Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays

Jonathan A. Bernstein*, Arkady B. Khodursky†‡, Pei-Hsun Lin§, Sue Lin-Chao§, and Stanley N. Cohen*¶

Departments of *Genetics and †Biochemistry, Stanford University School of Medicine, Stanford, CA 94305; and §Institute of Molecular Biology, Academia Sinica, Nankang, Taipei 115, Taiwan

Contributed by Stanley N. Cohen, May 27, 2002

Much of the information available about factors that affect mRNA decay in *Escherichia coli*, and by inference in other bacteria, has been gleaned from study of less than 25 of the ≈4,300 predicted *E. coli* messages. To investigate these factors more broadly, we examined the half-lives and steady-state abundance of known and predicted *E. coli* mRNAs at single-gene resolution by using two-color fluorescent DNA microarrays. An rRNA-based strategy for normalization of microarray data was developed to permit quantitation of mRNA decay after transcriptional arrest by rifampicin. We found that globally, mRNA half-lives were similar in nutrient-rich media and defined media in which the generation time was approximately tripled. A wide range of stabilities was observed for individual mRNAs of *E. coli*, although ≈80% of all mRNAs had half-lives between 3 and 8 min. Genes having biologically related metabolic functions were commonly observed to have similar stabilities. Whereas the half-lives of a limited number of mRNAs correlated positively with their abundance, we found that overall, increased mRNA stability is not predictive of increased abundance. Neither the density of putative sites of cleavage by RNase E, which is believed to initiate mRNA decay in *E. coli*, nor the free energy of folding of 5′ or 3′ untranslated region sequences was predictive of mRNA half-life. Our results identify previously unsuspected features of mRNA decay at a global level and also indicate that generalizations about decay derived from the study of individual gene transcripts may have limited applicability.

RNA degradation | RNA half-life | transcript abundance

The rate of mRNA degradation plays a central role in the metabolism of nucleic acids in both prokaryotic and eukaryotic cells (for reviews see refs. 1–5). mRNA decay has been studied in a range of organisms, and much has been learned about the substrate features and ribonucleolytic enzymes that influence mRNA stability. In the case of *Escherichia coli*, experimenters have demonstrated that extensive variation exists in the rate of decay of individual mRNAs and that specific features of the mRNA sequence as well as transcript secondary structure can be important determinants of such variation (1, 3). Additionally, although the half-lives of some RNAs depend on the physiological state of the cell, as influenced by genetic and environmental factors, the half-lives of other RNAs seem to be independent of cell physiology (6, 7).

Much of the information available about RNA decay has been derived from study of RNAs encoded by a very limited number of genes. In fact, we can find published reports of RNA half-life for less than 25 of the 4,288 predicted ORFs in the *E. coli* genome. Current models of mRNA decay are based on this relatively limited experimental sampling of message turnover. To gain a broader understanding of the fate of mRNAs in *E. coli* we adapted DNA microarray methodology to measure the decay of each chromosomally encoded mRNA simultaneously.

DNA microarrays containing sequences derived from thousands of individual genes enable mRNA abundance to be measured in parallel on a genome-wide scale. Chromosomal DNA sequences are displayed on a solid substrate for hybridization with labeled cDNA probes corresponding to the collective complement of mRNA transcripts. One commonly used approach for microarray analysis employs PCR products spotted on a polylysine-coated glass microscope slide (8, 9). mRNA hybridization is quantitated by comparing signals generated during concurrent hybridization of two differentially labeled mRNA pools with the target DNAs arrayed on the surface of the slide. This strategy has been widely used to compare steady-state levels of RNAs present in experimental and control samples and to evaluate events and conditions that may perturb these levels (for examples, see refs. 10–15). However, there has been only limited application of DNA microarrays to the study of kinetic events such as mRNA decay (16–18).

Here we report the results of investigations of the individual half-lives and relative abundance of the 4,288 predicted and documented mRNAs encoded by the *E. coli* genome by using two-color DNA microarrays. We had three goals in carrying out these studies: (*i*) to determine whether transcript stability is predictive of transcript abundance, (*ii*) to examine the effects of growth media on the decay of mRNAs, and (*iii*) to investigate relationships between the mRNA decay rate and certain transcript features including length, G/C content, the predicted secondary structure of untranslated regions (UTRs), and the density of potential sites of cleavage by RNase E, an endonuclease hypothesized to initiate degradation of most *E. coli* transcripts (19–21). Gene function was examined also as a possible predictor of mRNA half-life.

Materials and Methods

Strains and Media. All experiments were conducted using *E. coli* strain NCM3416 obtained from the S. Kustu Lab (University of California, Berkeley, CA). This strain is a derivative of the wild-type strain MG1655 in which a point mutation in the *rph* locus has been repaired by P1 transduction. M9 and LB media were prepared as described (22, 23). All cultures were grown at 30°C in a reciprocating water bath.

Procedures for Measurement of RNA Decay: Data Collection. An overnight culture was diluted 1:100, and cells were grown to OD_{600} = 0.8 in either LB or M9 + 0.2% glucose medium. A reference sample was drawn from culture, and then rifampicin dissolved in DMSO was added to arrest transcription (final concentration = 500 μg/ml). After rifampicin treatment, cells were harvested from culture at serial time points (2, 4, 6, and 8 min) for analysis of RNA decay. To preserve cellular RNA intact, culture samples were mixed with a one-tenth volume of a ''stop solution'' composed of 10% buffer-saturated phenol in ethanol and then chilled rapidly. RNA was prepared from

Abbreviation: UTR, untranslated region.

