JOURNAL OF BACTERIOLOGY, Mar. 2004, p. 1388–1397 Vol. 186, No. 5
0021-9193/04/$08.00+0 DOI: 10.1128/JB.186.5.1388–1397.2004

Transcriptional Organization and Regulation of the L-Idonic Acid Pathway (GntII System) in *Escherichia coli*

Christoph Bausch,† Matthew Ramsey, and Tyrrell Conway*

Department of Botany and Microbiology, University of Oklahoma, Norman, Oklahoma 73019

Received 23 July 2003/Accepted 19 November 2003

The genetic organization of the *idn* genes that encode the pathway for L-idonate catabolism was characterized. The monocistronic *idnK* gene is transcribed divergently from the *idnDOTR* genes, which were shown to form an operon. The 215-bp regulatory region between the *idnK* and *idnD* genes contains promoters in opposite orientation with transcription start sites that mapped to positions −26 and −29 with respect to the start codons. The regulatory region also contains a single putative IdnR/GntR binding site centered between the two promoters, a CRP binding site upstream of *idnD*, and an UP element upstream of *idnK*. The genes of the L-idonate pathway were shown to be under catabolite repression control. Analysis of *idnD*- and *idnK-lacZ* fusions in a nonpolar *idnD* mutant that is unable to interconvert L-idonate and 5-ketogluconate indicated that either compound could induce the pathway. The L-idonate pathway was first characterized as a subsidiary pathway for D-gluconate catabolism (GntII), which is induced by D-gluconate in a GntI (primary gluconate system) mutant. Here we showed that the *idnK* and *idnD* operons are induced by D-gluconate in a GntI system mutant, presumably by endogenous formation of 5-ketogluconate from D-gluconate. Thus, the regulation of the GntII system is appropriate for this pathway, which is primarily involved in L-idonate catabolism; the GntII system can be induced by D-gluconate under conditions that block the GntI system.

For three decades, there was thought to be two systems for D-gluconate catabolism, GntI and GntII (1). The GntI system consists of *gntT*, *gntU*, and *gntK*, which encode high- and low-affinity D-gluconate transporters and a thermoresistant gluconate kinase, respectively (19–21, 29). GntR negatively controls the GntI genes, as well as *edd* and *eda* of the Entner-Doudoroff pathway. The GntII system is composed of a thermosensitive gluconate kinase and a gluconate transporter, which provide for subsidiary catabolism of gluconate in GntI mutants (3, 16). Recently, we discovered that the GntII system is, in fact, a pathway for catabolism of L-idonate, which proceeds via a D-gluconate intermediate (3). The discovery of this novel pathway solved the longstanding question of why there are two pathways for gluconate; GntI is primarily involved in gluconate catabolism, and GntII is responsible for idonate catabolism.

The catabolic sequence for L-idonate is as follows: L-idonate is transported by the L-idonate transporter, IdnT; L-idonate is oxidized to 5-keto-gluconate (5KG) by L-idonate 5-dehydrogenase, IdnD; 5KG is reduced to D-gluconate by 5-keto-D-gluconate 5-reductase, IdnO; and D-gluconate is phosphorylated by a thermosensitive gluconate kinase, IdnK, to make 6-phosphogluconate (6PG), which is further catabolized via the Entner-Doudoroff pathway. Thus, IdnD and IdnO allow for the redox-coupled interconversion of L-idonate to D-gluconate via 5KG. The L-idonate catabolic pathway overlaps D-gluconate catabolism through the common intermediates D-gluconate and 6PG.

While the biochemistry of the L-idonate pathway is firmly established, the organization and regulation of the corresponding genes have not been characterized. Sequence annotation indicates that *idnK* is monocistronic and is divergently transcribed from a putative operon consisting of *idnD*, *idnO*, and *idnT* along with *idnR*, which encodes a repressor of the GalR-LacI family. In this report, we confirm the operon arrangement, transcription start sites, and regulation of the L-idonate genes. The results indicate that both L-idonate and 5KG act as inducers of the idonate pathway. Furthermore, the subsidiary role of the GntII system for gluconate catabolism was investigated in a GntI system mutant and shown to result from induction of the *idnD* operon and *idnK* by gluconate, presumably caused by accumulation of an endogenous inducer (e.g., 5KG). Lastly, functional genomic analyses with DNA arrays and two-dimensional (2-D) protein gels were used to characterize the global gene expression—and hence the physiology—of cells grown with L-idonate as the sole carbon source.

MATERIALS AND METHODS

Bacterial strains and growth conditions. The *Escherichia coli* strains used in this study are listed in Table 1. *E. coli* W1485 was the wild-type strain. All mutant and chromosomal *lacZ* fusion strains were derived from *E. coli* W1485 (2). *E. coli* DH5α and XL1-Blue were used for propagation of plasmids. Strains were grown at 37°C in Luria broth (LB) (13) with or without added carbohydrate (0.4%), in morpholinepropanesulfonic acid (MOPS) minimal medium with added carbohydrate (0.2%) (17), or in MOPS complete medium with added carbohydrate (0.2%) (32). MOPS complete medium contains amino acid, vitamin, purine, and pyrimidine supplements. When appropriate, ampicillin (100 μg/ml) and kanamycin (25 μg/ml) were included in the growth medium. All cultures (50-ml volume) were grown in 250-ml Erlenmeyer flasks and aerated by gyratory shaking at 300 rpm. Cell growth was monitored spectrophotometrically at 600 nm with a DU 530 Life Science UV/Vis spectrophotometer (Beckman Coulter, Inc., Fullerton, Calif.). Cultures in the early and late logarithmic phases of growth were harvested at optical densities of 0.3 and 0.7, respectively. Phenotypes of *E.*

* Corresponding author. Mailing address: Advanced Center for Genome Technology, OU Microarray Core Facility, Department of Botany and Microbiology, 770 Van Vleet Oval, University of Oklahoma, Norman, OK 73019-0245. Phone: (405) 325-1683. Fax: (405) 325-7619. E-mail: tconway@ou.edu.

† Present address: Stowers Institute for Medical Research, Kansas City, MO 64110.

TABLE 1. *E. coli* strains used in this study

Strain	Relevant genotype, phenotype, or characteristics	Reference or source
W1485	K-12 wild type (λ⁻ *rph-1*)	CGSC[a]
BW25113	*lacI*^{q} *rrnB*$_{T14}$ Δ*lacZ*$_{WJ16}$ *hsdR514* Δ(*araBAD*)$_{AH33}$ Δ(*rhaBAD*)$_{LD78}$	6
BW25141	*lacI*^{q} *rrnB*$_{T14}$ Δ*lacZ*$_{WJ16}$ *hsdR514* Δ(*araBAD*)$_{AH33}$ Δ(*rhaBAD*)$_{LD78}$ *galU95* *endA*$_{BT333}$ *uidA* [(Δ*mluI*)::*pir*$^+$ *recA1*]	6
CB130	λ⁻ *rph-1* Δ(*lacI′*^{b}-P$_{lacZ}$)^c P$_{idn}$ *idnK-lacZ*^{d}; Km^r	This study
CB131	λ⁻ *rph-1* Δ(*lacI′*-P$_{lacZ}$)^c P$_{idn}$ *idnD-lacZ*^{d}; Km^r	This study
CB360	λ⁻ *rph-1* *idnD*::*kan*	This study
CB361	λ⁻ *rph-1* Δ*idnD*	This study
CB361Z	λ⁻ *rph-1* Δ*idnD* Δ(*lacI′*-P$_{lacZ}$)^c *idnD-lacZ*^{d}	This study
CB366	λ⁻ *rph-1* Δ*idnR*	This study
CB370	λ⁻ *rph-1* Δ*crp*	This study
CB371	λ⁻ *rph-1* Δ*gntR*	This study
DH5α	*supE44* Δ*lacU169*(*lacZ*Δ*M15*Δ)*hsdR17* *endA1* *gyrA96* *thi-1* *relA1*	BRL[e]
MD5	λ⁻ *rph-1* *idnK*::*kan*	Fernando Valle
MD5E	λ⁻ *rph-1* *idnK*::*kan* *gntK*::*cat*	Fernando Valle
NP202	λ⁻ *rph-1* Δ(*gntRKU*)	This study
XL1-Blue	*recA1* *endA1* *gyrA96* *thi-1* *hsdR17* *supE44* *relA1* *lac* [F′ *proAB* *lacI*^{q} *Z*Δ*M15* Tn*10* (Tet^r)]	Stratagene

[a] CGSC, *E. coli* Genetic Stock Collection.
[b] Truncated gene (′).
[c] Δ(*lacI′*-P$_{lacZ}$) removes the 3′ terminus of *lacI*, the entire intergenic regulatory region upstream of *lacZ*, and the 5′-terminal region of *lacZ*.
[d] The *lacZ* coding region is fused in frame to the 5′ region of the indicated *idnK* or *idnD* gene, forming a translation fusion. (see Materials and Methods).
[e] BRL, Bethesda Research Laboratories.

coli strains were determined on MacConkey indicator medium (14), tryptone-yeast extract agar (6), or LB plates (13).

