Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 4558–4562, June 1991
Biochemistry

Cytochrome P450 family 4 in a cockroach: Molecular cloning and regulation by hypertrehalosemic hormone

(insect/fat body/metabolism/adipokinetic hormone/molecular evolution)

JAMES Y. BRADFIELD, YING-HUE LEE*, AND LARRY L. KEELEY

Laboratories for Invertebrate Neuroendocrine Research, Department of Entomology, Texas A&M University, College Station, TX 77843-2475

Communicated by Ronald W. Estabrook, March 4, 1991 (received for review January 8, 1991)

ABSTRACT **Hypertrehalosemic hormone (a carbohydrate-mobilizing neuroendocrine decapeptide) and starvation markedly increased levels of a cockroach (*Blaberus discoidalis*) fat body cytochrome P450 message. The gene represented by the cloned P450 cDNA has been named *CYP4C1* (cytochrome P450 family *4*, subfamily *C*, gene *1*), a newly identified member of the ubiquitous cytochrome P450 monooxygenase gene superfamily. *Blaberus* CYP4C1 (511 amino acids, M_r = 58,485) has a hydrophobic NH_2 terminus and a sequence near the COOH terminus that is homologous to the cysteine-containing heme-binding region definitive of cytochromes P450. The cockroach sequence is 32–36% identical to mammalian family 4A and 4B enzymes. It contains a 13-residue sequence characteristic of family 4 but not other P450s. This study suggests that CYP4C1 is hormonally regulated in association with energy substrate mobilization and supports the idea that family *4* is an old and widespread gene family.**

The cytochrome P450 gene (*CYP*†) superfamily encodes NAD(P)H-dependent, heme-containing monooxygenases that metabolize numerous endogenous and exogenous substrates such as steroids, fatty acids, drugs, carcinogens, and pesticides (2, 3). According to amino acid sequence analysis, ≈150 genes from the *CYP* superfamily have been assigned to 27 families (1), and a phylogeny of the superfamily has been proposed (1, 4, 5). Most P450 sequences have been obtained from vertebrates (especially mammals) but bacterial, yeast, fungal, invertebrate [including one insect (6)], and plant sequences are also described. The fact that cytochromes P450 are ubiquitous argues for descent of these proteins from a sequence that predated divergence of prokaryotes and eukaryotes.

We have identified a cytochrome P450 sequence (deduced from cDNA) from the neotropical cockroach *Blaberus discoidalis*.‡ The cDNA was obtained from a fat body (analogous to vertebrate liver) library by differential hybridization designed to select gene sequences regulated by the hypertrehalosemic hormone (HTH) (7). HTH is a neuroendocrine decapeptide produced by the corpora cardiaca—paired neurosecretory glands attached to the brain. A major action for HTH is the stimulation of fat body glycogenolysis for the production of precursors for synthesis of trehalose, the main circulating carbohydrate of insects (8). HTH is a member of a large group of structurally related neuropeptides—the adipokinetic hormone/red pigment-concentrating hormone family (9)—that regulate energy substrate mobilization and metabolism in arthropods. Levels of the P450 mRNA are low in the fat body of decapitated (gland-free) adult male cockroaches, but strongly stimulated in decapitated animals by HTH administration. The P450 message is also stimulated by starvation in the presence but not in the absence of the head.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Comparison of the 511-amino acid *Blaberus* P450 with other P450s revealed highest similarity (32–36% positional identity) with family 4 proteins, a group of microsomal enzymes that primarily catabolize fatty acids and prostaglandins (10–16). The most striking region of homology was a 13-amino acid peptide found in all family 4 sequences examined but in no other P450. The cockroach P450 gene has been named *CYP4C1* by a widely recognized committee on P450 nomenclature (1) and is thus deemed a member of a P450 family known previously from mammals.

MATERIALS AND METHODS

Insects. Adult male *B. discoidalis* were maintained on Purina Dog Chow and water as described (17). Ages of the insects were measured from the day of the adult molt (day 0).

Hormone Replacement Therapy. Cockroaches were decapitated on day 0 to remove the corpora cardiaca and ensure nutritional uniformity, and the wound was sealed with a melted beeswax/petrolatum mixture. On days 1 and 2, either 10 or 100 pmol (1 pmol = 1 ng) of synthetic HTH dissolved in 10 μl of Ephrussi–Beadle Ringer solution was injected into the hemocoel through a ventral abdominal intersegmental membrane. Control decapitated insects received 10 μl of the Ringer solution on the same days.

