Characterization of a CO-responsive Transcriptional Activator from Rhodospirillum rubrum*

Yiping He, Daniel Shelvert, Robert L. Kerby†, and Gary P. Roberts‡
From the Departments of Bacteriology and Biochemistry, University of Wisconsin, Madison, Wisconsin 53706

In Rhodospirillum rubrum, CO induces the expression of at least two transcripts that encode an enzyme system for CO oxidation. This region is positively regulated by CooA, which is a member of the cAMP receptor protein family of transcriptional regulators. The transcriptional start site of one of the transcripts (cooSCTJ) has been identified by primer extension. The ability of CooA to bind to this promoter was characterized with DNase I footprinting experiments using extracts of a CooA-overproducing strain. CooA and CO-dependent protection was observed for a region with 2-fold symmetry (5′-TGTCA-N4CGACA) that is highly similar to the consensus core motifs recognized by cAMP receptor protein/FNR family. In vivo analysis in a heterologous background indicates that CooA is sufficient for CO-dependent expression, implicating it as the likely CO sensor.

Exposure of the purple nonsulfur bacterium Rhodospirillum rubrum to CO stimulates the expression of the coo regulon, which consists of at least two transcriptional units. Among the products of this regulon are a carbon monoxide dehydrogenase (CooS), an Fe-S protein (CooF), and a hydrogenase (CooH), where the two former proteins have been purified and characterized (1–6). This CO-oxidizing system functions under anaerobic conditions to oxidize CO to CO₂, allowing growth on CO as sole energy source (7). The cooSCTJ region has been cloned, sequenced, and mutationally characterized, verifying the requirement for the encoded products for oxidation of CO (7, 8).

The mutational analysis has also indicated that cooSCTJ is organized in a single transcriptional unit. CooH lies at the 3′ terminus of the other known CO-regulated transcript, but this transcript has not yet been fully sequenced at the 5′ end. CooH is located 5′ of cooF and is separated from it by 450 nucleotides of noncoding DNA (Fig. 1B).

Our previous mutational studies revealed that CooA, which is apparently encoded on its own transcript on the 3′ side of cooSCTJ, is essential for the expression of the coo regulon of R. rubrum in response to CO (9). The sequence of CooA predicts that it is a member of the CRP/FNR family of transcriptional regulators (9), with a putative DNA-binding domain that is highly similar to that found in CRP and FNR. Modeling the sequence of CooA on the known CRP crystal structure (10) predicts the presence of four Cys and one His residues adjacent to the region known to bind cAMP in CRP (9). These residues suggest the possibility that CooA contains a metal center at this position, which might be expected if CooA binds CO. A particularly interesting question in this area is how the binding of a molecule as small as CO might induce a similar conformational change in CooA as that caused by cAMP binding in CRP.

To test the model of CooA as a CO-binding transcriptional activator, we have sought evidence for CO- and CooA-dependent DNA binding. The results described herein support the above model, and the assay of DNA-binding activity of CooA will aid in the purification of CooA for more direct analysis.

EXPERIMENTAL PROCEDURES

Growth of Bacterial Strains—R. rubrum strains were grown photoheterotrophically in SMN (supplemented malate-ammonium) medium supplemented with 10 μM NiCl₂ in stoppered serum vials with an argon head space (8).

For RNA isolation, cultures to be CO-induced were grown photoheterotrophically to an optical density of 1 at 680 nm, whereupon CO was added to a final concentration of 30% uninduced culture received no additions. The cultures were agitated under illumination (8) for 6 h.

The coo-overexpressing strain (UR459) was grown under nitrate-repression conditions in malate-glutamate medium (11) to an OD₆₅₀ of 2.0, and the expression of nifH promoter was monitored by nitrogenase activity (12).

RNA Extraction—Total cellular RNA was isolated by repeated phenol extraction as described previously (13) with the following modifications. After addition of 1.5 ml of lysis solution to 6 ml of culture, the mixture was boiled for 90 s and extracted with phenol 3 times. Purified RNA was dissolved in 100 μl of 10 mM Tris-Cl (pH 8.0).

