Crystal structure and mechanism of CO dehydrogenase, a molybdo iron-sulfur flavoprotein containing S-selanylcytsteine

HOLGER DOBBEK*,†, LOTHAR GREMER‡, ORTWIN MEYER‡, and ROBERT HUBER*

*Max-Planck-Institut für Biochemie, Abteilung Strukturforschung, D-82152 Martinsried, Germany; and ‡Lehrstuhl für Mikrobiologie, Universität Bayreuth, D-95440 Bayreuth, Germany

Contributed by Robert Huber, May 27, 1999

ABSTRACT CO dehydrogenase from the aerobic bacterium *Oligotropha carboxidovorans* catalyzes the oxidation of CO with H2O, yielding CO2, two electrons, and two H+. Its crystal structure in the air-oxidized form has been determined to 2.2 Å. The active site of the enzyme, which contains molybdenum with three oxygen ligands, molybdopterin–cytosine dinucleotide and S-selanylcytsteine, delivers the electrons to an intramolecular electron transport chain composed of two types of [2Fe–2S] clusters and flavin-adenine dinucleotide. CO dehydrogenase is composed of an 88.7-kDa molybdoprotein (L), a 30.2-kDa flavoprotein (M), and a 17.8-kDa iron-sulfur protein (S). It is organized as a dimer of LMS heterotrimers and resembles xanthine dehydrogenase/oxidase in many, but not all, aspects. A mechanism based on a structure with the bound suicide-substrate cyanide is suggested and displays the necessity of S-selanylcytsteine for the catalyzed reaction.

CO is a trace gas and important scavenger for hydroxyl radicals in the atmosphere. About one fifth of CO [115–230 Tg yr⁻¹ (1)] is used by soil microorganisms (2, 3). CO has a short lifetime within the soil, and consumption of atmospheric CO occurs mainly in the top few centimeters of the humus layer (O horizon), most likely under aerobic conditions.

Under aerobic conditions, CO can be used as the sole source of carbon and energy by the carboxidotrophic bacteria, which comprise a taxonomically diverse group of obligate or facultative chemolithoautotrophic species (3, 4). Under anaerobic conditions, CO is used with low affinity by a number of methanogenic archaea and acetogenic, sulfate-reducing, or phototrophic bacteria (3, 5). Generally, anaerobic CO dehydrogenases are nickel-containing [4Fe-4S] proteins. CO dehydrogenases of methanogens or acetogens have acetyl-CoA synthetase activity. CO dehydrogenase (EC 1.2.99.2) of the carboxidotrophic bacterium *Oligotropha carboxidovorans* is a molybdenum-containing iron-sulfur flavoprotein that catalyzes the oxidation of CO (CO + H2O → CO2 + 2e⁻ + 2H+) and generates a proton gradient across the cytoplasmic membrane by channeling the electrons formed via cytochrome *b*₅₆₁ into a CO-insensitive respiratory chain (4). *O. carboxidovorans* uses the reactions of the Calvin cycle for CO₂ fixation (4). In the exponential growth phase, cells of *O. carboxidovorans* have 87% of CO dehydrogenase associated with the inner aspect of the cytoplasmic membrane and 13% located in the cytoplasm (6).

CO dehydrogenase is composed of the 88.7-kDa L (809 residues), 30.2-kDa M (288 residues), and 17.8-kDa S (166 residues) subunits and exists as a dimer of LMS heterotrimers (4). The L subunit carries the molybdenum cofactor, which is a mononuclear complex of Mo and molybdopterin–cytosine dinucleotide (MCD) (7). FAD-binding is on the M subunit and requires conformational changes of M introduced through the binding of M to LS (unpublished data). The S subunit harbors two [2Fe–2S] centers, which are proximal and distal, respectively, to the molybdenum center. They are traditionally designated type I and II according to spectroscopic properties, but the assignment is controversial (8, 9). Recent experiments with M subunit-depleted CO dehydrogenase preparations assign the proximal center to type I and the distal to type II (unpublished data).