‡Present address: Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota, 1479 Gortner Avenue, St. Paul, MN 55108.

¶To whom reprint requests should be addressed. E-mail: sncohen@stanford.edu.

BIOCHEMISTRY

culture samples by using RNeasy kits (Qiagen, Chatsworth, CA). RNA samples were treated on the Qiagen columns with DNase I according to manufacturer instructions (Qiagen). In preparation for microarray hybridization, RNA samples (15 μg per reaction) were labeled fluorescently by using reverse transcriptase (SuperScript II, Life Technologies, Rockville, MD) and 1 μg of random hexamer primers (Amersham Pharmacia). The reference sample was labeled by using Cy3 dye (Amersham Pharmacia), and subsequent time-point samples were labeled by using Cy5 dye as described in ref. 13. Reaction products were separated from reactants and enzyme by using a Qiagen Qiaquick spin column. Samples then were concentrated with a Microcon-30 (Millipore) concentrator. Purified Cy3- and Cy5-labeled cDNA was combined with standard saline citrate (1× SSC = 0.15 M sodium chloride/0.015 M sodium citrate, pH 7.0, 2.5× SSC final), SDS (0.25% final), and 20 μg of yeast tRNA (Roche Molecular Biochemicals) in a final volume of 20 μl.

Before hybridization, samples were denatured for 2 min at 100°C. Hybridization was performed at 65°C for 5 h. After hybridization, two washes were performed at room temperature: 1× SSC + 0.03% SDS and 0.05% SSC. Slides then were dried by brief centrifugation. The arrays were scanned by using an Axonscanner (Axon Instruments, Foster City, CA) under the control of GENEPIX 3.0 software (Axon Instruments) at a resolution of 10 μm per pixel.

Arrays based on previously described *E. coli* arrays (13) were augmented to contain stable RNAs as well as various other controls: yeast RNAs, *E. coli* genomic DNA, and a duplication of 10% of represented ORFs. Array manufacture and preparation were carried out as described in the electronic publication The MGuide at http://cmgm.stanford.edu/pbrown/mguide/index.html. A tab-delimited text file containing the sequences of primers used to amplify stable RNAs can be found in Table 4, which is published as supporting information on the PNAS web site, www.pnas.org. These primers were designed by using a local implementation of the PRIMER 3 software package (24). The incidence of nonspecific hybridization for these arrays was determined experimentally to be <0.02%.

Procedures for Microarray Data Analysis. Sixteen-bit TIFF images produced by the Axon scanner were analyzed by using the GENEPIX 3.0 software package. A normalization coefficient was computed by averaging in log space the fluorescence hybridization ratios for 64 rDNA spots on each array. The mean intensity of the rDNA spots used for normalization was 25,100 units with a standard deviation of 13,400 units. Data from spots that were damaged physically or gave signal within 150 units of background in the reference channel were excluded from further analysis. The half-life of an mRNA was determined from a least-squares linear fit to a semilog plot (log base 2) of mRNA abundance versus time. Fits with regression coefficients of less than 0.7 were excluded from further analysis. All half-life and normalization calculations were carried out by using Microsoft EXCEL.

Determination of RNA Abundance Including Preparation, Labeling, and Hybridization of Genomic DNA. Genomic DNA was prepared from stationary-phase cultures grown at 30°C by using a blood and cell culture DNA kit (Qiagen) and then reduced into 0.5–1.0-kbp fragments by *Sau*3A (New England Biolabs) partial digestion. After phenol/chloroform extraction and ethanol precipitation, 2 μg of digested DNA was combined with 5 μg of random hexamers (Amersham Pharmacia) in a volume of 9.5 μl and denatured at 95°C for 10 min. Denatured template and primers then were combined in a total volume of 15 μl with 10 units Klenow fragment of DNA polymerase (Life Technologies)/1.5 μl of 10× Klenow buffer/1.5 μl of a 10× dNTP (Life Technologies) mix consisting of 0.5 mM dATP, 0.5 mM dCTP, 0.5 mM dGTP, and 0.2 mM dTTP/1.5 μl of 0.2 mM Cy3 dUTP (Amersham Pharmacia). The reaction then was incubated for 2 h at 37°C. Cy5 labeling of $t = 0$ RNA, purification of labeling reactions, and hybridization were carried out as described above under microarray half-life determination. Normalization of DNA-versus-RNA hybridization data assumed equal log average signal intensity in each fluorescence channel and was carried out by using built-in functions of the Stanford Microarray Database (http://www-genome.stanford.edu/microarray). Calculated ratios, thus, were equivalent to relative transcript abundances.

