

- 349; S. McBrearty, *Man* **25**, 129 (1990); R. G. Klein, *The Human Career* (Univ. of Chicago Press, Chicago, 1989).
8. A. S. Brooks *et al.*, *Science* **268**, 548 (1995); See also discussion of large mammal remains in this paper.
9. Earlier bevel-shaped or pointed bone and ivory implements are known from several middle Pleistocene sites including Přezletice, Czech Republic [J. Fridrich, *Proceedings of the IXth International Union of Pre- and Proto-Historic Sciences Section VIII: Les Premières Industries d'Europe* (Nice, France, 1976), pp. 8–23]; Bilzingsleben, Germany [D. Mania and T. Weber, *Bilzingsleben III Veröffentlichungen des Landesmuseums für Vorgeschichte in Halle* **39**, 9 (1986)]; Kabwe, Zambia [J. D. Clark, in *From the Earliest Times to c. 500 BC*, vol. 1 of *The Cambridge History of Africa*, J. D. Clark, Ed. (Cambridge Univ. Press, Cambridge, 1982), pp. 248–341], and possibly the lower Pleistocene site of Swartkrans [C. K. Brain *et al.*, *S. Afr. J. Sci.* **84**, 828 (1988)], but the shaping is simple and may be due in part to use. For a summary, see (12).
10. A summary of research history to 1988 is provided by J. Verriers and J. de Heinzelin, in *Evolution of Environments and Hominidae in the African Western Rift Valley*, N. T. Boaz, Ed. (Virginia Museum of Natural History Mem. 1, Martinsville, VA, 1990), pp. 17–39.
11. A. S. Brooks and C. C. Smith, *Afr. Archaeol. Rev.* **5**, 65 (1987); A. S. Brooks and P. Robertshaw, in *Low Latitudes*, vol. 2 of *The World at 18 000 BP*, C. S. Gamble and O. Soffer, Eds. (Unwin Hyman, London, 1990), pp. 121–169. On the basis of both radiocarbon and amino acid racemization dating of ostrich eggshell, barbed bone points and LSA lithics at White Paintings Shelter, Tsodilo Hills, Botswana, may be associated with a date as old as or older than the barbed points from Ishango (L. H. Robbins, personal communication).
12. H. Knecht, in "Hunting and Animal Exploitation in the Later Palaeolithic and Mesolithic of Eurasia," *Archaeol. Pap. Am. Anthropol. Assoc.* **4**, G. L. Peterkin, H. M. Bricker, P. Mellars, Eds. (1993), p. 33; L. G. Straus, *ibid.*, p. 83; A. S. Brooks, in *Les Fouilles Moviuis a Pataud*, H. M. Bricker Ed. (Centre National de la Recherche Scientifique, Paris, in press); H. Knecht, *Sci. Am.* **271** (no. 1), 82 (1994).
13. J. Treenig, *Worldwide Chronology of Fifty-Three Prehistoric Innovations* (Acta Archaeol. Lund. Ser. B° 21, Stockholm, 1993).
14. D. Helgren, personal communication.
15. We recorded these materials within 20 cm by 20 cm squares, together with the topography of the concentration in each square. All matrix was screened through 3-mm mesh.
16. We prefer to use "pavement" rather than the more usual archaeological designation "living floor." This avoids a priori assumptions about possible aggregating agents and hominid role.
17. Site 440 located near Wadi Halfa, Sudan, contained at least six species of fish in association with a Mousterian industry [P. H. Greenwood, in *The Prehistory of Nubia*, F. Wendorf, Ed. (Southern Methodist Univ. Press, Dallas, 1968), vol. 1, pp. 100–109; J. L. Shiner, *ibid.*, vol. 2, pp. 630–650].
18. L. Robbins, paper presented at the 11th Biennial Meeting of the Society for Africanist Archaeologists, Los Angeles, 26 to 29 March 1992. Ongoing excavation at White Paintings Shelter, Tsodilo Hills, Botswana, indicates the presence of fish associated with a MSA industry.
19. Two notched bone fragments and one unbarbed bone point occur in MSA and Howieson's Poort Industries at Klasies River Mouth. Split and polished warthog or bush pig tusks have also been reported from Border Cave [R. Singer and J. Wymer, *The Middle Stone Age at Klasies River Mouth in South Africa* (Univ. of Chicago Press, Chicago, 1982); T. P. Volman, in *Southern African Prehistory and Palaeoenvironments*, R. G. Klein, Ed. (Balkema, Boston, 1984), pp. 169–220]. Middle Paleolithic bone or ivory awls are known from the Khormusan industry of Egypt [A. E. Marks, in *The Prehistory of Nubia*, F. Wendorf, Ed. (Southern Methodist Univ. Press, Dallas, 1968), vol. 1, pp. 315–391] and from site Kostenki 17 on the Russian plain [M. Anikovich, *J. World Prehist.* **6**, 205 (1992)]. The complexity and formal nature of the Katanda industry, however, place it, we believe, in a cognitively distinct category.
20. C. Krupsha, thesis, George Washington University, Washington, DC (1993).
21. J. de Heinzelin, *Les Fouilles d'Ishango: Exploration du Parc National Albert, Mission J. de Heinzelin 1950*, Fasc. **2** (1957).
22. At Ishasha, south of Lake Rutanzige, in a swampy gallery forest environment similar to that implied for the Katanda MSA, large numbers of *Clarias* are taken annually by harpoon during the few days of the spawning season (T. E. Mugangu, personal communication).
23. Correlation of number of in situ lithics per square and depth of square below datum: $r^2 = 0.0370$; $n = 369$; $P = 0.01$.
24. A number of correlation analyses were conducted to examine the relation between lithic maximum length and depth below datum. These confirm that when controlled by both lithic type and raw material type, larger pieces have been displaced downslope.
25. Distance between cluster "centers," as defined by maximum lithic and faunal density, is 4.2 m. Maximum cluster diameters are 4.0 and 3.4 m. For comparative data, see J. E. Yellen, *Archaeological Approaches to the Present: Models for Reconstructing the Past* (Academic Press, New York, 1977); R. A. Gould and J. E. Yellen, *J. Anthropol. Archaeol.* **6**, 77 (1987).
26. J. D. Clark, *J. World Prehist.* **2** (no. 3), 235 (1988).
27. We thank the people and the government of Zaire for facilitating this research, which was a collaborative effort involving almost 200 individuals from more than 15 countries. This study is based on work supported by the National Science Foundation under grants to N. T. Boaz, J. W. K. Harris, and A.S.B. (BNS85–07891, BNS86–08269, and BNS90–14092). Additional financial support was provided by the George Washington University Committee on Research, the L. S. B. Leakey Foundation, the Holt Family Foundation, the National Geographic Society, the National Science Foundation, the Smithsonian Institution, the Social Sciences and Humanities Research Council of Canada, and the Musée Royal de l'Afrique Centrale. Identification of the rocks used for grind stones was made by S. Sorensen. Artifact photographs by R. E. Clark Jr. (Smithsonian Institution).