CO dehydrogenase has been grouped into the sequence family of molybdenum hydroxylases (10). Other prominent representatives of that family catalyze the oxidative hydroxylation of a diverse range of aldehydes and aromatic heterocycles (9). The high-resolution structure of aldehyde oxidoreductase (Mop) from *Desulfovibrio gigas* has been solved and has provided an insight into an active site of a molybdenum hydroxylase (11). Mop and CO dehydrogenase both contain the MCD-molybdenum cofactor and the same [2Fe–2S] clusters. However, Mop is devoid of flavin and contains a single polypeptide chain, whereas CO dehydrogenase is a heterotrimetric flavoprotein.

The CO dehydrogenase structural genes (*cox*) of *O. carboxidovorans* are clustered in the transcriptional order 5′-coxM-coxS-coxL-3′ along with the Calvin cycle (*cbh*) and hydrogenase (*hox*) genes on a 30-kb DNA segment of the 128-kilobase megaplasmid pHCG3 (10, 12). CO dehydrogenase shows significant sequence similarities to the two domains of Mop (10, 11). CoxL is 57% similar (25% identical) to amino acids 175–907 of Mop, and CoxS is 74% similar (41% identical) to amino acids 1–166 of Mop.

MATERIALS AND METHODS

Protein Crystallization. The protein was crystallized under vapor diffusion conditions by using 0.8 M KH2PO4, 0.8 M NaH2PO4, 2% MPD, and 100 mM Hepes adjusted to pH 7.3. Crystals containing cyanide were cocrystallized in the presence of 4 mM potassium cyanide. For synchrotron data collection under low-temperature conditions, 15% glycerol was added to the crystallization buffer. Before measurements, the concentration of glycerol was gradually increased to 25%. The crystals had an orthorhombic space group *P*2₁2₁2₁ with cell dimensions of *a* = 129.14 Å, *b* = 140.34 Å, and *c* = 164.44 Å and two monomers per asymmetric unit.

Structure Determination. With a first data set to 3.0 Å, a Patterson search was carried out. The program AMORE (13) gave a clear solution for a search model constructed from homologous parts of Mop, which comprised a total of about 10% of the scattering mass of the asymmetric unit [R factor of 30.8%, with a correlation coefficient (cc) of 29.4 after trans-

Abbreviations: MCD, molybdopterin–cytosine dinucleotide; VAO, vanillyl–alcohol oxidase.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID code 1QJ2).
†To whom reprint requests should be addressed. E-mail: dobbek@biochem.mpg.de.
lational search for the second monomer, compared with the next highest peak with an $R$ factor of 52.1% and a cc of 24.7 in a resolution range of 14.0–5.0 Å. Because it was not possible to interpret the calculated density, we initiated multi-wavelength anomalous diffraction data collection at two wavelengths (remote wavelength and iron $K$-edge wavelength) by using synchrotron radiation at BW6, DESY, Hamburg. Data sets native1 and native2 were collected on a MAR-Research (Hamburg, Germany) charge-coupled device detector at $-165^\circ$C at BW6, DESY and processed by using DENZO and SCALEPACK (14). By using an anomalous difference Fourier map calculated with the phases obtained from the Patterson search solution, the positioning of the four [2Fe–2S] clusters could be achieved. The phasing was carried out by using MLPHARE (15) and SHARP (16) and allowed after solvent flattening and histogram matching the recognition of the molecular boundaries. A phase combination with subsequent torsion-angle dynamics, simulated annealing, and maximum likelihood targets for positional refinement as provided by CNS (18).

The cyanide-cocrystallized crystal was measured on a MAR-Research (Hamburg, Germany) 300 image plate detector at room temperature by using monochromatized CuK-α radiation produced by a conventional rotating anode (Table 1).

## RESULTS

### Overall Structure

CO dehydrogenase is dimeric in solution, and two monomers were found in the asymmetric unit. The butterfly-shaped dimer (Fig. 1) has overall dimensions of 145 × 100 × 70 Å with an accessible surface area of 78,076 Å². The two monomers are basically identical, with an rms deviation of 0.30 Å for the Ca atoms.

