Biochemical and Structural Basis for Partially Redundant Enzymatic and Transcriptional Functions of DCoH and DCoH2†‡

Robert B. Rose,*,§ Kristi E. Pullen,§ J. Henri Bayle,‖ Gerald R. Crabtree,‖ and Tom Alber§
Department of Molecular and Cell Biology, University of California, Berkeley, California 94720-3206, and Howard Hughes Medical Institute, Beckman Center for Molecular and Genetic Medicine, Stanford University, Stanford, California 94305

Received February 23, 2004; Revised Manuscript Received April 12, 2004

ABSTRACT: An inherited form of diabetes, maturity-onset diabetes of the young type 3 (MODY3), results from mutations in the transcriptional activator, hepatocyte nuclear factor-1α (HNF1α). Transcription by HNF1α is stimulated by the bifunctional coactivator DCoH (dimerization cofactor of HNF1). Strikingly, an HNF1α deletion in mice causes more severe phenotypes than a DCoH deletion. It has been hypothesized that a DCoH homolog, DCoH2, partially complements the DCoH deletion. To test this idea, we determined the biochemical properties and the 1.6-Å-resolution crystal structure of DCoH2. Like DCoH, DCoH2 forms a tetramer, displays pterin-4α-carbinolamine dehydratase activity, and binds HNF1α in vivo and in vitro. DCoH and DCoH2 adopt identical folds with structural differences confined largely to the protein surfaces and the tetramer interface. In contrast to the hyperstable DCoH tetramer, DCoH2 readily disproportionates and forms a 2:2 complex with HNF1 in vitro. Phylogenetic analysis reveals six major subfamilies of DCoH proteins, including unique DCoH and DCoH2 branches in metazoans. These results suggest distinct roles for DCoH and DCoH2. Differences in conserved surface residues could mediate binding to different effectors. We propose that HNF1α binding kinetics may distinguish regulation by DCoH2, under thermodynamic control, from regulation by DCoH, under kinetic control.

DCoH1 is a small, conserved protein that functions both as a transcriptional co-activator and a metabolic enzyme. The pterin-4α-carbinolamine dehydratase (PCD) activity of DCoH helps to recycle tetrahydrobipterin, a cofactor for the amino acid hydroxylases and nitric oxide synthase (helps to recycle tetrahydrobiopterin, a cofactor for the amino pterin-4 as a transcriptional co-activator and a metabolic enzyme. The 5 HNF1 mutations disrupt the interaction with DCoH, as well as diabetes of the young type 3 (MODY3) (RHNF1HNF1R associates with the transcriptional activators HNF1β and stimulates HNF1 activity (6). Mutations in HNF1α cause a monogenic form of diabetes, maturity onset diabetes of the young type 3 (MODY3) (7). Two MODY3 mutations disrupt the interaction with DCoH, as well as HNF1α dimerization (8). Recently, DCoH has been reported to be overexpressed in colon cancer carcinomas (9) and malignant melanomas (10).

Structural studies indicated that the mammalian DCoH exists in two oligomerization states in its nuclear and cytoplasmic forms (1). Cytoplasmic DCoH forms a tetramer consisting of two saddle-shaped dimers (11). The tetramer is enzymatically active, with an independent active site in each monomer (12). In the nucleus, DCoH forms a 2:2 complex with HNF1 proteins (6). The structure of the DCoH dimer bound to the HNF1α dimerization domain demonstrated that DCoH tetramerization and HNF1 binding are mediated by the same surface. This shared site implies that HNF1α competes with DCoH homotetramerization to interact with DCoH (13). The DCoH tetramer is hyperstable, and this stability explains the required cofolding of DCoH and HNF1α to form the 2:2 complex (14, 15).

Despite the concerted functions of the HNF1α/DCoH complex, deletions of these genes in mice produce different phenotypes. HNF1α null mice display hyperphenylalaninemia, Laron-type dwarfism, and early onset diabetes mellitus (16–18). In contrast, DCoH null mice exhibit hyperphenylalaninemia, a predisposition to cataract formation, and only mild disruption of HNF1α function (19). This relatively mild syndrome led to a search of the mouse genome sequence for DCoH homologues that might partially complement the loss of the DCoH protein. This search identified the sequence of DCoH2, which shows 68% sequence identity to DCoH (also referred to as DCoH1 in this paper). DCoH2 orthologs also have been reported in chickens (20) and humans (21).

The sequence similarities of DCoH1 and DCoH2 suggested that they share similar biochemical properties. Here
we test this idea by comparing the structures and activities of DCoH1 and DCoH2. Phylogenetic analysis of the DCoH family shows that these proteins represent the only two subfamilies of DCoH proteins in vertebrates. We show that DCoH2 mediates the same enzymatic and coactivator activities as DCoH1 but with different surface residues and different requirements for binding HNF1α in vitro. We propose that these differences may contribute to distinct mechanisms of regulation of HNF1 function.