Northern Blot Half-Life Determination. Cells were cultured in M9 + 0.2% glucose medium, treated with rifampicin, and collected at time points as described above. Total cellular RNA was extracted as described in refs. 25–27. After precipitation at −20°C, RNA was centrifuged at 4°C for 10 min at 3,000 × g and then washed with 70% ethanol and vacuum-dried. RNA sample preparation for 1% agarose/formaldehyde gel electrophoresis, Northern blotting, and hybridization were performed as described (23). DNA templates for probes corresponding to the full-length *fbaA*, *rpsA*, and *guaC* ORFs and to an internal *ompC* fragment (nucleotide position +7 to 1,006) were synthesized by PCR amplification. The products were labeled by using a random primer DNA-labeling kit (Roche Molecular Biochemicals) with [α-^{32}P]dCTP. Radioisotope signals were quantified with either a Molecular Dynamics PhosphorImager or an FLA-5000 imager (Fuji). Abundance of 16S rRNA in each sample was used as a loading control for Northern blot hybridizations. 16S rRNA abundance was determined by scanning densitometry of ethidium bromide-stained gels using the Kodak Digital Science electrophoresis documentation and analysis system 120 package. RNA half-lives were calculated from normalized hybridization-intensity data by using linear fits to a semilog plot as described above.

Results

Global Analysis of mRNA Half-Life in Cells Growing at Different Rates. There is experimental evidence that mRNA half-life can be affected by the composition of culture medium or other factors that alter the cell-growth rate (6, 7, 28). For example, turnover of *ompA* mRNA as measured by Northern blotting varies in response to medium-induced changes in the rate of cell growth; the *ompA* half-life in L broth was reduced by two thirds in bacteria growth in Mops/acetate medium, where the cell-doubling time (200 min) was approximately five times longer than in L broth (7). The stability of other messages including *lpp* and *bla*, however, showed no response to medium or growth rate, indicating that the effects of growth rate on mRNA decay are not universal.

We followed the decay of transcripts encoding *E. coli* ORFs growing at 30°C in M9 + glucose or LB medium at 2, 4, 6, and 8 min after pharmacological arrest of transcription initiation by rifampicin. All time courses were begun immediately after cultures reached $OD_{600} = 0.8$. In preliminary experiments, we found that the extent of inhibition of transcription of different *E. coli* genes varied widely and that it was necessary to use a rifampicin concentration of 500 μg/ml to achieve genome-wide transcriptional inhibition. The abundance of Cy5-labeled transcripts at each time point was determined relative to the amount of Cy3-labeled transcript harvested at $t = 0$ before transcriptional arrest. Normalization of Cy3 and Cy5 hybridization signals from the arrays was accomplished by exploiting the stability of ribosomal RNAs, the abundance of which should remain constant during the period that mRNA decay was followed; the raw fluorescence-intensity data were normalized to achieve a one-to-one ratio of signal from the Cy3-labeled rRNA taken from $t = 0$ to Cy5-labeled rRNA from each successive time point.

The half-life of *E. coli* mRNAs was examined in cells growing in M9 + glucose medium (cell-doubling time, 90 min) or LB medium (cell-doubling time, 30 min), at 30°C (Fig. 1 and Table 1). We required that the fluorescence signal in the reference

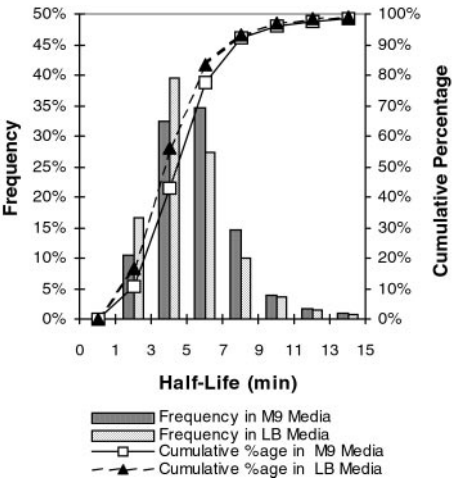

Fig. 1. Histogram showing the relative frequency expressed as a fraction of the total number of half-lives measured and the cumulative percentages of various mRNA half-lives in cultures grown in M9 + 0.2% glucose and LB media at 30°C and harvested at OD_{600} = 0.8. mRNA half-lives were determined as described in *Materials and Methods*. Source data for this histogram are available in Table 5, which is published as supporting information on the PNAS web site.

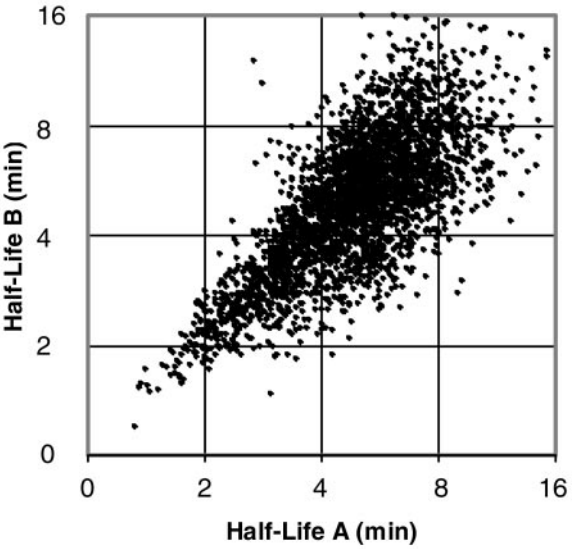

Fig. 2. Log-scale scatterplot comparing two measurements of mRNA half-lives in cultures grown in M9 + 0.2% glucose medium at 30°C. mRNA half-lives were determined as described in *Materials and Methods*. The Pearson correlation coefficient between the log-scaled repeated measurements was 0.7 for half-lives less than 15 min.

channel be above the background at all time points in the decay series for a half-life determination to be included in our analysis. A goodness-of-fit criterion required that all half-lives we reported were determined from linear fits to time-series data with a regression coefficient greater than 0.7. Applying these filters, half-lives were obtained for 3,835 and 2,267 *E. coli* transcripts synthesized by cells grown in M9 + glucose and LB media, respectively (Table 5). These half-lives were used in all subsequent analyses.

