

Complete coding sequence of human tryptophan hydroxylase

Sylviane Boularand, Michèle C. Darmon, Yves Ganem¹, Jean-Marie Launay¹ and Jacques Mallet*
Laboratoire de Neurobiologie Cellulaire et Moléculaire, C.N.R.S. 91198 Gif-sur-Yvette and ¹Service de Biochimie et Neuroendocrinologie, Hôpital Saint-Louis, 75475 Paris Cedex 10, France

Submitted June 11, 1990

EMBL accession no. X52836

We report here the isolation of a cDNA clone containing the complete coding sequence of human tryptophan hydroxylase (TPH). A cDNA library from a serotonin secreting carcinoid tumor was constructed in lambda-ZAP vector and was screened with a rat TPH cDNA probe (1). The nucleotide coding sequence of 1,332 bp encodes a protein of 444 amino-acids, as in the rat, with a predicted molecular weight of 50,952 daltons. The figure presents the comparison of the human TPH coding region with that of rat TPH. The identity between these two sequences is

86.8% in nucleotides and 91.2% in amino acids. Interestingly, three putative phosphorylation sites previously identified in rat TPH protein are also conserved in human TPH protein and are located at positions 58, 260 and 443.

REFERENCE

1. Darmon, M.C., Guibert, B., Leviel, V., Ehret, M., Maitre, M. and Mallet, J. (1988) *J. Neurochem.* 51, 312-317.

```

10      20      30      40      50      60      70      80      90      100     110     120     130     140     150     160     170     180
Met I LeG luAspAsnLysGluAsnLysAspHisSerLeuG luArgG luArgH laSerLeuI lePheSerLeuLysAsnGluVa IG lYg lYleuI leLysAlaLeuLysI lePheGInG luLysH isVa lAsnLeuLeuHisI leG luSerArgLysSerLysArgArgAsnSerG luPhe HU
ATGATTGAAGCAATAAGGAGCAAAAGCATTCTTAGAAGGGGAGAGCAGCTCATTTTCTTAAGAATGAAGTTGGAGGACTATAAAGCCCTGAAATCTTCAGGAGAACATGTGAATCTGTACATATCCAGTCCGAAAATCAAAAAGCAAACTCAGAATTT
.....
ATGATTGAAGCAATAAGGAGCAAAAGCATTCTCAGAAAGGGGAGAGTGAATCTCTTCTTGAAGAATGAAGTTGGAGGACTATAAAGCCCTGAAATCTTCAGGAGAACACGTGAACCTGTACATATGAGTCCCGAAAATCAAAAAGCAAACTCAGAATTT
Met I LeG luAspAsnLysGluAsnLysAspHisSerSerG luArgG luArgValThrLeuI lePheSerLeuLysAsnGluVa IG lYg lYleuI leLysAlaLeuLysI lePheGInG luAsnH isVa lAsnLeuLeuHisI leG luSerArgLysSerLysArgArgAsnSerG luPhe RAT
190     200     210     220     230     240     250     260     270     280     290     300     310     320     330     340     350     360
GluI lePheVa lAspCysAspI leAsnArgG luGInLeuAsnAspI lePheHisLeuLeuLysSerHisThrAsnVa lLeuSerVa lAsnLeuProAspAsnPheThrLeuLysG luAspG lYMetG luThrVa lProTrpPheProLysI leSerAspLeuAspHisCysAlaAsn HU
GAGATTTTGTGACTGTGACATCAACAGAGAACATGAATGATATTTTTCATCTGCTGAAGTCTCATCAATATGTTCTCTCTGTGAATCTACAGATAATTTACTTGAAGGAGATGGTATGGAACTGTCTTGGTTCCTCCAAAGAGATTCTGACCTGGACCATTTGCCAAC
.....
GAGATTTTGTGACTGTGACATCAACAGAGAACATGAATGATATTTTTCATCTGCTGAAGTCTCATCAATATGTTCTCTCTGTGAATCTACAGATAATTTACTTGAAGGAGATGGTATGGAACTGTCTTGGTTCCTCCAAAGAGATTCTGACCTGGACCATTTGCCAAC