The redox components of one LMS-structured monomer are the MCD–molybdenum cofactor, composed of a molybdenum ion with two oxo- and one hydroxoligand, complexed by the enedithiolene group of MCD, [2Fe–2S] clusters of type I and type II, and a noncovalently bound FAD molecule (Fig. 4). The iron-sulfur protein of CO dehydrogenase and is clearly divided into a C- and an N-terminal domain, each binding a [2Fe–2S] cluster, which can be distinguished by EPR (8). The C-terminal domain (residues 77–161) carries the proximal [2Fe–2S] cluster. The cluster is buried in CO dehydrogenase $-11$ Å below the protein surface at the interface between the S and the L subunit and is adjacent to the MCD–molybdenum cofactor. The [2Fe–2S] cluster is located at the N terminus of two α-helices that participate in a four-helix bundle of twofold symmetry, an architecture first described for the Mop protein (11).

The N-terminal domain (residues 3–76) is similar to those of plant-type [2Fe–2S]-ferredoxins (20). It carries the distal [2Fe–2S] cluster, which is adjacent to the FAD residing on M. Although the cluster is positioned at the interface between S and M, it is exposed to the solvent and mediates electron transfer from the proximal [2Fe–2S] cluster to the isoalloxazine ring of FAD.

### Structure of the Three Subunits

Subunit S represents the iron-sulfur protein of CO dehydrogenase and is clearly divided into a C- and an N-terminal domain, each binding a [2Fe–2S] cluster, which can be distinguished by EPR (8). The C-terminal domain (residues 77–161) carries the proximal [2Fe–2S] cluster. The cluster is buried in CO dehydrogenase $-11$ Å below the protein surface at the interface between the S and the L subunit and is adjacent to the MCD–molybdenum cofactor. The [2Fe–2S] cluster is located at the N terminus of two α-helices that participate in a four-helix bundle of twofold symmetry, an architecture first described for the Mop protein (11).

Subunit M is the flavoprotein of CO dehydrogenase. It binds FAD in a 1:1 molar stoichiometry (Fig. 5 A and B) and exhibits the motifs 52MAGH53M and 111MTI14M typical of the FAD-binding site of the vanillyl-alcohol oxidase family of flavoproteins. Both motifs are conserved in the molybdenum hydroxylase sequence family (Fig. 5C). M is arranged in three distinct domains (Fig. 5 B and C), designated the N-terminal domain (residues 1–54), the middle domain (residues 60–174), and the C-terminal domain (residues 180–285). The N-terminal domain is composed of a three-stranded parallel β-sheet ($β_1$–$β_3$) flanked by two helices (a1–a2). A loop between a2 and β2 of a βαβ-unit (P loop, residues 32–36) contains one of the two double-glycine motifs, which interact with the pyrophosphate moiety of the FAD molecule. The main structural element of the middle domain is a five-stranded antiparallel β-sheet ($β_4$–$β_8$) that contains an iner-
The C-terminal domain contains a three-stranded antiparallel β-sheet (β₁₀–β₁₂) that merges in a bundle of three α-helices (α₈–α₁₀). It is obvious that the isoalloxazine ring is accessible only from one side of the M subunit near the subunit interface. The central part of the isoalloxazine ring is shielded from the solvent by the side chain of Tyr₁⁹₃M, which is part of an extended loop created between β₁₀ and β₁₁ (Q loop, residues 192–195) of the C-terminal domain (Fig. 5B and C). The distal parts of the dimethyl benzene ring remain solvent accessible. This applies to the C7 and C8 positions with their methyl groups as well as for the C9 position of the isoalloxazine ring. The C4a and N5 atoms may become accessible by movements of the C-terminal domain or the side chain of Tyr₁⁹₃M. The C-terminal domain interacts only with the isoalloxazine ring at its O₄ position.