Library Construction and Screening. Total fat body RNA was isolated (18) from day-3 HTH-treated decapitated cockroaches and the poly(A)$^+$ fraction was prepared (19). Double-stranded cDNA (20) was ligated to *Eco*RI/*Not* I linker-adaptors and cDNAs ≥250 nucleotides were selected with a Sephacryl S400 column (Pharmacia). The cDNA was ligated to bacteriophage λZAP arms and packaged with Gigapack Plus (Stratagene). Recombinant plaques were propagated on *Escherichia coli* PLK-F lawns and transferred in duplicate to nylon filters at a density of 30 plaques per cm^2. To detect HTH-regulated gene sequences, filters were hybridized with [^{32}P]cDNA (21) generated to total fat body poly(A)$^+$ RNA from HTH-treated decapitated males versus [^{32}P]cDNA from the fat body of decapitated controls. Hybridization (10^6 cpm/ml) was in 50% (vol/vol) formamide/0.9 M NaCl at 42°C and washes were in 10 mM NaCl at 60°C as described (22).

Northern Blot. Total fat body RNA was denatured with methylmercuric hydroxide (23), separated by electrophoresis in 1.2% agarose, and transferred to positively charged nylon membranes. Blots were hybridized as described (22) with a nick-translated 2328-nucleotide bacteriophage insert selected to represent an HTH-stimulated fat body message.

Abbreviation: HTH, hypertrehalosemic hormone.

*Present address: The Beatson Institute for Cancer Research, Garscube Estate, Switchback Road, Bearsden, Glasgow G61 1BD, Scotland.

†We use the nomenclature recommended by Nebert *et al.* (1). *CYP* refers to a cytochrome P450 gene(s) and cDNA(s). CYP refers to P450 mRNA(s) and protein(s).

‡The sequence reported in this paper has been deposited in the GenBank data base (accession no. M63798).

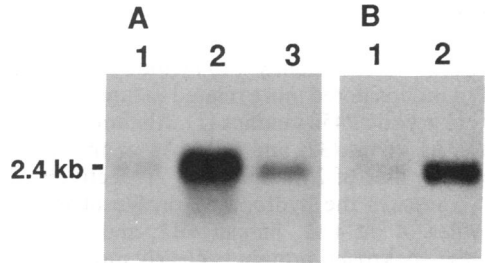

FIG. 1. Northern hybridization showing increase of a cockroach 2.4-kb fat body message by HTH (*A*) and starvation (*B*). Lanes contain 5 μg of total fat body RNA (pooled from six to eight males 3 days after the adult molt) probed with a 2328-nucleotide cloned cDNA. (*A*) Animals decapitated on day 0 and injected with Ringer solution (lane 1), 100 pmol of HTH (lane 2), or 10 pmol of HTH (lane 3) on days 1 and 2. (*B*) Day-3 males allowed (lane 1) or deprived of (lane 2) food and water from day 0.

Sequence Analysis. The 2328-nucleotide insert was subcloned into pBluescript (Stratagene) and a set of nested deletions from each end of the insert was made with an exonuclease III/S1 nuclease system (Promega). Double-strand DNA sequencing was with deoxyadenosine 5′-[α-[^{35}S]thio]triphosphate and Sequenase 2.0 (United States Biochemical). The sequence of the 2328-nucleotide fragment indicated an incomplete coding region. Therefore we constructed a 20-nucleotide primer complementary to a 5′ region of the cloned insert and determined 42 upstream nucleotides, using 8 μg of poly(A)$^+$ RNA and dideoxynucleotide chain termination (24).

RESULTS

Isolation of cDNA to an HTH-Regulated Message. Three thousand plaques from a once-amplified library of 2×10^6 primary recombinants were screened with radiolabeled fat body cDNA from HTH-treated decapitated males versus cDNA from decapitated control males. We detected one plaque that hybridized moderately with cDNA from HTH-treated cockroaches and weakly with cDNA from controls. We used the 1500-nucleotide insert from the selected recombinant to screen another 3000 plaques from the cDNA library. A 2328-nucleotide fragment was the longest insert found among 10 hybridizing plaques.

Fig. 1*A* shows Northern hybridization with the 2328-nucleotide cDNA and total fat body RNA from day-3 males that had been decapitated on day 0. A 2.4-kilobase (kb) hybridizing message was at a low level in Ringer solution-injected control cockroaches (lane 1) but stimulated in a dose-responsive manner in decapitated insects given 100 pmol (lane 2) or 10 pmol (lane 3) of HTH on days 1 and 2. We estimate by dot-blot titration (not shown) that the amount of 2.4-kb RNA was increased 50-fold by 100 pmol of HTH administered to decapitated cockroaches. Fig. 1*B* shows hybridization of the cloned cDNA with fat body RNA from intact day-3 adult males provided (lane 1) or deprived of (lane 2) food and water from day 0. The 2.4-kb RNA was clearly more abundant in cockroaches deprived of food and water. These results suggest that both HTH and starvation up-regulate levels of a 2.4-kb message in *Blaberus* fat body.