GluI lePheVa lAspCysAspI leAsnArgG luGInLeuAsnAspI lePheProLeuLeuLysSerHisThrThrVa lLeuSerVa lAspSerProAspGInLeuProLysG luAspVa lMetG luThrVa lProTrpPheProLysI leSerAspLeuAspHisCysAlaAsn RAT
370     380     390     400     410     420     430     440     450     460     470     480     490     500     510     520     530     540
ArgVal lLeuMetTyrG luSerG luLeuAspAlaAspHisProG lYPheLysAspAsnValTyrArgLysArgArgLysTyrPheAlaAspLeuAlaMetAsnTyrLysHisG lYAspProI leProLysVa lG luPheThrG luG luG luI leLysThrTrpG lYThrVa lPheGInG lu HU
AGAGTTCTGATGTATGATGACTGACATGATGACAGACATCTGGCTTCAAGACAATGCTACCGTAAGCTCGAAGATTTTGGCGACTGGCTATGACTATAAATCGGAGACCCCATTCANAGGTTGAATCTCAGCAGGAGGATTAAGACCTGGGAACTGATTCAGAG
.....
AGAGTTCTGATGTATGATGACTGACATGATGACAGACATCTGGCTTCAAGACAATGCTACCGTAAGCTCGAAGATTTTGGCGACTGGCTATGACTATAAATCGGAGACCCCATTCANAGGTTGAATCTCAGCAGGAGGATTAAGACCTGGGAACTGATTCAGAG
ArgVal lLeuLeuTyrG luSerG luLeuAspAlaAspHisProG lYPheLysAspAsnValTyrArgArgArgArgLysTyrPheAlaG luMetAsnTyrLysHisG lYAspProI leProLysI leG luPheThrG luG luG luI leLysThrTrpG lYThrI lePheArgG lu RAT
550     560     570     580     590     600     610     620     630     640     650     660     670     680     690     700     710     720
LeuAsnLysLeuTyrProThrHisAlaCysArgG luTyrLeuLysAsnLeuProLeuLeuSerLysTyrCysG lYArgG luAspAsnValIleProGInLeuG luAspVa lSerAsnPheLeuLysG luArgThrG lYPheSerI leArgProVa lAlaG lYlYrLeuSerProArgAsp HU
CTCAACAACCTACCCCAACCCATCTGACAGAGACTTCAAAAACCTTGGCTTCAAAATATTCGGATATCGGAGAGATAAATCCCAACATTTGGAGATGCTCCAACCTTTTAAAGAGCTACAGGTTTTCATCCGCTCGCTGGCTGGTACCCTCACAGAGAT
.....
CTCAACAACCTACCCCAACCCATCTGACAGGAGACTTCAAAAACCTTGGCTTCAAAATATTCGGATATCGGAGAGACTACCCCAACATTTGGAGATGCTCCAACCTTTTAAAGAGCTACAGGTTTTCATCCGCTCGCTGGCTGGTACCCTCACAGAGAT
LeuAsnLysLeuTyrProThrHisAlaCysArgG luTyrLeuLysAsnLeuProLeuLeuSerLysTyrCysG lYArgG luAspAsnValIleProGInLeuG luAspVa lSerAsnPheLeuLysG luArgThrG lYPheSerI leArgProVa lAlaG lYlYrLeuSerProArgAsp RAT
730     740     750     760     770     780     790     800     810     820     830     840     850     860     870     880     890     900
PheLeuSerG luLeuAlaPheArgVa lPheHisCysThrGInTyrVa lArgH isSerSerAspProPheTyrThrProG luProAspThrCysH isG luLeuLeuG lYHisVa lProLeuLeuAlaG luProSerPheAlaGInPheSerGInG luI leG lYLeuAlaSerLeuG lYAla HU
TCTTATACAGTGTAGCCCTTCAGGTTTTCAGTGCACCTCAATAATGAGAGACAGTTCAGATCCCTTCTATACCCAGAGCAGGATACCTGCCATGACTCTTAGTCTAGTCCCGCTTGGCTGAACCTATTGGCCAACTTCCCAAGAAATGGCTGGCTTCTCTGGCCCT
.....
TCTTATACAGTGTAGCCCTTCAGGTTTTCAGTGCACCTCAATAATGAGAGACAGTTCAGATCCCTTCTATACCCAGAGCAGGATACCTGCCATGACTCTTAGTCTAGTCCCGCTTGGCTGAACCTATTGGCCAACTTCCCAAGAAATGGCTGGCTTCTCTGGCCCT
TTCCTGTCAGGGTTAGCCTTTCAGTCTTTCAGTGCACCTCAATAATGAGAGACAGTTCAGATCCCTTCTATACCCAGAGCAGGATACCTGCCATGACTCTTAGTCTAGTCCCGCTTGGCTGAACCTATTGGCCAACTTCCCAAGAAATGGCTGGCTTCTCTGGAGCT
