A previously undescribed pathway for pyrimidine catabolism

Kevin D. Loh†, Prasad Gyaneshwar†*, Eirene Markenscoff Papadimitriou†*, Rebecca Fong*, Kwang-Seo Kim*, Rebecca Parales§, Zhongrui Zhou, William Inwood†, and Sydney Kustu‡,§

*Department of Plant and Microbial Biology, 111 Koshland Hall, University of California, Berkeley, CA 94720-3102; †Section of Microbiology, 1 Shields Avenue, University of California, Davis, CA 95616; ‡College of Chemistry, 8 Lewis Hall, University of California, Berkeley, CA 94720-1460

Contributed by Sydney Kustu, January 19, 2006

The b1012 operon of Escherichia coli K-12, which is composed of seven unidentified ORFs, is one of the most highly expressed operons under control of nitrogen regulatory protein C. Examination of strains with lesions in this operon on Biolog Phenotype MicroArray (PM3) plates and subsequent growth tests indicated that they failed to use uridine or uracil as the sole nitrogen source and that the parental strain could use them at room temperature but not at 37°C. A strain carrying an ntrB(Con) mutation, which elevates transcription of genes under nitrogen regulatory protein C control, could also grow on thymidine as the sole nitrogen source, whereas strains with lesions in the b1012 operon could not. Growth-yield experiments indicated that both nitrogen of uridine and thymidine were available. Studies with [14C]uridine indicated that a three-carbon waste product from the pyrimidine ring was excreted. After trimethylsilylation and gas chromatography, the waste product was identified by mass spectrometry as 3-hydroxypropionic acid. In agreement with this finding, 2-methyl-3-hydroxypropionic acid was released from thymidine. Both the number of available nitrogens and the waste products distinguished the pathway encoded by the b1012 operon from pyrimidine catabolic pathways described previously. We propose that the genes of this operon be named rutA–G for pyrimidine utilization. The product of the divergently transcribed gene, b1013, is a tetracycline repressor family regulator that controls transcription of the b1012 operon negatively.

Escherichia coli | GC/MS | pyrimidine degradation | αM | ORFs

Uracil and thymine are catabolized by two known pathways called reductive and oxidative (Fig. 1A) (1). The reductive pathway occurs in a wide variety of organisms, from archaea and bacteria to humans. The oxidative pathway, less well studied, has been found only in a variety of bacteria (2, 3). Although there is also a pathway for degradation of orotic acid (4), an intermediate in pyrimidine biosynthesis, there is no means reported for carboxylating uracil to orotic acid.

The reductive pathway for pyrimidine degradation yields NH3 and CO2 from both uracil and thymine (Fig. 1A upper pathway). The third product from uracil is β-alanine, and that from thymine is β-aminoisobutyric acid. Escherichia coli B, which has the reductive pathway for pyrimidine degradation, obtains only one utilizable nitrogen (N) per mole of uracil and thus apparently fails to obtain additional N from β-alanine (5).

The oxidative pathway for pyrimidine degradation yields urea from both uracil and thymine (Fig. 1A lower pathway). The second product from uracil is malonic acid, and that from thymine is methylnalonic acid. E. coli lacks urease and hence would obtain no utilizable N from the oxidative pathway.

In previous studies of all genes under control of the transcriptional activator nitrogen regulatory protein C (NtrC, also called GlnG) (6) in E. coli K-12, we found that the b1012 operon, which is composed of seven unidentified ORFs, was one of the most highly expressed. The product of the b1006 gene was predicted to be a nucleobase transporter (7), and most operons under NtrC control are involved in transport and/or catabolism of alternative N sources. Here we present evidence that the b1012 operon codes for proteins that constitute a previously undescribed pathway for pyrimidine degradation and thereby confirm the view of Simaga and Kos (8, 9) that E. coli K-12 does not use either of the known pathways.

Results

Behavior on Biolog Phenotype MicroArray Plates. We tested our parental strain NCM3722 and strains with mini Tn5 insertions in several genes of the b1012 operon on Biolog (Hayward, CA) Phenotype MicroArray plates (PM3) for nitrogen catabolism, which couple utilization of various N sources to reduction of a tetrazolium dye and production of purple color (10, 11). We also tested a strain carrying an ntrB(Con) mutation, NCM3876, and corresponding mutant derivatives. Strains NCM3722 and NCM3876 [ntrB(Con)] could respire with uracil or uridine as N source at room temperature, whereas mutant derivatives could not (Fig. 2 and data not shown). Parental strain NCM3876 [ntrB(Con)] could also respire with thymidine or thymine as N source (data not shown). None of the strains could respire with pyrimidines as N source at 37°C.