We observed a wide range of half-lives (Fig. 1) in both types of media. Although the most rapidly decaying mRNAs had half-lives as short as 1–2 min, others had half-lives that were 10 min or longer. Approximately 80% of half-lives ranged between 3 and 8 min in M9 + glucose medium, and 99% of the half-lives measured were between 1 and 18 min. In LB, 99% of half-lives in LB were between 1 and 15 min (Fig. 1). The distribution of half-lives was spread around a single mean for each type of culture condition; mean half-life was 5.7 min in M9 + glucose and 5.2 min in LB (Table 1). We saw no global shift toward greater or lesser stability in mRNA half-lives measured in the two media. Although most messages that had a relatively long half-life in LB medium also showed a long half-life in M9 + glucose, we also observed that several messages showed differential stability in the two media. These messages include the *gltJKL* operon transcripts, the half-lives of which were longer in M9 medium, and transcripts from the *entCEBA* and *hisG-I* operons, the half-lives of which were longer in LB medium.

Reliability of Microarray Measurements. Whereas the ordinarily small number of determinations made during analysis of RNA half-life by Northern blotting imposes practical limitations on statistical evaluation, the ability to carry out many half-life determinations concurrently with DNA microarrays provided an opportunity to quantitatively assess the reliability and accuracy

Table 1. mRNA half-lives and culture generation times

Media	Generation time	Mean half-life	Median half-life
LB	≈30′	5.2 ± 0.3	4.7
M9 + Glucose	≈90′	5.7 ± 0.3	5.4

of $t/2$ measurements. The Pearson correlation coefficient determined for the log-transformed half-lives measured in two experiments in M9 + glucose medium was 0.7 for half-lives of less than 15 min (Fig. 2). A second means of expressing the extent of experimental variability was to quantify the average difference between repeated measurements as a fraction of the mean between experiments (half-life$_a$ − half-life$_b$)/[(half-life$_a$ + half-life$_b$)/2]. The average across all genes measured was 0.23 (23%).

Two approaches were taken to investigate the extent of agreement between half-lives determined by using microarrays to those obtained by other analyses. First, a compilation of published half-life data was assembled and then compared on a gene-by-gene basis to the microarray results (Table 2). Second, the half-life of four selected messages found by our microarray analysis to have half-lives between 4.2 and 9.7 min was determined also by us using Northern blotting. Each Northern blot half-life determination was performed in duplicate; the average resulting half-lives for the *rpsA*, *fba*, *guaC*, and *ompC* transcripts in M9 medium are indicated in Table 2. The variation between the duplicate experiments averaged 0.26 (26%). In all four cases, the blot- determined half-lives were equivalent to those obtained by microarray analysis within the average observed experimental variation. As has been observed in prior studies of mRNA decay in wild-type cells, no detectable accumulation of decay intermediates was seen on the Northern blots.

Relationship of Transcript Abundance and Decay. Mathematically, RNA abundance is a function of both transcription and decay; abundant transcripts have a high transcription rate relative to their rate of decay. We examined our data set for evidence of a direct relationship between the relative abundance of an mRNA and its half-life. To determine the abundance of specific transcripts, fluorescent cDNA made from total RNA was hybridized against a reference of labeled stationary-phase (nonreplicating) genomic DNA in a manner similar to that described in ref. 29. Relative transcript abundances were determined in duplicate from log-phase cultures grown in LB and M9 + glucose media (Table 6, which is published as supporting information on the PNAS web site). Abundances determined by our analysis cor-

Table 2. mRNA half-lives as determined by DNA microarray and Northern blot analyses

Gene	Half-life determinations		
	Microarray M9 media	Northern blot (current study) M9 media	Northern blot (published) various media
fba	7.2 ± 1.9	6.9 ± 1.7	—
guaC	8.4 ± 2.2	7.4 ± 1.0	—
lpp	14.1	—	9.6[a], 13.3[b], 21[c]
ompA	13.1	—	14[a,c], 10.6[b], 8.1[d]
ompC	9.7 ± 1.2	8.4 ± 0.8	—
pnp	2.4 ± 0.1	—	1.0[b]
rnb	2.2 ± 0.2	—	1.2[e]
rplY	6.5 ± 1.0	—	3.5[b]
rpsA	4.2 ± 0.2	4.8 ± 0.2	—
rpsO	4.9 ± 0.3	—	1.0, 1.5[b]
rpsT	2.4 ± 0.1	—	1.6[b]
trxA	2.8 ± 0.1	—	1.8[a], 1.9[b], 3.5[f]

Half-life measurements are shown as mean ± standard error or mean alone where the standard error was not estimable. Superscript letters indicate the source of previously published half-life data presented and the conditions used in the study cited as follows: a, L broth at 30°C (53); b, L broth at 37°C (52); c, LB medium at 30°C (58); d, LB medium at unspecified temperature (54); e, LB medium at 30°C (59); f, K medium at 30°C followed by shift to 44°C for 80 sec (60).

responded closely to relative transcript abundances reported previously (29).