18 August 1994; accepted 15 February 1995

Crystal Structure of DCoH, a Bifunctional, Protein-Binding Transcriptional Coactivator

James A. Endrizzi, Jeff D. Cronk, Weidong Wang, Gerald R. Crabtree, Tom Alber

DCoH, the dimerization cofactor of hepatocyte nuclear factor-1, stimulates gene expression by associating with specific DNA binding proteins and also catalyzes the dehydration of the bipterin cofactor of phenylalanine hydroxylase. The x-ray crystal structure determined at 3 angstrom resolution reveals that DCoH forms a tetramer containing two saddle-shaped grooves that comprise likely macromolecule binding sites. Two equivalent enzyme active sites flank each saddle, suggesting that there is a spatial connection between the catalytic and binding activities. Structural similarities between the DCoH fold and nucleic acid-binding proteins argue that the saddle motif has evolved to bind diverse ligands or that DCoH unexpectedly may bind nucleic acids.

Transcription depends on interactions among numerous proteins, including regulatory proteins bound at specific DNA sites. Homeodomain transcription factors, for example, are recognized by protein-binding coactivators (1, 2). Little is known, however, about the molecular nature of the activation signals or the structures of the interacting protein motifs. One such coactivator is DCoH, the dimerization cofactor of hepatocyte nuclear factor-1 (HNF-1) (1). In vertebrates, DCoH associates with the HNF-1 proteins, which regulate tissue-specific genes by binding to DNA as dimers (3–5). The dimerization and transcriptional activities of HNF-1 α are stimulated in vitro and in vivo by DCoH, but DCoH itself contains no DNA

binding activity or activation domain. These properties imply that DCoH stimulates transcription through protein-protein interactions.

DCoH presumably binds other ligands, because it also functions in contexts devoid of HNF-1 α . In *Xenopus*, for example, DCoH is maternally encoded in oocytes and localizes to cell nuclei at a time when the primary germ layers are determined during the midblastula transition (6). A widespread transcriptional role for DCoH was suggested by the discovery of a bacterial homolog, phhB, with ~30% identity to the rat protein (7). Mutations in the gene *phhB* block expression of other genes in the *phh* operon, including the gene encoding phenylalanine hydroxylase. Expression of *phhB* in mammalian cells facilitates activation of HNF-1-dependent genes, and mammalian DCoH complements *phhB* mutations in bacteria by supporting transcription of the *phh* operon (8). This complementation suggests that the mechanisms of transcriptional activa-

J. A. Endrizzi, J. D. Cronk, T. Alber, Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720–3206, USA.

W. Wang and G. R. Crabtree, Department of Developmental Biology and Howard Hughes Medical Institute, Unit in Molecular and Genetic Medicine, Stanford University School of Medicine, Stanford, CA 94305–5425, USA.

tion by DCoH may be similar in prokaryotes and eukaryotes.

