Volume 15 Number 9 1987 Nucleic Acids Research

Nucleotide and deduced amino acid sequence of rat liver 17β-hydroxysteroid UDP- glucuronosyltransferase

David Harding, Stuart M.Wilson, Michael R.Jackson, Brian Burchell, Mitchell D.Green* and Thomas R.Tephly*

Department of Biochemistry, The University, Dundee DD1 4HN, UK and *Department of Pharmacology, University of Iowa, Iowa City, IA 52242, USA

Submitted February 6, 1987; Reformatted April 23, 1987 Accession no. Y00156

We have isolated a cloned cDNA (rlug 38) from a rat liver cDNA library in λgt11, which contained the full length coding sequence of a hepatic microsomal UDP-glucuronosyltransferase (UDPGT, see 1). Comparison of the amino acid sequence obtained by translation of the cDNA with the amino-terminal sequence of purified 17β-hydroxysteroid UDPGT protein showed an identical 38 amino acid sequence (underlined) starting at residue 23. This suggests that rlug38 cDNA encodes 17β-hydroxysteroid UDPGT and that this pre-protein contains a cleaved 23 residue signal sequence.

```
                                            GAAAGCTTACTAGATAAAAGGTCAAGCAGTGACAGAAAGGAACACACTGTGAACAGAAGGTTTTTGATTTTTAAG

ATG CCT GGG AAG TGG ATT TCT GCT CTT CTG CTG CTA CAG ATA AGT TGC TGT TTC CAA TCT GGG AAC TGT GGG AAG GTG TTG GTG TGG CCG      90
Met Pro Gly Lys Trp Ile Ser Ala Leu Leu Leu Leu Gln Ile Ser Cys Cys Phe Gln Ser Gly Asn Cys Gly Lys Val Leu Val Trp Pro
                                     10                                      20                                      30
ATG GAA TTC AGC CAC TGG ATG AAT ATA AAA ACA ATA CTG GAT GAA CTT GTA CAG AGG GGC CAT GAA GTC ACT GTT CTG AAA CCT TCG GCT     180
Met Glu Phe Ser His Trp Met Asn Ile Lys Thr Ile Leu Asp Glu Leu Val Gln Arg Gly His Glu Val Thr Val Leu Lys Pro Ser Ala
                                     40                                      50                                      60
TAC TAT GTT CTT GAT CCG AAA AAA TCG CCA GAT CTT AAG TTT GAA ACT TTC CCT ACA TCT GTC AGT AAA GAT GAA CTG GAA AAT TAT TTC     270
Tyr Tyr Val Leu Asp Pro Lys Lys Ser Pro Asp Leu Lys Phe Glu Thr Phe Pro Thr Ser Val Ser Lys Asp Glu Leu Glu Asn Tyr Phe
                                     70                                      80                                      90
ATA AAG CTT GTG GAT GTG TGG ACT TAT GAG TTG CAA AGA GAT ACA TGT TTG TCT TAT TCT CCT TTA CTA CAA AAT ATG ATA GAT GAA TTT     360
Ile Lys Leu Val Asp Val Trp Thr Tyr Glu Leu Gln Arg Asp Thr Cys Leu Ser Tyr Ser Pro Leu Leu Gln Asn Met Ile Asp Glu Phe
                                    100                                     110                                     120
TCT GAT TAC TAT CTA AGT CTT TGT AAA GAC ACC GTT TCA AAC AAG CAG CTC ATG GCA AAA CTA CAG GAA TCC AAG TTT GAT GTT CTT TTG     450
Ser Asp Tyr Tyr Leu Ser Leu Cys Lys Asp Thr Val Ser Asn Lys Gln Leu Met Ala Lys Leu Gln Glu Ser Lys Phe Asp Val Leu Leu
                                    130                                     140                                     150
TCA GAT CCT GTG GCT GCC TGT GGG GAG CTG ATA GCC GAA GTG CTC CAC ATT CCT TTT CTG TAC AGT CTT CGC TTC TCT CCA GGC TAC AAA     540
Ser Asp Pro Val Ala Ala Cys Gly Glu Leu Ile Ala Glu Val Leu His Ile Pro Phe Leu Tyr Ser Leu Arg Phe Ser Pro Gly Tyr Lys
                                    160                                     170                                     180
ATT GAA AAG TCC AGT GGA AGA TTT ATA TTA CCT CCA TCT TAT GTA CCA GTA ATT TTG TCA GGA ATG GGT GGC CCA ATG ACA TTC ATA GAC     630
Ile Glu Lys Ser Ser Gly Arg Phe Ile Leu Pro Pro Ser Tyr Val Pro Val Ile Leu Ser Gly Met Gly Gly Pro Met Thr Phe Ile Asp
                                    190                                     200                                     210
AGG GTA AAA AAT ATG ATA TGT ACG CTT TAT TTT GAC TTT TGG TTC CAT ATG TTT AAT GCC AAG AAA TGG GAT CCA TTT TAC AGT GAG ATT     720