In comparing transcript stability and abundance data obtained for cells growing in M9 medium we found no evidence for a positive correlation between half-life and relative transcript abundance. Rather, linear regression analysis showed a statistically significant inverse relationship between relative transcript abundance in minimal medium and transcript stability (Fig. 3, $P < 0.001$ for linear regression). Thus, overall, *E. coli* mRNAs that showed greater cellular abundance were less stable. However, in contrast to this overall inverse relationship, certain transcripts were both highly abundant and highly stable, and expression of these genes may be controlled significantly at the level of mRNA decay. Such genes include *cfa*, *oppA*, *osmE*, *rplV*, *rpmD*, *rpmI*, *rpsL*, *sodA*, and *tpx*.

Transcript Features and mRNA Stability. The data presented above indicate that the half-lives of *E. coli* transcripts extend over a wide range. Earlier investigations have shown that discrete determinants of mRNA decay exist and can be localized to particular mRNA sites rather than being dispersed widely and nonspecifically throughout the transcript (30). Consistent with such evidence for the site specificity of RNA decay determinants, our global analysis of RNA decay showed no correlation between ORF or operon length and half-life (Fig. 3).

The current model of mRNA decay in *E. coli* holds that the degradation of transcripts is initiated by endonucleolytic cleavage, and the resulting fragments are subsequently degraded by either 3′-to-5′ exonucleolytic attack or secondary endonucleolytic cleavages (for reviews see refs. 1 and 3). Multiple lines of evidence point to RNase E as carrying out the rate-limiting endonucleolytic step in mRNA decay. Although a particular nucleotide consensus sequence that defines an RNase E target sequence has not been identified (27, 31), RNase E cleavages occur site-specifically in single-stranded A + U-rich regions. We used this knowledge about RNase E specificity to search the *E. coli* MG1655 genomic DNA sequence (32) on file at GenBank for potential RNase E cleavage sites with the aim of correlating cleavage-site density with transcript stability.

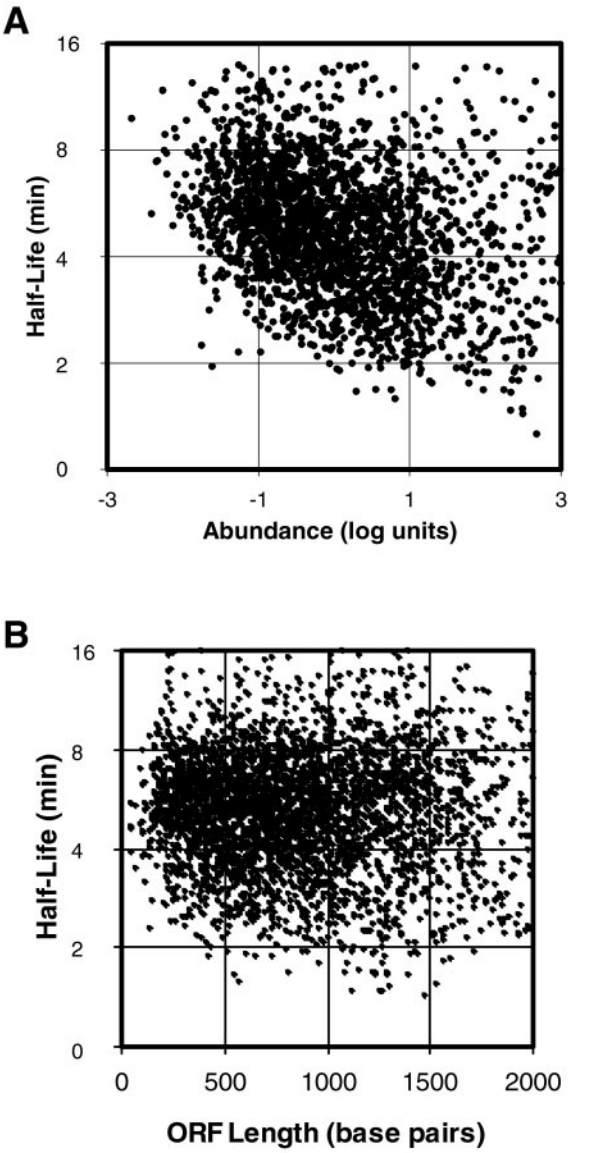

Fig. 3. (*A*) Scatterplot comparing $\log_2$ (transcript relative abundance) and log-scale mRNA half-life in M9 + 0.2% glucose medium at 30°C. Linear regression analysis of the log-scaled data indicates a statistically significant inverse relationship between relative transcript abundance and transcript stability ($P < 0.001$; correlation coefficient, $R = -0.31$). (*B*) Scatterplot comparing ORF length and log-scale mRNA half-life under the same growth conditions as described for *A*.

