SPECIFIC DELETION OCCURRING IN THE DIRECTED EVOLUTION OF 6-PHOSPHOGLUCONATE DEHYDROGENASE IN *ESCHERICHIA COLI*

RAYMOND D. MILLER, DANIEL E. DYKHUIZEN, LOUIS GREEN AND DANIEL L. HARTL

Department of Genetics, Washington University School of Medicine, St. Louis, Missouri 63110

Manuscript received April 24, 1984
Revised copy accepted July 30, 1984

ABSTRACT

A novel genetic change leading to increased activity of 6-phosphogluconate dehydrogenase (6PGD) in *E. coli* has been observed. The mutation is a deletion of approximately 0.4 kilobase pairs occurring between the structural gene of 6PGD (*gnd*) and one copy of an insertion element (IS5) found normally in *E. coli* K12 a few hundred base pairs upstream (counterclockwise) from *gnd* at 44 minutes on the conventional genetic map. The deletion is associated with a threefold higher activity of 6PGD and a 57% increase in the maximum growth rate when cells are grown in gluconate.

EXPERIMENTS in directed evolution usually recover bacterial strains that are able to express a novel metabolic function or that have an increased activity or efficiency of an existing function (reviewed in CLARK 1978; MORTLOCK 1982; HALL 1983). The mutations that are selected in such experiments are typically of three types: (1) mutations that cause constitutive expression of an otherwise inducible enzyme; (2) mutations that alter the substrate specificity of an enzyme; (3) mutations in which enzyme activity is increased as a result of duplication of the structural gene of the enzyme. In this paper we describe a novel type of mutation affecting the *gnd* gene of *Escherichia coli*, which arose during a directed evolution experiment. The *gnd* gene codes for 6-phosphogluconate dehydrogenase (6PGD), and in a chemostat in which the growth of the cells was limited by the amount of gluconate, a strain arose and was selected that had an increase in 6PGD activity resulting from the deletion of a DNA sequence located slightly upstream from the *gnd*-coding region.

Strains of *E. coli* that carry a deletion of the *edd* gene grow poorly in gluconate medium. The *edd* gene codes for phosphogluconate dehydratase, an enzyme that normally serves to route a significant fraction of phosphorylated gluconate through the Entner-Doudoroff pathway (BACHMANN 1983; INGRAHAM, MAALOE and NEIDHARDT 1983). In the absence of *edd*, all gluconate must be routed through the energetically less efficient pentose shunt by means of 6PGD. The *gnd* gene is a typical "housekeeping" gene in *E. coli*, which is expressed constitutively but at a level that varies according to growth rate (WOLF, PRATHER and SHEA 1979). Strains that are edd^- gnd^+ have a low

equilibrium cell density in chemostats in which the limiting nutrient is gluconate. However, the density of cells often increases abruptly after a few days, because mutations that increase growth rate in gluconate occur and are selected.

This paper presents the analysis of one strain isolated from a chemostat limited for gluconate. The evolved strain has improved growth in gluconate resulting primarily from a threefold increase in the activity of 6PGD. Cloning, restriction mapping and subcloning of the *gnd* region indicate that the *gnd*-coding region in the mutant is apparently unaltered. However, the evolved gene has a deletion of approximately 400 base pairs located immediately upstream from the *gnd* promoter. This deletion might result from the activity of the transposable element IS*5* (BLATTNER *et al.* 1974) that occurs upstream from *gnd* in some strains of *E. coli* (SCHONER 1983), including the parental strain that produced the *gnd* mutation.

MATERIALS AND METHODS

Chemostat methods and methods of bacteriophage P1-mediated transduction have been described prevously (DYKHUIZEN and HARTL 1980). Relevant strains are listed in Table 1. Chemostat medium consisted of Davis salts [40 mM K_2HPO_4, 15 mM KH_2PO_4, 7.6 mM $(NH_4)_2SO_4$, 1.7 mM sodium citrate and 0.8 mM $MgSO_4$] containing 0.4 g/liter of the sodium salt of gluconate.

Maximum growth rates of strains were estimated at 37° in aerated Davis minimal medium supplemented with 2.0 g/liter of the sodium salt of gluconic acid. Cell density was monitored with a Klett-Summerson colorimeter.