**EXPERIMENTAL PROCEDURES**

**DCoH2 Purification.** Mouse DCoH2 cDNA was subcloned into a pGEX-2T vector (Pharmacia) and expressed in *Escherichia coli* BL21(DE3) cells (Stratagene). Cells were grown at 37 °C, harvested, and sonicated in phosphate-buffered saline (PBS) at pH 7.0. After centrifugation, the supernatant was loaded onto a glutathione affinity column (Pharmacia) preequilibrated in PBS. After extensive washing with PBS, DCoH2 was eluted by adding thiomarin directly to the column and eluting according to the manufacturer's instructions. DCoH2 was further purified using an FPLC 15Q cation-exchange column (Pharmacia) preequilibrated in PBS. The protein was eluted with a 0–1 M NaCl gradient. A colloidion concentrator (Schleicher & Schuell) was used to concentrate and buffer-exchange the protein into PBS.

**Analytical Ultracentrifugation.** Protein samples (50–70 μM in PBS) were centrifuged in a Beckman X-LA centrifuge at 15 000 rpm at 25 °C for 40 h to reach equilibrium. Absorbance data were fit to a model containing a single species. The observed small, random residuals validated this assumption.

**Bacterial Complementation Assay.** *E. coli* strain JP2255 thiA, aroC, pheA, tyrA was transformed with pJMZ3a (expressing *Pseudomonas aeruginosa* PhhA [phenylalanine hydroxylase] and PhhB (PCD/DCoH) in cis from the phh operon), pJMZ4 (PhhB alone), or pJMZ5 (PhhB alone)). pMB1 was derived from pACYC184 (New England Biolabs) expressing the PhhA gene from pJMZ5 (22). This plasmid has non-ColE1 origin of replication and imparts chloramphenicol resistance. DCoH1 and DCoH2 were inserted into the Bluescript cloning vector downstream of a Lac promoter (pBS–DCoH). A Shine–Dalgarno sequence was added to increase translation efficiency. JP2255/pMB1 was transfected with pJMZ4 (phhB alone) or with pBS–DCoH1, pBS–DCoH2, or pBS–DCoH1 mutated at the active site for PCD into the Bluescript cloning vector downstream of a Lac promoter (pBS–DCoH). A Shine–Dalgarno sequence was added to increase translation efficiency. JP2255/pMB1 was transfected with pJMZ4 (phhB alone) or with pBS–DCoH1, pBS–DCoH2, or pBS–DCoH1 mutated at the active site for PCD activity (His63Leu (15)). Transformants were selected for tyrosine prototrophy on minimal media agar plates containing M9 salts, glucose, thiamine, and phenylalanine.

**Electrophoretic Mobility Shift Assays.** Nuclear extracts from transfected CHO cells were prepared by the method of Lassar (23). Gel shift assays were performed as described using a labeled oligonucleotide derived from the human albumin promoter (19). Truncated HNF1α (residues 1–280) was subcloned into the pQE-60 expression vector (Qiagen) in-frame with the C-terminal 6-His tag and purified according to the manufacturer's instructions (HNF1αtr-His). For mixed oligomer experiments, 10 ng of HNF1αtr-His was mixed with the nuclear extract on ice for 30 min prior to the addition of the oligonucleotide.

**Immunoprecipitation of DCoH2 Complexes.** CHO cells (lacking expression of either HNF1 or DCoH) were grown on plates in RPMI 1640 media supplemented with 10% fetal calf serum and antibiotics. All proteins were derived from the pBJ5 plasmid in which expression is directed by the SRα promoter. FLAG-tagged DCoH1 and DCoH2 with a 5′-tag (FLAG–DCoH) were generated as previously described (19). FKBP–DCoH1 is a fusion between the 12 kDa human FK506-binding protein (FKBP) cDNA at the 5′ end of the DCoH cDNA yielding an amino-terminal fusion lacking a FLAG tag. Mouse HNF1α was fused to a C-terminal nine amino acid hemagglutinin (HA) tag (HNF1α–HA). HNF1αΔ-dim-HA lacks the amino terminal 32 amino acids of the HNF1α dimerization domain and retains the HA tag (6). Electroporations were performed as described (19). Forty-eight hours posttransfection, cells were lysed on ice in TNEN (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40, 1 mM PMSF, 1 mM leupeptin, 1 mM pepstatin). Samples (30–100 μg protein) were incubated overnight at 4 °C with 2 μg of M2 antibody (Kodak), IgG1 isotype control (Becton Dickinson), or 12CA5 (0.1 mL (net) of ascites fluid) and 5–10 mL of 50% slurry of protein-G Sepharose. Western blots were performed as described (19) except 2 μg/mL M2 antibody was used to develop the anti-FLAG Western blot.

**Affinity Chromatography Assay.** The N-terminal 280 residues of mouse HNF1α were subcloned into pET24b (Novagen) with a His6 tag added at the C-terminus (HNF1αtr-His). This construct retains the DNA-binding elements of HNF1α, including the dimerization and DNA-binding domains (24). The protein was expressed in *E. coli* BL21 cells and purified from cell lysate in PBS using a Ni2+-chelating resin (Pharmacia). The fractions containing HNF1α were loaded onto a 15Q ion-exchange resin (Pharmacia) preequilibrated in PBS and eluted with a 0.15–1 M NaCl gradient. Further purification was achieved using gel filtration on a S300 column (Pharmacia) equilibrated with PBS. Interaction between DCoH2 and HNF1α was measured with His-tagged HNF1α immobilized on Ni2+-chelating resin. Beads (30 μL) were equilibrated in PBS and loaded with 100 μg of purified His6-tagged HNF1α. The beads were washed extensively in PBS containing 30 mM imidazole. Purified DCoH1 or DCoH2 (100 μg) was added to the beads and incubated for 45 min. The beads were washed three times with 1 mL of PBS plus 30 mM imidazole and mixed with 30 μL of 2× SDS loading buffer. This reaction (20 μL) was separated using SDS PAGE and visualized with Coomassie Blue (Sigma).

