Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 11891–11894, December 1992
Biochemistry

Identity of 4a-carbinolamine dehydratase, a component of the phenylalanine hydroxylation system, and DCoH, a transregulator of homeodomain proteins

(phenylalanine hydroxylase/4a-hydroxytetrahydropterin dehydratase/tetrahydrobiopterin/hepatocyte nuclear factor 1α/homeodomain regulatory factor)

BRUCE A. CITRON*†, MICHAEL D. DAVIS*, SHELDON MILSTIEN*, JOANNE GUTIERREZ‡, DIRK B. MENDEL§, GERALD R. CRABTREE§, AND SEYMOUR KAUFMAN*

*Laboratory of Neurochemistry and ‡Laboratory of Cell Biology, National Institute of Mental Health, Bethesda, MD 20892; and §Beckman Center for Molecular and Genetic Medicine, Howard Hughes Institute, Stanford University Medical School, Stanford, CA 94305

Contributed by Seymour Kaufman, September 25, 1992

ABSTRACT **The principal pathway for the metabolism of phenylalanine in mammals is via conversion to tyrosine in a tetrahydrobiopterin-dependent hydroxylation reaction occurring predominantly in the liver. Recently, the proposal that certain hyperphenylalaninemic children may have a deficiency of carbinolamine dehydratase, a component of the phenylalanine hydroxylation system, has widened the interest in this area of metabolism. Upon cloning and sequencing the dehydratase, we discovered that this protein is identical to DCoH, the cofactor which regulates the dimerization of hepatic nuclear factor 1α, a homeodomain transcription factor. The identity of the nuclear and cytoplasmic proteins is demonstrated by size, immunoblotting, stimulation of phenylalanine hydroxylase, and dehydratase activity. The evolution of the dual functions of regulation of phenylalanine hydroxylation activity and transcription activation in a single polypeptide is unprecedented.**

During the phenylalanine hydroxylation reaction (1), tetrahydrobiopterin is stoichiometrically oxidized to a carbinolamine, 4a-hydroxytetrahydrobiopterin, which is then converted by 4a-hydroxytetrahydrobiopterin dehydratase to quinonoid dihydrobiopterin; the latter compound is subsequently reduced back to tetrahydrobiopterin by NADH-dependent dihydropteridine reductase (2–4). 7-Substituted pterins have been found in some children with hyperphenylalaninemia, and this may be explained by decreased carbinolamine dehydratase activity (5–7). The observations reported here extend the function of the dehydratase protein to the nucleus, where it plays a role in transcriptional activation.

Homo- and heterodimerization of a variety of transcription factors now appear to be common means of regulatory control allowing many different factors to act through a single DNA binding site (8). The homeodomain-containing transcription factor hepatocyte nuclear factor 1α (HNF-1α) regulates the expression of a large number of genes in the liver, intestine, and kidney (9, 10). The homodimerization of HNF-1α is required in order for this protein to bind to DNA. This dimerization has recently been shown to be controlled by a protein cofactor (DCoH), copurified with nuclear HNF-1α, that has been cloned and characterized (11).¶

MATERIALS AND METHODS

Protein Size Analysis. SDS/PAGE was performed with pre-cast 10% or 16% minigels (Novex, Encinitas, CA). Antibody-positive bands were visualized after transblotting by reaction with alkaline phosphatase-conjugated goat anti-rabbit Fc (Promega) as described (12).

Peptide Sequencing. Dehydratase protein isolated from rat liver (13) and further purified to essential homogeneity by micropreparative high-performance electrophoresis/chromatography (Applied Biosystems model 230A) was digested with trypsin (1%, wt/wt) for 8 hr at 37°C. Peptides were separated by the microbore reverse-phase HPLC (Applied Biosystems model 130A). Peptide sequences were determined with an Applied Biosystems 475A sequencer and a 120A phenylthiohydantoin analyzer.

