JOURNAL OF BACTERIOLOGY, Jan. 1993, p. 568–571
0021-9193/93/020568-04$02.00/0
Copyright © 1993, American Society for Microbiology

Vol. 175, No. 2

Characterization of the *gcd* Gene from *Escherichia coli* K-12 W3110 and Regulation of Its Expression

MAMORU YAMADA,[1]* SHINTA ASAOKA,[1] MILTON H. SAIER, JR.,[2] AND YASUE YAMADA[3]

Department of Biological Chemistry, Faculty of Agriculture, Yamaguchi University, Yamaguchi 753,[1] *and Department of Pharmacology, Yamaguchi University School of Medicine, Yamaguchi 755,*[3] *Japan, and Department of Biology, University of California at San Diego, La Jolla, California 92093-0116*[2]

Received 18 August 1992/Accepted 3 November 1992

DNA sequence and expressional analyses of the *gcd* gene of *Escherichia coli* K-12 W3110 revealed that two promoters that were detected were regulated negatively by cyclic AMP and positively by oxygen. Sequence conservation of the *gcd* gene between *E. coli* K-12 W3110 and PPA42 suggests that glucose dehydrogenase is required for the *E. coli* cells, even though it ordinarily exists as an apoprotein.

In *Escherichia coli*, glucose is usually imported by the glucose-specific permease of the phosphoenolpyruvate-dependent phosphotransferase system (7, 14, 16). Membrane-bound glucose dehydrogenase (EC 1.1.99.17) is synthesized in aerobically grown cells, but its biological role for glucose utilization is not clear because of its occurrence as an apoprotein under ordinary conditions (1, 8). When pyrroloquinoline quinone is supplied as the prosthetic group, the holoenzyme becomes functional and feeds electrons into the respiratory chain to form the proton motive force (20). Here, we determined the nucleotide sequence of the coding region of the *gcd* gene as well as that of the upstream operator-promoter region, and we examined *gcd* transcriptional regulation. The strains and plasmids used are listed in Table 1.

Organization of the *gcd* gene. Although the coding region of glucose dehydrogenase has been cloned from *E. coli* PPA42 (4), the 5′-flanking region, including the operator-promoter region, has not been cloned. We therefore tried to clone the entire *gcd* gene for analyzing the regulation of its expression. The gene was subcloned from a λ phage clone, 4E11, from the Kohara library, which was constructed from *E. coli* K-12 W3110 (9) on the basis of previous work (4). The cloning was confirmed by measuring glucose dehydrogenase activity. Cells harboring pUCGCD1 exhibited a nearly 100-fold increase in activity compared with that of cells harboring the vector.

The entire nucleotide sequence of the *gcd* gene was determined, and the sequence around the 5′-flanking region is presented in Fig. 1. To define the promoter region of the *gcd* gene, the mRNA initiation site was determined by reverse transcriptase mapping. Two predominant bands were detected under various conditions (Fig. 2, lanes 2, 5, and 7). There are typical −10 and −35 sequences of the *E. coli* RNA polymerase bearing σ^{70} (15) for the lower band, but the sequence for the upper band was less homologous to the consensus sequence (Fig. 1). The lower band may not be the degradation product of the upper one, because it possesses a typical promoter sequence and it is under fine regulation, as will be described later. These results suggest that there are two promoters (P1 and P2 in Fig. 1) preceding the *gcd*-coding region. The sequence (5′-AATTGTGAT GACGATCACACAT-3′) homologous to the binding site (consensus sequence: 5′-AAATGTGATCTAGATCACAT TT-3′) of the cyclic AMP (cAMP) receptor protein (2, 6, 19) was found as described in the legend to Fig. 1. Two strongly conserved 5-bp regions (TGTGA and TCACA) were perfectly conserved in the possible cAMP receptor protein binding site. The binding site, which mainly plays a positive role in the expression of a variety of genes in *E. coli*, overlapped with the *gcd* promoters.

