Volume 16 Number 16 1988 Nucleic Acids Research

Nucleotide sequence of the rat ornithine decarboxylase gene

Harry van Steeg, Conny Th.M.van Oostrom, Henk J.van Kranen and Coen F.van Kreijl

National Institute of Public Health and Environmental Protection, PO Box 1, 3720 BA Bilthoven, The Netherlands

Submitted July 20, 1988 Accession no. X07944

Ornithine decarboxylase (ODC; EC 4.1.1.17) is the first and rate-limiting enzyme in the pathway of polyamine biosynthesis in eukaryotes. Previously genomic ODC sequences have been reported for yeast (1), trypanosome (2) and mouse (3-5). Here we report the complete nucleotide sequence of a 7776 bp BamHI fragment, carrying the complete rat ODC gene (clone pODC821). Comparison of this genomic DNA with rat ODC cDNA (6), reveals that the gene consists of 12 exons and 11 introns, as was also found for mouse genomic ODC (3,5). As in mouse, 2 introns are found in the 5' untranslated leader sequence, which spans 303/304 nucleotides (7). The first of these also represents the largest of all introns, comprising 1877 bp (nt 1344-3220). The organization of the protein coding part of the ODC gene is strongly conserved between mouse and rat. Intron/exon boundaries are found at identical positions in the coding sequence, giving rise to exons of identical length. Intron lengths, however, vary between rat and mouse. At the 3' end of the gene 2 alternate polyadenylation sites are found (see fig.1 and also ref. 6). Besides the TATA-box several other putative promotor- and enhancer elements are discernable: 9 GC-boxes, 5 upstream of the capsite (nt 1157/1158; ref.7), 1 in exon 1 and 3 in intron 1 (see fig.1 and ref. 8); 2 AP-1 bindingsites, 1 upstream of the capsite having a single nt mismatch and one perfect copy in intron 1 (see fig.1 and ref. 9); 1 Ap-2 bindingsite having 2 mismatches in a conserved 13 nucleotide sequence (see fig. 1 and ref.10);1 cAMP responsive element with 2 mismatches (see fig.1 and refs.3,11).

```
   1 GGATCCCGGC ATCCGGGCTT GGAAAGGGGT TTTAGTTGAG AATTAACAAA CCATTAACAG CGTGTACATA GTAACAAGTT GGAGTTACTC CTAAGTCTTC CCAATACTGA ATATGTAAGT
 121 AACCAGTCAG TCCCACTTAT TCTTTAACAA TGTAACTTTG GTGGGGGCTG GGGATTTAGC TCAGTGGTAG AGCGCTTACC TAGGAAGCGC AAGGCCCTGG GTTCGGTCCC CAGCTCCAAA
 241 AAAAAAAAAA ACCAAAAAAA AAAAACAAAA AAAACAATGT AACTTTGGTG GATACAAATT CACTGCCCAG ACTCTTTTCC GGTCCAGACT TCTCAAAGTG ACTTAGTTGT AGCTACCCTG
 361 AAAGATTTAA CAGAATAAGG GACAGCTGTT GGCTTGAAAC CCCTCAACCT CTTCCCTAAC ATGCGACTCT TGCTTAGATA AAAGGCATTA TCTTAGCATG CAACTGGTGC CTAGAATCTA
 481 GTTGGAAGCT ACCAACGTAA AACGGAGTGG CTCTCCAGCC TCTGCTGGCT TGTCTCTCAG CAAGTTTTCA GGCATAGGCA CACAGGGTGG TGAAGCAGGG TCAATCTACA CCACGAAGTC
 601 TAGTCTACGT GGCCCGCACA CAACTCAAGG CCCATCCATC CTCGGCTCGC ACCTGCGATC CCGCCAGGGA ATAAATACTA CGCGCTGAAT GCAAAGCCGA GGAAGAGTTT TGCTAGTCTT