By using the program Dali (21), two proteins were identified with similarity to the M subunit of CO dehydrogenase with respect to the architecture of structural elements interacting with FAD. MurB (UDP-N-acetylenopyruvylglucosamine reductase) of Escherichia coli (22) showed a Z score of 8.9 by using the fold of M as a search model. The vanillyl-alcohol oxidase (VAO) of Penicillium simplicissimum (23) gave a Z score of 6.0. VAO is a member of a new oxidoreductase family sharing a conserved FAD-binding domain (24). The fold of MurB and VAO involved in binding the dinucleotide moiety of the FAD cofactor resembles the N-terminal and middle domain of the M subunit of CO dehydrogenase. The residues forming two loops responsible for binding the FAD in the CoxM polypeptide are mainly conserved in the molybdenum hydroxylase enzyme family (Fig. 5C) and can also be found as conserved fingerprints in the new VAO family. Both loops have in their central portion two small amino acids, often a double glycine. The first binding motif lying on the P loop can be defined as aGgt, with the sequence 3₂AGGH₃₅ in CoxM. The second motif on α₆ and α₇ can be defined as AxqiqRnx-axtGgn, equivalent to 1₀²ADPOIRYMGTTGGN₁₁₅ in CoxM (Fig. 5C). In the VAO family, a sGHs and a shsG motif are responsible for binding the pyrophosphate moiety of FAD (24). It is likely that other flavin-containing molybdenum hydroxylases exhibit flavin-binding sites similar to that observed in CO dehydrogenase.
The C-terminal domains of the CO dehydrogenase M subunit and its counterparts in MurB and VAO are structurally unrelated and may convey class-specific properties.

Subunit L has a heart-like shape with dimensions of 60 × 83 × 70 Å and contains two subdomains (Fig. 2C). Its N-terminal domain (residues 10–439) interacts with the M and S subunits in the same monomer. Its fold is dominated by a five- and a four-stranded mixed β-sheet. The C-terminal domain (residues 440–809) can be further subdivided into two parts. The first part consists of a central α-helix of 30 aa that runs into a small three-stranded β-sheet followed by an α-helix and a four-stranded parallel–antiparallel β-sheet, which interacts with its symmetrmate of the other monomer, providing the main dimer contact. The second part of the C-terminal domain consists of a long three-stranded antiparallel β-sheet surrounded by seven α-helices.

**Active Site with S-Selanylcyamine.** The active site is accessible from the outside through a narrow channel that is 17 Å deep with an average diameter of 7 Å, determined from the Cβ of the surrounding residues. A restriction of the active-site channel may exclude substrates larger than CO and may prevent dissociation of the cofactor from the protein. The character of the channel is hydrophobic, so that the first charged residues an incoming molecule can interact with are the protein ligands around the molybdenum in its second coordination sphere.

The MCD–molybdenum cofactor is buried at the center of the L subunit (Fig. 2C) and is ligated through a dense network of hydrogen bonds originating from both domains of L. The geometry of the first coordination sphere around the Mo ion can be described as a distorted square pyramid (Fig. 3A). The dithiolene group of MCD is positioned in the equatorial plane of the substrate-analog isopropanol bound in the Mop structure.

Trans to the apical oxo group is a conserved glutamate residue (Glu763L) with a Mo–Oе distance of 3.1 Å. The second coordination sphere contains a modified cysteine residue (Cys388L, Fig. 3 A–C). Strong additional density between Cys388L and the equatorial oxo group was modeled as a SeH bond (ref. 25; unpublished results). The selenium atom of S-selanylcyamine is located in a distance of 3.7 Å from the Mo ion. It is near the equatorial oxo and hydroxo group of the Mo ion. The modified residue is situated at a loop that shows an amino acid composition unique for the CO dehydrogenases, because its sequence of VAYRC388LSFR is also present in the sequence of CO dehydrogenase from *Pseudomonas thermocarboxydovorans* or *Hydrogenophaga pseudoflava* but absent in other enzymes of the molybdenum hydroxylase family. The active-site loop occupies the position of the substrate-analog isopropanol bound in the Mop structure.

**FIG. 3.** The molybdenum site of the L subunit. (A) The Mo ion exhibits one hydroxo and two oxo ligands and is bound to MCD through its enedithiolene group. The second coordination sphere of the Mo ion is composed of the residues Glu763L, Glu763L, and S-selanyl-Cys388L. B and C show the active site together with the adjacent [2Fe–2S] cluster. Difference Fourier maps were calculated with the phases from the final model shifted by 90° (imaginary map) and with amplitudes from the Bijvoet differences at a wavelength λ₁ = 1.7316 Å (B) shown in red and calculated with data obtained at a wavelength of λ₂ = 0.9761 Å (C) shown in yellow. The maps are contoured at 3σ. Calculated f² values using the program ABSORB for selected free atoms are Fe: λ₁, 3.9 e⁻¹ and λ₂, 1.5 e⁻¹; Se: λ₁, 1.41 e⁻¹ and λ₂, 3.84 e⁻¹; Si: λ₁, 0.69 e⁻¹ and λ₂, 0.23 e⁻¹; Mo: λ₁, 3.36 e⁻¹ and λ₂, 1.23 e⁻¹.