Plasmid construction, DNA modification, and transformation. Standard methods were used for DNA restrictions, ligations, and transformations and other DNA manipulations (23). PCR amplification was performed with Platinum high-fidelity *Taq* DNA polymerase (Invitrogen Life Technologies, Carlsbad, Calif.). The plasmids used in this study are listed in Table 2. Primers specific to this study are listed as supplementary material found on our website (http://www.ou.edu/microarray). Gene-specific deletions were carried out with *E. coli* W1485 by the method reported by Datsenko and Wanner (6). PCR-generated products were purified with a QIAquick PCR Purification Kit (Qiagen Inc., Valencia, Calif.). Electroporation was performed on a Gene Pulser II with 0.2-cm-gap cuvettes (Bio-Rad Laboratories, Hercules, Calif.). Colony PCR was accomplished by scraping cells from agar plates, thoroughly washing the cells five times with water, and amplifying PCR products by using sequence-specific primers and HotStar*Taq* DNA polymerase (Qiagen Inc., Valencia, Calif.).

Construction of *idnK-lacZ* and *idnD-lacZ* gene fusions. Single-copy *idnK-lacZ* and *idnD-lacZ* gene fusions were constructed by a recombinase-assisted *lacZ* fusion system developed for this work by combining the *lacZ* fusion system described by Simons et al. (27) and the system for allele replacement described by Wanner et al. (32). The fusion vectors used for recombinase-assisted *lacZ* fusion included pCB551, pCB552, and pCB577 (Table 2). These plasmids contain the *ori* and *bla* genes from pSP72, 542 nucleotides of the 3′ end of the *E. coli* W1485 *lacI* gene cloned from pCB108, and functional elements common to pRS551, pRS552, and pRS577 , including a selectable kanamycin resistance gene; four tandem copies of the *T1* terminator from the *E. coli rrnB* operon; the unique multiple cloning site (MCS) containing *Bam*HI, *Sma*I, and *Eco*RI; and the 5′ region of the *lacZ* gene. A DNA fragment containing the untranslated region between the *idnK* and *idnD* genes and a terminal *Bam*HI or *Eco*RI site located upstream of the *idnK* and *idnD* start codons, respectively, was amplified by PCR. This DNA fragment was cloned into the *Bam*HI-*Eco*RI sites of the protein fusion vector pCB552 and the transcription fusion vector pCB551, cre-

TABLE 2. Plasmids used in this study

Plasmid name	Description	Reference or source
pBluescriptIISK$^+$	Cloning vector; Ap^r	Stratagene
pSP72	Cloning vector; Ap^r	Promega
pRS551	Transcription fusion vector containing the MCS, *Eco*RI-*Sma*I-*Bam*HI-*lacZYA′*[a]	27
pRS552	Translation fusion vector containing the MCS, *Eco*RI-*Sma*I-*Bam*HI-*lacZYA′*	27
pRS577	Translation fusion vector containing the MCS, *Bam*HI-*Sma*I-*Eco*RI-*lacZYA′*	27
pNP202	Replaced a 1.3-kb *Stu*I-*Bgl*II fragment on pTC221 with the Kan^r gene cassette	This study
pCB92	pBSKs$^+$ containing 300-bp *idnK′* *Pst*I-*Kpn*I DNA fragment generated by PCR	This study
pCB100	pBSKs$^+$ containing 300-bp *idnD′* *Pst*I-*Kpn*I DNA fragment generated by PCR	This study
pCB200	pBSKs$^+$ containing 300-bp *idnO′* *Pst*I-*Kpn*I DNA fragment generated by PCR	This study
pCB620	pBSKs$^+$ containing 300-bp *idnT′* *Pst*I-*Kpn*I DNA fragment generated by PCR	This study
pCB700	pBSKs$^+$ containing 300-bp *yigR′* *Pst*I-*Kpn*I DNA fragment generated by PCR	This study
pCB900	pBSKs$^+$ containing 300-bp *idnR′* *Pst*I-*Kpn*I DNA fragment generated by PCR	This study
pCB108	pSP72 containing 542-bp *lacI′* *Pst*I DNA fragment generated by PCR	This study
pCB551	pCB108 containing 2,223-bp *neo* *rrnB* T1$_4$ *lacZ′* DNA fragment from pRS551	This study
pCB552	pCB108 containing 2,223-bp *neo* *rrnB* T1$_4$ *lacZ′* DNA fragment from pRS552	This study
pCB577	pCB108 containing 2,223-bp *neo* *rrnB* T1$_4$ *lacZ′* DNA fragment from pRS577	This study
pCB120	pCB552 containing *Eco*RI-*Bam*HI fragment of 274-bp *idn* regulatory region	This study
pCB121	pCB577 containing *Bam*HI-*Eco*RI fragment of 274-bp *idn* regulatory region	This study
pCB220	pCB551 containing *Eco*RI-*Bam*HI fragment of 274-bp *idn* regulatory region	This study

[a] Truncated gene (′).

ating *idnK-lacZ* fusion plasmids pCB120 and pCB220, respectively. The same fragment was cloned in the opposite orientation into the protein fusion vector pCB577 at the *Eco*RI-*Bam*HI site, generating an *idnD-lacZ* fusion (pCB121).

Integration of the *lacZ* fusions into the chromosome of *E. coli* W1485 was achieved by allelic replacement by homologous recombination of the fusion construct into the *lacI-lacZ* region of the genome. Linear DNA fragments for allelic replacement were amplified by PCR with pCB120 and pCB121 as the templates. This method eliminated native *lacZ* regulation and generated a *lacZ* fusion in one step. Bacterial colonies with the desired phenotype on tryptone-yeast extract agar-kanamycin plates were transferred to MacConkey plates, and the cells were screened for the lactose-negative phenotype. PCR was used to verify correct allelic replacement of the native *lacZ* regulatory region with the *lacZ* fusion; all of the *lacZ* fusion constructions were confirmed by DNA sequence analysis (25).

β-Galactosidase measurements. β-Galactosidase activity was measured with the yeast β-galactosidase assay kit from Pierce Biotechnology, Inc. (Rockford, Ill.). Cell cultures were grown in triplicate, and each culture was assayed in triplicate. A 70-μl aliquot was taken from each culture, mixed with an equal volume of β-galactosidase assay solution, placed in individual wells of a 96-well assay plate (Falcon Software, Inc., Wellesley, Mass.), and then held at 4°C until the assay was performed. The β-galactosidase assay solution was a 1:1 mixture of Y-PER (yeast protein extraction reagent) and 2× β-galactosidase assay buffer. Before initiation of the assay, spectrophotometric measurements at 590 nm were made with a PowerWave X 96-well Microplate Spectrophotometer (Bio-Tek Instruments, Inc., Winooski, Vt.) to determine relative cell densities. The 96-well plate was incubated in the plate reader at 37°C. Measurements were made spectrophotometrically at 420 nm every 4 min for 1 h, and the data were analyzed with the KC4 kinetics software package (Bio-Tek Instruments, Inc.). β-Galactosidase activity was calculated when the reaction was linear and expressed in Miller units (15). The values reported for each sample are the means ± the standard deviations for nine independent measurements.

RNA isolation. Total RNA for Northern blot assays and primer extension analysis was isolated by the hot-phenol method as described previously (21). Total RNA for gene expression profiling and reverse transcriptase PCR (RT-PCR) was isolated by pipetting an equal volume of an actively growing cell culture into ice-cold RNA*later* (Ambion, Inc., Austin, Tex.). The RNA was then purified and treated with DNase with RNeasy mini kits and RNase-free DNase kits (Qiagen Inc.). RNA concentrations were determined by spectrophotometric measurements at 260 nm. RNA was stored in ethanol at −80°C.

Primer extension analysis. Oligonucleotides complementary to the mRNA sequences upstream of the *idnK* and *idnD* start codons were end labeled by using T4 polynucleotide kinase (Invitrogen Life Technologies) and [γ-^{32}P]ATP (>5,000 Ci mmol^{-1}) as previously described (23). Each 5′-end-labeled primer (0.5 pmol [~1.5 × 10^6 cpm]) was annealed to 30 μg of total RNA in a 10-μl reaction mixture by heating to 94°C for 2 min, followed by slow cooling to 42°C. The primers were then extended at 42°C for 5 h by using Moloney murine leukemia virus RT (Ambion, Inc.). The reaction was stopped by addition of 10 μl of sequence loading buffer. The reaction mixtures were boiled for 3 min, and 4-μl aliquots were run on 6% polyacrylamide gels with size reference ladders generated by dideoxy sequencing of pNP204 with the same primers used for primer extension.