The effects of starvation may be mediated by HTH. Decapitated control animals (Fig. 1*A*, lane 1) were obviously unable to eat and drink, yet they did not show the starvation-related increase in 2.4-kb transcript observed with intact animals. This suggests that the starvation effect required presence of the head. Possibly, starvation induced a stress-related secretion of HTH.

Sequence Analysis. The sequence of the 2328-nucleotide cDNA plus 42 upstream nucleotides determined by RNA sequencing is displayed in Fig. 2. The first ATG (position 34) begins an open reading frame for 511 amino acids (M_r =

```
ctaattctacgtaacttctgtcagacagtaaacATGGAATTCATCACCATCCTGCTGAGTACTGCACTCTTCATAGTCACCTTCCTGTTCCTGTTCCGTCAGGGCGCTAAGAGAGCTCGC  120
                          M  E  F  I  T  I  L  L  S  T  A  L  F  I  V  T  F  L  F  L  F  R  S  G  A  K  R  A  R    29
TTCGTGTACCTGGTTAACAAATTGCCCGGCCCCACCGCCTACCCCGTCGTCGGCAATGCCATCGAGGCCATCGTTCCAAGAAACAAACTGTTCCAGGTATTTGATAGAAGAGCCAAACTA  240
 F  V  Y  L  V  N  K  L  P  G  P  T  A  Y  P  V  V  G  N  A  I  E  A  I  V  P  R  N  K  L  F  Q  V  F  D  R  R  A  K  L   69
TATGGTCCTCTGTATCGAATTTGGGCTGGTCCTATAGCTCAAGTCGGTCTCACACGACCTGAACATGTCGAGCTGATTCTTCGAGACACGAAACATATCGACAAGTCTCTTGTCTATTCA  360
 Y  G  P  L  Y  R  I  W  A  G  P  I  A  Q  V  G  L  T  R  P  E  H  V  E  L  I  L  R  D  T  K  H  I  D  K  S  L  V  Y  S  109
TTTATACGTCCATGGCTTGGAGAAGGGCTTCTTACTGGAACAGGGGCCAAATGGCACTCCCATAGGAAGATGATCACCCCTACATTCCACTTCAAGATCCTCGACATATTTGTTGACGTA  480
 F  I  R  P  W  L  G  E  G  L  L  T  G  T  G  A  K  W  H  S  H  R  K  M  I  T  P  T  F  H  F  K  I  L  D  I  F  V  D  V  149
TTTGTTGAAAAAAGTGAAATCCTAGTGAAAAAATTACAGTCAAAAGTTGGAGGGAAGGATTTTGACATCTACCCATTCATTACACATTGTGCTCTAGATATTATCTGTGAAACTGCAATG  600
 F  V  E  K  S  E  I  L  V  K  K  L  Q  S  K  V  G  G  K  D  E  D  I  Y  P  F  I  T  H  C  A  L  D  I  I  C  E  T  A  M  189
GGGATCCAAATGAATGCACAGGAAGAATCAGAATCTGAATATGTGAAGGCAGTCTATGAGATCAGTGAACTGACAATGCAACGTTCAGTTCGGCCATGGCTTCACCCAAAAGTAATATTT  720
 G  I  Q  M  N  A  Q  E  E  S  E  S  E  Y  V  K  A  V  Y  E  I  S  E  L  T  M  Q  R  S  V  R  P  W  L  H  P  K  V  I  F  229
GATTTAACGACAATGGGGAAGAGGTATGCAGAATGCCTGAGGATTCTTCATGGTTTCACGAATAAAGTTATTCAAGAAAGGAAGAGCTTGAGACAAATGACTGGGATGAAGCCTACAATT  840
 D  L  T  T  M  G  K  R  Y  A  E  C  L  R  I  L  H  G  F  T  N  K  V  I  Q  E  R  K  S  L  R  Q  M  T  G  M  K  P  T  I  269
TCTAATGAAGAAGATGAACTTCTTGGAAAGAAAAAGAGATTGGCATTCTTGGATTTACTTCTGGAGGCTTCTGAGAATGGGACAAAGATGTCAGACACTGATATCAGAGAGGAAGTAGAC  960