**FIG. 4.** Hypothetical reaction mechanism. Product CO₂ may remain in the metal coordination sphere such that the number of oxygen ligands is unchanged during the catalytic steps.
the N-terminal domain is shown in yellow, the middle domain in red, and the C-terminal domain in blue. Tyr193M shields the isoalloxazine ring and interacts also with the riboflavin part. The C-terminal domain (blue box) shows only a single hydrogen bond to FAD. (C) Sequence alignment of the M polypeptide of CO dehydrogenase with the corresponding representative parts from other members of the molybdenum hydroxylase family: CODH, CO dehydrogenase from *Oligotropha carboxidovorans* (coxM, Swall: Q51323); QuinOR, quinoline 2-oxidoreductase from *Pseudomonas putida* (qorM, Swall: P72222); NicotDH, nicotine dehydrogenase from *Arthrobacter nicotinovorans* (ndhA, Swall: Q59127); XDHRC, xanthine dehydrogenase from *Rhodobacter capsulatus* (xdhA, Swall: O54050); XDHBM, xanthine dehydrogenase from *Bombylia mori* (Swall: Q17209); XDHCV, xanthine dehydrogenase from *Calliphora vicina* (xdh, Swall: XDH CALVI); XDHDM, xanthine dehydrogenase from *Drosophila melanogaster* (xdh, Swall: XDH DROME). Two boxes highlighting residues responsible for FAD binding in the M subunit are shown in red and blue. They correspond to fingerprints for FAD-binding by enzymes of the molybdenum hydroxylase family and define a FAD-binding motif similar to the motifs found for the vanillyl alcohol oxidase family of flavoproteins. The secondary structure is represented by arrows for β-sheets and bars for α-helices, in domain colors (B). Residues appearing at least three times in a row are colored in green. The position of Tyr193M is marked by a cyan circle below the aligned sequences.

**DISCUSSION**

**Hypothetical Reaction Mechanism.** The chemistry of carbon oxide selenide (O=C=Se) and the selenium-catalyzed synthesis of carbonates from CO (26) show remarkable analogies to the CO dehydrogenase-catalyzed oxidation of CO to CO$_2$. SeCO is a gas, which is easily prepared through the oxidation of CO with elemental selenium in the presence of a strong base at normal pressure and room temperature (26). When SeCO is treated with a sodium alcolholate (RONa) under mild conditions, the corresponding sodium O-alkyl selenocarbonate (ROCOSeNa) is formed, which undergoes nucleophilic attack by the alcohol, yielding the corresponding dialkyl carbonate (RO$_2$C==O). In aqueous sodium hydroxide (r = H), carbon oxide selenide reacts to carbonate and selenide (COSe + 4 NaOH → Na$_2$Se + Na$_2$CO$_3$ + 2 H$_2$O).

The active site of CO dehydrogenase literally assembles all components essential in the selenium-catalyzed synthesis of carbonates (Fig. 3). The selenium of the S-selenylcysteine, is formally elemental Se with the redox state 0 (Fig. 4). It reacts with cyanide, yielding selenocyanate, with the concomitant irreversible inactivation of CO dehydrogenase under oxic or anoxic conditions (half-life of 1 h, complete inactivation after 10 h). Enzyme crystallized in the presence of 4 mM potassium cyanide showed a reduced occupancy, with selenium at the S-selenylcysteine residue in the structure, and the remaining selenium now appears in a small linear molecule interpreted as selenocyanate. The molecule lies in the equatorial plane of the oxygen ligands, and the selenium is in hydrogen-bond distance to the amide nitrogen atoms of Cys388L and Ala385L. Because CO and cyanide are isoelectronic, isosteric, and show a similar chemical reactivity, we assume a similar reaction between CO and selenium, leading to a carbonyl selenide that reacts with water to yield CO$_2$ (Fig. 4).