**DCoH2 Structure Determination.** DCoH2 (~10 mg/mL) was crystallized in 10% 2-propanol, 0.1 M Na Hepes pH 7.5, and 20% PEG 4000 by hanging-drop vapor diffusion at 4 °C. The crystals were frozen in liquid nitrogen in mother liquor containing 10% PEG 400 as a cryoprotectant.

Data were collected at the Advanced Light Source and indexed with MOSFLM (25) using the EMLVES interface (Table 1) (26). The data were scaled with Scala, and F′s were calculated with Truncate (27–29). The structure was solved by molecular replacement using a search model containing a DCoH dimer with residues differing between DCoH1 and DCoH2 replaced by alanines. After the rotation and translation search at 8–4 Å using CNS (30), the initial correlation coefficient was 0.62. After rigid body refinement, the R-factor was 44% for all data from 30 to 2 Å resolution. Starting with these phases, the model was rebuilt using ARP/
Final rounds of refinement were accomplished using CNS alternated with manual model building using the program O (33). The model included 198 amino acids and 163 water molecules. Residues 1–3 of each monomer and residues 102 and 103 of one monomer were disordered and omitted from the final model. CNS was used to calculate buried surface area.

Phylogenetic Analysis. DCoH homologues were identified using PSI-BLAST (34) with the human DCoH2 sequence as the query. Five iterations were performed until no new sequences with e-values < 0.001 were identified. Forty-two distinct sequences from the nonredundant database at the National Center for Biotechnology (NCBI) were retrieved. Fewer orthologs were identified using the human DCoH1 sequence as the query. A phylogenetic tree was constructed using the PHYLIP Phylogeny Inference Package (35) implemented at the San Diego Super Computer Center (http://workbench.sdsc.edu). The tree was manually manipulated to eliminate branch crossings while preserving branch length and connectivity.

DCoH homologues in the human genome were identified using tblastn at the website of the NCBI. Searches were conducted using the human DCoH1 and DCoH2 sequences. Exon boundaries were identified using Genscan (36).

RESULTS

Alignment of the amino acid sequences of DCoH proteins from human, mouse, rat, and chicken revealed that the sequences could be clustered into two classes that define DCoH paralogs (Figure 1). The sequence of human DCoH2 is 61% identical to human DCoH1 and 88% identical to
mouse DCoH2. Rather than reflecting random drift of a duplicated gene, the evolutionary conservation of DCoH2 implies that DCoH2 and DCoH1 may perform distinct functions.

Mouse DCoH2 was expressed and purified to compare its properties with those of DCoH1. Analytical ultracentrifugation demonstrated that both DCoH1 and DCoH2 form homotetramers in solution (Figure 2A,B). DCoH1 and DCoH2 share extensive similarity in helix 2, the recognition helix (HNF1) binding and DCoH homotetramerization. To address whether the two DCoH proteins can form hetero-oligomers, DCoH2 was fused to an N-terminal FLAG tag and coexpressed in CHO cells with a DCoH1 N-terminal fusion to FKBP. Immunoprecipitation of DCoH2 co-immunoprecipitated FKBP-DCoH1 indicating that mixed oligomers formed in vivo (Figure 2C).

*P. aeruginosa* encodes a phenylalanine hydroxylase, PhhA, and a DCoH homolog, PhhB (22). The PCD activity of PhhB is essential for growth on minimal medium (38). Expression of *phha* and *phhb* can complement *E. coli* deficient in tyrosine biosynthesis, but deletion of *phhb* makes the strain auxotrophic for tyrosine. The loss of *phhb* was complemented by DCoH1 or DCoH2 but not the enzymatically inactive, His63Leu DCoH1 mutant (Figure 3A). This complementation implies that DCoH1 and DCoH2 exhibit similar PCD activities. This enzymatic activity is consistent with the nearly complete conservation of residues in the enzyme active site.