 S  N  E  E  D  E  L  L  G  K  K  K  R  L  A  F  L  D  L  L  L  E  A  S  E  N  G  T  K  M  S  D  T  D  I  R  E  E  V  D  309
ACATTCATGTTTGAGGGTCATGACACAACATCTGCAGGAATATGCTGGGCTCTTTTCCTTCTTGGATCTCATCCTGAAATTCAGGACAAAGTATATGAAGAATTGGACCACATATTCCAA 1080
 T  F  M  F  E  G  H  D  T  T  S  A  G  I  C  W  A  L  F  L  L  G  S  H  P  E  I  Q  D  K  V  Y  E  E  L  D  H  I  F  Q  349
GGCTCAGATCGATCCACAACAATGAGGGATTTAGCTGATATGAAATACCTTGAGAGGGTCATCAAAGAGAGTCTCAGACTGTTCCCCAGTGTACCATTCATCGGTAGAGTACTCAAGGAG 1200
 G  S  D  R  S  T  T  M  R  D  L  A  D  M  K  Y  L  E  R  V  I  K  E  S  L  R  L  F  P  S  V  P  F  I  G  R  V  L  K  E  389
GACACCAAGATAGGAGACTACTTGGTACCTGCAGGATGCATGATGAATTTACAGATCTACCACGTACATCGTAACCAAGACCAATACCCCAATCCTGAAGCTTTCAACCCTGACAACTTC 1320
 D  T  K  I  G  D  Y  L  V  P  A  G  C  M  M  N  L  Q  I  Y  H  V  H  R  N  Q  D  Q  Y  P  N  P  E  A  F  N  P  D  N  F  429
CTTCCTGAAAGAGTAGCTAAGAGACACCCCTACGCTTACGTTCCGTTCAGTGCTGGTCCTAGGAACTGCATTGGACAAAAGTTTGCAACACTGGAAGAGAAGACAGTATTGTCTAGCATC 1440
 L  P  E  R  V  A  K  R  H  P  Y  A  Y  V  P  F  S  A  G  P  R  N  C  I  G  Q  K  F  A  T  L  E  E  K  T  V  L  S  S  I  469
TTGCGCAACTTCAAGGTACGGTCAATAGAGAAGAGAGAGGACCTCACACTTATGAATGAGCTCATTCTTAGACCAGAGTCAGGCATCAAGGTAGAACTGATCCCAAGACTCCCAGCTGAT 1560
 L  R  N  F  K  V  R  S  I  E  K  R  E  D  L  T  L  M  N  E  L  I  L  R  P  E  S  G  I  K  V  E  L  I  P  R  L  P  A  D  509
GCATGTtgaattttaagttcctatcagaagacttataaatgtatgactaagttgaagaatgacaatgtttagtgatggtacagtagaatacttattacaaagaacatggttgctataaat 1680
 A  C end                                                                                                                  511
taagtcaataaaaccacgcataatttgtgcatgaaaatctgcgatagacatagaaagaaaaatggtacaatggatttcagggcaacttcctacctacatttttttttattatgtttgttga 1800
aaagttcttagtcacctcatcagccttgtcaccttaaaataaatttcttttgtattaattcatagtttaccacactgtttctacttttggaagcagttgtggtaagttttcttttgagt 1920
tacatggccctggaatactatttcaaaagttctgtgaatttttaactcatatacaaatcaatttaatctccttcctgatactgtaatgctactgtactgaaagtgtaatgttgtttgaat 2040
atgagcctaactttaatttatgtaatagtattttataacttgtgaggaggtgcgaagttgttagattcctattgaataattatagactgaatgtggaagcacaatagtggattcattgca 2160
gtacaacatgtgaaaattggtcattagtctgacagactcattcagtaatctgcgatggccaaaatattatcactattttgtatttacttttacacatagcttacttgttatgtgctaaa 2280
ttatttagtgcatctgtaaattactgaaatgatttgtaaatacgtatttttattatagtaataatgtgttaataaacagttttatttct(a)n                          2370
```