†To whom correspondence should be addressed. Tel.: 608-262-3567; Fax: 608-262-9865; E-mail: groberts@bact.wisc.edu.
‡R. L. Kerby and G. P. Roberts, unpublished data.
§The abbreviations used are: CooA, CO oxidation activator; CRP, cAMP receptor protein; bp, base pair(s); MOPS, 5-morpholinepropane-sulfonic acid.

* This work was supported by the College of Agricultural and Life Sciences, University of Wisconsin, Madison; by Department of Energy Grant 94ER13691; and by National Institutes of Health Grant GM53226. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

†R. L. Kerby and G. P. Roberts, unpublished data.
‡L. Leman and G. P. Roberts, unpublished data.
§The abbreviations used are: CooA, CO oxidation activator; CRP, cAMP receptor protein; bp, base pair(s); MOPS, 5-morpholinepropane-sulfonic acid.
CooA, a CO-responsive Activator

**Results**

Location of a CO-induced 5'-End mRNA of cooA — In order to identify the coo promoter upstream of cooA and to determine the effect of CO on its expression, we performed primer extension analysis on coo mRNA from both CO-induced and uninduced wild-type R. rubrum (strain UR2). Initially, we used a primer (Primer 1) that hybridized near the beginning of cooF coding region. The extension product suggested that the 5'-end of the cooF mRNA was approximately 250 nucleotides upstream from the translational start site of cooF. This signal was present only in the CO-induced sample, and no other primer extension products were observed (data not shown).

In order to more precisely identify the 5'-end of the cooF mRNA, a second primer (Primer 2) was designed that hybridized about 150 nucleotides upstream of cooF coding region. Results with this primer showed that the major transcript from the cooF promoter initiates with the A nucleotide positioned 257 bp upstream from the start codon of cooF. A minor product starting six nucleotides upstream of that site was also ob-

**Fig. 1. Identification of the translational start site for cooF.** Panel A shows the result of primer extension of the region upstream of cooF. A 22-mer oligonucleotide (Primer 2), which is complementary to the coding strand in region 154–133 relative to the translational start site of cooF was 5'-end-labeled and used to prime the reverse transcriptase reaction. The sequencing ladder (G, A, T, C) used the medium supplemented with 10 mM MgCl₂, 1 mM dithiothreitol, 1 mM diithiothreitol and passed through a French press at 16,000 p.s.i. The mixture was centrifuged at 13,200 × g for 1 h to remove the debris. The supernatant, to which glycerol was added to 5% (v/v), was used for the DNA-binding assay. The total protein concentration was determined by the Bradford dye binding procedure (22) using the protein assay dye reagent concentrate (Bio-Rad).

DNase I Footprinting — A 294-bp EcoRV-EagI fragment (Fig. 1B), which contains the PcooF region, was used for both DNase I footprinting and gel retardation assays. This fragment was isolated from a pBSKⅡ(−)-derivative (Stratagene), pC017, which contains a portion of the cooF region from pl. C24 (23). The fragment was uniquely 3'-end-labeled in the coding strand by filling the EagI end with [-32P]dGTP and Sequenase, followed by purification via polyacrylamide gel electrophoresis (14) and an Elutip Minicolumn (Schleicher & Schuell). DNase I footprinting analysis of the PcooF region was performed as described previously (24) with the following modifications. The sealed tubes used for the protein-DNA binding reaction were degassed, and the head space was filled with argon. The DNA binding reactions were done under stringently anoxic conditions in the buffer B (20 mM Tris-HCl (pH 7.6), 7 mM MgCl₂, 50 mM KCl, 7 mM diithiothreitol, 50 μM bovine serum albumin, 5% (w/v) glycerol) supplemented with 1.7 mM sodium dithionite in the presence or absence of CO. After 5–10 min of DNA was incubated with various amounts of extracts in a 20-μl volume at room temperature for 20 min, the mixtures were treated with 2 units/ml RNase-free DNase I (Promega) for 30 s. DNase I cleavage products were separated on a 6% (w/v) polyacrylamide-urea gel.