Northern blot analysis. Total cellular RNA (5 μg) was denatured by incubation for 10 min at 68°C in formaldehyde-MOPS gel loading buffer (Ambion, Inc.) and electrophoresed through a 1.5% agarose gel containing formaldehyde and MOPS buffer. RNA was transferred to Nytran SuPerCharged superior nylon transfer membranes (Schleicher & Schuell, Inc., Keene, N.H.) by using a rapid downward transfer system. Antisense RNA probes were generated by reverse transcription from plasmids pCB92, pCB100, pCB200, pCB620, pCB700, and pCB900, containing the truncated genes *idnK′*, *idnD′*, *idnO′*, *idnT′*, *yjgR′*, and *idnR′*, respectively. These plasmids were constructed by cloning PCR products generated with nested gene-specific primers into pBluescript II SK+ (Table 2). All plasmids were linearized at the 3′ end of the truncated gene at the *Bam*HI site, and a ^{32}P-labeled RNA probe was synthesized by transcription with T7 RNA polymerase (Cloned; Ambion, Inc.) in the presence of [α-^{32}P]UTP (23). Probe hybridization to the membrane-bound RNA and stripping from the membranes were done as described previously (29). Hybridized membranes were visualized by exposure to X-ray film or phosphorimaging screens, which were scanned with a STORM 820 PhosphorImager (Molecular Dynamics, Sunnyvale, Calif.).

RT-PCR. RT-PCR products were prepared by using the SuperScript One-Step RT-PCR system with Platinum *Taq* DNA polymerase (Invitrogen Life Technologies) as instructed by the manufacturer. Total RNA was isolated at an optical density of 0.7 from *E. coli* W1485 grown in MOPS complete medium containing 0.2% L-idonic acid. The primers were checked for performance in PCRs by using *E. coli* W1485 genomic DNA as the template. RNA samples were tested for contaminating genomic DNA by using each RNA sample as a template for PCR; RNA samples contaminated with DNA were not used. The RT-PCR products were separated by electrophoresis through 1% agarose gels stained with ethidium bromide and documented with an Epi Chemi II Darkroom (UVP, Inc., Upland, Calif.).

Transcriptome profiling and treatment of data. The methods used to handle whole-genome *E. coli* arrays and data analysis are described in detail on our website (http://www.ou.edu/microarray) and by Conway et al. (5). The C-terminal primer set (Sigma-GenoSys, The Woodlands, Tex.) was used to transcribe radioactively labeled cDNA (first-strand synthesis) with [α-^{32}P]dCTP and SuperScript II RNase H$^-$ RT (Invitrogen Life Technologies) from samples of total cellular RNA. Duplicate Panorama *E. coli* Gene Array membranes (Sigma-GenoSys) from consecutive printings were used. Hybridization and stripping of membranes were done as described previously (28). Phosphorimages of hybridized membranes were analyzed with ArrayVision (Imaging Research Inc., St. Catharines, Ontario, Canada) to obtain raw spot intensity data. The raw data were normalized by expressing individual spot intensities as a fraction of the sum of all gene-specific spot intensities in each image, and the data were analyzed as previously described by using semiautomated Microsoft Visual Basic programs in Microsoft Excel (5).

2-D polyacrylamide gel electrophoresis (PAGE). Cells were harvested by centrifugation and washed twice in a 10 mM MgCl–50 mM HEPES solution at pH 6.5 and then transferred to a lysis buffer that contained 9 M urea, 40 mM Tris-HCl, 4% 3-[(cholamidopropyl)-dimethylammonio]-1-propanesulfonate (CHAPS), and 1% dithiothreitol (DTT). After sonication on ice for 5 × 1 min with 30-s cooling intervals, cell debris was removed by centrifugation at 3,000 × *g* for 10 min at 4°C. The protein concentration of the supernatant was determined by the Bradford assay (4).

A 200-μg sample of cell extract was loaded onto 7-cm immobilized pH gradient strips that had a nonlinear pH range of 3 to 10 (Amersham Biosciences, Uppsala, Sweden). A rehydration solution that contained 8 M urea, 2% CHAPS, 1% DTT, and 0.5% immobilized pH gradient buffer (Amersham Biosciences) was added to the extract to a final volume of 120 μl. Rehydration was carried out for 10 h at 20°C as described by Sanchez et al. (24). Isoelectric focusing (IEF) was carried out with an Ettan IPGphor IEF unit (Amersham Biosciences) for 1 h at 100 V, 30 min at 500 V, 30 min at 1 kV, 1 h at 3 kV, 1 h at 5 kV, and 2 h at 8 kV. The temperature was held at 20°C throughout IEF. After IEF, the strips were incubated in a 50 mM Tris-HCl solution (pH 8.8) that contained 6 M urea, 30% glycerol, 2% sodium dodecyl sulfate, and 1% DTT for 30 min. The strips were then placed on top of 15% PAGE gels containing 2% sodium dodecyl sulfate and attached with a 0.5% agarose solution that contained a trace amount of bromophenol blue (American Bioanalytical, Natick, Mass.). Electrophoresis was then carried out with a Mini-PROTEAN II system (Bio-Rad Laboratories) at 20 mA for ~4 h, until the bromophenol blue front reached the bottom of the gel, and the gel was then stained with Coomassie brilliant blue.

In situ digestion, nano-electrospray MS-MS/MS, and data analysis. The stained gels were compared visually, and differentially expressed spots of interest were excised and prepared by trypsin digestion in accordance with the protocol of Devreese et al. (7). Nano-electrospray mass spectrometry (MS) and tandem mass spectrophotometry (MS/MS) were carried out on a Q-Tof mass spectrometer (Micromass, Manchester, United Kingdom) under conditions similar to these described by Devreese et al. (7). In situ digests were washed with C_{18} ZipTip pipette tips (Millipore Corp., Bedford, Mass.). Extracts thus prepared were loaded into a coated fused-silica capillary tip (New Objective, Inc., Woburn, Mass.) and then placed into the nanospray source on the mass spectrometer. The capillary tube voltage was held at 0.9 kV, and spraying was initiated with a flow of N_2 (~3 lb/in^2) at the back of the capillary tubing. Spectra were taken in the 100-to-2,000 mass range with 2-s scans, and data were collected for 2 min. Several of the most prominent doubly and triply charged molecular ions were manually identified and selected for collision-induced dissociation fragmentation with Ar as the collision gas, with the collision energy adjusted between 22 to 33 eV, depending on the optimum for fragmentation of the peptide.

The MS/MS spectra were interpreted with MassLynx 4.0 software as described by the manufacturer (Micromass). The MaxEnt3 tool was used to convert multiply charged fragment ions to singly charged species, and the PepSeq tool was used to determine the amino acid sequence after finding the fragment ion series. Sequences were matched to an *E. coli* protein database with version 3.4 of the FASTA search program (18).

Chemicals and enzymes. Restriction enzymes and DNA-modifying enzymes were purchased from Invitrogen Life Technologies, Qiagen Inc., and Promega Corp. (Madison, Wis.). The T7 Sequenase version 2.0 kit and radioactive [α-^{32}P]UTP and [γ-^{32}P]ATP were purchased from Amersham Biosciences, Inc.

(Piscataway, N.J.). Biochemicals were purchased from Sigma-Aldrich Corp. (St. Louis, Mo.). Panorama *E. coli* gene arrays were obtained from Sigma-GenoSys. Sodium L-idonate was received as a generous gift from Alisha S. Jarnagin (Genencor International, Inc., Palo Alto, Calif.).

RESULTS

Annotation of the *idn* promoter region. Examination of the 215-bp sequence between the *idnD* and *idnK* genes revealed two putative −10 and −35 RNA polymerase binding sites on opposite strands. Both genes contain conserved Shine-Dalgarno sequences located 4 and 7 bp upstream from the IdnD and IdnO translation start sites, respectively (26). In addition, a single putative cyclic AMP (cAMP) receptor protein (CRP) binding site (ATTTGTGA-TGAAGA-TCACGTCA) was identified upstream of the *idnD* gene. A putative IdnR operator site (ATGTTA-CGCA-TAACGT) with homology to the GntR consensus binding sequence (ATGTTA-[N_4]-TAACAT) (21) is centered between the two promoters, −78.5 and −83.5, with respect to the *idnK* and *idnD* transcription start sites, respectively, suggesting that this site may function as a regulatory element for both promoters. The position of the putative IdnR binding site is interesting because this location is atypical of negative control, despite the fact that IdnR and GntR belong to the GalR-LacI family of negative regulators (29). Slightly upstream of the *idnK* gene (−38 to −59) is a putative A-T-rich UP element sequence (8) that could be involved in stabilization of RNA polymerase-promoter interactions.

Transcription start sites for the *idn* genes. Primer extension analysis was used to map the transcription start sites for *idnK* and *idnD* with RNA extracted from cells grown in the presence of L-idonate (Fig. 1). Growth on 5KG resulted in the same transcription start sites (data not shown). The *idnD* transcript start site (PD1) was located 29 bp upstream of the *idnD* start codon (Fig. 1A), and the *idnK* transcript start site (PK1) was located 26 bp upstream of the *idnK* start codon (Fig. 1B). These transcription start sites are consistent with the locations of the putative *idnD* and *idnK* promoter sequence elements.