In addition to their regulatory activities, DCoH proteins (including phhB) catalyze the dehydration of the biopterin cofactor used by phenylalanine hydroxylase (7, 9, 10). The dehydration reaction can be rate limiting in phenylalanine utilization, and patients partially deficient in DCoH excrete chemically rearranged forms of the biopterin cofactor (11). Thus, DCoH proteins are both enzymes and transcriptional regulators of phenylalanine catabolism. A connection between these metabolic and transcriptional activities has not been established (12).

To investigate its dual activities, we determined the x-ray crystal structure of rat DCoH (Table 1) (13). The refined model contains residues 6 to 104 and one sulfate molecule for each of the eight polypeptide chains in the asymmetric unit. The 12-kD DCoH monomer is an α/β protein with a four-stranded, antiparallel β sheet flanked on one side by three α helices (Fig. 1, B and C). The arrangement of secondary structural elements, but not the topology, is similar to that seen in the TATA binding protein (TBP) (14), the ribonucleoprotein (RNP) RNA binding motif (15), and the "palm" domains of DNA and RNA polymerases (16). The eight independent monomers in the asymmetric unit are highly similar, with the largest positional differences in surface loops and at the chain termini (Fig. 1C).

DCoH forms a tetramer with approximate 222 symmetry and dimensions of 30 Å by 60 Å by 60 Å (Fig. 2A). This arrangement is consistent with sedimentation equilibrium experiments showing that the protein is tetrameric in solution (17). Approximately 1460 Å² of the surface area of each monomer are buried in the tetramer. Of the two distinct subunit interfaces, one is mediated by α helices, and the other is formed by helices and adjoining β strands. The first interface, containing paired, antiparallel H2 helices, comprises an up-and-down four-helix bundle. The pairs of antiparallel helices cross their symmetry-related counterparts at an angle of $\sim -48^\circ$, which is commonly observed in four-helix bundles (18).

The second interface forms the most striking feature in the structure, a saddle-shaped groove with four equivalent β strands from each monomer making a continuous, eight-stranded, antiparallel β sheet (Fig. 2). With a perpendicular two-fold rotation axis between the central S3 β strands, the halves of the saddle are antiparallel. The DCoH tetramer contains two saddles that are approximately perpendicular and ~ 30 Å apart (Fig. 2B). In the two independent tetramers, the saddle

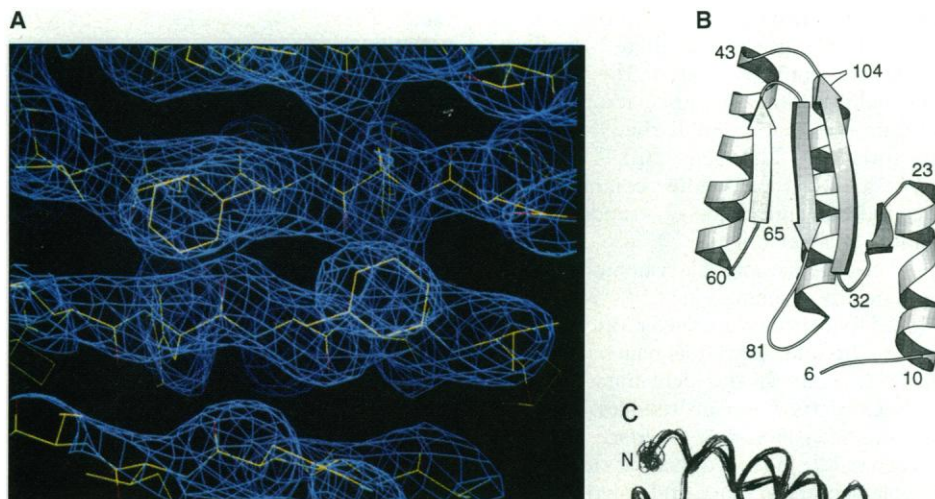


Fig. 1. Structure of DCoH. **(A)** Refined model superimposed on the solvent-flattened, SIRAS electron density map. The region near the molecular twofold axis perpendicular to the saddle is shown at a contour level of 1σ . The Phe⁶⁷ residues are paired in the center of the saddle. **(B)** Ribbon diagram (35) of the monomer. Helix 1 (residues 10 to 23) is followed by a β hairpin (residues 25 to 40, S1 and S2) that forms half of the four-stranded, antiparallel β sheet. Residues 43 to 60 form helix 2, a central element in the tetramer interface. The sequence His⁶²-His⁶³-Pro⁶⁴ connects helix 2 to the antiparallel β strands (S3, S4; residues 65 to 77). The COOH-terminal helix (H3, residues 87 to 104) packs over the middle of the β sheet. The NH₂-terminal five amino acids could not be positioned in the electron density map, presumably because of the flexibility of the chain. **(C)** Superposition of the main chains of the eight crystallographically independent monomers. The root-mean-square deviation of equivalent main chain atoms is 0.78 Å. This close similarity of the independently refined monomers supports the chain tracing and implies that the core of the monomer structure is relatively insensitive to crystal packing effects. C, COOH-terminus; N, NH₂-terminus.