The mild reduction of HNF1α activity in DCoH1 null mice and the biochemical observation of a complementing activity (19) strongly implied that DCoH2 could bind and stabilize HNF1α dimers. We used several assays to test the ability of DCoH2 to bind to HNF1α in vivo and in vitro. Both DCoH proteins bound to HNF1α when the proteins were coexpressed in vivo (Figure 3B,C). Gel mobility shift assays of HNF1α and FLAG-tagged DCoH1 in nuclear extracts of transfected CHO cells revealed an HNF1α complex that could be supershifted with a mouse anti-DCoH1 polyclonal antibody or anti-FLAG tag antibodies (Figure 3B). HNF1α/DNA complexes containing FLAG-tagged DCoH2 were only supershifted with anti-FLAG antibodies, because the mouse polyclonal antibody is specific for DCoH1 (19). DCoH1 and DCoH2 binding to HNF1α required the N-terminal dimerization domain of HNF1α (Figure 3C) (6). DCoH2 also bound HNF1α in vitro. In a gel mobility shift assay, a mixture of purified DCoH2 and HNF1α supershifted the HNF1α/DNA complex (Figure 3D). In contrast, purified DCoH1 did not efficiently supershift the HNF1α/DNA complex, consistent with the idea that the high stability of the DCoH1 tetramer inhibits complex formation with HNF1α (8). Similarly a glutathione-S-transferase (GST) pull-down assay demonstrated that purified DCoH2 protein but not DCoH1 bound efficiently to a column loaded with a large fragment (1–280) of HNF1α encompassing the dimerization, linker, and DNA-binding domains (Figure 3E). These results are consistent with previous work showing that DCoH1 does not readily interact with HNF1α upon mixing (14, 15) but folds together when the two proteins are coexpressed (6) or when DCoH1 and the HNF1α dimerization domain are renatured together from guanidinium hydrochloride (8).

DCoH proteins are thought to stimulate transcription by stabilizing HNF1α dimers (6). We tested the ability of DCoH2 to stabilize HNF1α dimers by measuring the exchange of monomers in preformed HNF1α dimers (Figure 3F). Full-length HNF1α was expressed in the presence or absence of DCoH1 or DCoH2. HNF1α dimers were challenged by mixing cell extracts with truncated HNF1α (HNF1αtr), which forms complexes that migrate more rapidly in gel-shift experiments. In the absence of DCoH proteins, HNF1α dimers exchanged readily with the challenging HNF1αtr, resulting in a band of intermediate...
**FIGURE 3:** DCoH2 is a pterin-4-carbinolamine dehydratase and interacts with HNF1α in vitro and in vivo. As seen in panel A, DCoH2 exhibits pterin-4-carbinolamine dehydratase (PCD) activity. DCoH1 and DCoH2 complement a mutation in the E. coli PCD ortholog, PhhB. Complementation was assayed in a tyrosine auxotroph of E. coli strain JP2255 with the phh operon of P. aeruginosa. PhhA encodes phenylalanine hydroxylase. JP2255 includes the following genotype: thiA, aroC, phe A, tyr A. Growth was assayed on minimal media supplemented with phenylalanine and thiamine. The His63Leu mutation inactivates the dehydratase activity of DCoH1 (15). In panel B, gel mobility shift assays demonstrate that DCoH2 interacts in vivo with HNF1α. Assays were performed with an HNF1α binding site oligonucleotide and nuclear extracts prepared from CHO cells cotransfected with HNF1α and either DCoH1 or FLAG-DCoH2 as indicated. Antibodies (e.g., αHNF1) were included as indicated to supershift the DNA–protein complexes. Note that the HNF1α/DCoH2 complex was supershifted with the M2 antibody (directed to the FLAG tag on DCoH2, lane 8) but not a murine polyclonal antibody to DCoH1 that appears to lack specificity for DCoH2 (lane 7). The HNF1α/DCoH1 complex was supershifted with the antibody to DCoH1 (lane 5). Panel C shows that DCoH2 binds to the dimerization domain of HNF1α. Lysates were prepared from CHO cells cotransfected with HA-tagged HNF1α or HNF1αΔ1–32 (HNF1αΔdiam) and FLAG-tagged DCoH1 or DCoH2 as indicated. Immunoprecipitations were performed with M2 antibody (directed to the FLAG tag on DCoH, lanes 1, 3, and 5) or 12CA5 (directed to the HA tag on HNF1α, lanes 2, 4, 6, and 7) or an unrelated rabbit antibody (CON, lane 8). DCoH proteins in the immunoprecipitates were analyzed by Western blot with the M2 (αFLAG) antibody. DCoH2 (lane 2) and DCoH1 (lane 6) coimmunoprecipitate with full-length HNF1α but not with HNF1α with the dimerization domain deleted (lanes 4 and 7). As shown in panel D, DCoH2, but not DCoH1, binds the HNF1α/DNA complex in vitro. The size of the HNF1α/DNA complex in the presence and absence of DCoH1 and DCoH2 was measured with an electrophoretic mobility shift assay. DCoH2 (lane 2), but not DCoH1 (lane 3), retarded the HNF1α/DNA complex. DNA binding assays were performed with 10 ng of purified recombinant HNF1α (HNF1αtr, truncated at amino acid 280). Purified recombinant DCoH (50 ng) was included in the indicated reactions. As shown in panel E, DCoH2 binds reversibly to HNF1α. Nickel sepharose beads were loaded with His6-tagged HNF1α (residues 1–280), and either purified DCoH2 or DCoH1 was added. After being washed in low or high imidazole buffers, the beads were loaded onto an SDS–polyacrylamide gel: (lane 1) molecular weight standards; (lane 2) purified DCoH1; (lane 3) DCoH1 loaded on nickel sepharose beads lacking HNF1α; (lane 4) nickel sepharose beads loaded with HNF1α; (lane 5) DCoH1 loaded on beads preloaded with HNF1α; (lane 6) beads loaded with HNF1α after 250 mM imidazole wash; (lane 7) purified DCoH2; (lane 8) DCoH2 loaded on nickel sepharose beads lacking HNF1α; (lane 9) DCoH2 loaded on nickel sepharose beads lacking HNF1α; (lane 10) nickel sepharose beads loaded with HNF1α; (lane 11) DCoH2 loaded on beads preloaded with HNF1α; (lane 12) beads loaded with HNF1α after 250 mM imidazole wash; (lane 14) purified HNF1α. Abbreviations in the “wash” row are as follows: n, no wash; l, low (30 mM) imidazole wash; h, high (250 mM) imidazole wash. DCoH2 binds efficiently to HNF1α (lane 11). In contrast, purified DCoH1 fails to bind HNF1α in vitro in the absence of denaturants (Lane 5 and refs 14 and 15). Panel F shows the stabilization of HNF1α dimerization by DCoH2 in vitro. HNF1α dimers formed in CHO cells were challenged by addition of purified, truncated HNF1α (HNF1αtr) and assayed by EMSA. In the absence of DCoH1 or DCoH2, HNF1αtr readily formed a mixed oligomer with HNF1α that displayed intermediate mobility (lane 2). Cotransfection of DCoH1 or DCoH2 inhibited the formation of mixed oligomers in the nuclear extract challenged with HNF1αtr.
mobility. In contrast, coexpression of either DCoH protein stabilized preformed HNF1α dimers. Thus, DCoH2 inhibited the exchange of subunits in HNF1α dimers in vitro.

