

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; and [¶]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 nitrogens 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 | σ^{54} | 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 methylmalonic 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 alterna-

tive 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 σ^{54} 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.

[†]Present address: 71-254 CHS, Department of Epidemiology, University of California School of Public Health, Los Angeles, CA 90095-1772.

^{**}Present address: Biotechnology Institute, University of Minnesota, 140 Gortner Laboratory, 1479 Gortner Avenue, St. Paul, MN 55108.

[§]Present address: Département de Microbiologie et Médecine Moléculaire, Université de Genève, 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

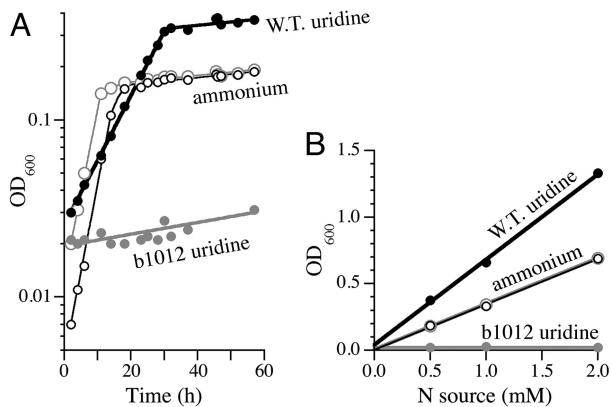


Fig. 3. Growth (A) and cell yield (B) of NCM3722 (wild-type, black circles) and NCM3996 (*b1012::Tn5*, gray circles) on NH_4Cl (open circles) or uridine (filled circles). (A) Cells were grown with NH_4Cl or uridine (1 mM) as N source and glycerol as C source at 22°C. Doubling times for NCM3722 on NH_4^+ and uridine were 3.1 and 7.6 h, respectively. The doubling time of NCM3996 on NH_4^+ was 3.2 h. (B) Final OD_{600} was determined as a function of the initial concentration of N source.

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 (*cdd*, *b2143*), which is not 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 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 growth 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

Table 2. Growth yield

Strain	Yield, mol of N per mol			
	NH_4^+	Uridine	Thymidine	Cytidine
NCM3722 (parental)	1	1.9	—	2.9
NCM3996 (lesion in <i>b1012</i>)	1	—	—	1.1
NCM3876 [<i>ntrB(Con)</i>]	1	2.1	2.0	3.1
NCM4011 [<i>ntrB(Con)</i> , lesion in <i>b1012</i>]	1	—	—	1.2

A growth yield of 1 corresponds to an OD_{600} of 0.35. —, No growth.

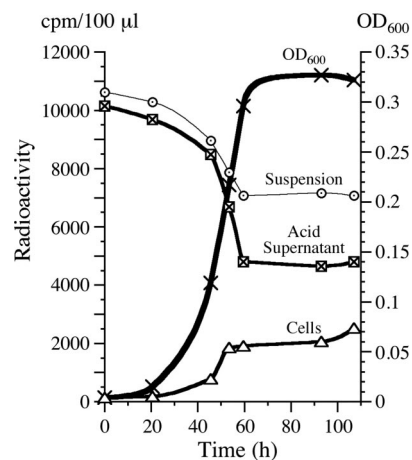


Fig. 4. Fate of $[\text{U-}^{14}\text{C}]$ uridine. Wild-type strain NCM3722 grown on glycerol and NH_4Cl (5 mM) was diluted 1/100 into medium containing glycerol and uridine (0.5 mM) containing radioactive tracer (see *Methods*). Radioactivity in cell suspensions (dotted circles), in cells (triangles), and in cell supernatants treated with acid to remove CO_2 (crossed squares) was measured as a function of OD_{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 ^{14}C -labeled waste product was excreted (see text).

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 ^{14}C from uniformly ^{14}C -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 ^{14}C -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_4Cl to the latter cultures, essentially all of the label remained in the supernatant. All experiments confirmed the view that ^{14}C -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.

Identification of Waste Products from Uridine, Cytidine, and Thymidine. Based on release of malonic acid, the three-carbon waste product produced from the oxidative pathway for uracil degradation (Fig. 1A), we reasoned that the waste product from the *b1012* pathway might be a carboxylic acid (Fig. 1B). To test this hypothesis, we extracted supernatants with ethyl acetate under conditions that favored efficient extraction of malonic acid, derivatized the extracts by trimethylsilylation, and performed gas chromatography/mass spectrometry (GC/MS) analysis (14). A compound identified as 3-hydroxypropionic acid was present in several supernatants of NCM3722 grown with uridine but not ammonium as N source (Fig. 5).

To further test whether 3-hydroxypropionic acid was the waste

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 A_{254} . [U - ^{14}C]Uridine (50 μ Ci/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 NCM3722 (wild-type) (19) and NCM3876 [*ntrB*(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 - ^{14}C]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 NH_4Cl as N source at the concentration indicated. Growth was at 37°C or room temperature (19–22°C). Cell yield was determined by OD_{600} by using initial concentrations of N sources between 0 and 2 mM. For growth on [U - ^{14}C]uridine, 9 μ l of stock solution was added to a 5-ml culture containing glycerol and 0.5 mM uridine, and the culture was incubated at 19°C with rapid shaking. Samples (700 μ l) were removed at various times to determine the radioactivity in cell suspensions (100 μ l), cell pellets, medium from which cells had been removed by centrifugation (100 μ l), and medium (500 μ l) through which air had been bubbled for 5–30 min after addition of 10 μ l of glacial acetic acid to remove $^{14}CO_2$. Growth (OD_{600}) was monitored for a parallel culture to which no radioactive tracer was added. Variations of this experiment are indicated in the text.