After the possible ribosome-binding site, 5′-ATGGTGT-3′ (18) at positions −13 to −7, was found, the open reading

TABLE 1. Strains and plasmids

Strain or plasmid	Genotype or description	Reference or source
Strains[a]		
W3110	Wild type	M. Yamada
CA8306	Hfr Hayes Sm^s B1$^-$ Δ*cya*	J. Beckwith
CC118	*araD139* Δ(*ara-leu*)*7697* Δ(*lac*)*X74 galK thi rpsE rpoB argE*(Am) *phoA*Δ*20 galE recA1*	11
JM103	Δ(*lac-pro*) *thi strA supE endA sbcB hspR4* F′ *traD36 proAB lacI*^{q}*Z*Δ*M15*	12
Plasmids		
pUC118	Amp^r	21
pUI310	pUC19 with *'phoA*	3
pUY311	pUC118 with *'phoA*	This study
pUCGCD1	pUC118 with *gcd*	This study
pGCDAP1	pUY311 with *gcd-Eco*RI-*Eco*RV fragment	This study
pGCDAP3	pUY311 with *gcd-Dra*I-*Sal*I fragment	This study
pGCDAP5	pUY311 with *gcd*-PCR[b] fragment (−104/+449)	This study
pGCDAP7	pUY311 with *gcd*-PCR fragment (−387/+449)	This study
pGCDCM1	pGCDAP1 with insertion of Cm^r gene into *gcd*	This study
pGCDCM3	pGCDAP3 with insertion of Cm^r gene into *gcd*	This study
pGCDCM5	pGCDAP5 with insertion of Cm^r gene into *gcd*	This study
pGCDCM7	pGCDAP7 with insertion of Cm^r gene into *gcd*	This study

[a] All were *E. coli* strains.
[b] PCR, polymerase chain reaction.

* Corresponding author.

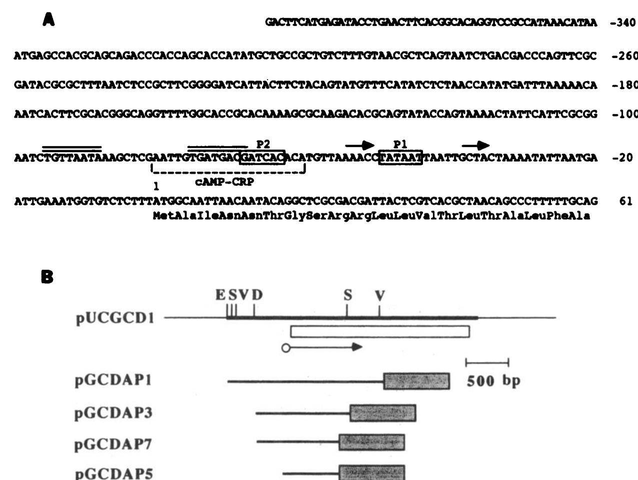

FIG. 1. Nucleotide sequence in the 5′-flanking region of the *gcd* gene and construction of gene fusions between *gcd* and *phoA* genes. (A) Nucleotide sequences of both strands were determined by the dideoxy-chain termination method (17) after DNA fragments were cloned into M13mp18 or M13mp19 (25). The deduced amino acid sequence at the N-terminal portion of glucose dehydrogenase is shown, and more of the sequence is available as described in the text. Promoter sequences (P1 and P2) of the gene are indicated by boxes (−10 sequences) and double overlines (−35 sequences). Arrows indicate the mRNA initiation sites and the direction of transcription. The sequence homologous to the cAMP-cAMP receptor protein binding site is indicated by a dotted bracket. (B) Gene fusions were constructed by insertion of the DNA fragments of the *gcd* gene from pUCGCD1 into pUY311. DNA manipulation and blunt-end formation of the staggered-end DNA fragments were carried out as described previously (10, 22). pGCDAP1 and pGCDAP3 included the *Eco*RI-*Eco*RV and *Dra*I-*Sal*I DNA fragments of the *gcd* gene, respectively, and pGCDAP5 and pGCDAP7 included polymerase chain reaction fragments corresponding to positions −104 to +449 and positions −387 to +449, respectively. Primers for polymerase chain reaction were 5′-ATGGATCCTTC ATGAGATACCTGA-3′ (positions −387 to −371), 5′-CCGGATCC GCGGAATCTGTTAATA-3′ (positions −104 to −88), and 5′-AAG GATCCAGGTGCCCGTTGATCT-3′ (positions +449 to +434). These primers contained a *Bam*HI site and two additional nucleotides at their 5′ portions. Restriction sites preceding and within the *gcd* gene are shown at the top. E, *Eco*RI; S, *Sal*I; V, *Eco*RV; D, *Dra*I. A box and an arrow represent the *gcd* gene and its transcriptional start site and direction, respectively. Bottom lines and shaded boxes represent portions of DNA fragments of the *gcd* gene and the *'phoA* gene, respectively, in constructed gene fusions.

