Volume 16 Number 10 1988 Nucleic Acids Research

The complete nucleotide sequence of the gene coding for the human aldolase C

P.Buono, G.Paolella, F.P.Mancini, P.Izzo and F.Salvatore

Istituto di Scienze Biochimiche, II Facoltà di Medicina e Chirurgia, Università degli Studi di Napoli, Via Sergio Pansini 5, 80131 Napoli, Italy
Submitted March 9, 1988 Accession no.X07292

A clone (lamba C1) containing the complete gene for the brain specific aldolase (aldolase C) has been isolated from a human genomic library (EMBL3), using a mouse aldolase C cDNA as probe (1). The cloned gene sequence, split into nine exons (capital letters), includes the complete coding region, all the introns and the 5'-and 3'-flanking regions, for a total of 5,200 bp. It codes for a polypeptide chain sequence (364 aa), reported below, with a Mr = 39,950.

The length of the first exon, based on sequence homology with mouse brain cDNA (1), has yet to be defined.

```
   1 tctagagaaatatccctagag gagaacttcggcgctctgac aaggagcttagaaaaactag atcgacatctgttcagccat atttatcaagcgtatacgca gacggacgatatcactaatg
 121 aacctggtacaaggcctggc cccaggaggtactaatgagc acagatctggttcccagcct gtgctgaagagctgggcctg aatcactattgcaggcttct ctgtggtgtctgttttctcc
 241 cctctaattcagcggctgcc tccttggttgggagccctgg gccttaagtctgtggtcctg agctgcagttttctgcctct cttcccagcccctgctcctca ttccagaggtggtgaggggg
 361 tgcaaagaactacagggatt gctggaatttctgagctaag aaactgaaagccagaatctg cttcacctcttttttacctgc aatacccccttaccccaata ccaagaccaactggcataga
 481 gccaactgagataaatgcta tttaaataaagtgtatttaa tgaatttctccaagcttacg gaatcttactttttgggtga gagggggcggatgatgaag tcagggaagaagcaggcttg
 601 tgcctcctgcctgggcaccc tccatagatcctctctccacc tcctctgcatccaaagcaga cgtgcctgccctttcctcct tcctcgtactttcaccaaaa actctgggcactggacgtac
 721 ttgttccctcgtccgcctgg aggaaacagaaagcttcatt ctctactttccccagcctcg gggtataccctggcaaatat ggggtgaggacagagaaggc tcaggttcattacttcctgc
 841 cctgaacttgggcaccttga agatgcttctgcccccacct actcaacctgttgttactgt tctggtgcctcctctttcacc agactaagcttaatccccag catccagcaagtgcccggca
 961 cgtggtaaagggattgtggc agaaatgaacgcaaggtcat ggccaggctccccgaggtgg gagggcagtccctaacagcc acggatgcctggcgctcact ccacctcgccgccgcctctg
1081 agggcgtggttctttgccccga caccagtcctgggggggggg tgtggtcagggcgggcatgc agccacgccccc ggaggagt cacgtagctctgcgacatcc gcagcctcatttaccagagg
1201 gagccaggctgcagcctcat ctgtttgcggatcagaaccc gagCTGTGCTTGTGGCTGCG GCTGCTAACTGGCTGCGCAC AGgtaaggccaggcaaggcg ggagccactcaacatctccg
1321 cttcatttccttggctcttc ccctgccatcctcgtcctgc tacctgggcactccgggatgt gccctttcgaccctttccta acatctttgctcctttccgc gtacttgaaacccccatggct
1441 caacctcttctgtttctatcc ctcttcagctctcccagctg aacctcagctttccaatccc aaatcctcctctcacttgttg gattttttccttggagtctgt ctctcctacccagagcatttt
1561 ccttctcagcctgctccctc tcctcctaggctaggtcctc tctccagttctcccgccttc tctgccccccggtctaggtc tctcccgctgactggttagc ctgcatcaccactagctccc
1681 ctcagtctcatctctctctc aggctctctactcctttcag ccttggtcctctgccctccc gtctggtgtctattgtggga attaatcttcctattgctat cctcctgccaaacagctaag
1801 tgcttcctggcacagagata cccggagtgtcactagccct atctccacacactaacacag ctgaccgctgggccctcttt cctacatcactgcacgactg gggccaacctctggtactgg
1921 gtggggaataggcaataggca taggcaggcaggtttgagta cagaaaagaagctgcaggag cctgtgactggtatttgtgc cactcctactccctacctgt tctttccaaccttttcctcta
2041 gaagctgagagaagaggggtg gcaataagtactttgcctc attctgaagccttggaagta agtacactttccctaggggt cctgtggaggatgagaaaag ggaagctggaaaggccagga
2161 cttttgcctactcaacaagg gaccaagttcagtgaaagaa gggttggcatccttgattgg gcagcagatttatcagaaga gctgtggcttcagggctgct cacctccccacccccaccct