 721 CCAGGCCAGG GCTCAGCCGT TAGGACCCGG GGAGAGACTG CGTGATTGAA CCGGTTGCAA GACGGACGAC CAAGCGCCTG TGCGCATGCA CCGGCCGAAT CCCGCCTGCT GCCCATAGGA
 841 CCCTGCGGCA AGCTGGTAGT CAGGACTGGT GGTGTGGTGC GCGTGCGCGG GCCTGCCGCA GGGAGTGTCC GGCACGAGGC CGGGCGGGGA GCGCGGGCGG ATGGGCGGGG GGCTGGGCAC
 961 GCGCGGCGCC CCGCCCCACT GACGCGCCGG GCCCGCGCCT CCCGCCTCCG GCGTCGGGAA CCGGGTTGGC CGCCAAGGAG TCCCCGCCCC TCCCCCGCGC CTCCCGGCCG GAACCGATCG
1081 GGGCTGGTTT GAGCTGGTGC GTCTCCATGA CGACGTGCTC CGCCTATAAG TAGCGGCGCG TCGCACCGTC GGGCTTTGTC AGTCCCTGCA GCCGCCGCCG CCGGCCGCCT TCAGTCAGCA
1201 GCTCGGCGCC ACCTCCGGTC GGCGACTGCG GCGGGCTCGA CGAGGCGGCT GACGGGGCGG CGGCGGGAAG ACGGCCGGGT GCGCCTTGGG GTTTAGTGGC GGCTTCTCCA TGGGTCCAGC
1321 CAGCCGCTTC CCTGTGCTGT GAGGTAAGCG GGGCGCGGGA GGATTGTGCG GCCACGTGTC GCGAGGCCCG GACTGCGGTC GCCGGCAGGG GACACGTGGC CGGAGGCTGG TGACCTTGCG
1441 AGCCTCCGCC CTGAAGGCAG CAGCTGTTCC TGGCCCTTGG CTGCAAGCCC TCGGATCGCC CCATGGGAGC CCGCATCGTG TCAGCCGGGC TCGGCGCTCC ACTTGCCCGG GCGGTGATGG
1561 AGGGCAGGTG CAGGGGAGCC CGGCCGCTAG CAGCCGCCCT GCAGTTTCCT GCTTTCTTTA GTTTTGCTAC GTGTACAGAG CAGAGCTGAT GACTCGAAGT GGAAGTCGAT AGACCTAATT
1681 TAGTCAGGGA GTCCCTGCTA ATGAATACTC ACCGGCTGTG ACTTGGATTT AGAACTGAAC TCTTTGAATG ACTTAGGGGT TTACGGGAAA CGGATTCCCG GCTGTCTTGG AAGCCCGACT
1801 CAGAATCCTC AATAACTGGT AGAACTGGGG AGCAGATCCT GTCAGGAATG TCTGGGAGAG GACAAAAGAA TGCAGTGGCT TGTTTTATTC AAACCTAACA TCGGGCAGAA TTTGACTGGA
1921 GTTCATTTTT GTTTTAATTG GCTTGTATCT CATCCAGATT TAACTTCTTG TGCTAATTGT TTAACAGGAT TTAGCTGCTT CCCATGCCAA AGTGTTAGCA TTTTTTTTCT TTTATCTTTT
2041 TTTTCTTTTT CTTTTTTTTT TTTTTTTTT TTGTATTGAG GCAAAGATAG GGAAAGGCAA AATAAGTGGT ATTTCAATGA AAGTTTTCCA GGTTTAAGCC GGAAGGCTGC CTGGAAAATA
2161 CTTAACAAGA CCTAAAATTA CAAAGTTACT GCCAAAGCCC ATTCATCAGT TCTAAGTGAT ATACAACTGA GTTGGCTGGC GCTGGTCTGT AGTCTTCACT TAAGAGAGGT TTATGCAGCT
2281 AACGAAGTTG ACCAGTGCCA AAAGTCTCCT TAAAGCTTAA AGTGTGCTTT ATCATGCATG GATTTGGATT TATGACAGGC TCATCCCCGT GGCCTCTCAT AGTGTCCCGT GCTAGAGCAA
2401 ACTGGCTCCT AACCATTGCC CAGCCTTGGT GCATGTAGGC TGCTGGCACT GAAGTGGGTC ACACAATGGA AAAAGTTTT ACCTGGCCAC TTGCTGCTGG TGTCCATTTC ATAAGACCCA
2521 GGCAGGAGCT TTTATTAAGC GTGGTGATCT TGCTTATTTT TGTTAGTAAA AAAGTTGGCC AAAAATTTCT TCCATATGGT TATTATCCAT AAGTGGATAT TAGAAGGCAC TTGAGGACTG
2641 TTGGATTCCC AGAAGTGCGC ATTTGGTAGG CAATCCAGGC TCTAGAATAC TGATGTTTCA TAACAGTAAC AATATACTGT GAATTGAATT TTTTTAAATG CTAGTACTGC ATAAAAGTTA
2761 GGGAAGTTTT GGTGGATTGC ACACTAACAT TGAGTGCTCT GGATACCTGA TATACTGTAT TCACTAAGTC ATTCAGCATA ACTCTGCAAG TGGGAGAAAA AGACCTTTTC AGGACTGCAT
2881 GTTACTGTAT GCCCATTAAA GTTTTTGAAG AGCTGGAATG GCTCAAATGA AGCAGATAGG AACATGTAAT AATGCCAGTA ATTAGACTGT TAACTCCGTT CCGTTTTGGA GAGGGGCAGG
3001 GATTGAGATG TTCTATGTGT ATAGAGCTGT GGCATTCAGT TTTGAACAGT GTGGAATCTA TTTGAGTATT GGAGATATTT GGGGGGGGGG TCTTCACGTA GTTCTGACCT CACAGTTGAT
3121 CTTCCTGTAC TGTGTACTCT CTGGTATACT GTTGCAGTGT GTAACCTTTA ATGGTAAGAT TTCAAGGGAG ATAGCAACAT GTTGTGTTTC TACTCACTAG TGTTTCCACC ACTCCAGGAG
```