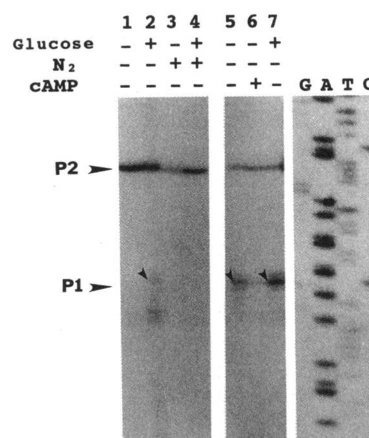

FIG. 2. Reverse transcriptase mapping of *gcd* gene with RNA from wild-type and Δ*cya* strains. RNA was isolated from the exponentially grown cells and subjected to the reverse transcriptase reaction as described previously (23). A primer (5′-CCTGTATTGT TAATTGCCAT-3′) corresponding to the complementary sequence of the coding strand around the initiation codon of the *gcd* gene was synthesized. At the same time, nucleotide sequencing was carried out with single-stranded DNA, including the 5′-flanking region of the *gcd* gene as a template and the same primer as that used in the reverse transcriptase reaction (G, A, T, and C). Both of the reaction products were run together in an 8 M urea–6% polyacrylamide gel (10). W3110 cells were aerobically grown in the absence (lane 1) or presence (lane 2) of glucose and under oxygen-reduced conditions (N_2) without (lane 3) or without (lane 4) glucose. CA8306 cells were aerobically grown in the absence (lane 5) or presence (lane 6) of cAMP and in the presence of glucose (lane 7). Arrowheads indicate the mRNA initiation sites for the promoters.

frame for the *gcd* structural gene, which consisted of 2,388 bases and encoded a 796-amino-acid residue protein, was found. The N-terminal 20-amino-acid sequence was identical to that determined with the purified glucose dehydrogenase from cells harboring pUCGCD1 (24). A possible rho-independent terminator consisting of a 14-bp stem and a 2-base loop followed by an AT cluster was found 38 bases after the stop codon of the open reading frame. Thus, the *gcd* operon appears to consist of one cistron.

Comparison with the *gcd* gene from strain PPA42. Previous reports demonstrated that glucose dehydrogenase in *E. coli* strains occurs as an apoenzyme and the cofactor rarely seems to be produced in the organism (1, 8). Under our assay conditions, the holoenzyme was found to exhibit nearly background activity. Thus, the biological function of the protein was questionable. If the enzyme is nonfunctional, the gene should be subject to random mutations in the coding region.

We compared the nucleotide and amino acid sequences from strains W3110 and PPA42. Out of 2,388 bases, 38 bases were substituted. Thirty-seven of the substitutions were found to be neutral changes which led to no amino acid alterations; only one amino acid substitution, L59R, was found. Furthermore, the glucose dehydrogenase from W3110 exhibited activity (Table 2) and was shown to feed into the electron transport chain when the cofactor was supplied (24). The *gcd* clone from PPA42 was shown to complement a *gcd* mutant of *E. coli* (4), and functional glucose dehydrogenases exist in various *E. coli* strains (8). These facts suggest that glucose dehydrogenase is conserved among *E. coli* strains because it is required for cell growth.