Organization of *idn* transcription. The organization of the *idn* genes suggested that *idnD*, *-O*, *-T*, and *-R* might be transcribed as a polycistronic message. The *idnK* transcript is monocistronic, as indicated by a 0.8-kb band of the expected length (Fig. 2). Northern blot analysis also suggested that *idnD*, *-O*, *-T*, and *-R* are cotranscribed (Fig. 2). Transcripts that hybridized with the *idnD* and *idnO* probes were observed at 1.9, 3.3, and 4.3 kb, although the latter transcript hybridized with very low intensity. An individual transcript for *idnD* was not observed, but there was an *idnO*-specific transcript of 0.8 kb. The most abundant transcript for *idnD* and *idnO* was 1.9 kb. The *idnT* probe hybridized to a 3.3-kb transcript, suggesting cotranscription with *idnD* and *idnO*. In addition, Northern hybridization revealed a 1.5-kb *idnT* transcript of sufficient length to encode *idnT* alone. The 4.3-kb transcript that hybridized to all four *idnD*, *idnO*, *idnT*, and *idnR* probes is consistent with the predicted transcript length of the *idnDOTR* operon. However, this transcript was apparently unstable and only a very faint band was observed. Overall, the results of Northern blot analysis supported the hypothesis that *idnD*, *idnO*, *idnT*, and *idnR* are cotranscribed and that the primary transcript is processed to form several gene-specific transcripts, which are more stable than the primary *idnDOTR* message.

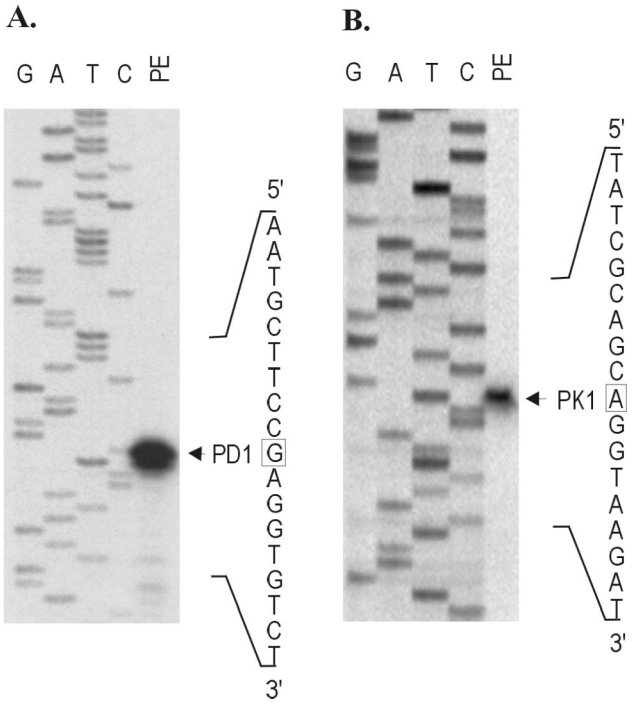

FIG. 1. Primer extension of the transcription start sites for *idnD* and *idnK*. (A) Extension of the *idnD* transcript (PD1). (B) Extension of the *idnK* transcript (PK1). Lanes: PE, primer extension products; G, A, T, and C, corresponding sequence ladders.

Computer analysis of predicted mRNA secondary structures in the *idn* regulatory region suggested the presence of stem-loop terminator-like structures at the 3′ ends of *idnO*, *idnT*, and *idnK*, but not *idnR* (data not shown). The strong intensity of the putative 1.9-kb *idnD-idnO* transcript in Northern blot assays implies that the predicted stem-loop structure at the end of *idnO* functions as a terminator. It is also likely that the stem-loop structures positioned after *idnT* and *idnK* function as transcription terminators, since transcripts ending after the predicted coding region of both genes were resolved in Northern blot assays.

RT-PCR with RNA obtained from cells grown on L-idonate confirmed cotranscription of the *idnDOTR* operon (lanes 2 to 5, 8 and 9, Fig. 3). The monocistronic *idnK* transcript observed by Northern blot analysis was also confirmed by RT-PCR (lanes 6 and 7, Fig. 3). RT-PCR indicated that transcription did not terminate immediately downstream of *idnR*, but rather extended at least 500 bp into the *yjgR* gene (lanes 10 and 11, Fig. 3). However, this transcript did not appear to extend beyond the carboxy terminus of the *yjgR* structural gene, as downstream primers failed to yield a product (lane 12, Fig. 3). Further, Northern hybridization with a probe specific for *yjgR* revealed a 1.5-kb transcript that was not induced by L-idonate or D-gluconate (data not shown). A *yjgR* knockout grew well on L-idonate, confirming that YjgR is not required for L-idonate catabolism (data not shown).

Transcription regulation of the *idn* genes. The enzymes of the L-idonate pathway are induced by L-idonate (3). To confirm

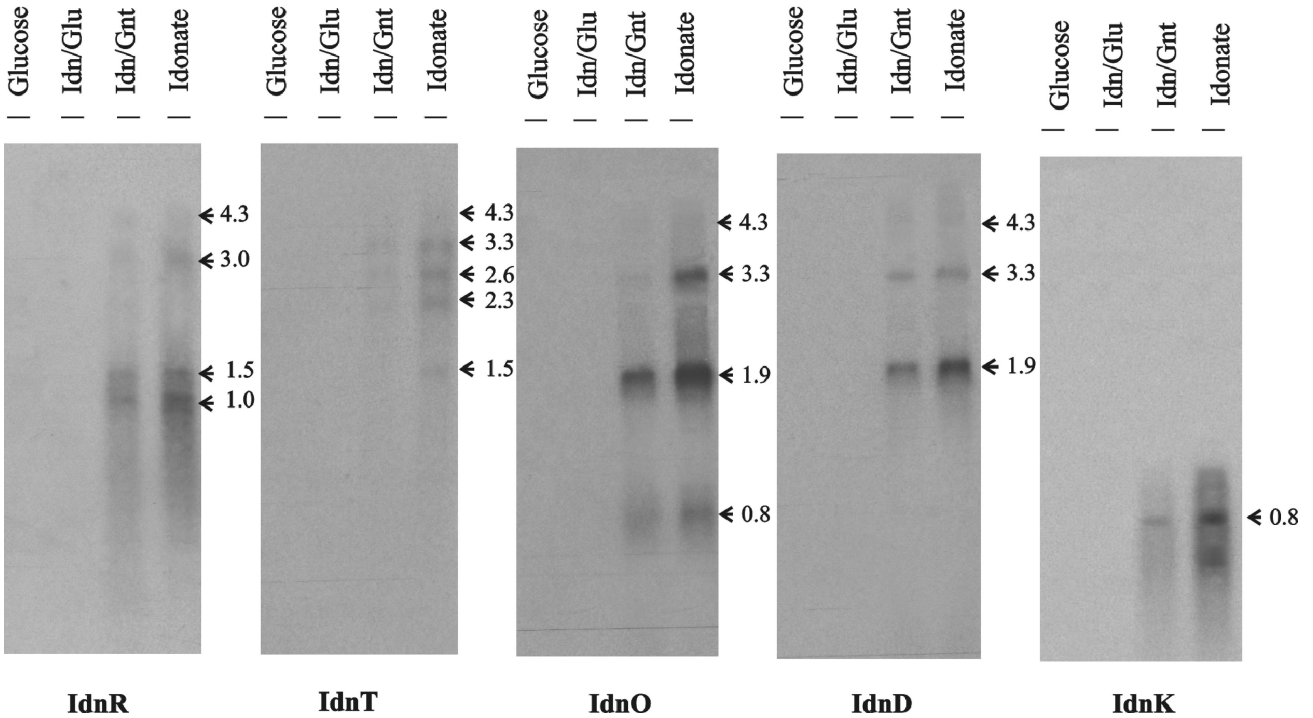

FIG. 2. Northern blot analysis of the *idnK*, *idnD*, *idnO*, *idnT*, and *IdnR* transcripts in *E. coli* W1485. Total RNA was isolated from late-log-phase cultures grown on MOPS minimal medium containing the carbohydrate listed above each lane. A total of 5 μg of RNA was loaded per lane. Estimated transcript sizes (in kilobases) are shown to the right of each blot and were determined from an RNA Millennium Marker (Ambion, Inc.) run with each independent RNA gel (data not shown). Hybridizations were carried out with 300-nucleotide probes specific for the gene encoding the protein indicated under each blot.

that the *idn* transcripts are similarly induced, we measured carbon source-dependent transcription of the L-idonate pathway genes (Fig. 2). Northern blot hybridization analysis indicated strong induction of *idn* transcripts in the presence of L-idonate and no induction with D-glucose. This result suggests that L-idonate functions to induce the *idn* genes.

To determine if 5KG also acts as an inducer, we tested induction of *idn* transcription in a strain containing a nonpolar *idnD* mutation that blocks the interconversion of L-idonate and 5KG without affecting expression of the other *idn* genes. An *idnD-lacZ* fusion in the *idnD* nonpolar mutant strain (CB361Z) was induced by 5KG and L-idonate, suggesting that both sugars can induce the L-idonate pathway (Table 3). This result was confirmed by Northern analysis, which showed that transcription of *idnO* was induced by growth on either 5KG or L-idonate in CB361Z (data not shown).