Comparison of DCoH2 and DCoH1 Structures. We determined the high-resolution crystal structure of DCoH2 to explore the basis for its biochemical properties. DCoH2 crystallized with a dimer in the asymmetric unit, and the structure was solved by molecular replacement using a DCoH dimer as the search model. The refined DCoH2 structure gave R/R_free values of 0.215/0.236 for all data to 1.6 Å resolution (Table 1, Figure 4A). The DCoH2 and DCoH1 dimers adopt the same fold with a root-mean-square deviation (rmsd) for the Cα atoms of 0.61 Å (residues 11–101 in each monomer, Figure 4B). A DCoH2 tetramer similar to the DCoH1 tetramer is generated through a crystallographic 2-fold rotation axis (Figure 4C). The DCoH2 dimer–dimer interaction is similar to that of DCoH1 with a Cα rmsd between the tetramers of 0.80 Å (residues 11–101 of each monomer). The interaction surface consists primarily of helix 2, residues 44–60, from each monomer. Upon tetramerization, 2434 Å² in DCoH2 and 2860 Å² in DCoH1 are buried. The difference in the buried surface area is primarily due to contacts between the C-termini; the C-terminus of one DCoH2 monomer is disordered in the crystal structure. Three sequence differences occur in the respective recognition helices of DCoH2 and DCoH1: Gln45Arg, Ser51Thr, and Asn61Asp (Figure 5A,E). The tetramers conserve the hydrophobic core but show altered H-bonding interactions. The major difference at the interface of the recognition helices occurs at position 51, which is a serine in DCoH2 and a threonine in DCoH1. A water molecule is buried at the center of the DCoH2 tetramer interface (blue). Fo–Fc difference electron density (3σ) for the water molecule is shown with the water deleted from the model. The water forms hydrogen bonds to four Ser51 side chains, one from each DCoH2 monomer. Two rotamers are seen for each Ser51 side chain. Distances are indicated to the serine side chain closest to the water molecule. No analogous water is seen in the DCoH1 tetramer (orange) because the larger Thr51 side chains fills the equivalent space and the Thr51 hydroxyl groups are not positioned correctly. 

Figure 4: Structure of DCoH2. Panel A shows a stereodiagram of active site residues. The model of DCoH2 (thick sticks) is shown with the final, 1.6-Å-resolution, 2Fo–Fc electron density. Superimposed are the corresponding residues from DCoH1 (thin sticks) in complex with the product analogue 7,8-dihydrobiopterin (12). The two structures were superimposed using the Cα atoms of the active site residues (residues 58, 61–63, 78, 80–81, 89) with a rmsd of 0.47 Å (wall-eyed stereo). As shown in panel B, DCoH2 (blue) and DCoH1 (orange) dimers adopt the same fold. The rmsd between Cα atoms of the dimers (residues 11–101 of each monomer) was 0.61 Å. Panel C shows a ribbon diagram of the DCoH2 homotetramer. The saddle-shaped dimers (above and below, with each monomer colored cyan and blue) interact primarily through helix 2 of each monomer (residues 44–59, magenta) to form the tetramer. This helix, called the recognition helix (8), also forms the interface for HNF1. A small dimer–dimer contact also occurs between the C-terminal residues (residues 102 and 103, magenta). Contact residues were defined as being within 4 Å of the neighboring dimer. As shown in panel D, a water molecule is buried at the center of the DCoH2 tetramer interface (blue). Fo–Fc difference electron density (3σ) for the water molecule is shown with the water deleted from the model. The water forms hydrogen bonds to four Ser51 side chains, one from each DCoH2 monomer. Two rotamers are seen for each Ser51 side chain. Distances are indicated to the serine side chain closest to the water molecule. No analogous water is seen in the DCoH1 tetramer (orange) because the larger Thr51 side chains fills the equivalent space and the Thr51 hydroxyl groups are not positioned correctly.
The "stirrups" are distinct within each subfamily (Figure 5C,F).