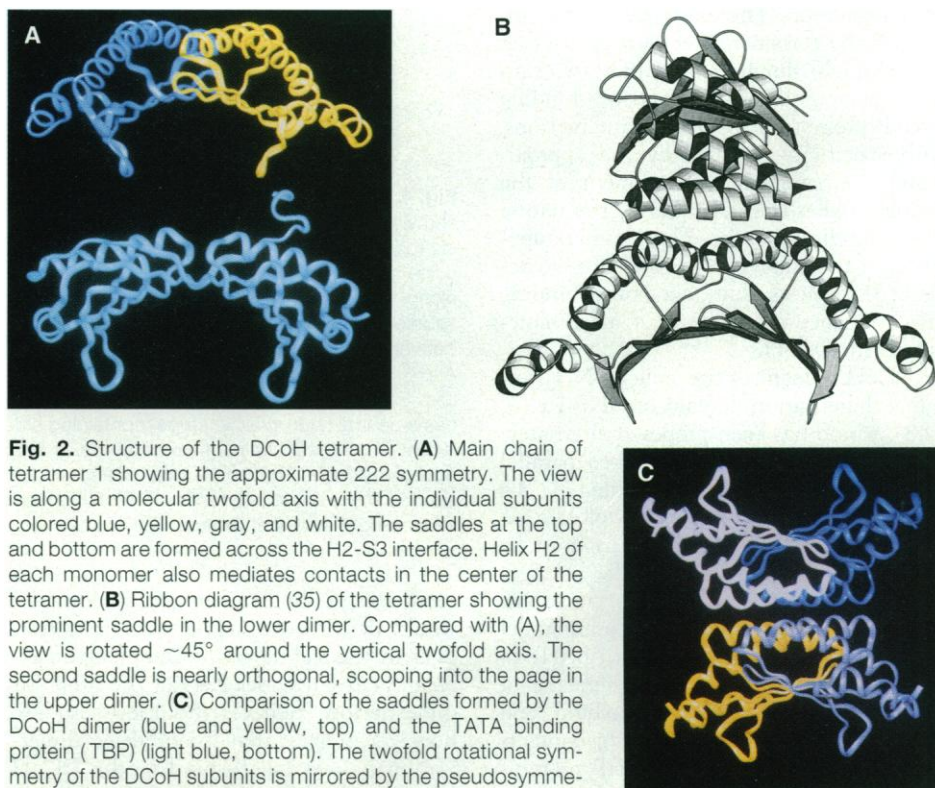


Fig. 2. Structure of the DCoH tetramer. **(A)** Main chain of tetramer 1 showing the approximate 222 symmetry. The view is along a molecular twofold axis with the individual subunits colored blue, yellow, gray, and white. The saddles at the top and bottom are formed across the H2-S3 interface. Helix H2 of each monomer also mediates contacts in the center of the tetramer. **(B)** Ribbon diagram (35) of the tetramer showing the prominent saddle in the lower dimer. Compared with (A), the view is rotated $\sim 45^\circ$ around the vertical twofold axis. The second saddle is nearly orthogonal, scooping into the page in the upper dimer. **(C)** Comparison of the saddles formed by the DCoH dimer (blue and yellow, top) and the TATA binding protein (TBP) (light blue, bottom). The twofold rotational symmetry of the DCoH subunits is mirrored by the pseudosymmetry of the domains of TBP. The TBP domains reflect a gene duplication event, whereas the two halves of the DCoH saddle are composed of identical, noncovalently associated monomers.

width varies from 24 to 27 Å (19). This 3 Å variation suggests that there is flexibility in the conformation of the grooves. The saddles form obvious macromolecule binding sites that resemble the DNA binding surface of TBP (Fig. 2C). TBP contains a 29.5 Å saddle comprising a pseudodyad-symmetric, 10-stranded, anti-parallel β sheet (14).

The sequences of DCoH homologs from bacteria to mammals (1, 9, 20) were aligned to identify conserved residues essential for the fold, as well as regions that are likely to contain the dehydratase active sites. Conserved surface residues occur at the monomer-monomer interface adjacent to each saddle in two equivalent clefts, each complementary in shape and polarity to the 4a-carbinolamine bioprotein substrate (Fig. 3, A and B). The invariant residues in the active sites include His⁶³ and the His⁶²-Asp⁸⁹ pair, which, like the His-Asp dyad in the serine proteases, may activate His⁶² to act as a general base in the catalytic mechanism (21). The arrangement of catalytic residues in DCoH closely resembles that in the trimeric enzyme scytalone dehydratase (22). Differences in the folds, the order of the catalytic residues in the respective sequences, and the number of subunits in the two structures argue that these dehydratases provide a classic example of convergent evolution.