Identification of Pyrimidine Waste Products. For identification of pyrimidine waste products, cells were grown with glycerol or ribose as C source and 5 mM N source. After addition of sodium sulfate and sulfuric acid, cell supernatants were extracted with

ethyl acetate. Extracts were completely dried under argon and then derivatized online with *N,O*-bis(trimethylsilyl)trifluoroacetamide (Sigma–Aldrich) (14). After GC/MS components were identified by comparing their electron impact spectra with a library of reference spectra (National Institute of Standards and Technology/Environmental Protection Agency/National Institutes of Health Mass Spectral Library for WINDOWS). To determine whether the yield of waste product was proportional to the starting concentration of pyrimidine (0.25–2 mM), malonic acid was added (to 1 mM) to cell supernatants (4 ml) as an internal standard immediately before they were extracted with ethyl acetate, and the ratio of waste product to malonic acid was determined as detailed below. Finally, the absolute amount of 3-hydroxypropionic acid formed from uridine was assessed by comparison to a chemical standard that was purchased.

For quantitative studies, 3 ml of supernatant was acidified to pH ≤ 0.5 by addition of 150 μ l 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)trifluoroacetamide, 25 μ l, was added to dissolve and derivatize dried samples.

Samples were analyzed by using an Agilent Technologies (Palo Alto, CA) 6890GC/5973MSD system installed with a J & W DB-XLB column (30 m \times 0.25 mm i.d. \times 0.25 μ m; 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 μ l), 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 (OD_{600} of ≈ 0.7) at 22°C with glycerol as C source and NH_4Cl , 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 RNeasy plant mini kit (Qiagen, Valencia, CA), and 5 μ g was subjected to electrophoresis on a formaldehyde RNA gel. After transfer to a nylon membrane, blots were probed with a ^{32}P -labeled fragment of 530 bp near the middle of the b1012 gene (forward primer 5'-ACTGCTGCCACCTTAACGTTACCT-3', position 313–336 from the first base of the ATG translational start, and reverse primer 5'-GCCCCGCTTTGTAGTGTTCCATTT-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 NH_4^+ -grown cells.

We thank Fred Blattner for strains and Barry Bochner, 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.).

1. Vogels, G. D. & Van der Drift, C. (1976) *Bacteriol. Rev.* **40**, 403–468.
2. Soong, C. L., Ogawa, J., & Shimizu, S. (2001) *Biochem. Biophys. Res. Commun.* **286**, 222–226.
3. Soong, C. L., Ogawa, J., Sakuradani, E., & Shimizu, S. (2002) *J. Biol. Chem.* **277**, 7051–7058.

4. Reynolds, E. S., Lieberman, I., & Kornberg, A. (1955) *J. Bacteriol.* **69**, 250–255.
5. West, T. P. (1998) *Can. J. Microbiol.* **44**, 1106–1109.
6. Zimmer, D. P., Soupene, E., Lee, H. L., Wendisch, V. F., Khodursky, A. B., Peter, B. J., Bender, R. A., & Kustu, S. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 14674–14679.

7. Saier, M. H., Jr., Eng, B. H., Fard, S., Garg, J., Haggerty, D. A., Hutchinson, W. J., Jack, D. L., Lai, E. C., Liu, H. J., Nusinew, D. P., *et al.* (1999) *Biochim. Biophys. Acta* **1422**, 1–56.
8. Simaga, S. & Kos, E. (1978) *Z. Naturforsch., C: Biosci.* **33**, 1006–1008.
9. Simaga, S. & Kos, E. (1981) *Int. J. Biochem.* **13**, 615–619.
10. Bochner, B. R., Gadzinski, P. & Panomitros, E. (2001) *Genome Res.* **11**, 1246–1255.
11. Zhou, L., Lei, X. H., Bochner, B. R. & Wanner, B. L. (2003) *J. Bacteriol.* **185**, 4956–4972.
12. Andersen, P. S., Frees, D., Fast, R. & Mygind, B. (1995) *J. Bacteriol.* **177**, 2008–2013.
13. Aufhammer, S. W., Warkentin, E., Ermler, U., Hagemeyer, C. H., Thauer, R. K. & Shima, S. (2005) *Protein Sci.* **14**, 1840–1849.
14. Docherty, K. S. & Ziemann, P. J. (2001) *J. Chromatogr. A* **921**, 265–275.
15. Kustu, S., North, A. K. & Weiss, D. S. (1991) *Trends Biochem. Sci.* **16**, 397–402.
16. Ma, D., Alberti, M., Lynch, C., Nikaido, H. & Hearst, J. E. (1996) *Mol. Microbiol.* **19**, 101–112.
17. Lu, C. D. & Abdelal, A. T. (1999) *J. Bacteriol.* **181**, 1934–1938.
18. McClelland, M., Florea, L., Sanderson, K., Clifton, S. W., Parkhill, J., Churcher, C., Dougan, G., Wilson, R. K. & Miller, W. (2000) *Nucleic Acids Res.* **28**, 4974–4986.
19. Soupene, E., van Heeswijk, W. C., Plumbridge, J., Stewart, V., Bertenthal, D., Lee, H., Prasad, G., Paliy, O., Charernnoppakul, P. & Kustu, S. (2003) *J. Bacteriol.* **185**, 5611–5626.