DISCUSSION

As demonstrated for the nuclear hormone receptor superfamily (reviewed in 39), diverse regulatory outcomes can be achieved through the interaction of transcription factors with multiple coactivators and corepressors. Here we compared the properties of two paralogs of the coactivator DCoH to test the idea that they fulfill distinct functional roles that may account for the phenotypes of DCoH1 knockout mice. The DCoH1 and DCoH2 subfamilies are more highly conserved from mice to humans than they are between each other, suggesting that their functions may have diverged. Table 2 summarizes a comparison between DCoH1 and DCoH2 paralogs. DCoH interacts specifically with the small family of HNF1 transcription factors comprising isoforms of HNF1β and HNF1δ. DCoH1 null mice retained significant HNF1β activity and a degree of phenylalanine hydroxylase activity, suggesting that DCoH2 partially complements the coactivator function of DCoH1 in vivo (19).

To determine the number of DCoH paralogs in humans, we searched the human genome using tblastn (www.ncbi.nlm.nih.gov/BLAST/). The search identified four sequences related to DCoH: the annotated DCoH gene on chromosome 10 (DCoH1), the annotated muscle DCoH on chromosome 5 (DCoH2), and two pseudogenes on chromosomes 2 and 17. The exon structures of DCoH1 and DCoH2 are identical. These gene structures were confirmed by Genscan. The sequence on chromosome 17 is located in intron 5 of the annotated CARKL gene. The DCoH coding region contains a frame shift at position 30 and stop codons at positions 8 and 70. The sequence on chromosome 2 contains two regions of low homology with DCoH2. Exon 2 of this aligned sequence contains a stop codon. In contrast with the expressed DCoH genes, the exon structures of the pseudogenes predicted by Genscan did not agree with the exons predicted from the alignment with DCoH.

FIGURE 5: Surface conservation patterns in the vertebrate DCoH2 (A–D) and DCoH1 (E–H) dimers. Colors indicate sequence conservation within the DCoH2 subfamily only (orange) or DCoH1 subfamily only (magenta), or within both DCoH1 and DCoH2 (black). Tan indicates residues not conserved within the DCoH2 subfamily; yellow indicates residues not conserved within the DCoH1 subfamily. Models A and E are viewed from the "top" (above the recognition helices). Most residues in helix 2 are conserved in both DCoH1 and DCoH2, except for residue 51, which is a serine in DCoH2 (A) and a threonine in DCoH1 (E). Models C and F show "side" views; models B and G show "front" views. Models D and H show views facing into the saddle. Each view is derived from the "side" views (models C and F) by a right-handed 90° rotation around the indicated axis. DCoH1 is more highly conserved than DCoH2. The "stirrups" and several surface residues, including in the saddle, distinguish the DCoH1 and DCoH2 subfamilies. This figure is based on the four species, human, mouse, rat, and chicken, known to encode both DCoH1 and DCoH2 paralogs. Conservation is defined according to similarity groups (FYW, IVLM, RK, DE, GA, TS, NQ, single letter amino acid code) except where residues are identical in DCoH1 or DCoH2 subfamilies and distinct from the other subfamily (e.g., both serine and threonine are found at residue 51 in DCoH2, Thr51 is conserved in DCoH1). The solvent-exposed surfaces were drawn using PyMol (Delano, W. L. (2002) The PyMol Molecular Graphics System, http://www.pymol.org).
four sequences (from Streptomyces coelicolor, Leishmania major, Vibrio cholerae, and Aribidopsis thaliana) lack the characteristic His62-His63-Pro64 tripeptide. This conservation suggests that the enzymatic function of DCoH predates the transcriptional function.

In mammals, DCoH catalyzes the dehydration of pterin-4α-carbinolamine, the first step in recycling tetrahydrobiopterin. The fully reduced tetrahydrobiopterin is an essential cofactor for the aromatic amino acid hydroxylases and the three nitric oxide synthases. It is generally accepted that bacteria lack tetrahydrobiopterin (42, 43). Nevertheless, bacterial homologues of mammalian phenylalanine hydroxylase (PAH) and nitric oxide synthase (NOS) have been identified. PAH has been reported in both α and γ divisions of Proteobacteria (38) with similar folds to the mammalian proteins (44). Bacterial PAH can utilize tetrahydrobiopterin as a cofactor in vitro (45). The NOS fold has been conserved as well, as demonstrated by the structure of a Staphylococcus aureus homologue of the oxygenase domain of mammalian NOS (46). Bacterial NOS lacks the reductase domain of the mammalian counterpart. A bacterial oxygenase fused to a mammalian reductase domain from neuronal NOS can use tetrahydrobiopterin or tetrahydrofolate to generate NO (47, 48). While the authentic cofactors are unknown for the bacterial enzymes, pteridine derivatives do function as cofactors in other electron-transfer reactions in prokaryotes and archaeabacteria (43). The conservation of the active site in all DCoH homologues suggests that the PCD activity may play roles in recycling pterin cofactors in all kingdoms of life.