We designed a computer-based algorithm to analyze protein coding and UTR sequences for potential RNase E cleavage sites. This algorithm scans transcripts and identifies regions containing 4 of 7, 5 of 7, or 5 of 9 A and U ribonucleotides with an AU dinucleotide pair within. The number of potential cleavage sites per transcript was divided by transcript length to provide a measure of their density, which we found to vary over a wide range in both ORF and UTR (0.07–1.1 sites per 10 base pairs using the 5 of 7 A + U criteria) regions of *E. coli* transcripts. However, this metric was not predictive of messenger half-life. Additionally, we carried out further sequence analyses by using a modified algorithm that required the presence of a G ribonucleotide one or two bases upstream of potential cleavage sites, because this feature has been shown to increase susceptibility to RNase E (33). A final modification to the algorithm allowed us

to exclude predicted double-stranded regions in UTR sequences from consideration (see below for details of secondary structure prediction; ref. 34). Inclusion of the upstream G and/or the single-stranded requirement(s) reduced the number of predicted RNase E sites per message; however, the density of these sites still was not predictive of half-life.

Some of the RNA features that affect stability are located in UTRs; for example, the UTRs of *E. coli* ompA (35, 36) and the human transferrin receptor protein (37) have been shown to play a role in determining RNA stability. We therefore examined various 5′- and 3′-UTR features for a possible relationship with measured half-lives; specifically, we looked at UTR length, G/C content, the degree of predicted single strandedness, and the thermodynamic stability of predicted secondary structure as predictors mRNA half-life.

Sequences of putative 5′ and 3′ UTRs were obtained from the GenBank repository. Upstream and downstream UTR boundaries for 5′ UTRs were read from the genomic annotation accompanying the GenBank genomic sequence (32, 38); the downstream end of 3′ UTRs was predicted by using the TRANSTERM program by Ermolaeva *et al.* (39) to locate probable *rho*-independent transcription-terminating sequences. Overall, 2,193 5′ UTRs and 688 3′ UTRs were predicted. The average lengths of 5′ and 3′ UTRs were 84 and 79 nucleotides, respectively. The web implementation of the MFOLD software package by Zuker *et al.* (40, 41) was used to predict the secondary structure of the UTRs and the corresponding ΔG values for folding. When multiple structures having similarly high stabilities were predicted to form, an average ΔG value for these structures was obtained. This analysis indicated that UTR length, G/C content, degree of predicted single strandedness, or predicted thermodynamic stability at the level of secondary structure were not predictive of mRNA half-life.

Earlier studies have suggested that ribosomes can shield *E. coli* mRNA from degradation during translation (42–45). These observations raise the possibility that occurrence in transcripts of codons with corresponding tRNA that is limiting may stimulate mRNA decay by hindering ribosome progression (46–49). However, we found no association between the frequency per unit length of a particular codon within a message and its measured half-life.

Gene-Product Function and mRNA Stability. Little is known about possible relationships between gene function and mRNA half-life or abundance. Recent evidence indicates that certain functional groups of genes may share the ability to be preferentially sensitive or resistant to *E. coli* endoribonuclease E (50). Moreover, it might be predicted that transcripts of housekeeping genes would have longer half-lives than transcripts synthesized in response to acute stimuli or needed only at a particular point in the cell-division cycle. Our global analysis provided the opportunity to address this question quantitatively, and thus we examined the relationships between the measured half-lives and published predictions of gene function.

Multiple groups have annotated the putative and documented ORFs of *E. coli*. Our analysis used the gene-family memberships predicted by Riley (51) to investigate relationships between gene function and mRNA half-life. A one-way ANOVA test conducted to identify category differences indicated that significant differences existed in the mean half-life of various functional groups ($P < 0.001$). Post hoc analyses of the between-group differences are summarized in Table 3. Notably, transcripts of genes assigned by Riley (51) to the categories of amino acid synthesis or macromolecule synthesis/modification (including, for instance, the *thrLABC* and *trpLEDCBA* operons) have shorter half-lives than average, whereas those in the cell-envelope maintenance category or recycling of small molecules

Table 3. Mean mRNA half-life by gene functional class

Gene class	No. of genes in category	Mean half-life
Amino acid synthesis	96	3.8*
Macromolecule synthesis	206	4.4*
Nucleotide biosynthesis	30	4.6
Synthesis of cofactors/carriers	106	4.8
Cell division	37	5.0
Macromolecule degradation	60	5.0
Global regulatory functions	49	5.1
Fatty acid synthesis	20	5.1
Adaptation	25	5.1
Ribosome constituents	50	5.4
Central intermediary metabolism	170	5.6
Protection responses	44	6.0
Transport/binding proteins	393	6.2**
Energy metabolism, carbon	173	6.3**
Degradation of small molecules	150	6.3**
Cell envelope	109	6.4**

A mean half-life in M9 + glucose medium was calculated for each functional class of genes. As described in *Results*, a one-way ANOVA indicated a significant between-groups difference: $P < 0.001$. Post-hoc analysis identified a subset of functional classes with significantly shorter mean half-lives indicated by *, and a subset with significantly longer mean half-lives indicated by **, $P < 0.05$.

category (including, for instance, the *flgB–K* and *mhpA–F* operons) have longer-than-average half-lives.

Discussion

Prior studies of RNA half-life in *E. coli* have been confined to less than 25 of the 4,288 predicted genes in this bacterial organism. The experiments presented here have increased the number of genes for which the half-life has been measured in *E. coli* by more than 100-fold. This advance was accomplished by the adaptation of DNA microarrays to the study of mRNA decay.

Our microarray-determined mRNA half-lives agree well with those determined by Northern blot-based methods in previous work and the present study, which is consistent with the premise that most of the RNA hybridizing to the microarray spots represents full-length transcripts and not decay intermediates. Accordingly, our own Northern blot results and those of others who have studied mRNA decay in wild-type bacteria do not show significant accumulation of decay intermediates after the addition of rifampicin (52–54).