Bioinformatic Analysis. A BLASTP search indicated that genes of the b1012 operon were not homologous to those of known pathways for pyrimidine degradation, but products of the operon were similar to proteins of known function (Table 1). Genes in the b1012 operon were homologous to clustered genes in only eight other organisms, all proteobacteria (Table 3, which is published as supporting information on the PNAS web site). Each organism contained homologues of b1009–b1012, indicating that products of these genes were probably central to function. With one exception, the b1012-like operon appeared to have a consensus αM promoter and at least one upstream NtrC binding site. With a different exception, a homologue of b1013, which is predicted to be a TetR (tetracycline repressor) family regulator, was adjacent to the operon. The b1013 homologue was also found adjacent to genes of the reductive pathway for pyrimidine catabolism in bacteria that have this pathway. These findings provided a hint that the product of b1013, which appears to be transcribed divergently from the b1012 operon in E. coli, might be an auxiliary regulator for the operon.

Initial Growth Tests. To confirm results on Biolog plates, we grew NCM3722, NCM3876 [ntrB(Con)], and congenic strains carrying

Conflict of interest statement: No conflicts declared.

Abbreviation: NtrC, nitrogen regulatory protein C.

1Present address: 71-254 CHS, Department of Epidemiology, University of California School of Public Health, Los Angeles, CA 90095-1772.
2Present address: Biotechnology Institute, University of Minnesota, 140 Gortner Laboratory, 1479 Gortner Avenue, St. Paul, MN 55108.
3Present address: Département de Microbiologie et Médecine Moléculaire, Université de Geneva, 1, Rue Michel-Servet, CH-1211 Geneva 4, Switzerland.

*To whom correspondence should be addressed. E-mail: kustu@nature.berkeley.edu.

© 2006 by The National Academy of Sciences of the USA
mini-Tn5 insertions in six of the seven genes of the b1012 operon (a lesion in b1006, the presumed uracil transporter, was not available) on N/C minimal medium, with glycerol as carbon source, and ammonium, uridine, or thymidine as sole N source. (In general we used ribonucleosides or deoxyribonucleosides rather than free bases because they are more soluble.) Uridine is converted to uracil by uridine phosphorylase (udp, b3831) and thymidine is converted to thymine by thymidine phosphorylase (deoA, b4382). All strains grew on ammonium at room temperature (22°C) and at 37°C. Only NCM3722 and NCM3876 grew on uridine and only at room temperature (Fig. 3 and data not shown). Only NCM3876 grew on thymidine at room temperature. As expected from DNA microarray studies indicating that NtrC was required for expression of the b1012 operon, strain NCM3877 (ntrC) failed to grow on uridine as N source (data not shown). Strains NCM3991 (b1007::Tn5) and NCM4023 [ntrB(Con) b1007::Tn5] yielded revertants that grew on uridine. These revertants were sensitive to kanamycin, and gel electrophoresis of an appropriate DNA fragment followed by DNA sequencing indicated that the mini-Tn5 insertion had precisely excised (data not shown). Strain NCM3876 [ntrB(Con)] yielded suppressor strains able to grow on uridine at 37°C (K.D.L. and S.K., unpublished data). Lesions in these strains were not linked to the b1012 operon by phage P1-mediated transduction and have not been studied further.

Growth Yields. To determine how many moles of utilizable N strain NCM3722 could obtain per mole of pyrimidine we first

![Fig. 1. Pyrimidine catabolic pathways. Known reductive (1) and oxidative (2) pathways for catabolism of pyrimidine rings (A, upper and lower, respectively) and the pathway described in this work (B). Note that ureidomalonate acid and ureidomalonase are not analogous to ureidopropionic acid and ureidopropionase and should probably be renamed.](image)

![Fig. 2. Respiration by wild-type (A) and a strain carrying a lesion in the b1012 operon (B) in medium containing various N sources. Biolog Phenotype MicroArray (PM3) plates were inoculated at an OD400 of 0.012 and incubated for 4 days at room temperature. Apart from differences in respiration with uracil or uridine as N source, a difference with cytosine (F5) was most notable.](image)

![Table 1. Conserved domain types for gene products of the b1012 operon](image)
compared its growth yield on limiting amounts of uridine to that on ammonium (Fig. 3 and Table 2). As assessed by OD_{600}, the yield per mole of uridine was approximately twice that on ammonium. Although strain NCM3996, which carries a mini-Tn5 insertion in b1012, had both the same growth rate and cell yield as NCM3722 on ammonium, it failed to grow on uridine. Next, we compared growth yields on cytidine to those on ammonium (Table 2). Cytidine carries a third N outside the pyrimidine ring. This N is released as ammonium by cytidine deaminase (add, b2143), which is known to be under NtrC control, to yield uridine. The yield of NCM3722 on cytidine was approximately three times that on ammonium, whereas the yield of NCM3996 (b1012::Tn5) was approximately equal to that on ammonium.