FIG. 2. Nucleotide sequence of the *Blaberus* CYP4C1 cDNA and deduced 511-residue protein (M_r = 58,485). Lowercase nucleotide symbols indicate noncoding regions. Doubly underlined residues near the NH_2 terminus indicate a hydrophobic region characteristic of microsomal CYPs. Doubly underlined residues around the heme-binding cysteine (position 452) are conserved in the CYP superfamily. The polyadenylylation signal is underlined.

58,485). The open reading frame is followed by an 801-nucleotide sequence that includes a consensus polyadenylylation signal (AATAAA) closely followed by poly(dA) at the 3′ terminus.

A search [WORDSEARCH program of the University of Wisconsin Genetics Computer Group (UWGCG)] of 12,000 protein sequences in National Biomedical Research Foundation Release 20.0 selected microsomal cytochromes P450 as most similar to the deduced *Blaberus* protein. Closer analysis of the cockroach sequence suggested that it is a microsomal CYP on the basis of its hydrophobic NH_2-terminal region, a number of amino acids near its COOH terminus that are invariant in the heme-binding region of cytochromes P450, and its size (Fig. 2 and ref. 25).

Fig. 3 shows an alignment of the cockroach P450 protein sequence with representatives from two subfamilies of cytochrome P450 family 4 from mammals: 4A1 [rat liver lauric acid ω-hydroxylase (12)] and 4B1 [human lung P450 (26)]. Among interspersed residues shared by the three proteins are two regions of conspicuous homology. One of these, located around Cys-452, identifies the proteins as members of the heme-binding CYP superfamily (see Fig. 2). Of equal interest is a 13-residue peptide (positions 307–319) that appears to be wholly conserved around position 315 in family 4 (12–15, 26–30).

Global alignments using the UWGCG GAP program based on the algorithm of Needleman and Wunsch (31) indicated 32–36% amino acid positional identity between the cockroach sequence and 10 mammalian CYP4s representing subfamilies A and B (12–15, 26–30). Although this identity is lower than the 40% usually required for inclusion of a sequence in an existent CYP family (1), it is proposed that the *Blaberus* protein be assigned to family 4 on the basis of presence of the 13-amino acid peptide found exclusively in CYP4s and in consideration of the evolutionary distance between insects and mammals. Family 4 is believed to have originated more than a billion years ago (4)—long before the divergence of invertebrates and vertebrates. Having been provided the cockroach protein sequence by our personal communication, a committee dedicated to evolutionary classification of P450s has assigned the name *CYP4C1* to the *Blaberus* gene in a recent update of the *CYP* superfamily (1). The cockroach sequence thus extends family 4 to include three subfamilies.

Family 4 is considered more related to families 3 and 6 than to other eukaryotic P450 families (1). Alignments of CYP4C1 with CYP3A1 of rat (32) and CYP6A1 of the house fly (6) indicated identities of 26% and 25%, respectively.

Fig. 4 compares the hydropathy profile of *Blaberus* 4C1 with profiles of rat 4A1, human 4B1, and house fly 6A1. Overall, the cockroach protein is clearly more similar to the mammalian enzymes than to the 6A1 sequence from another insect. The family 4 proteins are strikingly similar along the 200 residues (positions 250–450) encompassing the conserved 13-amino acid peptide and the cysteinyl heme-binding region.

DISCUSSION

The insect hyperglycemic and adipokinetic hormones are well recognized for their ability to promote conversion of stored fat body metabolites into circulating metabolites. We have determined that HTH, the hyperglycemic hormone of *B. discoidalis*, also causes a marked increase in levels of a cytochrome P450 message in the fat body. This was clearly a major response to HTH. The increase of the P450 message in response to starvation and the loss of the starvation effect in decapitated animals indicated that the P450 RNA levels are modulated according to the physiological state of the animal and that starvation likely promotes HTH release to stimulate P450 activity.