Synthesis and Subcloning of Dehydratase cDNA. cDNA was synthesized in a 20-μl reaction mixture containing total rat liver RNA (100 μg/ml), 0.5 μM 3′ oligomer, 1 mM each dNTP, 1.5 mM $MgCl_2$, 10 mM Tris·HCl (pH 8.3), 50 mM KCl, gelatin (10 μg/ml), and 200 units of RNase H^- Moloney murine leukemia virus reverse transcriptase (Superscript; BRL) incubated at: 23°C for 10 min, 42°C for 30 min, and 95°C for 5 min. The cDNA was amplified in a 100-μl volume with the same 3′ oligomer (0.5 μM), 5′ oligomer (0.5 μM), 0.2 mM each dNTP, 1.5 mM $MgCl_2$, 10 mM Tris·HCl (pH 8.3), 50 mM KCl, gelatin (10 μg/ml), and 0.5 unit of *Taq* DNA polymerase (Amplitaq; Perkin–Elmer) for 45 cycles of denaturation at 95°C for 30 sec, annealing at 55°C for 1 min, and extension at 72°C for 1 min with primers selected according to back-translations considering the most prevalent mammalian codons while minimizing potential 3′ end mismatches. The final products were ligated to *Sma* I-digested pUC18.

DNA Sequence Analysis. Candidates were sequenced across the entire insert with DNA primers matching the vector sequence ≈50 bases from each end of the insert. Sequenase (modified T7 DNA polymerase, United States Biochemical) and reaction conditions were adjusted according to EG&G Berthold (Natick, MA) for analysis on an Acugen automated DNA sequencer. The University of Wisconsin Genetics Computer Group sequence-analysis software (14) was used to analyze DNA sequences.

Construction and Expression of DCoH Fusion Protein. The glutathione *S*-transferase (GST)–DCoH fusion protein was prepared by making an in-frame fusion of the coding sequence of DCoH to the sequence encoding GST protein, using the pGEX-2T expression vector (Pharmacia). The construct was used to transform MC1061 by electroporation, and individual transformants were screened by culturing in the presence of 1 mM isopropyl β-D-thiogalactopyranoside for 4 hr. A clone producing high levels of the fusion protein

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.

Abbreviations: HNF, hepatocyte nuclear factor; GST, glutathione *S*-transferase.

†To whom reprint requests should be addressed.

¶The sequence reported in this paper has been deposited in the GenBank data base (accession no. L04537).

was selected for large-scale culture. The soluble fraction from a bacterial lysate was placed over a glutathione-agarose column, and the retained protein was eluted with 50 mM Tris·HCl, pH 7.5/5 mM glutathione. Yields of the protein were about 15 mg per liter of bacterial culture. The GST–DCoH fusion protein was digested with thrombin, and the transferase was removed by an additional pass over the same column. This left a protein that was estimated to be about 98% pure with the only additional visible band on Coomassie blue-stained polyacrylamide gels being the thrombin used to cleave the fusion protein (data not shown).

Phenylalanine Hydroxylase Stimulation Assay. The reaction mixture was incubated at 25°C and contained 30 mM potassium phosphate (pH 8.3), catalase (1 mg/ml), 100 μM NADH, 1 mM phenylalanine, excess dihydropteridine reductase, 12 μg of phenylalanine hydroxylase, and 2.9 μM (6*R*)-tetrahydrobiopterin. Approximately 1 min after the reaction was started, either nothing or 0.75 μg of pure rat liver dehydratase, or 0.5 μg of cloned DCoH, or 2.0 μg of the fusion protein was added as indicated. The reaction was monitored at 340 nm for the oxidation of NADH by dihydropteridine reductase as quinonoid dihydrobiopterin was recycled to tetrahydrobiopterin (2, 15).

4a-Carbinolamine Dehydratase Assay. The oxidation of tetrahydrobiopterin by phenylalanine hydroxylase was monitored at 245 nm for 4a-carbinolamine formation (3, 16). The reaction mixture contained 30 mM potassium phosphate (pH 8.3), catalase (100 μg/ml), 10 mM glucose 6-phosphate, 1 μM NADH, 12 μg of phenylalanine hydroxylase, 1 mM phenylalanine, and an excess of dihydropteridine reductase and glucose-6-phosphate dehydrogenase. The reaction was started with the addition of 28 μM (6*R*)-tetrahydrobiopterin following a 5-min preincubation of the mixture. At the indicated time, DCoH (3 μg/ml) was added to the reaction mixture.