For NCM3876 [ntrB(Con)] we compared growth yields on thymidine, uridine, and cytidine to those on ammonium. Although strain NCM3876 grows slower than its parental strain NCM3722 on ammonium as N source for reasons that are not defined (6), it has approximately the same yield (Table 2). The yield of NCM3876 per mole of thymidine was twice that per mole of ammonium. As was true of NCM3722, the yields of NCM3876 per mole of uridine and cytidine were, respectively, approximately twice and three times that on ammonium. Congenic strain NCM4011, which carries a lesion in b1012, failed to grow on thymidine or uridine, and its yield on cytidine was approximately equal to that on ammonium (data not shown).

**Evidence for a Three-Carbon Waste Product from the Uracil Ring.**

Based on the evidence that C-2 of uridine was released as CO_{2} (8, 9) and that both N atoms could be used and were likely to be released as ammonium, we postulated that *E. coli* K-12 might produce and excrete a three-carbon waste product from positions 4–6 of the uridine ring (Fig. 1). To test this hypothesis, we first monitored 14C from uniformly 14C-labeled uridine (nine carbons total) as NCM3722 degraded the pyrimidine ring (Fig. 4; see Methods). Approximately 40% of the label remained in an acidified supernatant when both of the ring N atoms had been incorporated into cell material. This amount was slightly more than the 33% predicted.

To obtain further evidence that the residual 14C-labeled material in the supernatant was derived from the uracil ring rather than the ribosyl moiety of uridine, we performed two additional experiments. First, we added more glycerol and ammonium to both experimental and control cultures and found that 90% of the initial label remained in the supernatant when growth ceased (data not shown). Thus, the labeled material could not be assimilated. We also repeated the experiment with ribose rather than glycerol as the carbon source so that the cells would have only a single carbon source. As in the previous experiment, 43% of the label remained in the supernatant when the 0.5 mM uridine provided had been consumed (data not shown). Moreover, when we added additional unlabeled uridine, ribose, and finally excess NH_{4}Cl to the latter cultures, essentially all of the label remained in the supernatant. All experiments confirmed the view that 14C-labeled material present in cell supernatants after both N atoms of uridine had been assimilated represented a three-carbon waste product derived from the uracil ring.

**Table 2. Growth yield**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Yield, mol of N per mol</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NH\textsubscript{4}\textsuperscript{+}</td>
</tr>
<tr>
<td>NCM3722 (parental)</td>
<td>1</td>
</tr>
<tr>
<td>NCM3996 (lesion in b1012)</td>
<td>1</td>
</tr>
<tr>
<td>NCM3876 [ntrB(Con)]</td>
<td>1</td>
</tr>
<tr>
<td>NCM4011 [ntrB(Con),</td>
<td>1</td>
</tr>
<tr>
<td>lesion in b1012]</td>
<td></td>
</tr>
</tbody>
</table>

A growth yield of 1 corresponds to a OD_{600} of 0.35. – No growth.

**Fig. 3.** Growth (A) and cell yield (B) of NCM3722 (wild-type, black circles) and NCM3996 (b1012::Tn5, gray circles) on NH\textsubscript{4}Cl (open circles) or uridine (filled circles). (A) Cells were grown with NH\textsubscript{4}Cl or uridine (1 mM) as N source and glycerol as C source at 22°C. Doubling times for NCM3722 on NH\textsubscript{4}Cl and uridine were 3.1 and 7.6 h, respectively. The doubling time of NCM3996 on NH\textsubscript{4}Cl was 3.2 h. (B) Final OD_{600} was determined as a function of the initial concentration of N source.