The DCoH structure is a rich source of new hypotheses about the mechanisms of gene regulation. The existence of two saddles in the tetramer directly suggests that DCoH might simultaneously bind two similar macromolecules. Bipartite binding would promote cooperative interactions, with specificity imposed by the approximately perpendicular arrangement of the DCoH saddles (Fig. 2B) and by the nature of the binding surfaces. The close juxtaposition of the catalytic sites and the saddles (Fig. 3) suggests that metabolite binding may influence associations of macromolecules with DCoH.

DCoH recognizes the helical, NH₂-terminal dimerization domain of HNF-1α (1, 5, 8), which has been proposed alternately to form a coiled coil (4), kinked parallel helices (23), or a four-helix bundle (24). To make symmetric contacts, the twofold rotation axes of DCoH and HNF-1α must be superimposed. Consequently, the arrangement of subunits in DCoH is consistent with "lengthwise" binding of antiparallel helices in the dimerization motif of HNF-1α. Alternatively, the 222 symmetry of DCoH restricts a parallel motif to bind "end on," with diverging helices making contacts in much the same way that bZIP proteins bind DNA (25).

The DCoH saddle is a potential binding site for the dimerization helices of

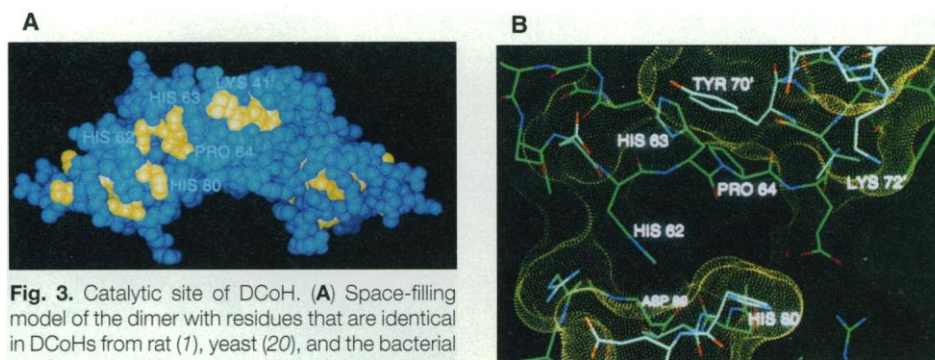


Fig. 3. Catalytic site of DCoH. (A) Space-filling model of the dimer with residues that are identical in DCoHs from rat (1), yeast (20), and the bacterial phhB protein (7) shown in yellow and variable residues in blue. The yeast and bacterial sequences have ~30% identity with the mammalian protein. Strongly conserved regions include helix 2 in the tetramer interface and the presumptive catalytic site including His⁶²-His⁶³-Pro⁶⁴, His⁸⁰, and Asp⁸⁹. The residue Tyr^{70'} (Trp in phhB) also occurs in the active site cleft. Insertions in the yeast protein occur in the "stirrup" of the saddle and in the loop preceding the COOH-terminal helix. (B) Presumptive active site cleft is next to the saddle. Key residues, including the conserved His⁶²-His⁶³-Pro⁶⁴, His⁸⁰, and Asp⁸⁹, are shown behind the solvent-accessible surface. Tyr^{70'} and Lys^{72'} from the neighboring subunit also line the pocket, suggesting that dimerization across the H2-S3 interface is required to form both the saddle and the active sites.

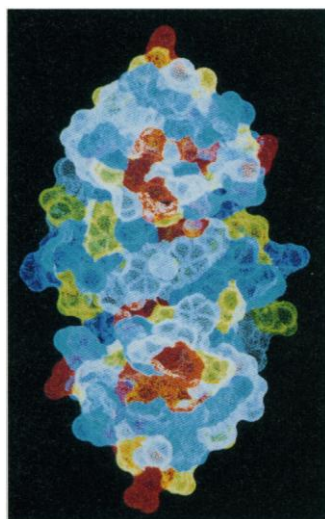


Fig. 4. Van der Waals surface of the DCoH dimer showing the distribution of nonpolar (white), basic (blue), acidic (red), and polar (yellow) amino acids in the saddle. The view is into the saddle along the central molecular twofold axis. The saddle opens to the left and right in the drawing, between the red acidic residues, Glu²⁹ and Asp³², in the stirrup in the center of each subunit. Like RNP RNA binding motifs (15), the saddle has a central hydrophobic stripe (containing adjacent Phe⁶⁷ side chains) flanked by basic and acidic residues. The active site clefts can be seen in the subunit interface, off-center on the left and right edges of dimer.