Previous theories about the relationship between mRNA half-life and transcript function and/or abundance have focused on the cellular "need" to maintain a proper steady-state level of transcripts (i.e., stability is associated with transcripts that produce large amounts of protein, whereas instability is associated with less abundant transcripts; refs. 55–57). It is axiomatic that mRNA stability contributes directly to mRNA abundance; however, our data indicate that stability is not a general predictor of abundance. In fact, although certain abundant transcripts were found to be stable in our experiments, in analyzing these new measurements we uncovered a surprising overall inverse relationship between mRNA abundance and half-life. This finding suggests that globally, transcription is the dominant factor in determining the mRNA steady-state level in *E. coli* and that variations in half-life may have an alternative biological role, perhaps to facilitate transient changes in mRNA abundance in response to specific environmental perturbations or the progress of bacteria through the cell-division cycle. Our results thus imply that relative instability of an mRNA may facilitate its removal from the translation apparatus in response to specific cellular signals. Because the cell populations in our experiments were

neither synchronous nor responding to environmental stresses, the effects of such removal would not be reflected in the average steady-state transcript levels we measured. Supporting the notion that modulation of mRNA degradation in *E. coli* may principally be a mechanism for controlling the ability of cells to respond dynamically to acute events, we found that functionally related groups of genes, as defined by Riley (51), have the likelihood of having similar half-lives.

On the basis of studies of selected individual mRNAs, a variety of factors previously have been identified as predictors or nonpredictors of mRNA stability, which include transcript length, codon composition, and G/C content. Analysis of each of these parameters as well as of the density of potential RNase E cleavage sites showed no global correlation with measured half-lives. Similarly, estimation of the free energy of folding and the degree of single strandedness in a catalog of 5′- and 3′-UTR sequences that we assembled showed no correlation with transcript half-lives by using the analytical methods we applied.

Two hypotheses about the mechanisms of mRNA decay in *E. coli* are suggested by our observations. First, as noted above, there are evident trends in our data set relating mRNA stability to both transcript abundance and gene-product function. It appears that cells quickly degrade certain abundant transcripts. Thus, the cell consumes additional energy to turn over abundant mRNAs rapidly at rates that would require increased transcription to maintain the observed steady-state message levels. Second, although we have not been able to identify transcript features predictive of mRNA half-life, the wide range of stabilities observed for *E. coli* mRNAs implies that such elements exist.

The methods presented here establish a framework for further investigation of mRNA decay on a genome-wide basis. Beyond evaluation of the effects of growth medium, it should be possible to use microarrays to examine the effects of a variety of environmental stimuli on RNA decay including heat shock, cold shock, and nutritional factors. These methods also are readily adaptable to study the effects of various mutations on mRNA decay, thereby enabling further elucidation of enzyme-target specificity.

We thank Drs. P. Fawcett and C. Scafe for assistance with data preparation and J. G. Belasco and P. J. Green for comments on the manuscript. We acknowledge E. Soupene of the S. Kustu Lab for providing *E. coli* strain NCM3416 and the cooperation of the Bay Area bacterial genomics group in preparation of materials for microarray analysis. A.B.K. also thanks Dr. P. Brown for support. These studies were supported by National Institutes of Health Grant GM 54158 (to S.N.C.) and National Science Council of Taiwan Research Grant for Frontier Science NSC 87-90-2321-B-001 (to S.L.C.). J.A.B. was supported by National Institutes of Health Training Grant GM 07790, and A.B.K. was supported by National Science Foundation Postdoctoral Fellowship DMS-9406348.