**Fig. 4.** Fate of [U-14C]uridine. Wild-type strain NCM3722 grown on glycerol and NH\textsubscript{4}Cl (5 mM) was diluted 1:100 into medium containing glycerol and uridine (0.5 mM) containing radioactive tracer (see Methods). Radioactivity in cell suspensions (dotted circles), in cells (triangles), and in cell supernatants treated with acid to remove CO\textsubscript{2} (crossed squares) was measured as a function of OD\textsubscript{600}, which was measured for a parallel culture that did not contain radioactive tracer. Residual radioactivity in the acid supernatant after both N atoms of the uridine ring had been consumed indicated that a 14C-labeled waste product was excreted (see text).
product, we performed additional qualitative and quantitative experiments. Authentic 3-hydroxypropionic acid had the same retention time and mass spectrum as the material identified by electron impact spectrum of a peak eluted on GC/MS at −7.75 min. This peak was found in supernatants of uridine but not NH₄⁺-grown cells. (B) Electron impact spectrum of 3-hydroxypropionic acid from the National Institute of Standards and Technology Library.

Effects of a b1013 Lesion. Introducing a mini-Tn5 lesion in b1013 into strains NCM3722 or NCM3876 [ntrB(Con)] by P1-mediated transduction did not appear to affect their growth on uridine or cytidine as N source (plate tests). Growth of strain NCM4063 [ntrB(Con) b1013::Tn5], which carries the b1013 lesion, on thymidine was also equivalent to that of its parent strain NCM3876 [ntrB(Con)]. However, unlike strain NCM3722, which fails to grow on thymidine as N source, strain NCM4062 (b1013::Tn5) acquired the ability to grow, although not as well as NCM3876 or NCM4063. The latter result provided evidence that the b1013 product, a putative TetR family regulator, controlled expression of the b1012 operon negatively. This finding was confirmed by Northern blot analysis (Fig. 6). Levels of mRNA for b1012 were higher in strain NCM4062 than strain NCM3722 when cells were grown on arginine as the limiting N source (~5-fold). When they were grown on uridine, mRNA levels were similar in the two strains and similar to those of NCM4062 on arginine. Thus, a pyrimidine or intermediate in the pyrimidine degradation pathway probably serves as a ligand for the protein product of b1013 and relieves its negative effects. As expected for an operon under NtrC control, levels of b1012 mRNA were very low in both strains on NH₄Cl as N source.

Discussion

The evidence presented indicates that E. coli K-12 can obtain both N atoms from the pyrimidine ring by using a catabolic pathway not previously characterized in any organism. The pathway, which involves six gene products and a predicted uracil transporter, degrades the uracil ring to 3-hydroxypropionic acid and the thymine ring to 2-methyl-3-hydroxypropionic acid, both of which are waste products for E. coli K-12. (Lactic acid is 2-hydroxypropionic acid.) The N atoms from the pyrimidine ring are probably released as ammonium. Pyrimidine rings can be used as sole N source at room temperature but not at 37°C, which would account for why their catabolism was missed in earlier tests of the metabolic capacities of E. coli K-12 (10, 11).

Because the designation put has already been taken for proline utilization, we propose that the genes of the b1012 operon be named rutA through G (b1012–b1006) (Table 1), with the rut acronym indicating pyrimidine utilization. The adjacent gene, b1013, should be rutR for rut repressor. If the perfect inverted repeat just upstream of the rut promoter serves as a binding site for RutR, this auxiliary regulator may hinder contact between NtrC and σ⁵-holoenzyme by stiffening the DNA between the enhancer and the promoter (15). Presumably, RutR would fail to bind in the presence of its pyrimidine-related ligand. A homologue of RutR, AcrR, also functions as a secondary modulator of gene expression by σ⁵-holoenzyme (16), as does ArgR, an unrelated protein, at the σ⁵-dependent ast promoter (17).

The b1012 (rut) operon is absent in Salmonella typhimurium LT2, a close relative of E. coli K-12. The genes that flank the operon, wrhA and rutR (b1013), are the same in the two organisms, but in S. typhimurium they are separated by ~700 bp. The same is true for
four other salmonellae whose genome sequences have been completed. The fact that rutR is retained indicates that the rut operon has probably been deleted in salmonellae rather than inserted in E. coli, as does the fact that the operon is present in Klebsiella pneumoniae (Table 3). Whereas small differences between the genome of salmonellae and E. coli are common, it is seldom possible to infer the direction of gene loss (18).

Although the Rut pathway is required for E. coli K-12 to use pyrimidines as sole N source at room temperature, this is unlikely to be its physiological role. Under N-limiting conditions, the entire rut operon is also highly transcribed at 37°C, and, at both temperatures, pyrimidines probably serve as one of several alternative N sources, i.e., E. coli probably encounters and degrades mixtures of N sources (6). In addition to allowing catabolism of external pyrimidines, the Rut pathway may also lower the internal pyrimidine concentration to help coordinate the slowing of growth under N-limiting conditions. Pathway intermediates and gene functions remain to be determined.