HNF-1α. A helix pair (~20 Å wide) would fit into the saddle, and basic residues in the saddle could complement exterior acidic residues (24) of the HNF-1α dimerization helices. Solution studies, however, suggest that the HNF-1α dimerization helices fail to bind to the DCoH tetramer, but instead form a heterotetrameric complex containing only two DCoH subunits (1, 8, 26). In contrast, the

Table 1. Data collection, phasing, and refinement statistics. PHMB refers to the *p*-hydroxymercuribenzoate derivative.

| Item | Native | PHMB |
|--------------------------------------|---|---|
| Space group | <i>P</i> ₃ ₂ ₁ | <i>P</i> ₃ ₂ ₁ |
| Unit cell dimensions (Å) | <i>a</i> = <i>b</i> = 105.65, <i>c</i> = 196.23 | <i>a</i> = <i>b</i> = 105.33, <i>c</i> = 196.17 |
| Resolution | 3.0 Å | 3.0 Å |
| Unique/measured reflections | 21,436/ 86,670 | 21,416/ 97,074 |
| Completeness | 82.2% | 82.1% |
| <i>R</i> _{merge} * | 6.0% | 5.2% |
| <i>R</i> _{iso} † | | 18.6% |
| <i>R</i> _{Cullis} ‡ | | 0.55 |
| Phasing power§ | | 1.7 |
| Figure of merit | | 0.37 |
| After solvent flattening | | 0.72 |
| <i>R</i> _{cryst} * (20–3 Å) | 0.184 | |
| Rms Δ bonds# | 0.019 Å | |
| Rms Δ angles | 2.9° | |

**R*_{merge} = $\sum |I - \langle I \rangle| / \sum I$; *I*, intensity. †*R*_{iso} = $\sum |F_{PH} - F_P| / \sum F_P$; *F*_{PH} and *F*_P, derivative and native structure factor amplitudes. ‡*R*_{Cullis} = $\sum |F_{PH} \pm F_P| - F_{H(calc)}| / \sum |F_{PH} - F_P|$; *F*_{H(calc)}, calculated heavy atom structure factor amplitude; where the sum is taken over centric reflections. §Phasing power = $\{ \sum |F_H|^2 / \sum |E|^2 \}^{1/2}$, with $\sum |E|^2 = \sum (|F_{PH(obs)}| - |F_{PH(calc)}|)^2$. ||Mean figure of merit = $\langle |\Sigma P(\alpha) e^{i\phi} / \Sigma P(\alpha)| \rangle$; α, phase; *P*(α), phase probability distribution. ¶*R*_{cryst} = $\sum |F_P - F_{calc}| / \sum F_P$; *F*_{calc}, calculated structure factor amplitude. #Root-mean-square deviations from ideal bond lengths and angles.

recombinant DCoH protein described here and cytoplasmic rat DCoH that is not associated with HNF-1α are homotetramers in solution (17). This difference in oligomerization states suggests that the tetramer-dimer equilibrium of DCoH may influence its transcriptional function.

The 2:2 stoichiometry of the DCoH-HNF-1α complex implies that HNF-1α binding blocks DCoH tetramerization. The present data leave open the possibility that conformational changes might ac-

count for both the difference in DCoH oligomerization and the failure of HNF-1 α to bind the DCoH tetramer. Nonetheless, the most direct mechanism of trapping a 2:2 complex entails association of HNF-1 α with the "back side" of the DCoH dimer, covering the interface formed by the H2 helices. Such a complex, which would contain a four-helix bundle with complementary charges, would leave the DCoH saddle available to bridge HNF-1 α to other ligands.

If not HNF-1 α , what might bind in the saddle? The structural similarity of the saddle to TBP (Fig. 2C) invites speculation that DCoH or its homologs may bind nucleic acids. This possibility is in accord with the size, antiparallel symmetry, and surface properties of the saddle. As yet, however, experiments have failed to detect an interaction between rat DCoH and DNA (1, 8). The possibility of RNA binding is supported by the similarity of the DCoH fold (Fig. 1B) to the RNP RNA binding motif (15). Moreover, the basic residues flanking the DCoH saddle and the paired Phe⁶⁷ residues at the center of the saddle (Figs. 1A and 4) have structural analogs in the RNA binding sites of the proteins U1A, ROP, and Sex lethal (27). Interactions with RNA also may rationalize the regulatory activities of DCoH in cells and organisms that lack HNF-1 α (28). At a minimum, the DCoH structure implies that the saddle motif shared with TBP has adapted to mediate diverse macromolecular interactions.

Note added in proof: Preliminary studies in which the murine DCoH gene was introduced into bacterial cells containing a truncation mutation in the *PhhB* gene resulted in restoration of levels of steady-state mRNA for the *PhhA* gene. Although such cross-species complementation for transcriptional control proteins is unexpected, these results indicate that DCoH may use the same mechanism of transcriptional control in mammals and bacteria.