The functional and phylogenetic evidence suggests that adaptation of the helical surface of the DCoH dimer for protein–protein interactions evolved after the enzymatic function (Figure 4C). The transcriptional coactivator function of DCoH appears to be unique to vertebrates. DCoH enhances the activity of the HNF1 transcription factors, which are unique to vertebrates (49). There are indications that DCoH may interact with a second vertebrate transcription factor (10, 50–53). No evidence for a transcriptional role for DCoH in prokaryotes has been found (38). The prokaryotic DCoH is dimeric, as demonstrated by the P. aeruginosa PhhB structure, indicating that a DCoH dimer is enzymatically active (54). The recent structure of DCoH from Thermus thermophilus (RCSB Protein Data Bank ID 1USO) (55) indicates that the archaeal homologues are dimers. The rmsd between Cα atoms of the archaeal and mouse dimers (residues 12–75 from both monomers of T. thermophilus DCoH aligned with 34–97 from both monomers of mouse DCoH) was 1.6 Å.

The interface that mediates tetramer formation in vertebrate DCoH proteins is composed primarily of residues from helix 2, residues 43–59, which we will refer to as the recognition helix (11). This same interface interacts with HNF1α (8). In contrast to the active site residues, the recognition helices are not conserved across the kingdoms of life. While tetramerization does not interfere with the enzymatic activity of DCoH, it does compete with HNF1 binding (8). The implication is that tetramerization coevolved with the

Table 2: Comparison between DCoH1 and DCoH2 Paralogs

<table>
<thead>
<tr>
<th>property</th>
<th>DCoH1</th>
<th>DCoH2</th>
</tr>
</thead>
<tbody>
<tr>
<td>human gene locus conservation, percent identity</td>
<td>10q22 mouse vs human, 99%</td>
<td>5q31.2 mouse vs human, 88%</td>
</tr>
<tr>
<td>dehydratase activity</td>
<td>complementation assay (22, 38)</td>
<td>complementation assay (3A)</td>
</tr>
<tr>
<td>in vitro activity assay (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>binding to HNF1 in vivo</td>
<td>EMSA</td>
<td>EMSA (Figure 3B)</td>
</tr>
<tr>
<td></td>
<td>Co–IP (6)</td>
<td>Co–IP (Figure 3C)</td>
</tr>
<tr>
<td>binding to HNF1 in vitro</td>
<td>no</td>
<td>EMSA supershift (Figure 3D)</td>
</tr>
<tr>
<td>stabilizing HNF1 dimer</td>
<td>competition assay (6)</td>
<td>transient transfection assay (21)</td>
</tr>
<tr>
<td>coactivator activity</td>
<td>transient transfection assay (6)</td>
<td>increased dissociation rate:</td>
</tr>
<tr>
<td>homotetramer stability</td>
<td>complex with HNF1</td>
<td>complex with HNF1 upon mixing (Figure 3D,E)</td>
</tr>
<tr>
<td></td>
<td>hyperstable:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>complex with HNF1 requires co-folding (14,15)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>melting temperature &gt; 90° (40)</td>
<td></td>
</tr>
<tr>
<td>tetramerization interface</td>
<td>TS1 at center of hydrophobic interface (11)</td>
<td>S51 and ordered water at hydrophobic interface (4D)</td>
</tr>
<tr>
<td>expression</td>
<td>liver, kidney, intestine, stomach, pancreas; low in skin, eye and brain (6, 19, 41)</td>
<td>intestine; muscle; low in kidney and liver (19, 21)</td>
</tr>
<tr>
<td>interactions</td>
<td>unknown</td>
<td>Dyrk1B kinase (21)</td>
</tr>
</tbody>
</table>

Table 3: The Human Genome Contains Two DCoH Paralogs and Two Pseudogenes

<table>
<thead>
<tr>
<th>chromosome/accession number</th>
<th>predicted exons</th>
<th>alignment</th>
<th>Genscan exons identified</th>
<th>$E$ value</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/NT_008583.13</td>
<td>1 promoter to Met1$^a$</td>
<td>$d$</td>
<td>63bp-Met1</td>
<td>0.565</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2–45$^b$</td>
<td></td>
<td>$6 \times 10^{19}$</td>
<td>0.997</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>46–72$^b$</td>
<td></td>
<td>$1 \times 10^{10}$</td>
<td>0.977</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>73–104$^a$</td>
<td></td>
<td>$2 \times 10^{10}$</td>
<td>0.950</td>
<td></td>
</tr>
<tr>
<td>5/NT_034776.2</td>
<td>1 promoter to Met1$^a$</td>
<td>$d$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2–45$^b$</td>
<td></td>
<td>$2 \times 10^{17}$</td>
<td>0.479</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>46–72$^b$</td>
<td></td>
<td>$3 \times 10^{10}$</td>
<td>0.468</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>73–103$^b$</td>
<td></td>
<td>$2 \times 10^{9}$</td>
<td>0.469</td>
<td></td>
</tr>
<tr>
<td>17/NT_010692.9</td>
<td>1 –43$^c$</td>
<td></td>
<td>$2 \times 10^{9}$</td>
<td>0.565</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>33–103$^c$</td>
<td></td>
<td>$1 \times 10^{11}$</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>2/NT_005403.11</td>
<td>1 17–47$^c$</td>
<td></td>
<td>$3 \times 10^{4}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>59–100$^c$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$ The $E$ values from the tblastn alignment and the $P$ values from Genscan (Burge, C. B., http://genes.mit.edu/GENSCAN.html) are shown. $^b$ Alignment with human DCoH1. $^c$ Alignment with mouse DCoH2. Not applicable.
transcriptional function of DCoH, and homotetramer formation contributes to the regulation of coactivation by DCoH. DCoH2, like DCoH1, forms homotetramers in solution. Conservation of the tetramerization interface suggests the possibility of mixed DCoH1/DCoH2 tetramers. Coexpression of DCoH1 and DCoH2 confirmed that they were capable of forming mixed tetramers and may do so in vivo.