PheLeuSerG luLeuAlaPheArgVa lPheHisCysThrGInTyrVa lArgH isSerSerAspProLeuTyrThrProG luProAspThrCysH isG luLeuLeuG lYHisVa lProLeuLeuAlaG luProSerPheAlaGInPheSerGInG luI leG lYLeuAlaSerLeuG lYAla RA
910     920     930     940     950     960     970     980     990     1000  1010  1020  1030  1040  1050  1060  1070  1080
SerG luG luAlaVa lGInLysLeuAlaThrCysTyrPhePheThrVa lG luPheG lYLeuLysCysGInAspG lYgInLeuArgVa lPheG lYAlaG lYLeuLeuSerSerI leSerG luLeuLysHisAlaLeuSerG lYHisAlaLysVa lLysProPheAspProLysI leThrCys HU
TCAGAGGAGGCTGTTCAAAAACCTGCAACCTGCTACTTTTTCAGTGCAGGTTGGCTATGTAACAAGATGGACACTCAAGACTTTCAGTGCAGGCTGCTTCTCTATCAGTGAATCAAAATGCACTTTCAGGACATCCAAAGTAAAGCCCTTGTATCCCAAGTACTGCT
.....
TCAGAGGAGGCTGTTCAAAAACCTGCAACCTGCTACTTTTTCAGTGCAGGTTGGCTATGTAACAAGATGGACACTCAAGACTTTCAGTGCAGGCTGCTTCTCTATCAGTGAATCAAAATGCACTTTCAGGACATCCAAAGTAAAGCCCTTGTATCCCAAGTACTGCT
TCAGAGGAGGCTGTTCAAAAACCTGCAACCTGCTACTTTTTCAGTGCAGGTTGGCTATGTAACAAGATGGACACTCAAGACTTTCAGTGCAGGCTGCTTCTCTATCAGTGAATCAAAATGCACTTTCAGGACATCCAAAGTAAAGCCCTTGTATCCCAAGTACTGCT
SerG luG luThrVa lGInLysLeuAlaThrCysTyrPhePheThrVa lG luPheG lYLeuLysCysGInAspG lYgInLeuArgVa lPheG lYAlaG lYLeuLeuSerSerI leSerG luLeuArgHisAlaLeuSerG lYHisAlaLysVa lLysProPheAspProLysVa lAlaLys RAT
1090    1100    1110    1120    1130    1140    1150    1160    1170    1180    1190    1200    1210    1220    1230    1240    1250    1260
LysGInG luCysLeuI leThrThrPheGInAspVa lTyrPheVa lSerG luSerPheG luAspAlaLysG luLysMetArgG luPheThrLysThrI leLysArgProPheG lYVa lLysTyrAsnProTyrThrArgSerI leGInI leLeuLysAspThrLysSerI leThrSerAla HU
AAACAGGAATGCTTATCACAACCTTTCAGATGCTACTTTGATCTGAAAGTTTGAAGATGCAAAAGGAGAGATGAGAGAAATTTGCCAAAACCTGAAAGCCCGCTTGGAGTGAAGTACAATCCGTCACACAGAGCATTCAGGTTCTGAGAGCAGCAGAGCATACCAAGTGC
.....
AAACAGGAATGCTTATCACAACCTTTCAGATGCTACTTTGATCTGAAAGTTTGAAGATGCAAAAGGAGAGATGAGAGAAATTTGCCAAAACCTGAAAGCCCGCTTGGAGTGAAGTACAATCCGTCACACAGAGCATTCAGGTTCTGAGAGCAGCAGAGCATACCAAGTGC
LysGInG luCysLeuI leThrThrPheGInAspVa lTyrPheVa lSerG luSerPheG luAspAlaLysG luLysMetArgG luPheThrLysThrVa lLysArgProPheG lYVa lLysTyrAsnProTyrThrInSerI leGInI leLeuLysAspThrLysSerI leThrSerAla RAT
1270    1280    1290    1300    1310    1320    1330    1340    1350    1360    1370    1380    1390    1400    1410    1420    1430    1440
MetAsnG luLeuGInHisAspLeuAspAlaValSerAspAlaLeuAlaLysVa lSerArgLysProSerI le*** HUMAN TPH
ATCAATGACGTGCAGCATGATCCGATGTGTGACGTATGCCCTTGTCAAGTGCAGAGGAAGCCGATATCTaa
.....
ATCAATGACGTGCAGCATGATCCGATGTGTGACGTATGCCCTTGTCAAGTGCAGAGGAAGCCGATATCTaa
ATCAATGACGTGCAGCATGATCCGATGTGTGACGTATGCCCTTGTCAAGTGCAGAGGAAGCCGATATCTaa
MetAsnG luLeuArgHisAspLeuAspAlaValAsnAspAlaLeuAlaArgVa lSerArgTrpProSerVa l*** RAT TPH

```

* To whom correspondence should be addressed