Taken together, several features of the HTH-stimulated cockroach protein (deduced from cDNA) described here indicate it to be a microsomal cytochrome P450. These features include its mass, a strongly hydrophobic NH_2 terminus, and a number of residues that are conserved around the COOH-terminal heme-binding region and elsewhere (25). To our knowledge this paper reports the second complete P450 sequence from an insect, the first being house fly CYP6A1, whose expression may be associated with insecticide resistance (6).

```
                           *        * *             *          * *      *                             *
Roach  4C1         MEFITILLSTALFIVTFLFLFRQGA-------KRARFVYLVNKLPGPTAYPVVGNAIEAIVPRNKLFQVFDRRAKLYGPLYRIWAGP-I   81
Rat    4A1  MSVSALSSTRFTGSISGFLQVASVLGLLLLLVKAVQFYLQRQWLLKAFQQFPSPPFHWFFGHK-Q--FQGDKELQQIMTCVENFPSAFPRWFWGSK   93
Human  4B1       MVPSFLSLSFSSLGLWASGLILVLGFLKLIHLLLRRRTLAKAMDKFPGPPTHWLFGHALE--IQETGSLDLVVSWAHQFPYAHPLWFGQFI   89

                   *                 **     * * ***   *  *  **   ** **  *   *            *            *
       4C1  AQVGLTRPEHVELILRDTKHIDKSLVYSFIRPWLGEGLLTGTGAKWHSHRKMITPTFHFKILDIFVDVFVEKSEILVKKLQSKV-GGKDFDIYPFI  176
       4A1  AYLIVYDPDYMKVILGRSDPKANG-VYRLLAPWIGYGLLLLNGQPWFQHRRMLTPAFHYDILKPYVKNMADSIRLMLDKWEQLAGQDSSIEIFQHI  188
       4B1  GFLNIYEPDYAKAVYSRGDPKAPD-VYDFFLQWIGRGLLVLEGPKWLQHRKLLTPGFHYDVLKPYVAVFTESTRIMLDKWEEKAREGKSFDIFCDV  184

                *                      *  *      *   *          *      *          *  *  **  **   *
       4C1  THCALDIICETAMGI-QMNAQEESESEYVKAVYEISELTMQRSVRPWLHPKVIFDLTTMGKRYAECLRILHGFTNKVIQERKSLRQMTGMKPTISN  271
       4A1  SLMTLDTVMKCAFSHNGSVQVDGNYKSYIQAIGNLNDLFHSRVRNIFHQNDTIYNFSSNGHLFNRACQLAHDHTDGVIKLRKDQLQNAGEL-----  279
       4B1  GHMALNTLMKCTFGR-GDTGLGHRDSSYYLAVSDLTLLMQQRLVSFQYHNDFIYWLTPHGRRFLRACQVAHDHTDQVIRERKAALQDEKVR-----  274

                       * *** ** *  *     ** * *  *************   *  *        *  *    **      *         *
       4C1  EEDELLGKKKRLAFLDLLLEA-SENGTKMSDTDIRE[EVDTFMFEGHDTT]SAGICWALFLLGSHPEIQDKVYEELDHIFQGSDRSTTMRDLADM   363
       4A1  ---EKVKKKRRLDFLDILLLARMENGDSLSDKDLRA[EVDTFMFEGHDTT]ASGVSWIFYALATHPKHQQRCREEVQSVL-GDGSSITWDHLDQI   368
       4B1  ---LLIQNRRHLDFLDILLGARDEDDIKLSDADLRA[EVDTFMFEGHDTT]TSGISWFLYCMALYPEHQHRCREEVREIL-GDQDFFQWDDLGKM   363

             *    ***  ** * **   * *    *       * *      **  * *    * ** *    *       **     *** * *****  *
       4C1  KYLERVIKESLRLFPSVPFIGRVL-KEDTKIGDYLVPAGCMMNLQIYHVHRNQDQYPNEAFNPDNFLPERVAKRHPYAYVPFSAGPRNCIGQKF   457
       4A1  PYTTMCIKEALRLYPPVPGIVRELSTSVTFPDGRSLPKGIQVTLSIYGLHHNPKVWPNPEVFDPSRFAPDSP--RHSHSFLPFSGGARNCIGKQF   461
       4B1  TYLTMCIKESFRLYPPVPQVYRQLSKPVTFVDGRSLPAGSLISMHIYALHRNSAVWPDPEVFDSLRFSTENASKRHPFAFMPFSAGPRNCIGQQF   458

            *  * *      *  *                * *    *    *
       4C1  ATLEEKTVLSSILRNFKVRSIEKREDLTLMNELILRPESGIKVELIPRLPADAC                                             511
       4A1  AMSEMKVIVALTLLRFEL-LPDPTKVPIPLPRLVLKSKNGIYLYLKKLH                                                  509
       4B1  AMSEMKVVTAMCLLRFEF-SLDPSRLPIKMPQLVLRSKNGFHLHLKPLGPGSGK                                             511
```

FIG. 3. Alignment of cockroach 4C1 with rat 4A1 (12) and human 4B1 (26). Asterisks indicate identical residues in the three proteins. The bracketed 13-amino acid peptide is invariant in family 4.

