubiquitin and ubiquitin gene in higher plants. *Oxford Surv. Plant Mol. Biol.* **293**, 1-30.
- Callis, J., Carpenter, T., Sun, C. -W. and Vierstra, R. D. (1995) Structure and evolution of genes encoding poly ubiquitin and ubiquitin-like proteins in *Arabidopsis thaliana* ecotype Columbia. *Genetics*, **107**, 205-212.
- Ciechanover, A. and Schwartz, A. (1994) The ubiquitin mediated proteolytic pathway: mechanism of recognition of the proteolytic substrate and involvement in the degradation of native cellular proteins. *FASEB J.* **8**, 182-191.
- Cowland, J. B., Wiborg, O. and Vuust, J. (1988) Human ubiquitin genes: one member of the UbB gene subfamily is tetrameric non-processed pseudogene. *FEBS Lett.* **231**, 187-191.
- Hershko, A. (1991) The ubiquitin pathway for protein degradation. *Trends Biochem. Sci.* **16**, 265-268.
- Hershko, A. and Ciechanover, A. (1986) The ubiquitin pathway for the degradation of intracellular proteins. *Prog. Nucleic Acid Res. Mol. Biol.* **3**, 19-56.
- Hochstrasser, M. (1996) Ubiquitin-dependent protein degradation. *Annu. Rev. Genet.* **30**, 404-439.
- Jentsch, S., Seufert, W. and Hauser, H.-P. (1991) Genetic analysis of the ubiquitin system. *Biochim. Biophys. Acta* **1098**, 127-139.
- Mezquita, J. and Mezquita, C. (1991) Characterization of a chicken polyubiquitin gene preferentially expressed during spermatogenesis. *FEBS Lett.* **279**, 69-72.
- Mezquita, J., Pau, M. and Mezquita, C. (1997) Characterization and expression of two chicken cDNAs encoding ubiquitin fused to ribosomal proteins of 52 and 80 amino acids. *Gene* **295**, 313-319.
- Mori, H., Kondo, J. and Ihara, Y. (1987) Ubiquitin is a component of paired helical filaments in Alzheimers disease. *Science* **235**, 1641-1644.
- Nenoi, M., Mita, K., Ichimura, S. and Cartwright, I. L. (1994) Novel structure of a Chinese hamster polyubiquitin gene. *Biochim. Biophys. Acta* **1204**, 271-278.
- Okubo, K., Yamano, K., Qin, Q., Aoyagi, K., Ototake, M., Nakanishi, T., Fukuda, H. and Dijkstra, J. M. (2002) Ubiquitin genes in rainbow trout (*Oncorhynchus mykiss*). *Fish Shellfish Immunol.* **12**, 335-351.
- Perelygin, A. A., Kondrashov, F. A., Rogozin, I. B. and Brinton, M. A. (2002) Evolution of the mouse polyubiquitin-C gene. *J. Mol. Evol.* **55**, 202-210.
- Perry, G., Friedman, R., Shaw, G. and Chau, V. (1987) Ubiquitin is detected in neurofibrillary tangles and senile plaque neurites of Alzheimers disease. *Proc. Natl. Acad. Sci. USA* **84**, 3033-3036.
- Rechsteiner, M. (1991) Natural substrate for the ubiquitin proteolytic pathways. *Cell* **66**, 615-618.
- Schlesinger, M. J. and Bond, U. (1987) Ubiquitin genes. *Oxford. Surv. Euk. Genes* **4**, 77-89.
- Varshavsky, A. (1992) The N-end rule. *Cell* **69**, 725-735.
- Vierstra, R. D. (1996) Proteolysis in plants: mechanism and functions. *Plant. Mol. Biol.* **32**, 275-302.
- Wiborg, O., Pedersen, M. S., Wnd, A., Berglund, L. E., Marcker, K. A. and Vuust, J. (1985) The human ubiquitin multigene family: some genes contain multiple directly repeated ubiquitin coding sequences. *EMBO J.* **4**, 755-759.
- Wilkinson, K. (1994) Cellular Roles of Ubiquitin; in *Heat Shock Protein in the Nervous System*, Mayer, R. J. and Brown, I. R. (eds.) pp. 191-234, Academic Press, New York, USA.
- Wong, S. and Campbell, D. A. (1989) A polyubiquitin gene from *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **37**, 147-150.
- Xia, X. and Mahon, J. (1998) Pea polyubiquitin genes: (1) structure and genomic organization. *Gene* **215**, 445-452.