Volume 17 Number 15 1989 Nucleic Acids Research

Cytochrome P450CMF cDNA: nucleotide sequence of P450CMF1b

Nobuhiro Ishida, Chikako Inuzuka[1], Yasunori Tawaragi[1], Osamu Sugita[1], Hiroshi Nakazato[1], Teruhisa Noguchi[1], Shigeru Sassa and Attallah Kappas

The Rockefeller University Hospital, New York, NY 10021, USA and [1]Suntory Institute for Biomedical Research, Osaka, Japan

Submitted June 10, 1989 EMBL accession no. M25143

P450CMF is a newly defined cytochrome P450 which is constitutively present in the Sprague-Dawley rat in both male and female liver[1], and in male kidney[2], and is highly active in estradiol 2- and 16α-hydroxylations[1]. A cDNA library constructed from rat liver poly(A)$^+$ RNA in λgt10 was screened using rat P450CMF3 cDNA[3] under conditions of low stringency. Positive clones could be further classified into those which hybridized with the specific oligomers of P450CMF1a, CMF1b, CMF2 or CMF3[3]. The nucleotide sequence of a clone, λcrCMF35, which encoded the entire P450CMF1b peptide, is shown below. The amino acid sequence of P450CMF1b which was deduced from its nucleotide sequence was 96.0% similar to that of P450CMF1a[3] and 95.2% similar to that of P450db1[4], the gene which codes for the P450 isozyme with debrisoquine 4-hydroxylase activity. The availability of the P450CMF cDNA will facilitate study of the expression of the P450CMF gene in various cell types, and the role of this new P450 isozyme in the oxidation of endogenous as well as exogenous substrates.

```
 -760                -740                -720                -700                -680                -660
GAATTCAATTAAATATTTATTGAACAAATGCAGAATAAATGAACCAGGGCCGTTTTAACCTGAAGATACATTAACCAACTCAAATCTTCAAGGGCTGCTGCCATAGGAGAACAAAAAGAG
 -640                -620                -600                -580                -560                -540
ATGCACTGAGCTTTTCTACACCGGGTCTGGATGCGCAGCAGGAGCGATGCGCAGCAGGAGGGCTGCAGGCCGGCCCTCAGCTCAGGAGCTCCCTCAGCACACACTGCACTCCATGCTCCTC
 -520                -500                -480                -460                -440                -420
CCTTCCTCTGAAAGGAGGCATCAGAACTTCGAGTTTCCCATCTGCACCTAAGTTGCGTTCTCATTGCAGACTGTCTGTCTCATGCAATGGATGCTTGCACAATGGGAAGACAAACATAGC
 -400                -380                -360                -340                -320                -300
ATTGACAAGCTCCTCCCCAGGAGACACCTGCCACCATCGGCTCTTTCTGATTCTACATCACTGATTCCTTGTTCTGGGAAGGGAAGAAGAGCACATTCTCCAGAGCACGAGTGACAATGT
 -280                -260                -240                -220                -200                -180
GGGTAGGGCACCGCCCAGGGCAGTCCCCAAATCCCCATCACTTCAGCTAGAAGCCAGTGCTAGGCAGGAGTGTGTTGTATGCCAGAGCGCCCTCTGCTACTTGAAGCAGGACAGCTTGT
 -160                -140                -120                -100                -80                 -60
GTTCCGAGTCCGGTCCTTGACAGAAACTCTGTATAATACTGATGGCTCCAACCAAAGTGATCAATACAAAGGAAAACTGCTCCAAGGACATTCAAAAGTGTCTGAGGTGCTGTATTGAGT
 -40                 -20                 -1                  20                  40                  60
GTTGGTCATTGGTCTTTGGGAGCCTGGGCAGCAGCAGGGCAGTCATGGAGCTGCTGAATGGGACTGGCCTGTGGCCCATGGCCATATTCACAGTCATCTTCATATTACTGGTGGACCTGA
                                                M  E  L  L  N  G  T  G  L  W  P  M  A  I  F  T  V  I  F  I  L  L  V  D  L  M
 80                  100                 120                 140                 160                 180
TGCACCGGCACCAGCGCTGGACTTCTGCCTACCCTCCAGCCCCTGTGCCGTGGCCTGTGCTGGGCAACCTGCTGCAGGTGGACCCCAGTAACATGCCATACAGCATGTACAAGCTTCAAC
 H  R  H  Q  R  W  T  S  R  Y  P  P  G  P  V  P  W  P  V  L  G  N  L  L  Q  V  D  P  S  N  M  P  Y  S  M  Y  K  L  Q  H
 200                 220                 240                 260                 280                 300
ACCGCTATGGTGAGTGTTCAGCCTGCAGATGGGTTGGAAGCCCATGGTCATTGTCAACAGACTGAAAGCGGTGCAGGAAGTGCTGGTGACTCATGGAGAGGACACTGCTGACCGCCCTC
 R  Y  G  D  V  F  S  L  Q  M  G  W  K  P  M  V  I  V  N  R  L  K  A  V  Q  E  V  L  V  T  H  G  E  D  T  A  D  R  P  P