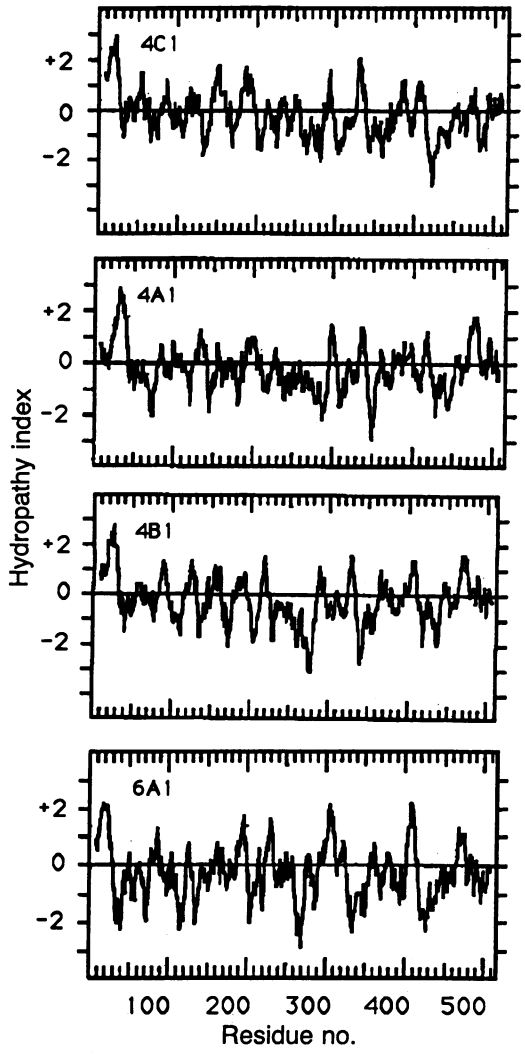

FIG. 4. Hydropathy indices of *Blaberus* 4C1, rat 4A1 (12), human 4B1 (26), and house fly 6A1 (6). Indices are based on a sliding window of 10 amino acids.

The *Blaberus* P450 shows 32–36% amino acid positional identity with family 4A and 4B enzymes and the cockroach gene has been named *CYP4C1*, establishing a third subfamily (*C*) within family *4* (1). Family 4 is considered to have appeared over one billion years ago (4) and, although we cannot rule out the possibility that the similarity between the cockroach protein and mammalian CYP4s is due to convergence or conversion, our data further support the notion of family 4 antiquity. Moreover, our observations suggest that family 4 is maintained across broad taxonomic boundaries. We will be able to draw further evolutionary inferences when we determine the organization of the *Blaberus CYP4C1* structural gene and compare it with that of other family *4* members.

The strict conservation of the 13-amino acid peptide around position 315 in family 4 suggests that substrates for these enzymes might be closely related. This is the case with subfamily 4A, whose members hydroxylate the ω and $\omega - 1$ positions of medium-chain-length fatty acids such as laurate, arachidonate, and palmitate and structurally related prostaglandins (10, 12–16). However, an attempt to demonstrate fatty acid hydroxylation by subfamily 4B has been unsuccessful (26), and compounds distinctly different from fatty acids such as aromatic amines and steroids serve as 4B substrates (11, 29, 33). These observations indicate catalytic variability among CYP4s despite conservation of the peptide near position 315. Nevertheless, the majority of experimental data and the apparent age and wide occurrence of family 4 suggest that the family functions primarily in endogenous metabolic pathways.

We speculate that there is a physiological role for CYP4C1 on the basis of a known action of HTH on the fat body and the elevation of 4C1 mRNA by both HTH and starvation. One of the principal functions of HTH is stimulation of fat body glycogen conversion to trehalose, the main circulating carbohydrate in insects. In cockroaches, the energy supporting the fat body during glycogen conversion to trehalose is apparently supplied by fatty acid oxidation controlled by a factor from the corpora cardiaca (34, 35). In the case of *Blaberus* we believe that the stimulatory factor for fatty acid oxidation is HTH and that the action of HTH may be mediated by increased 4C1, at least in part. If 4C1 were regulated to support glycogen conversion to trehalose, one would expect increased 4C1 activity during starvation when demand for glycogen mobilization is high. We consistently observe an increase of the 4C1 transcript in the fat body of starved cockroaches (an effect abolished by decapitation), and, similarly, increases in P450-dependent fatty acid hydroxylation, possibly hormonally regulated, are described for fasting rats (36). To further understand regulation of the cockroach *CYP4C1* gene by HTH and determine the catalytic activity of the 4C1 protein are our objectives.

We thank Dr. Tim Hayes for providing the synthetic HTH, Drs. Dan Nebert and Frank Gonzalez and their colleagues for classifying the cockroach P450 before its publication and for helpful discussion, and Dr. Bill Plapp for manuscript review. This work was supported by Grant DCB 8511058 from the National Science Foundation and Grant 5 RO1 AI28368 from the National Institutes of Health.

