Volume 16 Number 6 1988 Nucleic Acids Research

Nucleotide sequence of the mouse ornithine decarboxylase gene

Philip Coffino* and Emily Lichuan Chen[1]

Departments of Medicine and of Microbiology and Immunology and [1]Department of Anatomy, University of California, San Francisco, CA 94143, USA
Submitted January 25, 1988 Accession no.X06572

Ornithine decarboxylase (ODC) is a key enzyme for polyamine synthesis in eukaryotic cells, for it is the initial and usually rate-limiting step in that synthetic pathway. The nucleotide sequence of ODC cDNA's of mouse (1,2), hamster (3), human (4) and rat (5), and of ODC genomic DNA's of yeast (6) and trypanosome (7) have been reported. We describe here the sequence of recombinant clone pOD12.7 containing a functional mouse ornithine decarboxylase gene (8). The positions of exons were inferred by comparing the nucleotide sequences of ODC cDNA (1) and genomic DNA. Eleven introns are present, two in the sequence encoding the 312-313 nucleotide 5' untranslated leader and the other nine within the protein coding sequence. The first intron is the largest, approximately 2 kilobases in size. An 8-base palindrome within a cAMP-responsive genetic element of the somatostatin gene has been identified; similar or identical sequences exist in the promoter regions of other genes that respond to cAMP (9). A copy, differing from the consensus sequence in a single base, is present 48 nt before the site of transcription start and a fourfold imperfect iteration of the motif, having a repeat period of about ten nucleotides, spans the boundary between exon six and the subsequent intron.

```
        10
AGGTACATGT GCGCATGCAC CAGCCGACTC CCGCCCGCTG CCCATAGGGC CCTGCGGCAT GCTGGCAGCC AGGACTGGTG GTGTGGTGCG CGTGCGCAGG CCTGCCGCAG
       120
GGCGTGTCCG ACACGAGGCC CGGCGGGGAG CGCGGGGCGT ATGGGCGGGT GGGTGGGCAC GCGCTGCGCC CCGCCCCACT GACGCGCCCG GCCCGCGTCC CCGCCTCCGG
       230                                                                                                  >>>>>>>>>
CGCCGGGACC CGGGTTGGCC GCCACGGAGT CCCCGCCCCT CCCCCGCGCC TCCCGGCCGG AACCGATCGC GGCTGGTTTG AGCTGGTGCG TCTCCATGAC GACGTGCTCG
       340
GCGTATAAGT AGCGGCGCGT CGCACCGTCG GGCTTTGTCA GTCCCTGCAG CCGCCACCGC CGGCCGCCCT CAGCCAGCAG CTCGGCGCCA CCTCCGGTCG GCATCTGCGG
       450
CGGGCTCGAC GAGGCGGCTG ACGGGGCGGC GGCGGCGGGC GGACGGACGG ACGGACGGGC GCTCCTCGGG GTTTGGCGGC GGCGCCTCCA TGGGTCAGGC CAGCCGGGCC
       560                                                                                                  2530
ACCGTGCTGT GAGGTGAGGC GGGACGCGGG AGGGATGTGC          APROX. 2 KILOBASES NOT SEQUENCED                    CACATCTTGT
      2540
GTTTCTATTC ACTAGTGTTT CCACCACTCC AAGAAGGCAG CATTCAGAGT TCTTGGCTAA GTCGACCTTG TGAGGAGCTG GTGATAATTT GATTCCATCT CCAGGTTCCC
      2650                                                                                                  **
TGTGTAAGTA TCGGTTTACT TGTGGATTTG TGCTGGAGCA TTGCTACGAC TTTGAGTGTG TTCCCGCTCA CTTGCAGCTT CCTCTGTTTC AGAAGCACAT CGAGAACCAT
*     2760
GAGCAGCTTT ACTAAGGACG AGTTTGACTG CCACATCCTT GATGAAGGCT TTACTGCTAA GGACATTCTG GACCAAAAAA TCAATGAAGT CTCTTCCTCT GTAAGTACGG
      2870
GAAGCCCACG GAAGGCCGCA ACTCTGCTGA GAGCTCCTAG CACCATACAT GAGCCTGTCT TGCCAGAGAA TCTAGAATGT GACTGTGGAC TGTCTAGTGG TTGGTCATGC
      2980
CATGTCAGTG ACTGCTAACA GGGATAGAAT TTGATTAAGG AAGGGAAAAG GGTTTTCAGT GTGGCCAGCT GCTGGACAGC TAATGAGGTC CCTGAACCTG TTCTCATTCC
      3090
AGGACGATAA GGATGCGTTC TATGTTGCGG ACCTCGGAGA CATTCTAAAG AAGCATCTGA GGTGGCTAAA AGCTCTTCCC CGCGTCACTC CCTTTTACGC AGTCAAGTGT
      3200
AACGATAGCA GAGCCATAGT GAGCACCCTA GCTGCCATTG GGACAGGATT TGACTGTGCA AGCAAGGTAA GACTGCTCAC CCCGCCCCAA GGAGGCATCA GTTGTGTTAA
      3310
TAAGTGTTAT TAATAAGCTG AGGTGCACAT GACAACTTCA TGTGCTTTTG TTTGTCAGAC TTGGTCTGTA TAGAGCCCAA CACTGCTCTT CTCTTTCAGA CTGAAATACA
      3420
GTTGGTGCAG GGGCTTGGGG TGCCTGCAGA GAGGGTTATC TATGCAAATC CTTGTAAGCA AGTCTCTCAA ATCAAGTATG CTGCCAGTAA CGGAGTCCAG ATGATGACTT
      3530