The catalytic mechanism of CO dehydrogenase thus can be imagined to involve the formation of carbon oxide selenide with the subsequent nucleophilic attack by the Mo–OH group, yielding CO$_2$ (Fig. 4). Catalysis in CO dehydrogenase starts by the reaction of CO with the selenium of S-selenylcysteine, thereby forming a selenocarboxyl species (Fig. 4). The selenocarboxyl undergoes a nucleophilic attack by the Mo–OH group of the molybdenum, which is in the oxidized +VI state. Subsequently, two electrons are abstracted, and CO$_2$ is released. Electron transfer from the reduced Mo ion in the +IV state to the proximal [2Fe–2S] center and the addition of water regenerates the oxidized Mo–OH in the +VI oxidation state.

The transport of the electrons from the active-site Mo ion to the external electron acceptors is facilitated through a chain of
redox active cofactors composed of two [2Fe–2S] clusters and FAD. The MCD–molybdenum cofactor is adjacent to the proximal [2Fe–2S] cluster located in the C-terminal domain of the S subunit (Fig. 2D). The molybdopterin ring is positioned between the molybdenum and the [2Fe–2S] cluster and supports electron transfer through its conjugated double bonds. The N2’ amino group of the fused pyrimidin ring is at a distance of 5.5 Å from the Fe1 atom of the proximal cluster.

CO Dehydrogenase as a Member of the Xanthine Oxidase Family. The CO dehydrogenase reveals significant sequence homologies with other members of the molybdenum hydroxylase family, suggesting structural conservation and represents, therefore, a close structural model of the eukaryotic xanthine oxidase/dehydrogenase in tertiary and quaternary structure. Mop has served as a first approximation to the structure of the xanthine oxidase/dehydrogenase (11), but lacked the flavin domain that is now defined in CO dehydrogenase by the M subunit. Its location differs from the position of the connecting loop in Mop suggested to be a marker (11). Major gaps in sequence alignments between the enzymes are due to the fact that most of the xanthine oxidases/dehydrogenases, as well as Mop, contain connecting segments between their structural domains, which are missing in the heterotrimeric CO dehydrogenase.

The members between CO dehydrogenase and other members of the molybdenum hydroxylase family extends beyond to residues in the core of the domains/subunits to the interdomain/subunit contacts. In addition, the arrangement of the L and the S subunit of CO dehydrogenase is identical to the arrangement of the corresponding parts of Mop (Fig. 2).

The major functional difference between the interconvertible forms of xanthine oxidase and xanthine dehydrogenase is their usage of electron acceptors. These different reactivities may be correlated to the ability of binding NAD+/NADH. Affinity-labeling studies with the pyridine nucleotide analog 5′-p-fluorosulfonylbenzoyladenosine (5′-FSBA) (27) hinted to a tyrosine residue affecting NAD+ binding. Reactivity studies of xanthine dehydrogenase variants from rosy mutant strains of Drosophila melanogaster comprised mutants that showed a diminished affinity for NAD+ or were unable to use NAD+ as an electron acceptor (28). Through sequence comparisons between the xanthine oxidase/dehydrogenase sequences from different sources with the sequence of the CO dehydrogenase M subunit, we localized in our model the positions of these mutations and the labeled tyrosine residue. They correspond to Gly149M, Asn123M and Ala156M in CoxM (Fig. 5B and C). All of these residues cluster around two loops of the M subunit near the solvent-exposed side of the isoalloxazine ring of the FAD. Their mutation may alter the ability to bind NAD+. The likely docking site of NAD+ suggests hydride transfer from the isoalloxazine ring to a pyridine ring positioned at the si face. This unusual stereochemistry was also suggested for the interaction of MurB with NADP+ (29).

We thank Dr. H. Bartunik and G. B. Bourenkov at the BW6, DESY, Hamburg for their help in data collection and multilwavelength anomalous diffraction measurements. This work was supported by Grant Me 732/5-3 from the Deutsche Forschungsgemeinschaft (Bonn-Bad Godesberg, Germany) to O.M.