RESULTS AND DISCUSSION

To obtain cDNA clones that would facilitate the characterization of the dehydratase, we isolated and sequenced tryptic peptides prepared from pure rat liver enzyme (Fig. 1). The four isolated peptides were short and contained multiple ambiguous codons, rendering them unsuitable for the synthesis of a hybridization probe. We developed a PCR approach to surmount this problem, basing the design of six PCR primers on the three (of four initial) peptides which could be translated from the least ambiguous codon set. The putative dehydratase-specific PCR-generated fragment was ligated to a plasmid vector, and two inserts containing the primers straddling an identical 76-base-pair segment were sequenced. The correspondence between the DNA and peptide sequence (Fig. 2) confirmed the identification of our subclone as dehydratase cDNA. The new DNA sequence

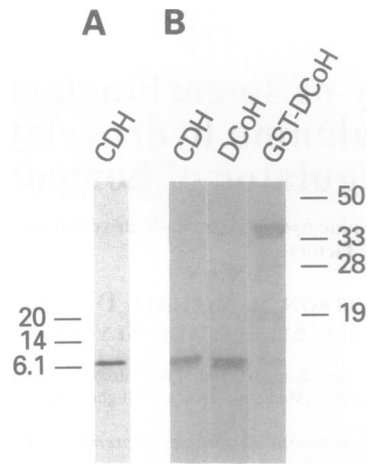

FIG. 1. Gel electrophoresis and immunoblot analysis of 4a-tetrahydrobiopterin dehydratase (CDH), its dimerization cofactor (DCoH), and GST–DCoH fusion protein. (*A*) Purified dehydratase (3 μg) stained with Coomassie blue R-250. (*B*) Immunoblots (12) of 7 μg of dehydratase or DCoH and 20 μg of GST–DCoH fusion protein. Dashes indicate the positions of molecular mass (kDa) markers. The molecular mass of the dehydratase/DCoH is 12 kDa, and that of the fusion protein is 38 kDa.

was then compared (17) with the sequence data bases; we found that GenBank (Release 71.0) contained three significant matches: rat, human, and mouse HomeoA loci. Translation of the reported coding sequence for the dimerization cofactor (DCoH) for HNF-1α indicated identity to our DNA insert and to all 46 residues coded by our sequence and present in the tryptic peptides. Additionally, the molecular mass of DCoH (12 kDa) and its amino acid composition are essentially identical to those previously reported for the dehydratase (13). Furthermore, the tissue distribution of both dehydratase activity and DCoH are very similar, with activity expressed mainly in liver and kidney, and with much lower amounts in other tissues (11, 12).

To determine whether the dehydratase activity and HNF-1α homodimerization activity are indeed present within the same polypeptide, the dehydratase activity of purified, recombinant DCoH protein (11) was examined. Dehydratase activity was determined indirectly by measurement of the stimulation of phenylalanine hydroxylase activity under conditions where the breakdown of the 4a-carbinolamine limits the rate of hydroxylation (2, 3, 15). Recombinant DCoH and the GST–DCoH fusion protein had essentially the same dehydratase specific activity as pure rat liver dehydratase (Fig. 3*A*). A more direct assay for the dehydratase activity entails monitoring the loss of the characteristic 245-nm absorbance of the 4a-carbinolamine that is formed during the tetrahydrobiopterin-dependent hydroxylation of phenylala-

FIG. 2. Sequence comparison. The DNA and protein sequence for DCoH (Genbank, RATHOMEOA) matches our insert DNA (---) and our translated and determined (boxed) peptide sequences. A single discrepancy at position 66 as depicted, a cytosine (C) in our sequence (clones pMDD1 and 2) vs. a thymine (T) in DCoH, is silent and most likely a strain difference.