Enzyme assays were carried out as follows. Overnight broth cultures, 1.5 ml, supplemented as appropriate with 25 μg/ml of ampicillin were centrifuged, resuspended in 0.3 ml of sonication buffer [10 mM Tris-Cl (pH 7.5), 10 mM $MgCl_2$, 1 mM 2-mercaptoethanol], sonicated on ice and centrifuged. 6PGD activity was monitored by the increase in absorbance at 340 nm due to formation of NADPH, after adding 5–20 μl of extract to 400 μmol Tris-Cl (pH 7.5), 2 μmol trisodium 6-phosphogluconate and 1.2 μmol NADP in a volume of 3 ml at 37°. Activities were adjusted for protein concentration as measured by the method of BRADFORD (1976). The extinction coefficient of NADPH at 340 nm is 6220 M^{-1} cm^{-1}.

DNA isolation and manipulation was carried out as described by HARTL *et al.* (1983). Analysis of plasmids was carried out by digestion to completion with appropriate restriction enzymes according to manufacturer's instructions, followed by electrophoresis in agarose gels in 89 mM Tris, 89 mM boric acid, 2.5 mM EDTA (pH 8.3) and staining with ethidium bromide.

RESULTS

The mutation in *gnd* resulting in an increased activity of 6PGD was recovered in a strain that had undergone 160 hr of growth in a bacterial chemostat in which the limiting nutrient was gluconate. The original strain was DD811 (Table 1), an *E. coli* K12 into which the *gnd* allele from strain RM72B had been introduced by bacteriophage P1-mediated transduction (DYKHUIZEN and HARTL 1980; HARTL and DYKHUIZEN 1981). Strain RM72B is a natural isolate of *E. coli* obtained from a lowland gorilla, and the *gnd* allele in this strain is designated *gnd*$^+$*(RM72B)*. The original strain DD811 also carried a small deletion at 41 minutes on the genetic map, eliminating the genes *edd* and *eda*. The *eda* gene codes for 2-keto-3-deoxygluconate 6-phosphate aldolase, which is required along with the product of *edd* for utilization of gluconate

TABLE 1

E. coli *strains and plasmids used in the study*

Strain or plasmid	Relevant genotype[a]	Source and comments
RM72B	Wild type [b]	MILKMAN (1973)
DD715	Δ^c *his gnd1*	DYKHUIZEN and HARTL (1980)
DD725	Δ^c	HARTL and DYKHUIZEN (1981)
DD811	Δ^c *gnd*$^+$(RM72B)	HARTL and DYKHUIZEN (1981)
DD812	*fhuA* Δ^c *gnd*$^+$(RM72B)	T5-resistant derivative of DD811
DD854	*fhuA gnd*$^+$(RM72B)	Edd$^+$ transduction of DD812
DD862	*fhuA gnd*$^+$(862)	From DD812 grown in chemostat
DD1520	Δ^c *gnd*$^+$(862)	His$^+$ transduction of DD715
RW181[d]	Δ(*edd*)^e Δ(*gnd*)^f	NASOFF and WOLF (1980)
pLX3	pBR322 plus *gnd*-bearing *Eco*RI fragment from DD725	
pLX5	pBR322 plus *gnd*-bearing *Eco*RI fragment from DD811	
pLX7	pBR322 plus *gnd*-bearing *Eco*RI fragment from DD862	
pLX3-1	*Acc*I deletion derivative of pLX3	
pLX5-1	*Acc*I deletion derivative of pLX5	
pLX7-1	*Acc*I deletion derivative of pLX7	

[a] All DD strains also carry *rpsL*.
[b] Electrophoretic mobility of 6PGD differs from K12.
[c] Deletion (*eda-edd-zwf*)1.
[d] Host for plasmids.
[e] Deletion (*edd-zwf*)22.
[f] Deletion (*attP2H-attHK139-sbcB-his-gnd-rfb*)1.

through the Entner-Doudoroff pathway (INGRAHAM, MAALOE and NEIDHARDT 1983). In the absence of these enzymes, gluconate is metabolized through the less efficient pentose shunt. Strain DD854 (Table 1), which carries *gnd*$^+$ *(RM72B)* and is wild type with respect to *edd* and *eda*, has a maximum growth rate in gluconate of 0.0096 min^{-1}, but strain DD811 has a maximum growth rate of 0.0047 min^{-1} (Table 2).