3241 ACAGCATTCA GAGTTGACCT TGTGAGAGCT GGCCATAATT TAATTCCATC TCTAGGTTTT CTGTGTAAGT ATCAGTTTAT ACTTGTGAAT TTGTGTTGAA GCAGTGCTAA GACTTTGAGT
3361 GCATTCACTT ATTATCTGCT TATTGTCAGC TTATTGTTTC AGAGGCACAT CGAGAACCAA CCATGGGCAG CTTTACTAAG GAAGAGTTTG ACTGCCATAT CCTCGATGAA GGTTTCACTG
3481 CTAAGGACAT TCTGGACCAA AAAATCAATG AAGTTTCTTC CTCTGTAAGT ACGGGAAGCC CACAGCCACA ACTCAAAAGT TCCATAAAAG AGCATGCCTT GCCAGAGAAT CTAGAATTTG
3601 ACTGACTGGG CTGTCTAGTG AGTGTCCATG TCAGTACTGC TAACAGGGAT AGGATTTTGT TAGGGAAGGG GGAAAAAGAT CTTCAGTGTG GCCAGGTGCA AGACTGTTAA TGAGGTCCCT
3721 GAACCCACTC ACTCTCATTC CAGGATGATA AGGATGCTTT CTATGTTGCG GACCTCGGAG ACGTTCTAAA GAAGCATCTG AGGTGGCTGA AAGCTCTTCC CCGTGTTACT CCCTTCTATG
3841 CTGTCAAGTG TAATGACAGC AGAGCCATAG TGAGCACCCT GGCTGCCATT GGGACAGGAT TTGATTGTGC AAGCAAGGTA AGACTGCTTA TCTCACCCCA AAAGGGTTAT CAGTTGATGT
3961 GCTTTTGTTT GTCAAACTCC GTCTATATAT ACTGATCTTC TCTTTCAGAC TGAAATACAG TTGGTGCAGG GGCTTGGGGT GCCTCCAGAG AGGATTATCT ATGCAAATCC TTGTAAGCAA
4081 GTGTCTCAGA TCAAGTATGC TGCCAGTAAT GGAGTCCAGA TGATGACTTT TGACAGTGAA ATTGAGTTGA TGAAAGTTGC CAGAGCACAT CCAAAGGCAA AGTAAGTCTT CTGATTGAGC
4201 ACAAAAAGGC TGGGCTTGCT AGGCAGACTA TAATCTTGGT TCCTTATTCC TATACATAGT AAGAACCAGG CTAAACCCTG GGTTAGAATC AGACCCACAA GCAGTAGTGC CTAGACAGGC
4321 TGATTGGATG AACGTGTACA GAGTGCTTAC TTGGTGCATG CACTTGTTTC TGTCCACCTA GGTTGGTTTT GCGGATTGCC ACTGATGATT CCAAAGCAGT TTGTCGGCTC AGTGTTAAGT
4441 TTGGTGCCAC ACTGAAAACC AGCAGGCTTC TCTTGGAACG GGCAAAAGAG CTAAATATTG ATGTCATTGG TGTCAGGTGA GGTCTCAGTG ATAGTTACAG ACTGAGACAT TAAATTTTAA
4561 GGCCCTTTTT TCCTCTGAGA ACTAGTCAAA GAACAGCATC TTGTTTTTAT TTCAGCTTCC ATGTGGGCAG TGGGTGTACT GACCCTGAGA CCTTCGTGCA GGCAGTGTCA GATGCCCGCT
4681 GTGTCTTTGA CATGGGAGTA AGTACATGAC TCCTGGAGGA GGGCTGTCTT GATTGAAATA ACCAAACCAC TGAATCTAGA ATTTGCCTCT TGTAGAAGCT GTCTGCATAA CAGATTTTAA
4801 AACCATCTGT CCTGTGCATT TAAACTTGTG AATCTAGCAA TTTGACTTCC AAATTTTCTT GATTCTAGAC AGAAGTTGGT TTCAGCATGT ATCTGCTTGA CATTGGTGGT GGCTTTCCTG
4921 GGTCTGAAGA CACGAAGCTT AAATTTGAGG AGGTAATTAT TACAACATTA ACCTACAGAC GGTGAGATAT GTATTTACCA TTGTGGTGTT TTTACTGATT AAAGGAGCCA ACATGCCTGG
5041 TGTTTTGTTC TCACTACTTG TTATTTGCAA CTGTAGCCCA GGCTGGCCTT GAACTTTCAG TTCTCCTGCC TGTTTCCCAT ATTTTAATGT TTAGCTCAGT GCTTATCACT ATTGTACTAT
5161 GAAGTAGACA GGGTGGAAGG AGTTATATCT TGTTGGCATG TATTCAGAAT TCAGCATTTG GCCCCTCTTG TAATCCATTT TGTGCTATTT TATTAAATAT TAATATAAAC TATTCATAAT
5281 GTGTTAGAAA GAGATGGTCC TGAGGGGAGT GGGAAGTGGT GTATGTACAG ACTAGTTTCC AGAGAGGTTT GTTGCTTCTG TTGCCTGATG ACAGACCTGC TATGCTAGGT CATGTCCTGT
5401 TCTTACACAC AGCAGTGTGA TCTTTCTCCT CAACTATAAC AAATGGCTGC TCCCTTCTCT CACTTAGATC ACCAGTGTAA TCAACCCAGC TCTGGACAAG TACTTCCCAT CGGACTCTGG