Transcriptional regulation of the *idn* regulon was further analyzed with *lacZ* gene fusions. Because the *idnD* and *idnK* genes are divergently transcribed from the same 215-bp region of DNA, gene fusions were constructed with the same promoter-containing fragment cloned in opposite orientations—one in the direction of *idnK* transcription and the other in the direction of *idnD* transcription. These fusions were integrated into the genome as single copies, because multicopy fusions expressed from plasmids did not appropriately reflect regulation. The *idnD*- and *idnK-lacZ* fusions were remarkably similar in expression, suggesting that regulation of the two promoters is coordinated (Table 4). Both fusions were induced by L-idonate and 5KG and slightly induced by D-gluconate, whereas D-glycerol, D-glucose, and succinate did not cause induction. Sugars related to the L-idonate pathway in other eubacteria (30), 2-ketogluconate, 2,5-diketogluconate, iduronate, and 2-ketogulonate, did not cause induction of the *idn* genes in *E. coli* W1485.

Catabolite repression of the *idn* genes was observed in cells growing on a combination of L-idonate and D-gluconate or L-idonate and D-glucose; greater repression was observed with the addition of D-glucose (Fig. 2 and Table 4). Moreover, addition of cAMP (4 mM) to cells harboring the *lacZ* reporter fusions caused a 2.5-fold increase in reporter activity when the cells were grown on L-idonate, and a similar response was also observed for cells grown on D-gluconate (Table 4). A *crp* mutant strain (CB370) was unable to grow on MOPS minimal medium containing L-idonate (Table 5). As reported previously, the *crp* mutant demonstrated very poor growth on D-gluconate (19). Taken together, these results indicate that the *idn* promoters are subject to cAMP-CRP-dependent catabolite repression.

Growth physiology of GntI and GntII system mutants. To understand the role of the GntI and GntII systems in growth on sugar acids, we used mutational analysis to evaluate growth on MOPS minimal medium supplemented with either D-glucose, D-gluconate, 5-ketogluconate, or L-idonate (Table 5). The wild-type *E. coli* strain, W1485, grew well on all of the carbon sources used except 5KG, which has been described previously (3). The *idnR* mutant (CB366) was unable to grow

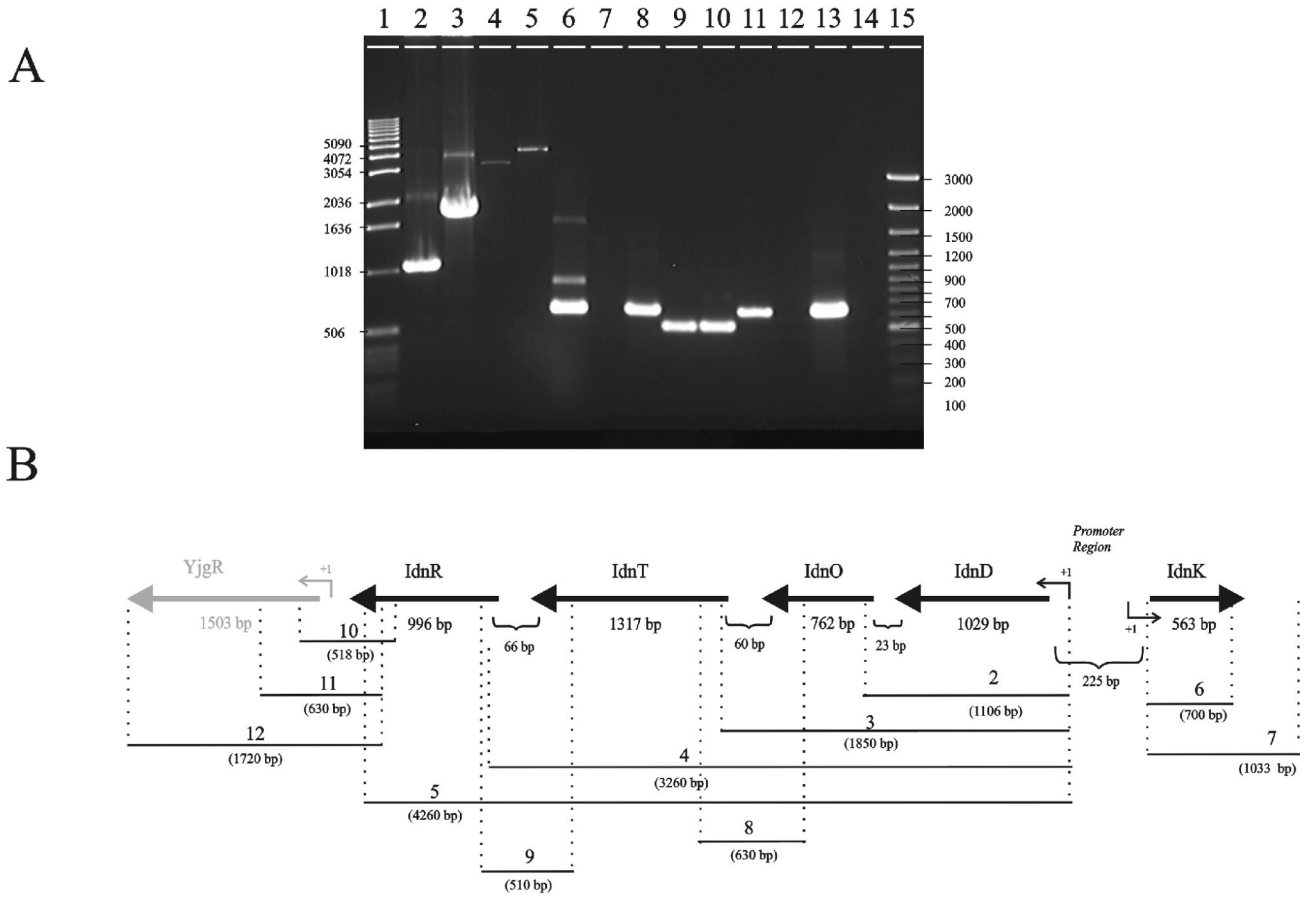

FIG. 3. RT analysis of the *idnDOTR* and *idnK* transcripts. (A) A 1.5% agarose gel showing the RT-PCR products with template RNA isolated from cells grown on MOPS compete medium containing 0.2% L-idonate. Lanes 2 to 12 correspond to regions 2 to 12 in the schematic representation (B). The RT-PCR products shown in lanes 2 to 12 were generated with primer pairs that flanked the corresponding regions depicted in the schematic. The length of each predicted RT-PCR product is shown in the schematic (in base pairs). Lanes: 1, 1-kb DNA ladder; 15, 100-bp DNA ladder; 13, control PCR product obtained from *E. coli* W1485 genomic DNA with a primer set that generated a 620-bp DNA fragment; 14, control PCR product obtained from total RNA with the same primer set as in lane 13. The values on the left and right are sizes in base pairs.

on L-idonate. By comparison, the *gntR* mutant (CB371) was unaffected for growth on L-idonate. Failure of the *idnK* mutant (MD5) to grow on MOPS minimal medium containing L-idonate is consistent with its role in phosphorylation of D-gluconate, an intermediate of the L-idonate pathway. The *idnK gntK* double mutant (MDE5) failed to grow on D-gluconate, as well as L-idonate. Interestingly, a *gntRKU* deletion mutant (NP202) can grow on D-gluconate after a lag phase of 24 h. Northern blot analysis of the *gntRKU* mutant, NP202, revealed that the *idnD* and *idnK* transcripts were fully induced when cells were grown on D-gluconate (Fig. 4). In the wild-type strain, induction of the *idn* genes by D-gluconate is minimal compared to that by L-idonate (Table 4). This result suggests that growth of the *gntK* mutant on D-gluconate causes endogenous accumulation of the inducer of *idnK* and *idnD*.

Functional genomic analysis of cells grown on idonate. Very little is known about the physiology of cells growing on rarely studied sugar acids, such as L-idonate. Therefore, we used whole-genome DNA arrays to identify genes induced by growth on MOPS complete medium containing L-idonate and D-glucose. These data sets are available on the Internet (http://www.ou.edu/microarray). The five *idn* genes were among the most strongly induced genes in cells grown on L-idonate compared to D-glucose, including *idnD* and *idnO*, which topped the

TABLE 3. Expression of *idnD-lacZ* in the *idnD* nonpolar mutant and the wild-type background

Strain	Relevant genotype	β-Galactosidase activity (Miller units[a]) with indicated carbon source[b]				
		None	Idonate	5-Ketogluconate	Gluconate	Glucose
CB130	Wild type	76 ± 3.5	930 ± 27	580 ± 33	260 ± 40	40 ± 9.7
CB361Z	Δ*idnD*	39 ± 1.2	1,477 ± 105	530 ± 20	255 ± 38	20 ± 7.7

[a] Reported values are means and standard deviations of three independent experiments performed in triplicate.
[b] Cells were grown in LB medium containing the carbon sources indicated at 0.2%.