When chemostats in which the limiting nutrient was gluconate were inoculated with DD811, the equilibrium density of cells was low, approximately 5 $\times$ 10^6 cells ml^{-1}. In one such chemostat we observed an abrupt increase in density to 7 $\times$ 10^8 cells ml^{-1} after 90 hr of growth. A strain designated DD862 was isolated from this chemostat after 160 hr, and its maximum growth rate in gluconate was found to be 0.0083 min^{-1} (Table 2). The *gnd* allele in this strain, designated *gnd862*, was transduced into a *gnd*$^-$ *edd*$^-$ *eda*$^-$ strain (DD715) to create strain DD1520, which carries *gnd862* in the genetic background of DD811. The maximum growth rate of strain DD1520 was 0.0074 min^{-1} (Table 2), indicating that 75% of the improvement in maximum growth rate of strain DD862 can be attributed to genetic changes in or near the *gnd* locus itself.

The characteristics of the 6PGD enzyme isolated from strains DD811 and DD862 were as follows. The enzymes were indistinguishable in their electrophoretic mobility as determined using cellulose acetate strips (MILKMAN 1973). No difference between the enzymes could be detected in K_m for 6-phospho-

TABLE 2

Growth rate of original and evolved strains

Strain	Allele			Genetic background	μ_{max} ± SEM (min^{-1})	Doubling time (min)
	gnd	*edd*	*eda*			
DD854	*RM72B*	+	+	K12	0.0096 ± 0.0001	72
DD811	*RM72B*	−	−	K12	0.0047 ± 0.0006	147
DD862	*gnd862*	−	−	DD862	0.0083 ± 0.0002	84
DD1520	*gnd862*	−	−	K12	0.0074 ± 0.0006	94

TABLE 3

Relative 6PGD activities

gnd location	Source of *gnd* allele[a]		
	DD725	DD811	DD862
Chromosomal	1.14 ± 0.02 (3)	1.00 ± 0.05 (3)	3.27 ± 0.19 (3)
Plasmid[b]	0.89 ± 0.09 (5)	1.00 ± 0.05 (7)	2.91 ± 0.20 (5)
Average	0.98 ± 0.07	1.00 ± 0.06	3.08 ± 0.16

[a] Absolute enzyme activity of DD811 extract was 0.152 μmol NADPH/min/mg of protein; that of pLX5 extract was 3.81 μmol NADPH/min/mg of protein. The numbers in parentheses are the number of separate determinations of enzyme activity.

[b] Plasmids pLX3, pLX5 and pLX7 carry the *gnd* gene from strains DD725, DD811 and DD862, respectively.

gluconate or the K_i for NADPH; in one experiment the K_m values were 52 and 50 μM 6-phosphogluconate, and the K_i values were 73 ± 17 and 124 ± 40 μM NADPH, respectively, which do not differ significantly. However, the total 6PGD activity per milligram of protein was threefold higher in the evolved strain DD862 than in the original strain DD811 (Table 3).

In *E. coli* K12 the entire *gnd* gene and nearby flanking sequences are contained within a single 3.9 kilobase (kb) *Eco*RI restriction fragment (NASOFF and WOLF 1980). This *gnd* fragment also contains 1.1 kb of an IS*5* element located counterclockwise from *gnd* and upstream from the *gnd* promoter (SCHONER 1983; NASOFF and WOLF 1980). When *Eco*RI-digested DNA from strains DD811 and DD862 were probed in SOUTHERN (1975) hybridizations with *gnd*-specific DNA, the *gnd*-bearing *Eco*RI restriction fragment in DD862 was observed to be approximately 0.4 kb smaller than the corresponding fragment in DD811.