REFERENCES AND NOTES

1. D. B. Mendel *et al.*, *Science* **254**, 1762 (1991).
2. R. S. Mann and D. S. Hogness, *Cell* **60**, 597 (1990); M. Tanaka, J. S. Lai, W. Herr, *ibid.* **68**, 755 (1992); Y. Luo, H. Fujii, T. Gerster, R. G. Roeder, *ibid.* **71**, 231 (1992); J. L. Pomerantz, T. M. Kristie, P. A. Sharp, *Genes Dev.* **6**, 2047 (1992); W. Zeng *et al.*, *Development* **118**, 339 (1993).
3. G. Courtois, J. G. Morgan, L. A. Campbell, G. Fourel, G. R. Crabtree, *Science* **238**, 688 (1987); G. Courtois, S. Baumhueter, G. Crabtree, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 7937 (1988); S. Cereghini, M. Blumenfeld, M. Yaniv, *Genes Dev.* **2**, 957 (1988); E. M. Haddon, M. Frain, G. Paonessa, R. Cortese, *EMBO J.* **7**, 1711 (1988); M. Frain *et al.*, *Cell* **59**, 145 (1989); T. Chouard *et al.*, *Nucleic Acids Res.* **18**, 5853 (1990); S. Baumhueter *et al.*, *Genes Dev.* **4**, 372 (1990); C. J. Kuo, P. B. Conley, C.-L. Hsieh, U. Francke, G. R. Crabtree, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 9838 (1990).
4. A. Nicosia *et al.*, *Cell* **61**, 1225 (1990).
5. D. B. Mendel, L. P. Hansen, M. K. Graves, P. B. Conley, G. R. Crabtree, *Genes Dev.* **5**, 1042 (1991).
6. E. P. von Strandmann and G. U. Ryffel, *Development*, in press.
7. G. Zhao, T. Xia, J. Song, R. A. Jensen, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1366 (1994).
8. L. P. Hansen, thesis, Stanford University, Stanford, CA (1994).
9. B. A. Citron *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 11891 (1992).
10. S. Kaufman, *J. Biol. Chem.* **245**, 4751 (1970); R. A. Lazarus, S. J. Benkovic, S. Kaufman, *ibid.* **258**, 10960 (1983).
11. B. A. Citron *et al.*, *Am. J. Hum. Genet.* **53**, 768 (1993).
12. M. D. Davis, S. Kaufman, S. Milstien, *FEBS Lett.* **302**, 73 (1992).
13. Recombinant rat DCoH was purified as a fusion protein with glutathione-S-transferase and isolated after cleavage with thrombin (9). The protein was judged to be >95% pure by SDS gel electrophoresis. The mass determined by electrospray mass spectrometry (12,143.2 daltons) indicated that the dipeptide Gly-Ser from the fusion protein preceded the complete DCoH sequence (1). Crystals grown by vapor diffusion from 15 to 20 mg of DCoH per milliliter in 5 mM KPO₄, 5 mM methionine, 1.2 M ammonium sulfate, 2% polyethylene glycol (PEG) 400, and 0.1 M Hepes, pH 7.5, diffracted x-rays to ~2.5 Å resolution. The crystals had the symmetry of space group P3₂21 and contained two tetramers in the asymmetric unit. A *p*-hydroxymercuribenzoate (PHMB) derivative was obtained by soaking a crystal for 30 days in PHMB-saturated mother liquor. X-ray data were collected with a Rigaku R-Axis II C image plate detector with Cu-K α radiation produced by a RU200 rotating anode generator equipped with a graphite monochromator. Data were processed with the R-Axis software, and merging and scaling were performed with the PROTEIN program package (29). The positions of Hg atoms were determined with the program SHELX-86 (30) and refined with PROTEIN. A single isomorphous replacement map incorporating anomalous difference information (SIRAS) was calculated to 3.2 Å resolution, and the phases were improved by solvent flattening (31). The solvent-flattened SIRAS map contained unambiguous electron density for residues 24 to 104 of each of the eight monomers in the asymmetric unit. An initial model for these residues was built into one monomer and used to construct the two independent tetramers with FRODO (32). The partial model (with a starting *R* factor of 46.6%) was refined with X-PLOR (33). Rigid body refinement, Powell minimization, and simulated annealing refinement starting at 5000 K reduced the *R* factor to 22.3%. The NH₂-termini of all eight subunits were built in stages into the 3 Å resolution 2F_o - F_c map and the SIRAS maps with and without solvent flattening (F_o and F_c are the observed and calculated structure factor amplitudes). At each stage, the model was refined against 20 to 3 Å resolution data with the program TNT (34). The final model contains residues 6 to 104 (792 amino acids) and one sulfate molecule in the active site of each monomer. The model has excellent geometry, a single overall B value of 11.5 Å² for all atoms, and a crystallographic *R* factor of 18.4% for all data from 6 to 3 Å resolution. The difference map calculated with coefficients F_{Hg} - F_{nat} (where F_{Hg} and F_{nat} are the Hg derivative and native structure factors) and model phases contains peaks of up to 19 σ at the positions of the Hg atoms, which occur at an average distance of ~3.0 Å from the sulfur atoms of the eight Cys⁶² side chains. The low *R* value in concert with the good geometry of the model (Table 1), the virtual absence of disallowed main chain torsion angles, the high quality of the electron density map (Fig. 1A), the similarity of the independently refined monomers (Fig. 2C), and the chemically reasonable binding sites for heavy atoms demonstrate the correctness of the fold. The coordinates have identification number 1DCH in the Protein Data Bank.