Transcriptional regulation of the *gcd* gene. In order to investigate expression regulation of the *gcd* gene, glucose dehydrogenase activity under various conditions was first measured as shown in Table 2. The activity in the wild-type strain, W3110, was induced 4-fold in the presence of glucose and decreased 10-fold under oxygen-reduced conditions. In Δ*cya*, CA8306, the activity was depressed 10-fold by the addition of cAMP. Thus, we analyzed the regulation of transcription with respect to glucose, cAMP, and oxygen.

Primer extension analysis was performed by using RNA isolated under various conditions (Fig. 2). In W3110, mRNA from P1 was detected only in the presence of glucose. Some bands were also detected in positions lower than the indicated band in Fig. 2, lane 2. These bands may be degradation products of mRNA from P1, because they were found only when glucose was added. The level of mRNA from P2 was

TABLE 2. Activities of glucose dehydrogenase, alkaline phosphatase, and CAT after growth under various conditions[a]

Compound and strain	Sp act (U/mg of protein) under the following conditions:					
	−Glu	+Glu	$+N_2$, −Glu	$+N_2$, +Glu	−cAMP, +Glu	+cAMP, +Glu
GDH						
W3110	4.6	19.0	1.0	2.0	—[b]	—
CA8306	—	—	—	—	36.0	3.5
AP						
CC118(pGCDAP1)	0.05	0.14	—	0.01	—	—
CC118(pGCDAP3)	0.05	0.13	—	0.01	—	—
CC118(pGCDAP5)	0.04	0.14	—	0.01	—	—
CC118(pGCDAP7)	0.04	0.11	—	0.01	—	—
CAT						
W3110(pGCDCM1)	1.53	4.56	—	0.66	—	—
W3110(pGCDCM3)	1.44	4.18	—	0.74	—	—
W3110(pGCDCM5)	1.30	3.24	—	0.68	—	—
W3110(pGCDCM7)	1.83	4.06	—	0.63	—	—
CA8306(pGCDCM1)	—	—	—	—	2.00	0.68
CA8306(pGCDCM3)	—	—	—	—	2.20	0.78
CA8306(pGCDCM5)	—	—	—	—	1.70	0.61
CA8306(pGCDCM7)	—	—	—	—	2.44	0.81

[a] Cells were grown to late exponential phase at 37°C in Luria-Bertani broth in the presence of ampicillin (100 μg/ml), chloramphenicol (20 μg/ml), or 0.5% glucose (+Glu), as required. CA8306 cells were grown in M9 medium (10) containing 0.03% Casamino Acids (Difco Laboratories, Detroit, Mich.) in the presence (+cAMP) or the absence (−cAMP) of 2 mM cAMP. To examine the effect of glucose, fructose was used as an alternative carbon source. When cells were grown under oxygen-reduced conditions ($+N_2$), air was replaced with nitrogen gas which was blown into a 25-ml-volume Erlenmeyer flask employing a pressure of 2 kg/cm^2 for 1 min. The flask was tightly sealed with a rubber cap before being incubated. Cells were harvested and washed twice with 0.85% NaCl. Crude extracts were prepared by sonic oscillation and then were centrifuged (8,000 × *g*, 5 min) to remove the unbroken cells. Glucose dehydrogenase (GDH) (1), alkaline phosphatase (AP) (11), and CAT (5) activities were measured as described previously. β-Lactamase (13) activity was measured for monitoring the copy numbers of plasmids. Reported values are the averages of at least three independent experiments carried out in triplicate. All strains were *E. coli*, and plasmids are shown in parentheses.