To investigate this difference between DD811 and DD862, *Eco*RI restriction fragments of the appropriate size were purified from agarose gels containing DNA from strains DD811 and DD862, and these fragments were ligated into the *Eco*RI-cloning site of pBR322. The resulting plasmids were used to transform a *gnd*$^-$ *edd*$^-$ *eda*$^-$ strain (RW181 in Table 1), and cells were plated on gluconate medium in order to select for *gnd*$^+$ transformants. By means of this procedure we obtained plasmids carrying both possible orientations of the *Eco*RI restriction fragment containing either *gnd*$^+$*(RM72B)* or *gnd862*. These

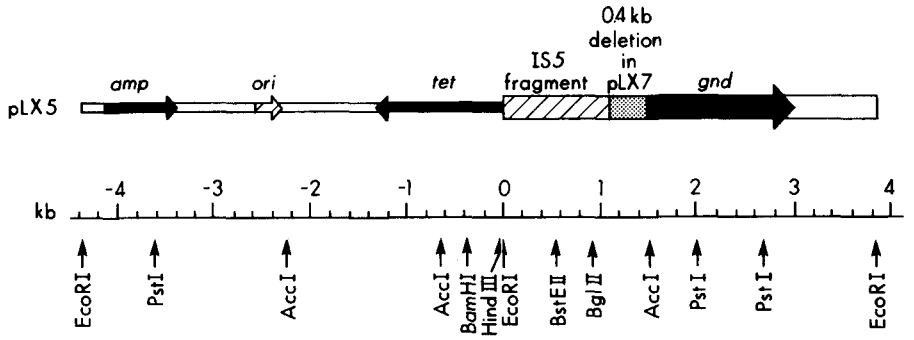

FIGURE 1.—Restriction map of pLX5 carrying *gnd*$^+$*(RM72B)*. Wide bar corresponds to *gnd*-containing *Eco*RI restriction fragment and narrow bar to pBR322. Map of pLX7, which carries the evolved allele *gnd862*, is identical except for the 0.4-kb deletion indicated. The uncloned part of IS5 includes the promoter for a long open reading frame oriented in the same direction as *gnd* (KROGER and HOBOM 1982).

plasmids were designated pLX plasmids, and the 6PGD activity of pLX-bearing strains is shown in Table 3. When *gnd862* is present at its normal chromosomal site, 6PGD activity is 3.27, relative to a value of 1.00 for DD811, which carries *gnd*$^+$*(RM72B)*. When present in pLX plasmids these alleles produce approximately the same relative levels of 6PGD activity. However, because of the high copy number of the plasmids, the absolute levels of 6PGD activity are increased approximately 30-fold. Strain DD725 in Table 3 is isogenic with DD811 except that it carries the *gnd*$^+$ allele from *E. coli* K12, and the 6PGD activity associated with the K12 allele is virtually identical with that of *gnd*$^+$*(RM72B)*.

Restriction mapping of the pLX plasmids revealed the location of the 0.4-kb deletion in the cloned fragment from DD862. Figure 1 is a restriction map of pLX5 in which the cloned *Eco*RI fragment derives from DD811. The restriction map of pLX3, which carries a comparable fragment from DD725, is identical except that it lacks the *Pst*I site at coordinate +2.00. However, pLX7, which carries the *Eco*RI fragment from DD862, has a deletion of 0.4 kb between the *Bgl*II site at coordinate +0.94 and the *Acc*I site at coordinate +1.51. The deletion is evident from the restriction fragments produced by *Bgl*II-*Acc*I double digests of the plasmids (Figure 2).

To determine whether the small deletion is associated with the increased 6PGD activity we took advantage of the observation that pBR322 contains a sequence strongly resembling a −35 promoter sequence exactly 17 base pairs to the left of the 5′ end of the *Acc*I site at coordinate −2.25 in Figure 1. Since that *Acc*I site at coordinate +1.51 represents the −10 promoter sequence of *gnd* (NASOFF, BAKER and WOLF 1984), we reasoned that we could create a new *gnd* promoter by deleting the region between the *Acc*I sites in question, and the spacing between the −10 and −35 sites would be optimal for promoter activity (HAWLEY and MCCLURE 1983). If the small deletion in the *gnd862* allele is responsible for its increased activity, then the reconstituted plasmids should have the same 6PGD activity as comparable constructs carrying *gnd*$^+$*(RM72B)* or *gnd*$^+$*(K12)*. However, if the increased 6PGD activity of