TABLE 4. β-Galactosidase activity of CB130 (*idnK-lacZ*) and CB131 (*idnD-lacZ*) grown in MOPS complete medium

Carbon source[b]	β-Galactosidase activity (Miller units[a]) with strain:	
	CB130	CB131
Idonate + cAMP[c]	1,667 ± 169	1,816 ± 146
Idonate	750 ± 28	714 ± 33
Idonate + gluconate	441 ± 40	418 ± 33
Gluconate	279 ± 20	342 ± 19
Gluconate + cAMP[c]	405 ± 31	383 ± 43
5-Ketogluconate	499 ± 62	495 ± 57
2,5-Diketogluconate	157 ± 51	119 ± 22
2-Ketogluconate	169 ± 25	108 ± 25
Iduronate	140 ± 10	155 ± 16
2-Ketogulonate	25 ± 13	45 ± 11
Glycerol	104 ± 3.6	115 ± 18
Glucose	123 ± 8.2	116 ± 25
Succinate	155 ± 18	123 ± 34

[a] Reported values are means and standard deviations of three independent experiments performed in triplicate.
[b] Cells were grown in MOPS complete medium containing the carbon sources indicated at 0.2%.
[c] Exogenous cAMP added to growth medium at 4mM.

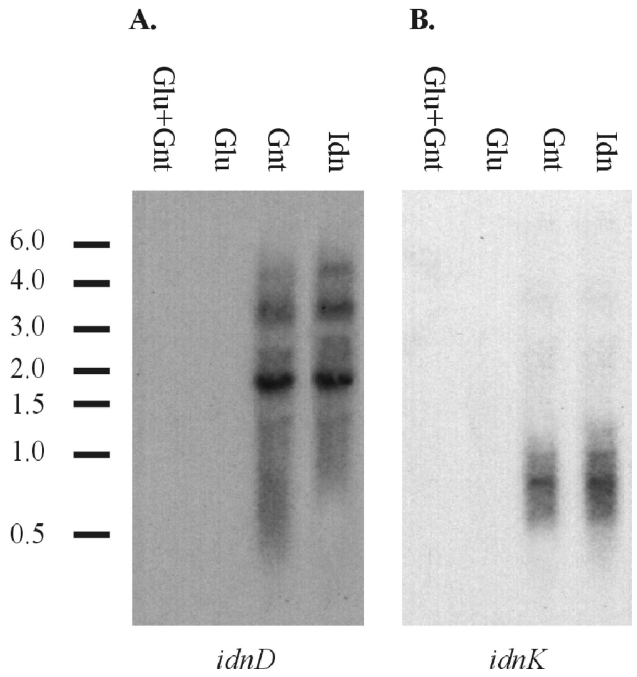

FIG. 4. Northern blot analysis of a GntI system mutant. *idnD* (A) and *idnK* (B) transcription is shown. Total RNA was isolated from late-log-phase cultures of *E. coli* NP202 (W1485 Δ*gntRKU*) grown on LB medium containing the carbohydrate listed above each lane. An aliquot of 5 μg of RNA was loaded per lane, and the bars and corresponding values to the left of each blot show the locations and sizes (in kilobases) of RNA standards.

list (Table 6). The expression profile of the *idn* genes in cells grown on L-idonate was qualitatively similar to the relative induction observed in Northern blot assays (compare Fig. 2 and Table 6). When cells were grown on L-idonate, the percentage of total transcripts in the cells was highest for *idnD*, followed by *idnO*, *idnT*, *idnR*, and *idnK* (Table 6). Of all of the transcripts in *E. coli* cells grown on L-idonate, the *idnD* and *idnO* transcripts were the 36th and 58th most highly expressed, respectively (data not shown). These levels are typical of highly expressed genes in fast-growing bacteria (12). To confirm that changes in the transcript levels of *idnD* and *idnO* directly correlated with the changes in the protein levels, proteins found to be specifically induced by growth on L-idonate were cut out of 2-D gels (Fig. 5).), digested with trypsin, and identified by MS/MS. Four spots thus analyzed were identified as being IdnD and IdnO. Two modified forms of each protein were present on the gels.

The 50 genes most highly induced on L-idonate compared to D-glucose are shown in Table 6. Only 19 of these genes encode proteins with known functions, 5 of which belong to the *idn* operon. The remaining 31 significantly induced genes encode products with unknown functions. The induction of these genes was not confirmed by other methods used for monitoring transcription, and it is not clear that their induction is relevant to growth on L-idonate. Thus, expression profiling did not shed any additional light on the physiology of growth on L-idonate.

TABLE 5. Specific growth rates of mutant strains on MOPS minimal medium

Strain	Genotype	Specific growth rate (h^{-1})[a]			
		Glu	Gnt	5KG	Idn
W1485	Wild type	1.3	1.4	<0.1	1.0
CB370	Δ*crp*	1.3	<0.1	NG[b]	NG
CB366	Δ*idnR*	1.3	1.3	NG	NG
CB371	Δ*gntR*	1.4	1.3	NG	1.1
MD5	*idnK*::*kan*	1.3	1.2	NG	NG
MDE5	*idnK*::*kan gntK*::*cat*	1.4	NG	NG	NG
NP202	Δ*gntRKU*	1.4	1.0[c]	NG	0.9

[a] Growth rate experiments were performed in triplicate, and all standard deviations were less than ± 0.1. Abbreviations: D-glucose, Glu; D-gluconate, Gnt; 5-ketogluconate, 5KG; L-idonate, Idn.
[b] NG, no growth.
[c] Growth after >24-h lag.

DISCUSSION

The organization of the genes of the L-idonate pathway, which is suggested by the arrangement of the pathway genes around a divergent regulatory region, was confirmed in these studies. Transcription start sites for the divergent promoters are positioned at −29 and −26 relative to the *idnD* and *idnK* start codons, respectively, consistent with the predicted −10 and −35 promoter elements (Fig. 1). The pathway genes are arranged in two transcription units, the *idnDOTR* operon, and the divergently transcribed, monocistronic *idnK* gene (Fig. 2 and 3). The putative regulatory elements identified within the *idn* regulatory region provide some interesting clues regarding the regulation of the *idn* genes. A putative CRP binding site is positioned at −41.5 relative to the *idnD* transcription start site, suggesting a CRP-dependent class II promoter (33). The UP element at −42.5 bp relative to the *idnK* transcription start site is in a position expected to improve transcription initiation at the *idnK* promoter (22). The location of the putative IdnR binding site centered between the *idnD* and *idnK* transcription start sites suggests that IdnR may coordinately regulate both promoters.

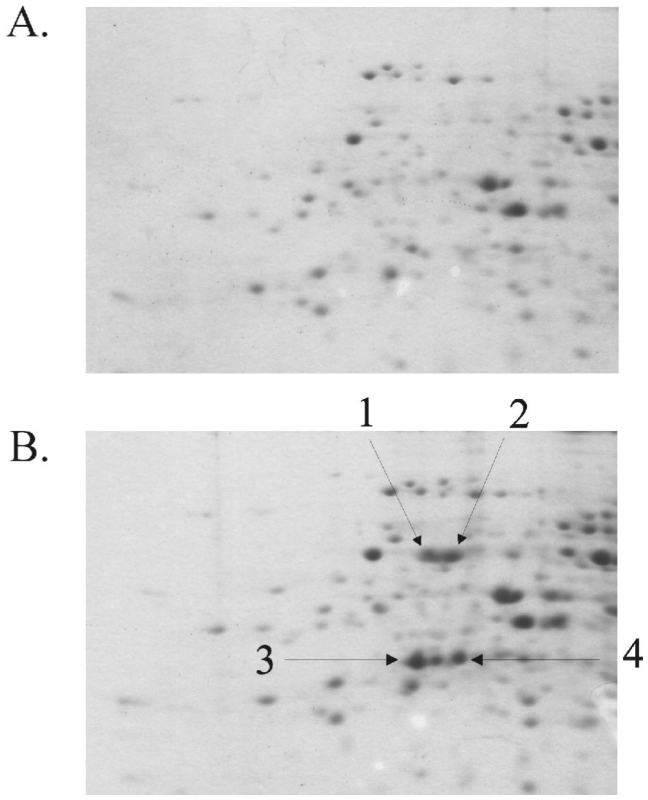

FIG. 5. 2-D PAGE of extracted proteins from cells grown in MOPS minimal medium containing 0.2% D-glucose (A) or L-idonate (B). The two modified forms of IdnD and IdnO were identified by MS/MS as described in the text and are indicated by arrows 1 and 2 and arrows 3 and 4, respectively.

L-Idonate and 5KG both induced the L-idonate pathway, as indicated by induction of *idnD* and *idnK* reporter fusions in an *idnD* nonpolar mutant that cannot interconvert L-idonate and 5KG (Table 3). The induction ratios of the *idnD* and *idnK* promoters were remarkably similar, indicating that transcription from the divergent promoters is, in fact, coordinated (Table 4). This coordinated expression apparently provides a mechanism by which to balance flux through the L-idonate pathway and maintain concentrations of the pathway intermediates at levels required for induction of the pathway genes and for appropriate regulation of the closely associated GntI pathway.