1. Nebert, D. W., Nelson, D. R., Coon, M. J., Estabrook, R. W., Feyereisen, R., Fujii-Kuriyama, Y., Gonzalez, F. J., Guengerich, F. P., Gunsalus, I. C., Johnson, E. F., Loper, J. C., Sato, R., Waterman, M. R. & Waxman, D. J. (1991) *DNA Cell Biol.* **10**, 1–14.
2. Ortiz de Montellano, P. R., ed. (1986) *Cytochrome P-450: Structure, Mechanism, and Biochemistry* (Plenum, New York).
3. Juchau, M. R. (1990) *Life Sci.* **47**, 2385–2394.
4. Nelson, D. R. & Strobel, H. W. (1987) *Mol. Biol. Evol.* **4**, 572–593.
5. Gonzalez, F. J. & Nebert, D. W. (1990) *Trends Genet.* **6**, 182–186.
6. Feyereisen, R., Koener, J. F., Farnsworth, D. E. & Nebert, D. W. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 1465–1469.
7. Hayes, T. K., Keeley, L. L. & Knight, D. W. (1986) *Biochem. Biophys. Res. Commun.* **140**, 674–678.
8. Keeley, L. L., Hayes, T. K., Bradfield, J. Y. & Sowa, S. M. (1991) *Insect Biochem.*, in press.
9. Gäde, G. (1990) *J. Insect Physiol.* **36**, 1–12.
10. Bains, S. K., Gardiner, S. M., Mannweiler, K., Gellet, D. & Gibson, G. G. (1985) *Biochem. Pharmacol.* **34**, 3221–3229.
11. Vanderslice, R. R., Boyd, J. A., Eling, T. E. & Philpot, R. M. (1987) *Mol. Pharmacol.* **31**, 320–325.
12. Hardwick, J. P., Song, B.-J., Huberman, E. & Gonzalez, F. J. (1987) *J. Biol. Chem.* **262**, 801–810.
13. Matsubara, S., Yamamoto, S., Sogawa, K., Yokotani, N., Fujii-Kuriyama, Y., Haniu, M., Shively, J. E., Gotoh, O., Kusunose, E. & Kusunose, M. (1987) *J. Biol. Chem.* **262**, 13366–13371.
14. Yokotani, N., Bernhardt, R., Sogawa, K., Kusunose, E., Gotoh, O., Kusunose, M. & Fujii-Kuriyama, Y. (1989) *J. Biol. Chem.* **264**, 21665–21669.
15. Johnson, E. F., Walker, D. L. & Griffin, K. J. (1990) *Biochemistry* **29**, 873–879.
16. Yoshimura, R., Kusunose, E., Yokotani, N., Yamamoto, S., Kubota, I. & Kusunose, M. (1990) *J. Biochem.* **108**, 544–548.
17. Keeley, L. L., Sowa, S. M., Hayes, T. K. & Bradfield, J. Y. (1988) *Gen. Comp. Endocrinol.* **72**, 364–373.
18. Chirgwin, J. M., Przybyla, A. E., Macdonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
19. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1408–1412.

20. Gubler, U. & Hoffman, B. J. (1983) *Gene* **25,** 263–269.
21. Schleif, R. F. & Wensink, P. W. (1981) *Practical Methods in Molecular Biology* (Springer, New York), pp. 141–143.
22. Bradfield, J. Y., Berlin, R. L. & Keeley, L. L. (1990) *Insect Biochem.* **20,** 105–111.
23. Bailey, J. M. & Davidson, N. (1976) *Anal. Biochem.* **70,** 75–85.
24. Fichot, O. & Girard, M. (1990) *Nucleic Acids Res.* **18,** 6162.
25. Gonzalez, F. J. (1989) *Pharmacol. Rev.* **40,** 243–288.
26. Nhamburo, P. T., Gonzalez, F. J., McBride, O. W., Gelboin, H. V. & Kimura, S. (1989) *Biochemistry* **28,** 8060–8066.
27. Kimura, S., Hanaoka, N., Matsunaga, E. & Gonzalez, F. J. (1989) *DNA Cell Biol.* **8,** 503–516.
28. Kimura, S., Hardwick, J. P., Kozak, C. A. & Gonzalez, F. J. (1989) *DNA Cell Biol.* **8,** 517–525.
29. Gasser, R. & Philpot, R. M. (1989) *Mol. Pharmacol.* **35,** 617–625.
30. Strömstedt, M., Hayashi, S.-I., Zaphiropoulos, P. G. & Gustafsson, J.-Å. (1990) *DNA Cell Biol.* **9,** 569–577.
31. Needleman, S. B. & Wunsch, C. D. (1970) *J. Mol. Biol.* **48,** 443–453.
32. Gonzalez, F. J., Nebert, D. W., Hardwick, J. P. & Kasper, C. B. (1985) *J. Biol. Chem.* **260,** 7435–7441.
33. Aoyama, T., Korzekwa, K., Nagata, K., Gillette, J., Gelboin, H. V. & Gonzalez, F. J. (1990) *Endocrinology* **126,** 3101–3106.
34. Wiens, A. W. & Gilbert, L. I. (1967) *J. Insect Physiol.* **13,** 779–794.
35. McDougall, G. E. & Steele, J. E. (1988) *Insect Biochem.* **18,** 591–597.
36. Imaoka, S., Yamaguchi, Y. & Funae, Y. (1990) *Biochim. Biophys. Acta* **1036,** 18–23.