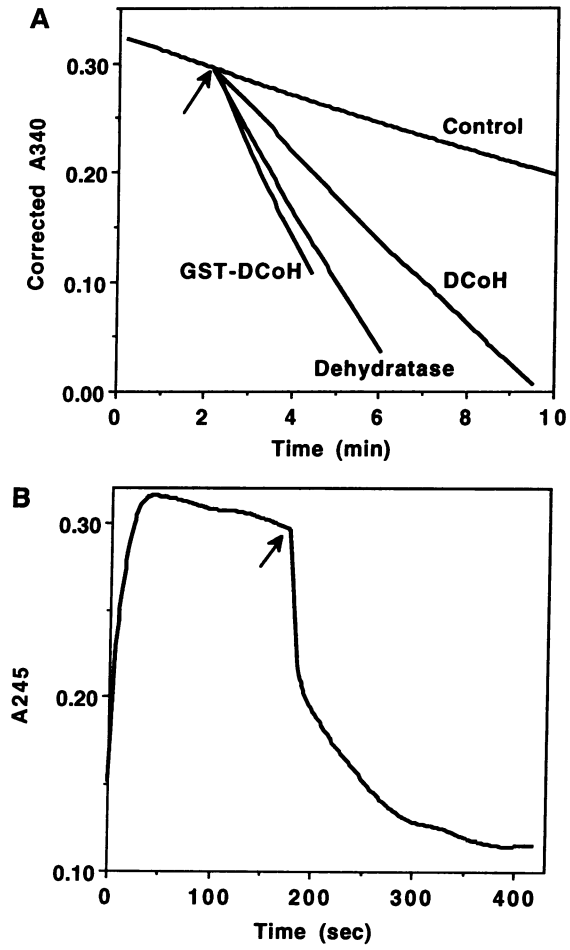

FIG. 3. 4a-Carbinolamine dehydratase activity. (*A*) Stimulation of phenylalanine hydroxylase catalyzed hydroxylation of phenylalanine with the addition (at arrow) of rat liver dehydratase, cloned DCoH, or the GST–DCoH fusion protein. Neither the addition of glutathione alkyltransferase nor glutathione alkyltransferase fusion protein with FK506-binding protein 12 stimulated this activity (data not shown). (*B*) Effect of the DCoH (added at arrow) on the UV absorbance changes during oxidation of tetrahydrobiopterin by phenylalanine hydroxylase.

nine by phenylalanine hydroxylase (3) (Fig. 3*B*). At the start of the reaction, there is a rapid increase in absorbance followed by a very slow decline. The increase is due to the enzymatic conversion of tetrahydrobiopterin to the 4a-carbinolamine, which rapidly reaches its steady-state concentration as the pterin is continuously recycled during the reaction (2, 3, 16). Upon addition of dehydratase (arrow, Fig. 3*B*), the concentration of the 4a-carbinolamine decreased as indicated by the disappearance of the characteristic absorbance at 245 nm. The addition of cloned DCoH resulted in a rapid decline in 245-nm absorbance (Fig. 3*B*), indicating not only that DCoH possesses the same enzymatic activity as carbinolamine dehydratase but that the specific activities of the two proteins are similar. Dehydratase activity was also found in whole cell extracts of Chinese hamster ovary cells transfected with DCoH or DCoH plus HNF-1α but was absent in transfectants containing only the parent vector or HNF-1α alone (data not shown).

Although multifunctional proteins are uncommon, other examples exist, such as the human glyceraldehyde-3-phosphate dehydrogenase subunit that is also a DNA glycosylase (18). It is tempting to speculate that the two biological roles of dehydratase/DCoH may be related. Cytoplasmic dehydratase activity is necessary to ensure that the breakdown of the 4a-carbinolamine to the corresponding quinonoid dihydropterin does not limit the rate of hepatic phenylalanine hydroxylation. The nuclear dimerization activity is required to stimulate HNF-1α-dependent transcription in the liver. Hyperphenylalaninemia has been associated with deficiencies in phenylalanine hydroxylase, dihydropteridine reductase, and *de novo* synthesis of tetrahydrobiopterin (19). Some cases of mild hyperphenylalaninemia may be due to dehydratase deficiency (5–7). The evidence in support of this conclusion includes markedly increased excretion of 7-biopterin, which is derived from tetrahydrobiopterin, and the *in vitro* inhibition by dehydratase of the conversion of tetrahydrobiopterin to 7-biopterin. If these cases represent genetic variants of carbinolamine dehydratase, then they may also be unique examples of a naturally occurring genetic defect in a transcriptional control factor. Decreased expression of phenylalanine hydroxylase and/or dihydropteridine reductase resulting from deficiency of DCoH could be a factor in the hyperphenylalaninemia in these patients. Indeed, a deficiency of hepatic dihydropteridine reductase in these patients could account for the unexplained finding that, unlike the situation with normal individuals, most of the biopterin found in several of patients' urine was not present as the tetrahydro derivative (20).