Arg Val Lys Asn Met Ile Cys Thr Leu Tyr Phe Asp Phe Trp Phe His Met Phe Asn Ala Lys Lys Trp Asp Pro Phe Tyr Ser Glu Ile
                                    220                                     230                                     240
TTT GGA AGG CCC ACC ACA TTA GCT GAG ACA ATG GGC AAA GCA GAA ATG TGG CTC ATT AGA TCC TAC TGG GAT TTG GAA TTT CCC CAC CCA     810
Phe Gly Arg Pro Thr Thr Leu Ala Glu Thr Met Gly Lys Ala Glu Met Trp Leu Ile Arg Ser Tyr Trp Asp Leu Glu Phe Pro His Pro
                                    250                                     260                                     270
ACA TTA CCA AAT GTT GAC TAT ATT GGA GGA CTC CAA TGC AGA CCT CCT AAA CCC TTG CCC AAG GAT ATG GAA GAT TTT GTC CAG AGC TCT     900
Thr Leu Pro Asn Val Asp Tyr Ile Gly Gly Leu Gln Cys Arg Pro Pro Lys Pro Leu Pro Lys Asp Met Glu Asp Phe Val Gln Ser Ser
                                    280                                     290                                     300
GGA GAG CAT GGC GTG GTG GTG TTT TCT CTG GGG TCA ATG GTC AGC AGC ATG ACA GAA GAA AAG GCC AAT GCA ATT GCA TGG GCC CTT GCC     990
Gly Glu His Gly Val Val Val Phe Ser Leu Gly Ser Met Val Ser Ser Met Thr Glu Glu Lys Ala Asn Ala Ile Ala Trp Ala Leu Ala
                                    310                                     320                                     330
CAG ATT CCA CAA AAG GTT CTT TGG AAA TTT GAT GGC AAA ACC CCA GCA ACC TTA GGA CCC AAT ACC AGA GTC TAC AAG TGG CTT CCC CAG    1080
Gln Ile Pro Gln Lys Val Leu Trp Lys Phe Asp Gly Lys Thr Pro Ala Thr Leu Gly Pro Asn Thr Arg Val Tyr Lys Trp Leu Pro Gln
                                    340                                     350                                     360
AAT GAC CTC CTT GGT CAT CCA AAA ACC AAA GCC TTT GTA ACT CAT GGT GGA GCC AAT GGT GTC TAT GAG GCC ATC TAT CAT GGA ATC CCT    1170
Asn Asp Leu Leu Gly His Pro Lys Thr Lys Ala Phe Val Thr His Gly Gly Ala Asn Gly Val Tyr Glu Ala Ile Tyr His Gly Ile Pro
                                    370                                     380                                     400
ATG GTT GGC ATT CCT ATG TTT GGA GAA CAA CAT GAT AAC ATT GCC CAC ATG GTG GCC AAA GGA GCA GCT GTT ACA CTG AAT ATC AGG ACA    1260
Met Val Gly Ile Pro Met Phe Gly Glu Gln His Asp Asn Ile Ala His Met Val Ala Lys Gly Ala Ala Val Thr Leu Asn Ile Arg Thr
                                    400                                     410                                     420
ATG TCA AAG ACA GAT TTG TTC AAT GCA CTA AAG GAA ATA ATA AAC AAT CCA TTC TAT AAA AAA AAT GCT GTG TGG TTG TCA ACC ATT CAC    1350
Met Ser Lys Thr Asp Leu Phe Asn Ala Leu Lys Glu Ile Ile Asn Asn Pro Phe Tyr Lys Lys Asn Ala Val Trp Leu Ser Thr Ile His
                                    430                                     440                                     450
CAT GAC CAA CCT ATG AAG CCC CTG GAC AAG GCT GTC TTC TGG ATT GAG TTT GTC ATG CGC CAC AAA GGG GCC AAG CAC CTG AGA CCA CTT    1440
His Asp Gln Pro Met Lys Pro Leu Asp Lys Ala Val Phe Trp Ile Glu Phe Val Met Arg His Lys Gly Ala Lys His Leu Arg Pro Leu
                                    460                                     470                                     480
GGA CAT GAC CTT CCC TGG TAC CAG TAC CAC TCT CTG GAT GTG ATT GGA TTC CTG CTC TCC TGT TCG GCA GTC ATT GCA GTC CTT ACT GTA    1530
Gly His Asp Leu Pro Trp Tyr Gln Tyr His Ser Leu Asp Val Ile Gly Phe Leu Leu Ser Cys Ser Ala Val Ile Ala Val Leu Thr Val
                                    490                                     500                                     510
AAA TGC TTC TTG TTT ATT TAC CGA CTC TTT GTG AAG AAG GAA AAG AAA ATG AAG AAT GAG TAG AGCTCATTTACAATGCACTACAGGAATGAAATTTC  1628
Lys Cys Phe Leu Phe Ile Tyr Arg Leu Phe Val Lys Lys Glu Lys Lys Met Lys Asn Glu
                                    520                                     530

AGCCACATTCTAA
```

REFERENCES

1. Jackson, M R, McCarthy, L R, Harding, D, Wilson, S, Coughtrie, M W H & Burchell, B (1987) Biochem. J. 242 581-588.