The relative order of *idn* transcript abundance in the Northern blot and DNA array experiments (Fig. 2 and Table 6, respectively) indicates that *idnD* and *idnO* are the most highly expressed *idn* transcripts (in that order), followed by *idnT*, *idnR*, and *idnK*. Thus, their relative expression levels are correlated with their proximity to the promoters. The low level of *idnR* expression is consistent with the known expression level of most regulators (9). The lower level of *idnT* and *idnK* expression suggests that flux through the pathway could be limited by L-idonate transport and phosphorylation. In addition to being highly induced by growth on L-idonate, *idnD* and *idnO* were among the most highly expressed genes in the *E. coli* transcriptome (Table 6) and their products were among the most abundant proteins (Fig. 5).

The relative levels of gene-specific *idn* gene transcripts appear to be controlled by posttranscriptional processing and/or mRNA secondary structures that could act as terminators. Under inducing conditions, there was a low level of the full-length *idnDOTR* transcript and shorter gene-specific transcripts were observed. The relatively high abundance of 1.9-kb *idnDO* and 3.3-kb *idnDOT* transcripts suggests that the predicted mRNA stem-loop structures located at the 3′ ends of the *idnO* and *idnT* genes may function as transcriptional terminators. The alternative possibility that the gene-specific transcripts correspond to promoters within the *idnDOTR* operon was not tested. The 3′ end of the *idnDOTR* transcript does not appear to contain any secondary structure indicative of a terminator, and transcription of the operon was found to extend into the 5′ end of the downstream *yjgR* gene (Fig. 3). However, *yjgR* knockout mutants grew normally on L-idonate and *yjgR* was not induced in cells grown on L-idonate, indicating that YjgR is not involved in L-idonate catabolism.

Catabolite repression of the L-idonate pathway indicates that glucose and D-gluconate are preferred over L-idonate (Table 4); the slower growth rate of cells on L-idonate seems to explain this hierarchy of nutrient choice (Table 5). Hogema et al. (10) demonstrated that D-gluconate is catabolite repressing because it lowers the intracellular cAMP and CRP concentrations through a mechanism that does not involve the phosphotransferase system (PTS) $EIIA^{Glu}$ enzyme. This explains why the addition of cAMP did not fully relieve the repression of the *idn* genes caused by D-gluconate (Table 4). In the presence of catabolite-repressing sugars such as D-glucose and D-gluconate, cAMP and CRP levels are low and transcription of the *idn* genes is not induced. Only in the absence of catabolite-repressing sugars, when L-idonate or 5KG is present, are the *idn* genes fully expressed.

Failure of the *idnK* mutant (MDE5) to grow on L-idonate indicates that the presumed intracellular accumulation of D-gluconate formed by IdnD and IdnO did not reach levels high enough to induce the GntI system for D-gluconate catabolism, specifically *gntK*, the *idnK* paralog. This result suggests that transcription of the GntI and GntII systems is tuned to the concentrations of inducers such that the D-gluconate and L-idonate pathways are regulated appropriately (i.e., GntI is induced by gluconate and GntII is induced by L-idonate). This possibility is being explored.

The operation of GntII as a subsidiary gluconate pathway was examined in a *gntRKU* mutant (Fig. 4 and Table 5) that exhibits a lag before initiating growth on D-gluconate (11). It was previously suggested that the physiological reason why 5KG functions as an inducer of the L-idonate pathway could be to act as an endogenous inducer of the GntII system for subsidiary D-gluconate catabolism (31). Since cells grow poorly on 5KG, it is unlikely that 5KG is physiologically relevant as a growth substrate. Induction of the GntII system in the GntI mutant can be attributed to accumulation of D-gluconate in mutants blocked in gluconate kinase (e.g., *gntK*); in turn, the accumulated D-gluconate could be converted to 5KG by the basal level of IdnO, a freely reversible enzyme that converts D-gluconate to 5KG with NAD as a cofactor (3). As 5KG accumulates, it would induce the subsidiary D-gluconate kinase

TABLE 6. The fifty most highly induced *E. coli* genes in cells grown on L-idonate

Gene	Gene product	Intensity[a]	Ratio[b]	*P* value[c]
idnD	L-Idonate dehydrogenase	0.254	1.6	1.4×10^{-6}
idnO	5-Ketogluconate:gluconate oxidoreductase	0.187	1.3	1.9×10^{-7}
b2790	ORF[d] hypothetical protein	0.172	0.97	2.4×10^{-4}
yheH	Putative export protein for general secretion pathway	0.041	0.93	9.1×10^{-4}
idnK	Gluconate kinase	0.039	0.83	2.1×10^{-3}
yfeK	ORF, hypothetical protein	0.055	0.81	4.8×10^{-4}
yhfP	ORF, hypothetical protein	0.022	0.79	5.9×10^{-4}
yicM	Putative transport protein	0.104	0.78	8.9×10^{-4}
yhdT	ORF, hypothetical protein	0.035	0.78	5.0×10^{-5}
narW	Cryptic nitrate reductase 2, delta subunit	0.019	0.73	5.4×10^{-4}
flhA	Flagellar biosynthesis; possible export function	0.108	0.72	2.9×10^{-3}
ymfL	ORF, hypothetical protein	0.021	0.72	1.6×10^{-4}
rimJ	Acetylation of N-terminal alanine protein S5	0.124	0.69	2.6×10^{-8}
ymfO	ORF, hypothetical protein	0.050	0.68	2.0×10^{-3}
b2459	ORF, hypothetical protein	0.108	0.66	6.2×10^{-4}
yggC	Putative kinase	0.031	0.60	3.8×10^{-4}
narJ	Nitrate reductase 1, delta subunit, assembly function	0.068	0.60	8.9×10^{-5}
narU	Nitrite extrusion protein 2	0.014	0.59	9.7×10^{-4}
yjiW	ORF, hypothetical protein	0.017	0.58	5.0×10^{-4}
yiaC	ORF, hypothetical protein	0.041	0.58	1.1×10^{-4}
thiM	Hydoxyethylthiazole kinase	0.089	0.57	2.0×10^{-3}
b1600	Possible chaperone	0.017	0.56	5.9×10^{-4}
ycal	ORF, hypothetical protein	0.032	0.56	7.3×10^{-4}
b1152	ORF, hypothetical protein	0.015	0.55	2.3×10^{-3}
hyaF	Nickel incorporation into hydrogenase 1 proteins	0.045	0.54	2.5×10^{-5}
b2460	ORF, hypothetical protein	0.061	0.54	2.5×10^{-3}
ymfA	ORF, hypothetical protein	0.020	0.53	4.7×10^{-4}
nrfE	Formate-dependent nitrite reductase, assembly function	0.381	0.53	2.3×10^{-3}
yibK	ORF, hypothetical protein	0.027	0.53	2.8×10^{-3}
aceE	Pyruvate dehydrogenase (decarboxylase component)	0.046	0.52	5.8×10^{-4}
evgS	Putative sensor for regulator EvgA	0.205	0.51	7.4×10^{-7}
ycjV	Putative ATP binding component of a transport system	0.088	0.51	6.9×10^{-5}
queA	Synthesis of queuine in tRNA	0.080	0.51	2.7×10^{-3}
idnR	L-Idonate operon regulator	0.045	0.51	1.2×10^{-3}
idnT	L-Idonate transporter	0.064	0.51	6.0×10^{-5}
yafV	Putative EC 3.5 amidase-type enzyme	0.023	0.51	1.7×10^{-3}
yi21-2	IS2 hypothetical protein	0.057	0.50	1.3×10^{-3}
cynT	Carbonic anhydrase	0.034	0.50	2.2×10^{-5}
yhgG	ORF, hypothetical protein	0.019	0.50	1.5×10^{-4}
b1565	ORF, hypothetical protein	0.011	0.50	6.6×10^{-10}
yi21-6	IS2 hypothetical protein	0.046	0.49	1.3×10^{-3}
yihN	Putative resistance protein (transport)	0.020	0.49	1.8×10^{-4}
b2931	Putative oxidoreductase	0.097	0.48	2.7×10^{-3}
ychM	ORF, hypothetical protein	0.160	0.48	3.2×10^{-5}
b1141	ORF, hypothetical protein	0.080	0.48	2.3×10^{-3}
trpD	Anthranilate synthase component II	0.177	0.48	1.2×10^{-5}
yrdB	ORF, hypothetical protein	0.013	0.47	1.2×10^{-3}
yhfY	ORF, hypothetical protein	0.019	0.47	4.0×10^{-5}
entA	2,3-dihydroxybenzoate dehydrogenase, enterochelin	0.017	0.47	2.4×10^{-4}
yjgH	ORF, hypothetical protein	0.085	0.46	1.4×10^{-4}

[a] Intensity normalized to transcript levels in cells grown on L-idonate and expressed as a percentage of the sum of the transcript levels of all of the genes in the arrays (5).
[b] Log_{10} ratio of normalized transcript levels on L-idonate compared to those on D-glucose.
[c] Each *P* value (scientific notation) indicates the probability that the reported $\log_{10}$ ratio is significant.
[d] ORF, open reading frame.

encoded by *idnK*, which can functionally substitute for GntK of the GntI pathway for D-gluconate catabolism. This same mechanism would also be expected to substitute for GntT in a *gntT* mutant by inducing the subsidiary D-gluconate transporter IdnT.