Although DCoH has been shown to interact with the dimerization domain of HNF-1α and -1β and to stabilize homodimers (11), there is as yet no clear explanation for the enhancement of transcriptional activity. The simple possibility that DCoH may stabilize the DNA-binding dimeric form of HNF-1α, giving rise to greater DNA binding and hence greater transcriptional activity, seems incompatible with present experimental evidence (D.B.M., W. D. Wang, and G.R.C., unpublished data). Studies described above suggest that carbinolamine dehydratase is also somehow involved in enhancing transcription. Another example of a transcriptional activity that involves a bifunctional enzyme is the biotin operator repressor, which both represses the biotin biosynthetic operon and also activates biotin in *Escherichia coli* (21). Additional studies will be necessary to determine whether dehydratase activity is essential for transcriptional enhancement by DCoH.

The identification of carbinolamine dehydratase/DCoH increases the number of sequences that can be examined for a common pterin-binding motif. Simultaneous sequence analysis identified two possible sites, an acidic region at positions 8–25 and an additional region at 82–89 (Fig. 4). The N-ter-

```
HumDHPR      73 KLLG.EEKVDAILC.V.AGGW  90
                :| : ||: |.:|. : |.||
CDH/DCoH      7 RLSA.EER.DQLLPNLRAVGW  25
                :.:: |  |. :|.| .|:
RatPAH      215 KYCGFRE..DN.IPQLEDVSQ 232


RatSR       161 ALQ.PF...KG.WGLYCAGKAARD 179
                :: .:   ..  :  ||| ..||
CDH/DCoH     72 KVHITL...ST.HE..CAGLSERD  89
                |  ||   :| |   | | |
Rat NOS     319 K..STLETGCTEHI..CMG.SIML 337
```

FIG. 4. Sequence similarity to other proteins that interact with pterins. Carbinolamine dehydratase (CDH)/DCoH was compared with other enzymes which are involved in tetrahydrobiopterin synthesis or utilization. These included rat GTP cyclohydrolase (22), rat sepiapterin reductase (SR) (23), human dihydropteridine reductase (DHPR) (24, 25), rat phenylalanine hydroxylase (PAH) (26), two putative sites for pterin binding (27, 28), and rat nitric oxide synthase (NOS) (29). Similarities were considered only if five or more target sequences matched a specific segment of CDH/DCoH. Only two regions met this requirement, and the two strongest matches of each of these subsequences are depicted.

minal site displays a stronger similarity to other pterin-binding proteins. Directed mutagenesis of dehydratase/DCoH will help identify the structure–function relationships in this compact protein. Sequence analysis of DNA from putative dehydratase-deficient patients will be necessary to confirm the association between these metabolic disorders. This is an important goal, since the transcription regulatory role of carbinolamine dehydratase/DCoH may also involve the tissue-specific control of many other liver-specific functions.

We thank C. Falke, K. Kenyon, and K. Park for expert technical assistance; M. J. Brownstein for oligomer synthesis; and K. Weeks for maintaining DNA sequence data bases.