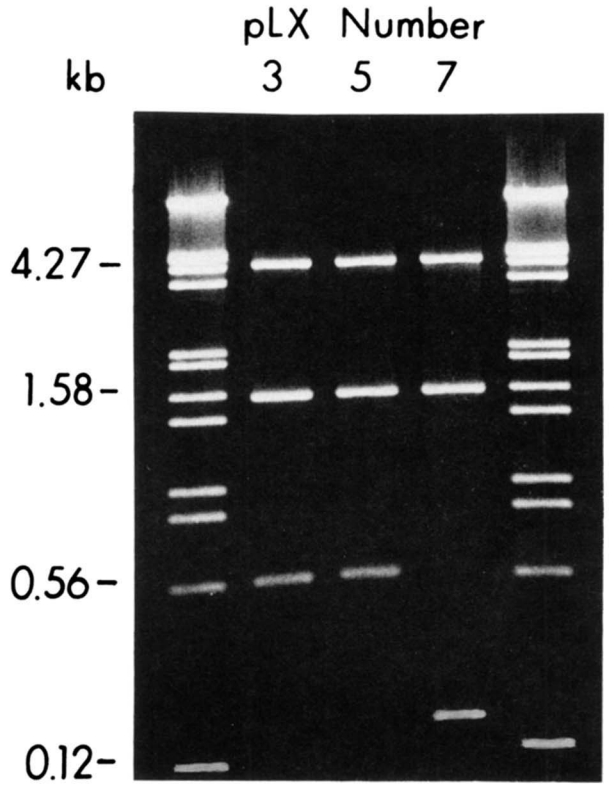

FIGURE 2.—Restriction fragments produced by *Bgl*II-*Acc*I double digestion of pLX plasmids carrying *gnd*$^+$*(K12)* (pLX3), *gnd*$^+$*(RM72B)* (pLX5) or *gnd862* (pLX7) showing 0.4-kb deletion in one fragment of pLX7. Size standards are *Eco*RI-*Hin*dIII double digests of λ DNA. Bands at 0.12 and 0.17 kb have been slightly enhanced to improve contrast. The band at 1.6 kb in all plasmids results from the superposition of two bands with the same molecular weight.

gnd862 is due to sequences downstream from the −10 promoter region, then the construct should retain its higher activity.

The pLX plasmids were digested with *Acc*I and religated, and after transformation those having a deletion between coordinates −2.25 and +1.51 were studied further. These new plasmids (pLX3-1, pLX5-1 and pLX7-1) contain the equivalent of an artificial promoter. Results regarding 6PGD activity are summarized in Table 4. Whereas plasmids pLX3-1 and pLX5-1, which carry *gnd* from DD725 and DD811, respectively, have about a 10% decrease in 6PGD activity as a result of having a new promoter, the plasmid pLX7-1, which carries the evolved *gnd* from DD862, has about a 67% decrease in 6PGD activity. The last line in Table 4 gives the 6PGD activities relative to that observed with the plasmid pLX5-1, which carries the *gnd* allele from DD811. Plasmids pLX3-1, pLX5-1 and pLX7-1 all have approximately the same 6PGD activity, and in particular the increased activity of the *gnd862* allele derived from DD862 has been eliminated in the construction of the new promoter.

TABLE 4

Relative 6PGD activities in plasmids

Plasmid	Source of *gnd* allele[a]		
	DD725	DD811	DD862
Original[b]	1.00 ± 0.08 (5)	1.00 ± 0.07 (5)	1.00 ± 0.01 (3)
Derived[c]	0.91 ± 0.03 (2)	0.91 ± 0.05 (6)	0.33 ± 0.03 (3)
Derived[c]	0.92 ± 0.03 (5)	1.00 ± 0.05 (6)	1.03 ± 0.08 (3)

[a] Absolute enzyme activity of pLX5 as in Table 3.

[b] Original plasmids pLX3, pLX5 and pLX7 carry *gnd* gene from strains DD725, DD811 and DD862, respectively.

[c] Derived plasmids pLX3-1, pLX5-1 and pLX7-1 were obtained from pLX3, pLX5 and pLX7, respectively, by *in vitro* deletion of the region between the *Acc*I sites at coordinates −2.25 and +1.51.