Gel Retardation Assay — Radiolabeled DNA fragments were prepared similar to those in the footprinting experiment. The DNA-binding reactions were performed anoxically in the same buffer and at the same temperature used in the footprinting assays. The samples were applied to a 6% (w/v) polyacrylamide-urea gel (37:1.5) that was run in standard 1 × Tris borate EDTA buffer (14) with 1.7 mM sodium dithionite. After prerunning at 180 V for 1.5 h, the upper running buffer was changed once in order to maintain strictly anoxic conditions. The electrophoresis was performed at 180 V for 3 h in 4°C. RESULT

Location of a CO-induced 5'-End mRNA of cooA — In order to identify the coo promoter upstream of cooA and to determine the effect of CO on its expression, we performed primer extension analysis on coo mRNA from both CO-induced and uninduced wild-type R. rubrum (strain UR2). Initially, we used a primer (Primer 1) that hybridized near the beginning of cooF coding region. The extension product suggested that the 5'-end of the cooF mRNA was approximately 250 nucleotides upstream from the translational start site of cooF. This signal was present only in the CO-induced sample, and no other primer extension products were observed (data not shown).

In order to more precisely identify the 5'-end of the cooF mRNA, a second primer (Primer 2) was designed that hybridized about 150 nucleotides upstream of cooF coding region. Results with this primer showed that the major transcript from the cooF promoter initiates with the A nucleotide positioned 257 bp upstream from the start codon of cooF. A minor product starting six nucleotides upstream of that site was also ob-

**Fig. 1. Identification of the translational start site for cooF.** Panel A shows the result of primer extension of the region upstream of cooF. A 22-mer oligonucleotide (Primer 2), which is complementary to the coding strand in region 154–133 relative to the translational start site of cooF was 5'-end-labeled and used to prime the reverse transcriptase reaction. The sequencing ladder (G, A, T, C) used the medium supplemented with 10 mM MgCl₂, 1 mM dithiothreitol, 1 mM diithiothreitol and passed through a French press at 16,000 p.s.i. The mixture was centrifuged at 13,200 × g for 1 h to remove the debris. The supernatant, to which glycerol was added to 5% (v/v), was used for the DNA-binding assay. The total protein concentration was determined by the Bradford dye binding procedure (22) using the protein assay dye reagent concentrate (Bio-Rad).

DNase I Footprinting — A 294-bp EcoRV-EagI fragment (Fig. 1B), which contains the PcooF region, was used for both DNase I footprinting and gel retardation assays. This fragment was isolated from a pBSKⅡ(−)-derivative (Stratagene), pC017, which contains a portion of the cooF region from pl. C24 (23). The fragment was uniquely 3'-end-labeled in the coding strand by filling the EagI end with [-32P]dGTP and Sequenase, followed by purification via polyacrylamide gel electrophoresis (14) and an Elutip Minicolumn (Schleicher & Schuell). DNase I footprinting analysis of the PcooF region was performed as described previously (24) with the following modifications. The sealed tubes used for the protein-DNA binding reaction were degassed, and the head space was filled with argon. The DNA binding reactions were done under stringently anoxic conditions in the buffer B (20 mM Tris-HCl (pH 7.6), 7 mM MgCl₂, 50 mM KCl, 7 mM diithiothreitol, 50 μM bovine serum albumin, 5% (w/v) glycerol) supplemented with 1.7 mM sodium dithionite in the presence or absence of CO. After 5–10 min of DNA was incubated with various amounts of extracts in a 20-μl volume at room temperature for 20 min, the mixtures were treated with 2 units/ml RNase-free DNase I (Promega) for 30 s. DNase I cleavage products were separated on a 6% (w/v) polyacrylamide-urea gel.

Gel Retardation Assay — Radiolabeled DNA fragments were prepared similar to those in the footprinting experiment. The DNA-binding reactions were performed anoxically in the same buffer and at the same temperature used in the footprinting assays. The samples were applied to a 6% (w/v) polyacrylamide-urea gel (37:1.5) that was run in standard 1 × Tris borate EDTA buffer (14) with 1.7 mM sodium dithionite. After prerunning at 180 V for 1.5 h, the upper running buffer was changed once in order to maintain strictly anoxic conditions. The electrophoresis was performed at 180 V for 3 h in 4°C.