TTGACAGTGA AATTGAATTG ATGAAAGTCG CCAGAGCACA TCCAAAGGCA AAGTGAGTCT TCTGATAGAG CACAAAAGGC CGGGCCTTGT TGGGCAGACT CATATCTTGG
      3640
TTCATTTATT TATTCCTATA CATAGTAGAA CTAGGCTAAA CCCTGTGTCA GACAAGCAGC AGCACCTACA CGTAGGCTCC TGAGTGGATG AGCATTATAG AGCACTTACA
      3750
CAGTGTACTT CCACCTAGGT TGGTTCTACG GATTGCCACT GATGATTCCA AAGCTGTCTG TCGCCTCAGT GTTAAGTTTG GTGCCACACT CAAAACCAGC AGGCTTCTCT
      3860          >>>>>>>>> >>>>>>>>  >>>>>>>>  >>>>>>>>>
TGGAACGGGC AAAAGAGCTA AATATTGACG TCATTGGTGT GAGGTGAGAT CTCAGTGATG TCATTACAGG CTGAGACATG AAATTTTAAG GCCCTTTCTC TTCCTGAGAA
      3970
CTAGTGAAAG ACCAGCTTCC TGTTTGTATT TCAGCTTCCA TGTGGGCAGT GGATGTACTG ATCCTGATAC CTTCGTTCAG GCAGTGTCGG ATGCCCGCTG TGTGTTTGAC
      4080
ATGGCAGTGA GTACACGGGA CTTGTTCAAG GGGAGGGAGG GGCTGTCTGA GATAATTAGA GTCTAGACTT TGTCTCTTGG GGAAGCCTTC TGCATGACAG ATTTTAAACC
      4190
CATCTGTCGT GTGCATTTAA ACTCTGGCAA TTTGACTTGA ATTTTCTTGG TTCTAGACAG AAGTTGGTTT CAGCATGCAT CTGCTTGATA TTGGTGGTGG CTTTCCTGGA
      4300
TCTGAAGATA CAAAGCTTAA ATTTGAAGAG GTAATTACAG CATTCATTAT TAATTAATGA CCTACAGAGG GTATTTTATA TCTAGTAGGT TCCATTTTGG TGTTTTTACT
      4410
GATATTAAAA GGTGCCAAAC AAACAAGTGG CCTGGCGCTG CAATCCCATT GACTGGTGTA CGGTAGCCCA GGCTGGCCCT GAACTTCAGT TCTGTCTCAG TTTGTACTGT
      4520
CATCCCTGTC TCCCATATAT TTTTAATGTC TCCTAGGAAA TGAAGCCATT GTTTAGTGCT TGTGTTATAT TTGTACAATT ATGTGAGCTA GGCAGGGTGG AAGGAGGTTT
      4630
ATCTTGGCAT GCATTATTTG TTAACAGAAT TATTTCAGCG TTTGTCCCTC TTTTGTAATT TATTTTGTGC TGTTTAATAA AAATATTCAT AATGTGTTGG AACAATTGAG
      4740
AGGGGAATGG GCAATGGTGT GCAGACTGGT TTCCAGGGAG AGGGGTGTTG GTGTTGCCTG GTGACAGACC TGCTGGGTCA TGTCCTGTTC CTTACACACC GCATAACATG
```

```
      4850
GCTGCTCCCT TCTCTCCCTC TTAGATCACC AGTGTAATCA ACCCAGCTCT GGACAAGTAC TTCCCATCAG ACTCTGGAGT GAGAATCATA GCTGAGCCAG GCAGATACTA
      4960
TGTCGCATCA GCTTTCACGC TTGCAGTCAA CATCATTGCC AAAAAAACCG TGTGGAAGGA GCAGCCCGGC TCTGACGGTA TGTGGTGGCA GGGTGAGTCA TGTAGGGTAA
      5070
CTGGAAGTTG ATATGCTGGG TGGTAATTAG GGTGATCTGT TTTTCTAGAT GAAGATGAGT CAAATGAACA AACCTTCATG TATTATGTGA ATGATGGAGT ATATGGATCA
      5180
TTTAACTGCA TTCTTTATGA TCATGCCCAT GTGAAGGCCC TGCTGCAGAA GGTAAGTTCT GAGCATGCTC TTTAGCAGTG AGAATGGTGG ACAGGATTCG GGGCTATTAA
      5290
AGAACAATGT CTTCTTCATT CAGAGACCCA AGCCAGACGA GAAGTATTAC TCATCCAGCA TCTGGGGACC AACATGTGAT GGCCTTGATC GGATCGTGGA GCGCTGTAAC
      5400
CTGCCTGAAA TGCATGTGGG TGATTGGATG CTGTTTGAGA ACATGGGTGC ATACACCGTT GCTGCTGCTT CTACTTTCAA TGGGTTCCAG AGGCCAAACA TCTACTATGT
      5510
AATGTCACGG CCAATGTGGT GAGTGAGATT GATTTTGCTT GCTTGGTGGT GGAATATTTG CCAACCAGGA GCCAGAAGCT ATCCCTGGTG CATACATACA CACATACTAT
      5620
GGGGAAAACA ATATGTGCTG AAGGGGAGGG ATCACTTGAG TGAGGGCCTT GATAGAATAA TAACTTGCTT GCCTGTCTCA AGAAAAGGAC TGAACTTGTA CTTTGGTTTT
      5730
TGTCTTTTTG ATATAAGAAA TTATTTTTGT ACCTTGATCT AAATACAGAT AAATGGAAGG GAGTTCTCCA ATAATACTGT TTGTTTACAG ATTCTTATAC TAGGAAAGGT
      5840
CTTAGAACTA AAAGCAATTA GATCCTTTGC AACTAAAATG TTAATTAATG CAACTAAAGT ATTCAGCTGG CATTTGTGAC CTGTGGTGCA TTGGATTGTT TCCTGGTGAT
      5950
GTAGTGACAA GGGTGAGGTG TCAGGAGACC TCTTGGGAGG CTGCCCAAAT TTGGAGACAC TTGGGTTTTG AATATATGTA CCTCTTTGTT TTCAGGCAAC TCATGAAACA
      6060
GATCCAGAGC CATGGCTTCC CGCCGGAGGT GGAGGAGCAG GATGATGGCA CGCTGCCCAT GTCTTGTGCC CAGGAGAGCG GGATGGACCG TCACCCTGCA GCCTGTGCTT
      6170      &&&
CTGCTAGGAT CAATGTGTAG ATGCCATTCT TGTAGCTCTT GCCTGCAAGT TTAGCTTGAA TTAAGGCATT TGGGGGGACC ATTTAACTTA CTGCTAGTTT GGGATGTCTT
      6280
TGTGAGAGTA GGGTTGGCAC CAATGCAGTA TGGAAGGCTA GGAGATGGGG GGTCACACTT ACTGTGTTCC TATGGAAACT TTGAATATTT GTATTACATG GATTTTTATT
      6390
CACTTTTCAG ACATGATACT AACGTGTGCC CCTCAGCTGC TGAGCAAGCG TTTGTAGCTT GTACATTGGC AGAATGGGCC AGAAGCTTAT GTTGTGACCC ATTGTGAAAA
      6500
TAAAGTATCT TGAAATAACT GGGCATCAGG GAATGTTTGC AAGTATCCTT AAAGAAGGCA CCAACATCTG CACAGTCTGC TGTGTCATGG AGAGACCCAC TGCCTGTGGA
      6610
TCTGAAGGTT GAGCTAGCCC CGCATAGCAC AGAGGAGAGG TGGATGGCAC AAGGCTGTGC CTCTCTGTAC AGCATCAGTC TGCTTAGCCC ATCCCAAGTG TGCAGTTGGC
      6720
TGAGAACTTT GTTGCCCAGA GTCTGTTGGT GAGGAATGTA CCTGCCTAGT GACCGGTTGG CATGGCCACT
```