[b] —, not determined.

nearly the same in the presence and in the absence of glucose. Reduction of the oxygen concentration brought about a reduction in the level of mRNA read from P2 and a slight enhancement of glucose induction for mRNA synthesis from P2. In Δ*cya*, the mRNA from P1 was diminished by the addition of cAMP, but the mRNA from P2 was not changed. These results suggest that promoter P1 is predominantly regulated by glucose and cAMP and that oxygen largely affects promoter P2. The activity of glucose dehydrogenase after growth under various conditions as shown in Table 2 appeared to be reflected by the levels of the mRNAs.

Analysis with fusion plasmids. For further analysis of transcriptional regulation of the *gcd* gene, fusion plasmids with two different kinds of reporter genes were constructed (Fig. 1B). The fusion plasmids with the *phoA* gene were designed to be under the control of the *gcd* operator-promoter and to produce fusion proteins between glucose dehydrogenase and alkaline phosphatase, so that alteration of the alkaline phosphatase activity would reflect the results of translational as well as transcriptional regulation of the *gcd* gene. The operon fusion with the chloramphenicol acetyltransferase (CAT) gene under the control of the *gcd* operator-promoter exhibited only transcriptional regulation.

To construct PhoA protein fusions, the reporter plasmid pUY311, which is a derivative of pUI310 and which has the *phoA* gene in the direction opposite to that of the *lac* promoter, was used. The activities of alkaline phosphatase from all of the fusion plasmids were nearly the same under the various conditions (Table 2). In the fusion constructs with the CAT gene lacking its own promoter (5), the gene was inserted into the *Nru*I site of all pGCDAPs corresponding to the 8th amino acid residue of glucose dehydrogenase. CAT activities from all fusions were also nearly the same (Table 2). These results indicated that expression of the *gcd* gene was regulated at the transcriptional level and that the regulatory elements responsible for glucose, oxygen, and cAMP responses appeared to be located up to position −104.

In conclusion, from the facts mentioned above, we more precisely defined the regulatory region of the *gcd* gene. Thus, it seems that the cAMP-cAMP receptor protein complex binds to the sequence depicted in Fig. 1 and represses the *gcd* promoter and that the positive regulation of the gene by glucose may be due to modulation of the cytoplasmic cAMP concentration. Even though we have no direct evidence for the mechanism of oxygen regulation, regulation of the *gcd* gene may be specific, because CAT activity was slightly reduced under the same nitrogen-substituted conditions when the CAT gene was placed under the control of the tetracycline resistance gene promoter (data not shown).

Nucleotide sequence accession number. Parts of the deduced amino acid sequence of glucose dehydrogenase appear in the DDBJ, EMBL, and GenBank data bases under accession no. D12651.

We thank O. Adachi and K. Matsushita for their helpful discussions and R. B. Gennis, Y. Kohara, and J. Beckwith for providing us with plasmids, phage clones, and bacterial strains.

This work was supported by the Ministry of Education, Science and Culture of Japan (grant no. 4808030) and by the Uehara Memorial Foundation and grants no. 5R01AI21702 and 2R01A I14176 from the National Institutes of Allergy and Infectious Diseases (Public Health Service).