We used functional genomic tools to ensure that nothing was overlooked regarding the physiology of growth on L-idonate. As predicted, the genes of the L-idonate pathway were induced by growth on idonate (Table 6). What was not anticipated was the induction of genes such as *araD*, *narW*, *thiM*, *hyaF*, and *nrfE*. The induction of these genes has not been confirmed by other methods used to monitor transcription, and it is not clear that their induction is relevant to growth on L-idonate. Thus, expression profiling failed to shed any additional light on the physiology of growth on L-idonate.

We investigated the translation of the *idnD* and *idnO* transcripts and determined that the protein level directly correlated with the transcript level, suggesting little, if any, translation control in expression of the *idnDOTR* transcript. The 2-D gel analysis revealed duplicate spots for both IdnD and IdnO,

suggesting that a charged group had modified these proteins and altered their mobility in the gel (Fig. 5). The only protein-modifying enzyme that was induced by growth on L-idonate was *rimJ*, which encodes an N-terminal acetyltransferase that modifies ribosomal protein S5 (Table 6). It is unlikely that RimJ modifies IdnD or IdnO, since acetyl groups are generally neutral in charge. Alternatively, the negatively charged molecules L-idonate and 5KG may have remained bound to the catalytic sites of these proteins during extraction, thereby changing their overall charge.

In summary, the results presented here indicate that the *idn* genes are organized in two coordinately regulated operons, *idnDOTR* and *idnK*. The *idn* genes are specifically induced by L-idonate and 5KG and are catabolite repressed by glucose and gluconate. Whole-genome expression profiling of cells growing on L-idonate indicated that the majority of the genes induced code for proteins of unknown function and thus reveal little about the physiology of growth on L-idonate. Lastly, D-gluconate does not normally induce the *idn* (GntII) genes unless the GntI system is nonfunctional and does so apparently by formation of the endogenous inducer 5KG.

ACKNOWLEDGMENTS

We thank April Anderson for critical reading of the manuscript.

Work on this project was supported by grants from the NSF (MCB-9723593) and NIH (AI48945), as well as a generous gift from Genencor International.

REFERENCES

1. **Bachi, B., and H. L. Kornberg.** 1975. Genes involved in the uptake and catabolism of gluconate by *Escherichia coli*. J. Gen. Microbiol. **90:**321–335.
2. **Bachmann, B. J.** 1996. Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12, p. 2460–2488. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.
3. **Bausch, C., N. Peekhaus, C. Utz, T. Blais, E. Murray, T. Lowary, and T. Conway.** 1998. Sequence analysis of the GntII (subsidiary) system for gluconate metabolism reveals a novel pathway for L-idonic acid catabolism in *Escherichia coli*. J. Bacteriol. **180:**3704–3710.
4. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248–254.
5. **Conway, T., B. Kraus, D. L. Tucker, D. J. Smalley, A. F. Dorman, and L. McKibben.** 2002. DNA array analysis in a Microsoft Windows environment. BioTechniques **32:**110, 112–114, 116, 118–119.
6. **Datsenko, K. A., and B. L. Wanner.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc. Natl. Acad. Sci. USA **97:**6640–6645.
7. **Devreese, B., F. Vanrobaeys, and J. Van Beeumen.** 2001. Automated nanoflow liquid chromatography/tandem mass spectrometric identification of proteins from *Shewanella putrefaciens* separated by two-dimensional polyacrylamide gel electrophoresis. Rapid Commun. Mass Spectrom. **15:**50–56.
8. **Estrem, S. T., T. Gaal, W. Ross, and R. L. Gourse.** 1998. Identification of an UP element consensus sequence for bacterial promoters. Proc. Natl. Acad. Sci. USA **95:**9761–9766.
9. **Gottesman, S.** 1984. Bacterial regulation: global regulatory networks. Annu. Rev. Genet. **18:**415–441.
10. **Hogema, B. M., J. C. Arents, T. Inada, H. Aiba, K. Van Dam, and P. W. Postma.** 1997. Catabolite repression by glucose 6-phosphate, gluconate, and lactose in *Escherichia coli*. FEMS Microbiol. **24:**857–867.
11. **Istúriz, T., E. Palmero, and J. Vitelli-Flores.** 1986. Mutations affecting gluconate catabolism in *Escherichia coli*. Genetic mapping of the locus for the thermosensitive gluconokinase. J. Gen. Microbiol. **132:**3209–3212.
12. **Karlin, S., J. Mrazek, A. Campbell, and D. Kaiser.** 2001. Characterizations of highly expressed genes of four fast-growing bacteria. J. Bacteriol. **183:**5025–5040.
13. **Luria, S. E., and M. Delbruck.** 1943. Mutations of bacteria from virus sensitivity to virus resistance. Genetics **28:**491–511.
14. **MacConkey, A.** 1905. Lactose fermenting bacteria in faces. J. Hyg. **5:**333–378.
15. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Springs Harbor Laboratory, Cold Springs Harbor, N.Y.
16. **Nagel de Zwaig, R., N. Zwaig, T. Istúriz, and R. S. Sánchez.** 1973. Mutations affecting gluconate metabolism in *Escherichia coli*. J. Bacteriol. **114:**463–468.
17. **Neidhardt, F. C., P. L. Bloch, and D. F. Smith.** 1974. Culture medium for enterobacteria. J. Bacteriol. **119:**736–747.
18. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85:**2444–2448.
19. **Peekhaus, N., and T. Conway.** 1998. Positive and negative transcriptional regulation of the *Escherichia coli* gluconate regulon gene *gntT* by GntR and the cyclic AMP (cAMP)-cAMP receptor protein complex. J. Bacteriol. **180:**1777–1785.
20. **Peekhaus, N., S. Tong, J. Reizer, M. H. Saier, Jr., E. Murray, and T. Conway.** 1997. Characterization of a novel transporter family that includes multiple *Escherichia coli* gluconate transporters and their homologues. FEMS Microbiol. Lett. **147:**233–238.
21. **Porco, A., N. Peekhaus, C. Bausch, S. Tong, T. Isturiz, and T. Conway.** 1997. Molecular genetic characterization of the *Escherichia coli gntT* gene of GntI, the main system for gluconate metabolism. J. Bacteriol. **179:**1584–1590.
22. **Ross, W., A. Ernst, and R. L. Gourse.** 2001. Fine structure of *E. coli* RNA polymerase-promoter interactions: alpha subunit binding to the UP element minor groove. Genes Dev. **15:**491–506.
23. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
24. **Sanchez, J. C., V. Rouge, M. Pisteur, F. Ravier, L. Tonella, M. Moosmayer, M. R. Wilkins, and D. F. Hochstrasser.** 1997. Improved and simplified in-gel sample application using reswelling of dry immobilized pH gradients. Electrophoresis **18:**324–327.
25. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
26. **Shine, J., and L. Dalgarno.** 1975. Terminal-sequence analysis of bacterial ribosomal RNA. Correlation between the 3′-terminal-polypyrimidine sequence of 16-S RNA and translational specificity of the ribosome. Eur. J. Biochem. **57:**221–230.
27. **Simons, R. W., F. Houman, and N. Kleckner.** 1987. Improved single and multicopy lac-based cloning vectors for protein and operon fusions. Gene **53:**85–96.
28. **Tao, H., C. Bausch, C. Richmond, F. R. Blattner, and T. Conway.** 1999. Functional genomics: expression analysis of *Escherichia coli* growing on minimal and rich media. J. Bacteriol. **181:**6425–6440.
29. **Tong, S., A. Porco, T. Isturiz, and T. Conway.** 1996. Cloning and molecular genetic characterization of the *Escherichia coli gntR*, *gntK*, and *gntU* genes of GntI, the main system for gluconate metabolism. J. Bacteriol. **178:**3260–3269.
30. **Truesdell, S. J., J. C. Sims, P. A. Boerman, J. L. Seymour, and R. A. Lazarus.** 1991. Pathways for metabolism of ketoaldonic acids in an *Erwinia* sp. J. Bacteriol. **173:**6651–6656.
31. **Tsunedomi, R., H. Izu, T. Kawai, K. Matsushita, T. Ferenci, and M. Yamada.** 2003. The activator of GntII genes for gluconate metabolism, GntH, exerts negative control of GntR-regulated GntI genes in *Escherichia coli*. J. Bacteriol. **185:**1783–1795.
32. **Wanner, B. L., R. Kodaira, and F. C. Neidhardt.** 1977. Physiological regulation of a decontrolled *lac* operon. J. Bacteriol. **130:**212–222.
33. **Zhou, Y., T. J. Merkel, and R. H. Ebright.** 1994. Characterization of the activating region of *Escherichia coli* catabolite gene activator protein (CAP). II. Role at class I and class II CAP-dependent promoters. J. Mol. Biol. **243:**603–610.