**Methods**

**Materials.** Stock solutions of pyrimidine nucleosides (Sigma–Aldrich) were filter-sterilized and stored at 4°C. Their concentrations were confirmed by retention times with [U-14C]uridine (50 μCi/ml, 468 mCi/mmol) (1 Ci = 37 GBq) (Amersham Pharmacia BioSciences) was stored at −20°C. Malonic acid and methylmalonic acid (Sigma–Aldrich) were made fresh before use or stored frozen in aliquots that were thawed once. A 3-hydroxypropionic acid standard was purchased from Ernesto Brunet (Departamento de Química Orgánica, Universidad Autónoma, Madrid).

**Strains.** Strains with lesions (mini-Tn5) in b1007–b1013 (FB22265–FB22271, respectively) were obtained from the E. coli K-12 MG1655 Genome Initiative, and lesions were transferred into NCYM3722 (wild-type) (19) and NCYM3876 [mbrB(Con)] (6) by phage P1-mediated transduction (Table 4, which is published as supporting information on the PNAS web site). Sequencing of PCR products obtained with appropriate primers verified that lesions in all strains were as expected.

**Growth, Determination of Cell Yield, and Excretion of a Waste Product from [U-14C]Uridine.** Strains were grown in N–C– minimal medium (6) with glycerol or ribose (0.4%, unless specified otherwise) as C source and a pyrimidine nucleoside, base, arginine, or NH₄Cl as N source at the concentration indicated. Growth was at 37°C or room temperature (19–22°C). Cell yield was determined by OD₆₀₀ of 0.5 by addition of 150 μl of concentrated sulfuric acid and was saturated with ~1.2 g of sodium sulfate to increase ionic strength of the aqueous phase. Organic acids were extracted into 1.5 ml of ethyl acetate by vigorous shaking by hand for 5 min. After allowing separation of the phases for 5 min, a 0.6-ml portion of the organic phase was removed and dried under a stream of argon gas. N-O-bis(trimethylsilyl) trifluorooracetamide, 25 μl, was added to dissolve and derivatize dried samples. The samples were analyzed by using an Agilent Technologies (Palo Alto, CA) 6890GC 5973MSD system interfaced with a J & W DB-XLB column (30 m × 0.25 mm i.d. × 0.25 μm; Delta Technical Products, Des Plaines, IL), with an interface temperature of 280°C and a helium carrier gas flow of 1.2 ml/min. The injector was used in the splitless mode at 220°C. After the injection of sample (1 μl), the temperature of the GC oven was maintained at 50°C for 3 min and then increased to 280°C at the rate of 20°C/min. Mass spectra were acquired with a mass selective detector, scanning from 50 to 400 atomic mass units. For each calibration standard, the data system software was used to calculate the response factor, which is the ratio of the integrated abundance of the quantitation ion (m/z 147) for 3-hydroxypropionic acid to the integrated abundance of the quantitation ion (m/z 147) for malonic acid.

**Northern Blot Analysis.** Cells were grown to late exponential phase (OD₆₀₀ of ~0.7) at 22°C with glycerol as C source and NH₄Cl, arginine, or uridine (10 mM N in each case) as N source. Total RNA was extracted from cells in 5–10 ml of culture by using the RNaseasy plant mini kit (Qiagen, Valencia, CA) and 5 μg was subjected to electrophoresis on a 1% formaldehyde RNA gel. After transfer to a nylon membrane, blots were probed with a 32P-labeled fragment of 530 bp near the middle of the b1012 gene (forward primer 5′-ACTGCTGCCCACCTTAAAGTACCT-3′, position 313–336 from the first base of the ATG translational start, and reverse primer 5′-GGCGGTGGTGTAGTGTTCCATT-3′, position 820–843). Blots were scanned with a Typhoon 8600 PhosphorImager (Molecular Dynamics), and quantification of the highest molecular weight band was done with the IMAGEQUANT program (version 5). Background values were calculated from other regions of the lanes containing RNA from NH₄Cl-grown cells.

We thank Fred Blattner for strains and Barry Boehner, Jared Leadbetter, Boris Magasanik, Andrei Osterman, and John Roth for criticisms of the manuscript. This work was supported by National Institutes of Health Grant GM38361 (to S.K.).

A Product from Thymidine

B 2-Methyl-3-Hydroxypropionic Acid