REFERENCES

1. **Ameyama, M., M. Nonobe, E. Shinagawa, K. Matsushita, K. Takimoto, and O. Adachi.** 1986. Purification and characterization of the quinoprotein D-glucose dehydrogenase apoenzyme from *Escherichia coli*. Agric. Biol. Chem. **50:**49–57.
2. **Berg, O. G., and P. H. von Hippel.** 1988. Selection of DNA binding sites by regulatory proteins. II. The binding specificity of cyclic AMP receptor protein to recognition sites. J. Mol. Biol. **200:**709–723.
3. **Chepuri, V., and R. B. Gennis.** 1990. The use of gene fusions to determine the topology of all of the subunits of the cytochrome *o* terminal oxidase complex of *Escherichia coli*. J. Biol. Chem. **265:**12978–12986.
4. **Cleton-Jansen, A.-M., N. Goosen, O. Fayet, and P. van de Putte.** 1990. Cloning, mapping, and sequencing of the gene encoding *Escherichia coli* quinoprotein glucose dehydrogenase. J. Bacteriol. **172:**6308–6315.
5. **Close, T. J., and R. L. Rodriguez.** 1982. Construction and characterization of the chloramphenicol-resistance gene cartridge: a new approach to the transcriptional mapping of extrachromosomal elements. Gene **20:**305–316.
6. **Ebright, R. H., Y. W. Ebright, and A. Gunasekera.** 1989. Consensus DNA site for the *Escherichia coli* catabolite gene activator protein (CAP): CAP exhibits a 450-fold higher affinity for the consensus DNA site than for the *E. coli lac* DNA site. Nucleic Acids Res. **17:**10295–10305.
7. **Erni, B.** 1989. Glucose transport in *Escherichia coli*. FEMS Microbiol. Rev. **63:**13–24.
8. **Hommes, R. J. W., P. W. Postma, O. M. Neijssel, D. W. Tempest, P. Dokter, and J. A. Duine.** 1984. Evidence of a quinoprotein glucose dehydrogenase apoenzyme in several

strains of *Escherichia coli*. FEMS Microbiol. Lett. **24:**329–333.

9. **Kohara, Y., K. Akiyama, and K. Isono.** 1987. The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell **50:**495–508.
10. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
11. **Manoil, C., and J. Beckwith.** 1985. TnphoA: a transposon probe for protein export signals. Proc. Natl. Acad. Sci. USA **82:**8129–8133.
12. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101:**20–79.
13. **Nakazawa, T., S. Inouye, and A. Nakazawa.** 1980. Physical and functional mapping of RP4-TOL plasmid recombinants: analysis of insertion and deletion mutants. J. Bacteriol. **144:**222–231.
14. **Postma, P. W., C. P. Broekhuizen, and R. H. Geerse.** 1989. The role of PEP:carbohydrate phosphotransferase system in the regulation of bacterial metabolism. FEMS Microbiol. Rev. **63:**69–80.
15. **Raibaud, O., and M. Schwartz.** 1984. Positive control of transcription initiation in bacteria. Annu. Rev. Genet. **18:**173–206.
16. **Saier, M. H., Jr.** 1985. Mechanisms and regulation of carbohydrate transport in bacteria. Academic Press, Inc., New York.
17. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
18. **Shine, J., and L. Dalgarno.** 1974. 3′-Terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71:**1342–1346.
19. **Stormo, G. D., and G. W. Hartzell III.** 1989. Identifying protein-binding sites from unaligned DNA fragments. Proc. Natl. Acad. Sci. USA **86:**1183–1187.
20. **van Schie, B. J., K. J. Hellingwerf, J. P. van Dijken, M. G. L. Elferink, J. M. van Dijl, J. G. Kuenen, and W. N. Konings.** 1985. Energy transduction by electron transport via pyrrol-quinoline quinone-dependent glucose dehydrogenase in *Escherichia coli*, *Pseudomonas aeruginosa*, and *Acinetobacter calcoaceticus* (var. *lwoffi*). J. Bacteriol. **163:**493–499.
21. **Vieira, J., and J. Messing.** 1987. Production of single-stranded plasmid DNA. Methods Enzymol. **153:**3–11.
22. **Yamada, M., and A. Nakazawa.** 1984. Factors necessary for the export process of colicin E1 across cytoplasmic membrane of *Escherichia coli*. Eur. J. Biochem. **140:**249–255.
23. **Yamada, M., and M. H. Saier, Jr.** 1987. Glucitol-specific enzymes of the phosphotransferase system in *Escherichia coli*: nucleotide sequence of the *gut* operon. J. Biol. Chem. **262:**5455–5463.
24. **Yamada, M., K. Sumi, K. Matsushita, O. Adachi, and Y. Yamada.** Unpublished data.
25. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.