5521 AGTGAGAATC ATAGCTGAGC CAGGCAGATA CTACGTCGCA TCAGCTTTCA CACTTGCAGT GAATATCATT GCCAAAAAAA CCGTGTGGAA GGAGCAGACC GGCTCGGACG GTATGTGGTG
5641 GCAGGATGAG TGAGTCATGT GGGATAACAG TTGATATGCT GGGTGGTAAT TAGGGTGATC TGTTTTTCTA GATGAAGATG AGTCAAACGA GCAAACTTTG ATGTATTACG TGAATGATGG
5761 AGTGTATGGG TCATTTAACT GCATTCTTTA TGACCATGCA CATGTGAAGG CCCTGCTGCA GAAGGTGAGT TCTGAGCGCA TGCTCTAGCA GTGAGTGATG CATAGGATAG AGAGAACATT
5881 GGGCCGAGAG ATGGCTCAGC AGTTAAGAGC ACTGACTGCT TCCAGAGGCC CTGAGTTCAA TCCCAGCAAC AACATGGTGG CTCACCACCA TCTGTATGTA ATGATGCCCT CTTCTGGTGT
6001 GTCTGAAGAA TAGCTACAGT GTATTTATAA TCTTTAAAAA AAAACAAACA AAAAACAAAC ATTGTCTTTA TACAGAGACC CAAGCCAGAT GAGAAGTATT ACTCATCCAG CATCTGGGGA
6121 CCAACATGTG ATGGCCTTGA TCGGATCGTC GAGCGCTGTA GCCTGCCTGA AATGCATGTG GGTGATTGGA TGCTGTTTGA GAACATGGGT GCATACACTG TTGCTGCTGC TTCTACTTTC
6241 AATGGGTTCC AGAGGCCAAA CATCTACTAC GTAATGTCAC GGTCAATGTG GTGAGTGGGG CTGTGCCTGC TCATGCTCTT TTGGTTGCTC AGTGTAATGT TTGCCTACTG TGAGCCAGAA
6361 GCCTGGGTTA TGCCTGATGC ACACACACAC ACATACTATG GGGGGAAGGC AGTATGTGCT GAGGGGAAGG ATCACCTGAC CTGTCTCAAA GGGACTGAAT TTGTGCTTGC TTTTTATGTT
6481 TGATAAGTTT TTTTTTTACC TTGAGCTAAA TACAGAAATG GAGTTTTTCC AAGTAATTAC TGCTCATCTT CAGATATTTA CTACTAGGAG AGGTCTTACA AGTAAGCCAT TAGATTCTTT
6601 TGCAACTAAA ATATTCAAGT TAGCATATGT GACCTGTGGG TTCCTTAATA ATGAAATTTT ACATTGGATT GTTTCCTGGT ATAGTGACAG GGTGAGGTGT CAGGAGATCT CTTGGGAGGC
6721 TGCCCAATTT GGAGGGATTT GGGTTTTGAA TATGCCATAT ATGCCTTTGT TTTTCAGGCA ACTCATGAAG CAAATCCAGA GCCATGGCTT CCCGCCAGAA GTGGAGGAGC AGGATGTTGG
6841 CACTCTGCCC ATGTCTTGTG CCCAGGAGAG CGGGATGGAC CGTCACCCTG CAGCCTGTGC TTCTGCTAGT ATCAATGTAT AGATGCCATT CTTGTAGCTC TTACCTGCAA GTTTAGCTTG
6961 AGTTCACGGC ATTTGGGGGG ACCATTTAAC TTAATTACTG CTAGTTTGGA ATGTCTTTGT AAGAGTAGGG TTGGCACCAA TGCAGTATGG AAAGACTAGG AGATGGGGGT CACACTTACT
7081 GTGTTCCTAT GGAAACTTTG AATATTTTAT ATGGATTTTT ATTCACTTTT CAGACCTGAT ACTAATGAGT GCCCCTCGGC TGCTGAGCAA GCATTTGTAG CTTGTACATT GGCAGAATGG
7201 GCTAAAGCT TATGTTGTGA CCCATTTTGA AATAAAGTA TCTTGAAATG ATTGGACATT GGAGAATGTG TGCAAGTATC CCTTACAGAA GGCACAAACT TCTGCACAGG CTGTGTGTTA
7321 CAGCAGTGAG TCTAGCCCAG CAGAGATGTG GATGATACAA AGCTGTGCCC CCTCTGTACA GCATCAATGT GCTTAGCCCA TCTCAAGTGT TTACTGTGAA CTTGGTGCCC AAAGTCTCTT
7441 AAGAGTGTCA TCTGCCTAGT GGCCTCTTGA CTTGGCCACT TCCTAAGGAG AGGGCATCTG AGGCTCTTTG AACCTTGCCT GCAGAAACCC TGACTGCTCC CTCAACCCTT GGCCTCAGCT
7561 GCATCACCAC CTAGTAACAG TTGAAGTATC GAACAGCACT TGGCAGTCAA CTTCCTGTAA TAAATTCAAC AACAGCAACT ACTGTGTTGT AAATTTTTAC CTTGACCTTT CATATTATAG
7681 TCCTGCTCAG TTTATACACC TATTGCATCC CCAAAGAAAT GCCAGACTCT TAACCTAGTT AGTATTGAAG ACCTGTAGGT CAGGAGGAAG GGATCC