DISCUSSION

The *gnd* system is unique in directed evolution experiments because it is a "housekeeping" gene, the 6PGD product of which is produced constitutively in amounts that are a function of growth rate of the cells (WOLF, PRATHER and SHEA 1979). Most genes that have been studied from the standpoint of directed evolution are negatively controlled by repressors, and regulatory mutations resulting in constitutive production of the enzymes are commonly observed (MORTLOCK 1982; HALL 1983). Since *gnd* is naturally constitutive, one might expect directed evolution of *gnd* to result in structural mutations affecting the catalytic activity of the enzyme. However, it is evident that the increased 6PGD activity of the evolved *gnd862* allele is a result of the 0.4-kb deletion illustrated in Figures 1 and 2. This deletion can be classified as a regulatory mutation because it is located upstream from the transcription initiation site (NASOFF, BAKER and WOLF 1984) and yet increases the amount of enzyme activity. Thus, regulatory mutations may also be preeminent in the evolution of genes that are constitutively expressed.

This work was supported by National Institutes of Health grant GM30201. We are grateful to R. E. WOLF, JR., for providing the initial *gnd* hybridization probe and to B. G. HALL for his helpful advice.

LITERATURE CITED

BACHMANN, B. J., 1983 Linkage map of *Escherichia coli* K-12, Edition 7. Microbiol. Rev. **47:** 180–230.

BLATTNER, F. R., M. FIANDT, K. K. HASS, P. A. TWOSE AND W. SZYBALSKI, 1974 Deletions and insertions in the immunity region of coliphage lambda: revised measurement of the promoter startpoint distance. J. Virol. **62:** 458–471.

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.

CLARKE, P. H., 1978 Experiments in microbial evolution. pp. 137–218. In: *The Bacteria*, Vol. 6, Edited by T. C. GUNSALOS, L. N. ORNSTON and J. R. SOKATCH. Academic Press, New York.

DYKHUIZEN, D. E. and D. L. HARTL, 1980 Selective neutrality of 6PGD allozymes in *E. coli* and the effects of genetic background. Genetics **96:** 801–817.

HALL, B. G., 1983 Evolution of new metabolic functions in laboratory organisms. pp. 234–257. In: *Evolution of Genes and Proteins*, Edited by M. NEI and R. K. KOEHN. Sinauer Associates, Sunderland, Massachusetts.

HARTL, D. L. and D. E. DYKHUIZEN, 1981 Potential for selection among nearly neutral allozymes of 6-phosphogluconate dehydrogenase in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **78:** 6344–6348.

HARTL, D. L., D. E. DYKHUIZEN, R. D. MILLER, L. GREEN and J. DE FRAMOND, 1983 Transposable element IS*50* improves growth rate of *E. coli* cells without transposition. Cell **35:** 503–510.

HAWLEY, D. K. and W. R. MCCLURE, 1983 Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11:** 2237–2255.

INGRAHAM, J. L., O. MAALOE and F. C. NEIDHARDT, 1983 *Growth of the Bacterial Cell.* Sinauer Associates, Sunderland, Massachusetts.

KROGER, M. and G. HOBOM, 1982 Structural analysis of insertion sequence IS*5*. Nature **297:** 159–162.

MILKMAN, R., 1973 Electrophoretic variation in *E. coli* from natural sources. Science **182:** 1024–1026.

MORTLOCK, R. P., 1982 Regulatory mutations and the development of new metabolic pathways by bacteria. Evol. Biol. **14:** 205–265.

NASOFF, M. S., H. V. BAKER, II and R. E. WOLF, JR., 1984 DNA sequence of the *Escherichia coli* gene, *gnd*, for 6-phosphogluconate dehydrogenase. Gene **27:** 253–264.

NASOFF, M. S. and R. E. WOLF, JR., 1980 Molecular cloning, correlation of genetic and restriction maps, and determination of the direction of transcription of *gnd* of *Escherichia coli*. J. Bacteriol. **143:** 731–741.

SCHONER, B., 1983 *Escherichia coli* K-12 F′ plasmids carrying insertion sequences IS*1* and IS*5*. Gene **21:** 203–210.

SOUTHERN, E. M., 1975 Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:** 503–517.

WOLF, R. E., JR., D. M. PRATHER and F. M. SHEA, 1979 Growth-rate-dependent alteration of 6-phosphogluconate dehydrogenase and glucose 6-phosphate dehydrogenase levels in *Escherichia coli* K-12. J. Bacteriol. **139:** 1093–1096.

Corresponding editor: J. R. POWELL