Figure 1. Exons are underlined, initiating at the site of transcription start (nt 365) and terminating at the first of two (ref. 10) polyadenylation signals (nt 6494). Approximately 2 kilobases in intron one (following nucleotide 590) was not sequenced. The ATG translation initiation codon is marked by asterisks (nt 2749) and the TAG translation termination codon by ampersands (nt 6178). Arrowheads mark an eight nucleotide sequence with homology to a conserved cyclic AMP responsive element (nt 317-324 and nt 3876-3913).

ACKNOWLEDGEMENT

This work was supported by NIH Grant CA29048 and ACS Grant NP-477B to P.C.

*To whom correspondence should be addressed at: Microbiology 4125, University of California, San Francisco, CA 94143, USA

REFERENCES

(1) Gupta, M. & Coffino, P., J. Biol. Chem. 1985, **260**, 2941; (2) Kahana, C. & Nathans, D., Proc. Natl. Acad. Sci. 1985, **82**, 1673; (3) Srinivasan, P.R. et al., J. Biol. Chem. 1987, **262**, 12871; (4) Hickok et al., DNA 1987, **6**, 179; (5) van Kranen et al., Gene 1988, in press; (6) Fonzi, W. A. & Sypherd, P. S., J. Biol. Chem. 1987, **262**, 10127; (7) Pillips et al., J. Biol. Chem. 1987, **262**, 8721; (8) Brabant et al., Proc. Natl. Acad. Sci. 1988, in press; (9) Montminy et al., Proc. Natl. Acad. Sci. 1986, **83**, 6682.; (10) Hickok et al., Proc. Natl. Acad. Sci. 1986. **83**, 594.