Figure 1. Rat ornithine decarboxylase genomic sequence: exons are underlined. Translation initiation - (ATG, nt 3423) and termination site (TAG, nt 6920) are indicated by closed circles. Conserved sequences are depicted by boxes: TATA-box (nt 1125); 9 GC-boxes (nt 922, 935, 943, 971, 1044, 1255, 1446, 1549 and 1594); c-AMP responsive element (nt 1108); 2 AP-1 binding sites (nt 856 and 1680); AP-2 binding site (nt 822); 2 polyadenylation signals (nt 7232 and 7619).

Acknowledgement

We are indebted to H.M. Hodemaekers for her skillful technical assistence. We thank Deborah Dil for typing and preparing the layout of the manuscript.

REFERENCES

1. Fonzi, W.A. & Sypherd, P.S., J. Biol. Chem. (1987) 262, 10127.
2. Pillips et al., J. Biol. Chem. (1987) 262, 8721.
3. Coffino, P. & Lichuan Chen, E. Nucl. Acids Res. (1988) 16, 2731.
4. Brabant et al., Proc. Natl. Acad. Sci. USA (1988) 85, 2200.
5. Katz, A. & Kahana, C., J. Biol. Chem. (1988) 263, 7604.
6. Van Kranen et al., Gene (1987) 60, 145.
7. Van Steeg et al., manuscript in preparation.
8. Dynan, W.S. & Tjian, R. Nature (London) (1985) 316, 774.
9. Chiu et al., Nature (London) (1987) 329, 648.
10. Imagawa et al., Cell (1987) 51, 251.
11. Montminy et al., Proc. Natl. Acad. Sci. USA (1986) 83, 6682.