RESULTS

Location of a CO-induced 5'-End mRNA of cooA — In order to identify the coo promoter upstream of cooA and to determine the effect of CO on its expression, we performed primer extension analysis on coo mRNA from both CO-induced and uninduced wild-type R. rubrum (strain UR2). Initially, we used a primer (Primer 1) that hybridized near the beginning of cooF coding region. The extension product suggested that the 5'-end of the cooF mRNA was approximately 250 nucleotides upstream from the translational start site of cooF. This signal was present only in the CO-induced sample, and no other primer extension products were observed (data not shown).

In order to more precisely identify the 5'-end of the cooF mRNA, a second primer (Primer 2) was designed that hybridized about 150 nucleotides upstream of cooF coding region. Results with this primer showed that the major transcript from the cooF promoter initiates with the A nucleotide positioned 257 bp upstream from the start codon of cooF. A minor product starting six nucleotides upstream of that site was also ob-
CooA, A CO-responsive Activator

served. These primer extension products were only detectable in the CO-induced culture (Fig. 1A), indicating that the effect of CO is on the accumulation of coo mRNA. Fig. 1B shows a schematic of the transcription start site relative to other features in the region, including the putative CooA-binding site (see below).

CooA Is Sufficient for CO-dependent Expression in a Heterologous System—As noted above, mutant analysis has shown that CooA is necessary for CO-dependent expression (9). To test the possibility that CooA was sufficient for the response, consistent with the hypothesis that it was actually the CO sensor, two derivatives of plasmid pRK were created. The first carried cooSCTJ (pCO46R), and the second contained cooSCTJ A (pCO47R). These two constructs were introduced into E. coli and R. sphaeroides and were tested for their ability to produce CO dehydrogenase in response to CO. The E. coli strains showed no detectable CO dehydrogenase activity by the plate overlay assay (data not shown). Likewise, the R. sphaeroides strain (UR453) with pCO46R and the parent strain (UR363) accumulated no CO dehydrogenase activity detectable by either the plate overlay assay or the in vitro spectrophotometric assay. In contrast, the R. sphaeroides strain with pCO47R, upon CO induction, produced detectable CO dehydrogenase (0.23 μmol of CO oxidized/min·OD₅₄₆) in the spectrophotometric assay. Because R. sphaeroides lacks an endogenous coo system, it is likely that CooA is the CO sensor in this heterologous system.

Overexpression of CooA in R. rubrum—We expected that CooA, as a regulatory protein, would be nonabundant and thus anticipated that its activity would be difficult to detect in crude extracts. Consequently, we overexpressed CooA in R. rubrum from the R. rubrum nifH promoter and ribosome binding site. This promoter, which is active when R. rubrum is grown under nitrogen-fixing conditions, has been successfully used in our lab to overproduce nonabundant regulatory proteins (19) and has the potential to produce 1% of total cell protein. Construction of the P₅₄₄::cooA fusion and its integration into R. rubrum, creating strain UR459, are described under “Experimental Procedures.”

Extracts of UR459 (P₅₄₄::cooA) were examined by SDS-polyacrylamide gel for the presence of a protein band corresponding to CooA. The extracts of UR2 and UR407 (cooA::aacC1) (9), grown under the same conditions as UR459, were used as controls, as CooA was deficient in UR407 and is not expected to be detectable in UR2. A band migrating at about 25 kDa, the predicted molecular mass of CooA, was significantly more intense in extracts of UR459 than in those of UR2 or UR407 (data not shown).