1. Kaufman, S. & Fisher, D. B. (1974) in *Molecular Mechanisms of Oxygen Activation*, ed. Hayaishi, O. (Academic, New York), pp. 285–369.
2. Huang, C. Y. & Kaufman, S. (1973) *J. Biol. Chem.* **248**, 4242–4251.
3. Kaufman, S. (1976) in *Iron and Copper Proteins: Advances in Experimental Medicine and Biology*, eds. Yasunobo, K. T., Mower, H. F. & Hayaishi, O. (Plenum, New York), Vol. 74, pp. 91–102.
4. Lazarus, R. A., Benkovic, S. J. & Kaufman, S. (1983) *J. Biol. Chem.* **258**, 10960–10962.
5. Curtius, H. C., Kuster, T., Matasovic, A., Blau, N. & Dhondt, J. L. (1988) *Biochem. Biophys. Res. Commun.* **153**, 715–721.
6. Davis, M. D., Kaufman, S. & Milstien, S. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 385–389.
7. Curtius, H. C., Adler, C., Rebrin, I., Heinzmann, C. & Ghisla, S. (1990) *Biochem. Biophys. Res. Commun.* **172**, 1060–1066.
8. Mendel, D. B. & Crabtree, G. R. (1991) *J. Biol. Chem.* **266**, 677–680.
9. Courtois, G., Morgan, J. G., Campbell, L. A., Fourel, G. & Crabtree, G. R. (1987) *Science* **238**, 688–692.
10. Nicosia, A., Monaci, P., Tomei, L., De, F. R., Nuzzo, M., Stunnenberg, H. & Cortese, R. (1990) *Cell* **61**, 1225–1236.
11. Mendel, D. B., Khavari, P. A., Conley, P. B., Graves, M. K., Hansen, L. P., Admon, A. & Crabtree, G. R. (1991) *Science* **254**, 1762–1767.
12. Davis, M. D., Kaufman, S. & Milstien, S. (1992) *FEBS Lett.* **302**, 73–76.
13. Huang, C. Y., Max, E. E. & Kaufman, S. (1973) *J. Biol. Chem.* **248**, 4235–4241.
14. Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
15. Kaufman, S. (1970) *J. Biol. Chem.* **245**, 4751–4759.
16. Davis, M. D. & Kaufman, S. (1989) *J. Biol. Chem.* **264**, 8585–8596.
17. Pearson, W. R. & Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448.
18. Meyer-Siegler, K., Mauro, D. J., Seal, G., Wurzer, J., deReil, J. K. & Sirover, M. A. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 8460–8464.
19. Scriver, C. R., Kaufman, S. & Woo, S. L. C. (1989) in *The Metabolic Basis of Inherited Disease*, eds. Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D. (McGraw Hill, New York), pp. 495–546.
20. Curtius, H. C., Matasovic, A., Schoedon, G., Kuster, T., Guibaud, P., Guidici, T. & Blau, N. (1990) *J. Biol. Chem.* **265**, 3923–3930.
21. Barker, D. F. & Campbell, A. M. (1981) *J. Mol. Biol.* **146**, 469–492.
22. Hatakeyama, K., Inoue, Y., Harada, T. & Kagamiyama, H. (1991) *J. Biol. Chem.* **266**, 765–769.
23. Citron, B. A., Milstien, S., Gutierrez, J. C., Levine, R. A., Yanak, B. L. & Kaufman, S. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6436–6440.
24. Lockyer, J., Cook, R. G., Milstien, S., Kaufman, S., Woo, S. L. & Ledley, F. D. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 3329–3333.
25. Dahl, H.-H. M., Hutchison, W., McAdam, W., Wake, S., Morgan, F. J. & Cotton, R. G. H. (1987) *Nucleic Acids Res.* **15**, 1921–1932.
26. Dahl, H.-H. M. & Mercer, J. F. B. (1986) *J. Biol. Chem.* **261**, 4148–4153.
27. Jennings, I. & Cotton, R. (1990) *J. Biol. Chem.* **265**, 1885–1889.
28. Gibbs, B. S. & Benkovic, S. J. (1991) *Biochemistry* **30**, 6795–6802.
29. Bredt, D. S., Hwang, P. M., Glatt, C. E., Lowenstein, C., Reed, R. R. & Snyder, S. H. (1991) *Nature (London)* **351**, 714–718.