The structure of DCoH2 is consistent with the functional similarities and differences compared to DCoH1. We demonstrated that DCoH2, like DCoH1, recycles 4a-hydroxytetrahydrobiopterin and binds HNF1α dimers. DCoH2 exhibits a number of properties that may distinguish it functionally from DCoH1. The expression pattern of DCoH2 differs from that of DCoH1. For example, DCoH2 has recently been identified in skeletal muscle (21), which does not contain DCoH1 (6, 56). DCoH1 and DCoH2 may interact with distinct sets of factors, since the sequence differences map primarily to the protein surfaces (Figure 5). The saddle of DCoH, which presents a large concave surface suitable for protein—protein interactions, is lined with different residues in the DCoH1 and DCoH2 subfamilies. Mutagenesis studies suggest that the coactivator function of DCoH involves recruitment of additional factors to the DCoH saddle (15). A candidate for a paralog-specific interacting protein is the serine/threonine kinase Mirk, which has been shown to interact with DCoH2 (21).

Although both DCoH1 and DCoH2 can interact with HNF1α, they differ in their availability for HNF1 binding (Figure 3). The DCoH dimer, and not the tetramer, associates with HNF1α in solution. In contrast, formation of the DCoH1/HNF1α complex requires cofolding of the two proteins. The hyper-stable DCoH1 tetramer sequesters the HNF1α binding surface. The primary sequence difference between DCoH2 and DCoH1 in the recognition helix is at residue 51, which is a threonine in DCoH1 and a serine in DCoH2. The smaller serine side chain allows binding of a buried water molecule upon DCoH2 homotetramerization. Water burial and ordering
are unfavorable (57–60), and this effect would be expected to facilitate DCoH2/HNF1 complex formation. As a result, DCoH2 may preferentially associate with HNF1α, instead of forming stable homotetramers. The relative instability of DCoH2 homotetramers may account for the low levels of cytoplasmic DCoH2 activity and the predominant nuclear association of DCoH2 with HNF1 observed in DCoH1 null mice (19).

The human DCoH2 sequence also encodes serine at residue 51 suggesting a less stable homotetramer than human DCoH1, similar to mouse. It remains to be seen whether decreased homotetramer stability is a property that distinguishes DCoH2 from DCoH1 in general. While both Gallus gallus and Rattus norvegicus DCoH2 sequences contain Thr51, they each possess nonconserved residues elsewhere in the recognition helix composing the tetramerization interface (at Asn59 and Val55, respectively) that may destabilize the homotetramer. Alternatively, these variants may alter the stability of the DCoH/HNF1 complex, since the same surface forms the interface with HNF1.

The decreased stability of the DCoH2 homotetramer relative to DCoH1 may underlie normal functional differences between the two paralogs. We have proposed that homotetramerization of DCoH1 autoinhibits coactivator function by inhibiting HNF1α binding (8). The phylogenetic data support the idea that tetramerization of DCoH coevolved with the coactivator function of DCoH. Thus, DCoH1 may function as a timer in which coactivator function is regulated kinetically. The dissociation rate of the DCoH1/HNF1 complex would determine the lifetime of the coactivator complex if the DCoH1 tetramer acts as a stable sink that lowers the concentration of DCoH1 dimers. In contrast, the less stable DCoH2 tetramer may provide a thermodynamic switch that affords more rapid changes in coactivator activity. Further experiments are required to test this model.

ACKNOWLEDGMENT

We thank James Holton for assistance in data collection and data processing. X-ray data were collected at Beamline 5.0.2 at the Lawrence Berkeley National Laboratory Advanced Light Source. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences Division, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098. We thank George Johnen for supplying rabbit anti-DCoH antibody.

SUPPORTING INFORMATION AVAILABLE

A list of 42 aligned DCoH sequences identified using PSI-BLAST, which was used to generate the phylogenetic tree (Figure 6). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES


lamine dehydratase from rat and human liver. Purification, characterization, and complete amino acid sequence, J. Biol. Chem. 4828–4831.
Enzymatic and Transcriptional Functions of DCoH and DCoH2


35. Felsenstein, J. (1993) PHYLIP (Phylogeny Inference Package), version 3.5c, Distributed by the author, Department of Genetics, University of Washington, Seattle, WA.


55. Tahirov, T., and Inagaki, E. DCoH, a bifunctional protein-binding transcriptional coactivator, Pro9Leu mutant, to be published.