CooA- and CO-dependent DNA Binding—To test the hypothesis that CooA is a CO-sensing transcriptional activator, in vitro interactions between CooA and the promoter region of cooF were investigated with or without CO. A 294-bp EcoRV-EagI fragment containing the promoter region of cooF was used in this assay. CooA− refers to extracts of UR407 (cooA::aacC1) and CooA+ refers to extracts of UR459 (the CooA-overproducer with cooA under PnifH control). The numbers reflect micrograms of protein in each assay, and the and + on the line labeled CO reflect the presence or absence of CO in the binding reaction; all experiments were performed anoxically. The G+A lane represents the Maxam-Gilbert sequencing marker. The box on the right side indicates the region protected in this experiment. Panel B shows a comparison of the CRP and FNR consensus binding sites with the detected CooA-binding site.

**Fig. 2. Identification of the CooA binding site.** Panel A shows the result of a DNase I footprinting experiment. A 294-bp EcoRV-EagI fragment containing the promoter region of cooF was used in this assay. CooA− refers to extracts of UR407 (cooA::aacC1) and CooA+ refers to extracts of UR459 (the CooA-overproducer with cooA under PnifH control). The numbers reflect micrograms of protein in each assay, and the and + on the line labeled CO reflect the presence or absence of CO in the binding reaction; all experiments were performed anoxically. The G+A lane represents the Maxam-Gilbert sequencing marker. The box on the right side indicates the region protected in this experiment. Panel B shows a comparison of the CRP and FNR consensus binding sites with the detected CooA-binding site.

Gel retardation analysis with the same DNA fragment and extracts also revealed a DNA-protein complex whose presence requires both CooA and CO (data not shown). This CooA- and CO-dependent complex is very large; it remained in the wells of a 5% polyacrylamide (19:1 acrylamide/bis ratio) gel but entered a 5% polyacrylamide (37:5:1 acrylamide/bis ratio) gel, suggesting the presence of additional proteins in this complex.

**DISCUSSION**

Our previous mutational studies and sequence analysis of cooA led us to predict that CooA is a CO-sensing transcriptional activator similar to CRP and FNR. The work presented in this paper and elsewhere strongly supports the hypothesis that CooA is a CO-sensing protein responsible for controlled expression of the coo region in a fashion reminiscent of the action of CRP: (i) Northern blot analysis and primer extension experiments demonstrate that CooA affects mRNA accumulation; (ii) CooA is sufficient for CO-dependent expression in R. sphaeroides; (iii) DNA binding appears to be CooA- and CO-dependent in vitro; (iv) the detected CooA target site is very similar to
the CRP/FNR consensus binding site; and (v) CooA is very similar to CPR and FNR in the helix-turn-helix DNA binding domain (9).

We initially looked for the CooA-binding site by both footprinting and gel retardation assays in the 250-bp Eagl-BsmI region immediately upstream of cooF (Fig. 1B), in part because a strain (UR284) with an insertion at the CooA with 25 and region immediately upstream of cooF displayed CooS activity in a CO-dependent manner (8). We now believe that the observed expression in UR284 reflects transcription from PcooF through the Kan insertion; the Kan gene is derived from pUC4K and apparently lacks transcriptional terminators. No CooA- or CO-dependent DNA binding could be found between the Eagl site and cooF (data not shown), however, and the transcriptional start site identified in this paper is clearly the physiologically significant one in vivo.

While the detected transcription start lacks the −10 and −35 sequences expected at a typical E. coli σ70 promoter, a typical E. coli σ54 recognition sequence is present, with GC and GG at −13 and −25, respectively (25) (Fig. 1B). An interaction of CooA with σ54 would be interesting as no CRP- or FNR-controlled promoters are known to be recognized by σ54.

The center of the two-fold symmetry of the CooA binding site is at −43.5 with respect to the transcriptional start point of cooF. This distance is similar to the location of the CRP sites in class II CRP-dependent promoters (e.g. galP1) and FNR sites in the FNR-dependent promoters (26, 27). In CRP, the specific interaction between the side chains of the protein and a given base within the core motif 5′-TGTGA-3′ have been reviewed (28, 29): Arg-180 and Glu-181 directly contact the 5′-G and the nucleotide of complementary to the 3′-G, respectively. It is possible that Arg-177 of CooA, which is in the homologous position of Arg-180 of CRP, contacts the 5′-G of the CooA target site 5′-TGTCA-3′. The absence of a Glu in CooA corresponding to Glu-181 in CRP is consistent with the fact that there is a 3′C instead of a 3′-G in the CooA target site (Fig. 2B).