2281 gcatctttccccagggctgg gaaggatgcctaccagggac aaaaggagatgtggaaactg gagccctaagcttgctagct gtcagaaggacctggtgcca cttgatgcccaggactcatg
2401 ccaaggactgctgccctgtt cccagccccttgcttgatgg ggaggccatttggcccatct ggcaggagaggcagcagagg gtgaggtctggcttttttat cttgtctccactccagGGAG
2521 CTGTCACCATGCCTCACTCG TACCCAGCCCTTCTCGCTGA GCAGAAGAAGGAGTTGTCTG ACATTGCCCTGCGGATTGTA GCCCCGGGCAAAGGCATTCT GGCTGCGGATGAGTCTGTAG
2641 gtaagtggacatctgtagcc aggtagggtacaggtggcta gggaccctggggatgttct cactgcctctctttgtttgc ccctagGCAGCATGGCCAAG CGGCTGAGCCAAATTGGGGT
2761 GGAAAACACAGAGGAGAACC GCCGGCTGTACCGCCAGGTC CTGTTCAGTGCTGATGACCG TGTGAAAAAGTGCATTGGAG GCGTCATTTTCTTCCATGAG ACCCTCTACCAGAAAGATGA
2881 TAATGGTGTTCCCTTCGTCC GAACCATCCAGGATAAGGGC ATCGTCGTGGGCATCAAGgt gcagccccctggccctgctct gaatggaagctgggtgtgaa aataagcttgtgtaggaggg
3001 gtagcaaggagaatcctcct ggattcaacccttcttttgta cttcctctacagGTTGACAA GGGTGTGGTGCCTCTAGCTG GGACTGATGGAGAAACCACC ACTCAAGgtataggatgggt
3121 gggcttgaggaccaaagaggg tgttagatagttgatgctgg taaaagaggggcagagtaat gaggttggccactgtgcttg cagGGCTGGATGGGCTCTCA GAACGCTGTGCCCAATACAA
3241 GAAGGATGGTGCTGACTTTG CCAAGTGGCGCTGTGTGCTG AAAATCAGTGAGCGTACACC CTCTGCACTTGCCATTCTGG AGAACGCCAACGTGCTGGCC CGTTATGCCAGTATCTGCCA
3361 GCAGgtgtgtgtgttgggag ggtggtgagctaggtgccct gtatgcctggtggggagaga gtacaaggctttcttcatct ccccctactgccccteccaag catctctgctcttgcctgca
3481 GAATGGCATTGTGCCTATTG TGGAACCTGAAATATTGCCT GATGGAGACCACGACCTCAA ACGTTGTCAGTATGTTACAG AGAAGgtggtccacacctg ggcacacaaacatactgcag
3601 ggacagctcggcaggagtgt ctgttccccagaacccccag cttagatccaggccaatttcc cctagcatttttcacttcat cccgcgcacaggcctgtgat ctcgagcctgtactagccct
3721 acagtctgtgcccatctacc cctacatagggagcatcgag cagtaaccagtgggggcccca gaccttagtaaacctcctct aatcccacccagGTCTTGGC TGCTGTGTACAAGGCCCTGA
3841 GTGACCATCATGTATACCTG GAGGGGACCCTGCTCAAGCC CAACATGGTGACCCCGGGCC ATGCCTGTCCCATCAAGTAT ACCCCAGAGGAGATTGCCAT GGCAACTGTCACTGCCCTGC
3961 GTCGCACTGTGCCCCCAGCT GTCCCAGgtactacccagct ccctaacctgctcctatccc taaggcccatctttcaggtcc ttctttgtggccttaggggggt tccctatcctggaaaaattg
4081 ggagtgaccagtcagtttgt cttctctcctccacactagG AGTGACCTTCCTGTCTGGGG GTCAGAGCGAAGAAGAGGCA TCATTCAACCTCAATGCCAT CAACCGCTGCCCCCCTTCCCC
4201 GACCCTGGGCGCTTACCTTC TCCTATGGGCGTGCCCTGCA AGCCTCTGCAGTCAATGCCT GGCGAGGGCAACGGGACAAT GCTGGGGCTGCCACTGAGGA GTTCATCAAGCGGGCTGAGg
4321 ttgggagctacaggtggtgg tgggtgggggcagcaccaga ggctatagctgggcagggct tggcagctgtgggctggcta ctgcttactccaggctccct tttgcagGTGAATGGGCTTG
4441 CAGCCCAGGGCAAGTATGAA GGCAGTGGAGAAGATGGTGG AGCAGCAGCACAGTCACTCT ACATTGCCAACCATGCCTAC TGAGTATCCACTCCATACCA CAGCCCTTGGCCCAGCCATC
4561 TGCACCCACTTTTGCTTGTA GTCATGGCCAGGGCCAAATA GCTATGCAGGAGCAGAGATG CCTTCACTGGCACAACTTGT CTTTCCTTTCTCTCTTCCCT TCCCCTCTCTCATTGCTGCA
4681 CTGAGACCATAGGATGGGAG GATAGGGAGCCCCTCATGAC TGAGGGCAGAAGAAATTGCT AGAAGTCAGAACAGGATGGC TGGGTCTCCCCCTACCTCTT CCAGCTCCCACAATTTTCCC
4801 ATGATGAGGTAGCTTCTCCC TGGGCTCTCCTTCTTGCCTG CCCTGTCTCCTGGGATCAGA GGGTAGTACAGAAGCCCTGA CTCATGCCTTGAGTACATAC CATACAGCAAATAAAtggta
4921 gcaaaacatttcactttgcc tgtctgttttacacatcaaa ttccagcctcccagtttctg atctctgctaatttctatctc tgggccctctgactctggag gtggagagggtgggatttga
5041 gtcttacgtgggcttcaagt tatggaggaagggcaatgca gtcaccatccccagctcagg ctcttgctctcttgactgtc caagtctggagtggggcaat gaggaagactgcaagtcttc
5161 tagggactgccacatcagtc gatgggctgcag
```

Acknowledgements. This work was supported by grants from C.N.R. (P.F. Ingegneria Genetica) and from M.P.I., Rome, Italy.

References.

1. Paolella, G., Buono, P., Mancini, F.P., Izzo, P. & Salvatore, F., (1986), Eur. J. Biochem., 156, 229-235.