14. D. B. Nikolov *et al.*, *Nature* **360**, 40 (1992); J. L. Kim, D. B. Nikolov, S. K. Burley, *ibid.*, p. 520 (1993); Y. Kim, J. H. Geiger, S. Hahn, P. B. Sigler, *ibid.*, p. 512; D. M. Nikolov and S. K. Burley, *Nature Struct. Biol.* **1**, 621 (1994); J. L. Kim and S. K. Burley, *ibid.*, p. 638.
15. C. C. Query, R. C. Bentley, J. D. Keene, *Cell* **57**, 89 (1989); K. Nagai *et al.*, *Nature* **348**, 515 (1991); C. G. Burd and G. Dreyfuss, *Science* **265**, 615 (1994).
16. M. R. Sawaya, H. Pelletier, A. Kumar, S. H. Wilson, J. Kraut, *Science* **264**, 1930 (1994).
17. DCoH was found to be tetrameric in solution by sedimentation equilibrium. With a Beckman Optima XL-A analytical ultracentrifuge, the radial distribution of protein at initial concentrations of 7 and 21 mM in 5 mM KPO₄, pH 7.4, was monitored by absorbance at 280 nm. With the XL-A software, the distribution curves at two rotor speeds fit to an average molecular weight of 48,200 \pm 1000. This result compares favorably with the molecular weight of 48,575 expected for a tetramer, and it agrees with earlier size determinations of DCoH purified from rat liver [C. Y. Huang, E. E. Max, S. Kaufman, *J. Biol. Chem.* **248**, 4235 (1973)].
18. N. L. Harris, S. R. Presnell, F. E. Cohen, *J. Mol. Biol.* **236**, 1356 (1994).
19. The width of the saddles was determined from the distance between the α carbons of the Gly³⁰ residues in the stirrups.
20. M. Johnston *et al.*, *Science* **265**, 2077 (1994).
21. On the basis of enzymatic precedents for proton abstraction, a carboxylate was expected to act as a general base. No invariant, exposed Asp or Glu residues, however, occur in the bioprotein binding site. Instead, the conserved Asp⁸⁹ is buried behind His⁶². An analogous Asp-His pair in the active site of scytalone dehydratase was proposed to function as a general base, and a second histidine analogous to His⁶³ in DCoH was proposed to position the hydroxyl leaving group (22).
22. T. Lundqvist *et al.*, *Structure* **2**, 937 (1994).
23. R. De Francesco, A. Pastore, G. Vecchio, R. Cortese, *Biochemistry* **30**, 143 (1991); A. Pastore *et al.*, *ibid.*, p. 148.
24. A. Pastore, R. De Francesco, M. A. C. Morelli, D. Nalis, R. Cortese, *Protein Eng.* **5**, 749 (1992).
25. T. E. Ellenberger, C. J. Brandl, K. Struhl, S. C. Harrison, *Cell* **71**, 1223 (1992).
26. J. D. Cronk and T. Alber, unpublished results.
27. A. L. Lee, R. Kanaar, D. C. Rio, D. E. Wemmer, *Biochemistry* **33**, 13775 (1994); C. Oubridge, N. Ito, P. R. Evans, C.-H. Teo, K. Nagai, *Nature* **372**, 432 (1994); P. F. Predki, L. M. Nayak, M. B. C. Gottlieb, L. Regan, *Cell* **80**, 41 (1995).
28. RNA binding by DCoH also would suggest an analogy to the human immunodeficiency virus (HIV) Tat protein, which activates transcription in concert with host factors by binding to RNA [Q. Zhou and P. A. Sharp, *EMBO J.* **14**, 321 (1995)].
29. W. Steigemann, *PROTEIN, Version 3.1: A Program System for the Crystal Structure Analysis of Proteins* (Max-Planck-Institut für Biochemie Computer Center, Martinsried bei Muenchen, Federal Republic of Germany, 1991).
30. G. M. Sheldrick, *Acta Crystallogr.* **A46**, 467 (1990).
31. B. C. Wang, *Methods Enzymol.* **115**, 90 (1985).
32. T. A. Jones, *J. Appl. Crystallogr.* **11**, 268 (1978).
33. A. Brunger, *X-PLOR, Version 3.1: A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, CT, 1992).
34. D. E. Tronrud, L. F. TenEyck, B. W. Matthews, *Acta Crystallogr.* **A43**, 489 (1987).
35. P. Kraulis, *J. Appl. Crystallogr.* **24**, 924 (1991).
36. We thank D. Mendel for advice and DNA constructs and M. Graves and M. Bogden for help with protein purification. We are indebted to D. King for mass measurements, H. Schachman and S. Marquese for use of equipment and L. Gonzalez Jr. for help using X-PLOR. Thanks to S. Burley for providing the coordinates of TBP. D. E. E. Baldwin, and members of the Alber lab critically reviewed the manuscript. W.W. was sponsored by a Damon Runyon-Walter Winchell postdoctoral fellowship. This work was supported by grants from NIH to G.R.C. and T.A.

15 December 1994; accepted 17 March 1995