We are currently employing the described gel shift as a functional assay for the purification of CooA. Analysis of the purified protein, together with the eventual determination of the CooA-regulated promoter upstream of cooH, will significantly increase our understanding of this regulatory response.

Acknowledgments—We thank Marcin Filutowicz for valuable discussions and Holly Simon, Jon Roll, Mary Homer, Yaoping Zhang, and Marjeta Urh for technical assistance and helpful discussions.

REFERENCES
1. Bonam, D., Lehman, L., Roberts, G. P., and Ludden, P. W. (1989) J. Bacteriol. 171, 3102–3107
2. Bonam, D., and Ludden, P. W. (1987) J. Biol. Chem. 262, 2980–2987
3. Bonam, D., McKeena, M., Stephens, P., and Ludden, P. W. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 31–35
4. Ensign, S. A., Bonam, D., and Ludden, P. W. (1989) Biochemistry 28, 4968–4973
5. Ensign, S. A., Hyman, M. R., and Ludden, P. W. (1989) Biochemistry 28, 4973–4979
6. Ensign, S. A., and Ludden, P. W. (1991) J. Biol. Chem. 266, 18395–18403
7. Kerby, R. L., Ludden, P. W., and Roberts, G. P. (1995) J. Bacteriol. 177, 2241–2244
8. Kerby, R. L., Hong, S. S., Ensign, S. A., Coppoc, L. J., Ludden, P. W., and Roberts, G. P. (1992) J. Bacteriol. 174, 5284–5294
9. Shelver, D., Kerby, R. L., He, Y.-P., and Roberts, G. P. (1995) J. Bacteriol. 177, 2157–2163
10. Weber, I. T., and Steltz, T. A. (1987) J. Mol. Biol. 198, 311–326
11. Lehman, L., and Roberts, G. P. (1991) J. Bacteriol. 173, 5705–5711
12. Kanemoto, R. H., and Ludden, P. W. (1984) J. Bacteriol. 158, 713–720
13. Sarmientos, P., Sylvester, J. E., Contente, S., and Cashel, M. (1986) Cell 32, 1337–1346
14. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
15. Josaitis, C. A., Gali, T., and Gourse, R. L. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 1117–1121
16. Schweiger, H. P. (1993) BioTechniques 15, 631–633
17. Simon, R., Priefer, U., and Hübner, A. (1983) Biol. Technology 1, 784–791
18. Mackenzie, C., Chidambaram, M., Sodergren, E. J., Kaplan, S., and Weinstock, G. M. (1995) J. Bacteriol. 177, 3027–3033
19. Grunwald, S. K., Lies, D. P., Roberts, G. P., and Ludden, P. W. (1998) J. Bacteriol. 177, 628–635
20. Lehman, L., Fitzmaurice, W. P., and Roberts, G. P. (1990) Gene (Amst.) 95,143–147
21. Win, P. D., and Nickoloff, J. A. (1992) Anal. Biochem. 200, 81–88
22. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254
23. Coppoc, L. J. (1991) Studies Concerning the Carbon Monoxide Dehydrogenase of Rhodospirillum rubrum. M.Sc. thesis, University of Wisconsin, Madison, WI
24. Filutowicz, M., Uhlenhopp, E., and Helsen, D. R. (1985) J. Mol. Biol. 187, 225–239
25. Morett, E., and Budz, M. (1989) J. Mol. Biol. 210, 65–77
26. Elbright, R. H. (1993) Mol. Microbiol. 8, 797–802
27. Elgmüller, K., Honore, N., Luchi, S., Lin, E. C. C., and Cole, S. T. (1989) Mol. Microbiol. 3, 869–878
28. Schultz, S. C., Shields, G. C., and Steltz, T. J. (1991) Science 253, 1001–1007
29. Kolb, A., Busby, S., Buc, H., Garges, S., and Adhya, S. (1993) Annu. Rev. Biochem. 63, 749–795