 320                 340                 360                 380                 400                 420
CAGTGCCCATCTTTAAGTGCCTGGGTGTGAAGCCCAGATCCCAAGGTGTGGTCTTTGCATCTTACGGGCCTGAGTGGCGAGACCAGAGGCGATTCTCTGTGTCCACCCTGCGCACCCTCG
 V  P  I  F  K  C  L  G  V  K  P  R  S  Q  G  V  V  F  A  S  Y  G  P  E  W  R  E  Q  R  R  F  S  V  S  T  L  R  T  F  G
 440                 460                 480                 500                 520                 540
GCATGGGCAAGAAATCCCTGGAGGAGTGGGTGACCAAGGAGCCCGCCACCTCTGTGATGCCTTCACTGCCCAGAATGGGCGGTCCATCAATCCCAAGGCCATGCTGAACAAAGCCTTGT
 M  G  K  K  S  L  E  E  W  V  T  K  E  A  G  H  L  C  D  A  F  T  A  Q  N  G  R  S  I  N  P  K  A  M  L  N  K  A  L  C
 560                 580                 600                 620                 640                 660
GTAATGTGATCGCATCCCTCATTTTTGCCCGTCGCTTTGAATATGAAGACCCTTACCTCATCAGGATGCTGACACTAGTGGAAGAAAGTTTGATAGAAGTCTCTGGTTTCATTCCTGAGG
 N  V  I  A  S  L  I  F  A  R  R  F  E  Y  E  D  P  Y  L  I  R  M  L  T  L  V  E  E  S  L  I  E  V  S  G  F  I  P  E  V
 680                 700                 720                 740                 760                 780
TTCTTAACAGTTCCCAGCACTGCTGCGCATCCCAGGGCTGGCTGACAAGGTCTTCCAAGGTCAGAAGACCTTCATGGCCTTCCTGGATAACCTGCTGGCTGAGAACAGGACCACCTGGG
 L  N  T  F  P  A  L  L  R  I  P  G  L  A  D  K  V  F  Q  G  Q  K  T  F  M  A  F  L  D  N  L  L  A  E  N  R  T  T  W  D
 800                 820                 840                 860                 880                 900
ACCCTGCCCAGCCACCCCGAAATCTGACTGATGCCTTCCTGGCTGAGGTGGAGAAGGCCAAGGGGATCCTGAGAGCAGCTTTAATGATGAGAACCTAGCCATGGTGGTGAGTTGACCTGT
 P  A  Q  P  P  R  N  L  T  D  A  F  L  A  E  V  E  K  A  K  G  N  P  E  S  S  F  N  D  E  N  L  R  M  V  V  V  D  L  F
 920                 940                 960                 980                 1000                1020
TCACTGCAGGGATGGTGACCACCGCCCACCACACTGACCTGGGCCCTACTGCTCATGATTCTATACCCGGATGTGCAGCGCAGAGTCCAACAAGAGATTGATGAGGTCATAGGGCAGGTGC
 T  A  G  M  V  T  T  A  T  T  L  T  W  A  L  L  L  M  I  L  Y  P  D  V  Q  R  R  V  Q  Q  E  I  D  E  V  I  G  Q  V  R
1040                1060                1080                1100                1120                1140
GGTGTCCAGAGATGACAGACCAGGCCCACATGCCCTACACCAATGCTGTCATCCATGAGGTACAGCGCTTTGGGGACATTGCTCCATTGAATTTGCCACGCATCAGGAGTTGTGACATTG
 C  P  E  M  T  D  Q  A  H  M  P  Y  T  N  A  V  I  H  E  V  Q  R  F  G  D  I  A  P  L  N  L  P  R  I  T  S  C  D  I  E
1160                1180                1200                1220                1240                1260
AAGTCCAGGACTTCGTCATCCCCAAGGGGACGACCCTCATCATCAACCTGTCGTCCGTGCTGAGGATGAGACTGTCTGGGACAAGCCCCTCCGCTTCCACCCTGACACTTCCTGGATG
 V  Q  D  F  V  I  P  K  G  T  T  L  I  I  N  L  S  S  V  L  K  D  E  T  V  W  E  K  P  L  R  F  H  P  E  H  F  L  D  A
1280                1300                1320                1340                1360                1380
CCCAGGGCAACTTTGTGAAGCATGAGGCCTTCATGCCATTCTCAGCAGGCCGCAGAGCATGCCTTGGGAGCCCCTGGCCCGCATGGAGCTCTTCCTCTTCTTCACCTGCCTCCTGCAGC
 P  G  N  F  V  K  H  E  A  F  M  P  F  S  A  G  R  R  A  C  L  G  E  P  L  A  R  M  E  L  F  L  F  F  T  C  L  L  Q  H
1400                1420                1440                1460                1480                1500
ACTTCAGCTTCTCCGTGCCGCCGGACAGCCCCGGCCCAGCACCCTTGGCAACTTTGCTATTTCAGTTGCCCCCTTGCCCTACCAGCTCTGTGCTGCGGTACGGGAGCAAGGACACTAAT
 F  S  F  S  V  P  A  G  Q  P  R  P  S  T  L  G  N  F  A  I  S  V  A  P  L  P  Y  Q  L  C  A  A  V  R  E  Q  G  H  *
1520                1540                1560
TCCAGTCCAGCTCTACAGGGCAAGCTGAGCCACGTTTAATAAACCAGTCTTGAGGCT(A)n
```

References

1. Sugita, O. et al. (1988) Biochemistry 27:678-686.
2. Sugita, O. et al. (1988) Biochem. Biophys. Res. Commun. 150:925-930.
3. Ishida, N. et al. (1988) Biochem. Biophys. Res. Commun. 156:681-688.
4. Gonzalez, F.J. et al. (1987) DNA 6:149-161.

© IRL Press