1. Grunberg-Manago, M. (1999) *Annu. Rev. Genet.* **33,** 193–227.
2. Steege, D. A. (2000) *RNA* **6,** 1079–1090.
3. Coburn, G. A. & Mackie, G. A. (1999) *Prog. Nucleic Acid Res. Mol. Biol.* **62,** 55–108.
4. Ross, J. (1996) *Trends Genet.* **12,** 171–175.
5. Beelman, C. A. & Parker, R. (1995) *Cell* **81,** 179–183.
6. Meyer, B. J. & Schottel, J. L. (1991) *J. Bacteriol.* **173,** 3523–3530.
7. Nilsson, G., Belasco, J. G., Cohen, S. N. & von Gabain, A. (1984) *Nature (London)* **312,** 75–77.
8. Schena, M., Shalon, D., Davis, R. W. & Brown, P. O. (1995) *Science* **270,** 467–470.
9. Shalon, D., Smith, S. J. & Brown, P. O. (1996) *Genome Res.* **6,** 639–645.
10. DeRisi, J., Penland, L., Brown, P. O., Bittner, M. L., Meltzer, P. S., Ray, M., Chen, Y., Su, Y. A. & Trent, J. M. (1996) *Nat. Genet.* **14,** 457–460.
11. Iyer, V. R., Eisen, M. B., Ross, D. T., Schuler, G., Moore, T., Lee, J. C., Trent, J. M., Staudt, L. M., Hudson, J., Jr., Boguski, M. S., *et al.* (1999) *Science* **283,** 83–87.
12. Spellman, P. T., Sherlock, G., Zhang, M. Q., Iyer, V. R., Anders, K., Eisen, M. B., Brown, P. O., Botstein, D. & Futcher, B. (1998) *Mol. Biol. Cell* **9,** 3273–3297.
13. Khodursky, A. B., Peter, B. J., Cozzarelli, N. R., Botstein, D., Brown, P. O. & Yanofsky, C. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 12170–12175.
14. Huang, J., Lih, C. J., Pan, K. H. & Cohen, S. N. (2001) *Genes Dev.* **15,** 3183–3192.
15. Perez-Amador, M. A., Lidder, P., Johnson, M. A., Landgraf, J., Wisman, E. & Green, P. J. (2001) *Plant Cell* **13,** 2703–2717.
16. Holstege, F. C., Jennings, E. G., Wyrick, J. J., Lee, T. I., Hengartner, C. J., Green, M. R., Golub, T. R., Lander, E. S. & Young, R. A. (1998) *Cell* **95,** 717–728.
17. Lam, L. T., Pickeral, O. K., Peng, A. C., Rosenwald, A., Hurt, E. M., Giltnane, J. M., Averett, L. M., Zhao, H., Davis, R. E., Sathyamoorthy, M., *et al.* (2001) *Genome Biol.* **2,** 1–11.
18. Wang, Y., Liu, C. L., Storey, J. D., Tibshirani, R. J., Herschlag, D. & Brown, P. O. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 5860–5865.
19. Carpousis, A. J., Vanzo, N. F. & Raynal, L. C. (1999) *Trends Genet.* **15,** 24–28.
20. Cohen, S. N. & McDowall, K. J. (1997) *Mol. Microbiol.* **23,** 1099–1106.
21. Rauhut, R. & Klug, G. (1999) *FEMS Microbiol. Rev.* **23,** 353–370.
22. Neidhardt, F. C., Bloch, P. L. & Smith, D. F. (1974) *J. Bacteriol.* **119,** 736–747.
23. Sambrook, J. & Russell, D. W. (2001) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
24. Rozen, S. & Skaletsky, H. (2000) *Methods Mol. Biol.* **132,** 365–386.
25. Lin-Chao, S. & Bremer, H. (1986) *Mol. Gen. Genet.* **203,** 143–149.
26. Lin-Chao, S. & Cohen, S. N. (1991) *Cell* **65,** 1233–1242.
27. Lin-Chao, S., Wong, T. T., McDowall, K. J. & Cohen, S. N. (1994) *J. Biol. Chem.* **269,** 10797–10803.
28. Georgellis, D., Arvidson, S. & von Gabain, A. (1992) *J. Bacteriol.* **174,** 5382–5390.
29. Wei, Y., Lee, J. M., Richmond, C., Blattner, F. R., Rafalski, J. A. & LaRossa, R. A. (2001) *J. Bacteriol.* **183,** 545–556.
30. Belasco, J. G., Nilsson, G., von Gabain, A. & Cohen, S. N. (1986) *Cell* **46,** 245–251.
31. McDowall, K. J., Lin-Chao, S. & Cohen, S. N. (1994) *J. Biol. Chem.* **269,** 10790–10796.
32. Blattner, F. R., Plunkett, G., 3rd, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., *et al.* (1997) *Science* **277,** 1453–1474.
33. Kaberdin, V. R., Walsh, A. P., Jakobsen, T., McDowall, K. J. & von Gabain, A. (2000) *J. Mol. Biol.* **301,** 257–264.
34. McDowall, K. J., Kaberdin, V. R., Wu, S. W., Cohen, S. N. & Lin-Chao, S. (1995) *Nature (London)* **374,** 287–290.
35. Arnold, T. E., Yu, J. & Belasco, J. G. (1998) *RNA* **4,** 319–330.
36. Diwa, A., Bricker, A. L., Jain, C. & Belasco, J. G. (2000) *Genes Dev.* **14,** 1249–1260.
37. Mullner, E. W. & Kuhn, L. C. (1988) *Cell* **53,** 815–825.
38. Lukashin, A. V. & Borodovsky, M. (1998) *Nucleic Acids Res.* **26,** 1107–1115.
39. Ermolaeva, M. D., Khalak, H. G., White, O., Smith, H. O. & Salzberg, S. L. (2000) *J. Mol. Biol.* **301,** 27–33.
40. Zuker, M., Mathews, D. H. & Turner, D. H. (1999) in *RNA Biochemistry and Biotechnology*, eds. Barciszewski, J. & Clark, B. F. C. (Kluwer, Dordrecht, The Netherlands) Vol. 70, pp. 11–43.
41. Mathews, D. H., Sabina, J., Zuker, M. & Turner, D. H. (1999) *J. Mol. Biol.* **288,** 911–940.
42. Bjornsson, A. & Isaksson, L. A. (1996) *Nucleic Acids Res.* **24,** 1753–1757.
43. Braun, F., Le Derout, J. & Regnier, P. (1998) *EMBO J.* **17,** 4790–4797.
44. Har-El, R., Silberstein, A., Kuhn, J. & Tal, M. (1979) *Mol. Gen. Genet.* **173,** 135–144.
45. Joyce, S. A. & Dreyfus, M. (1998) *J. Mol. Biol.* **282,** 241–254.
46. Hayes, C. S., Bose, B. & Sauer, R. T. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 3440–3445.
47. Zahn, K. (1996) *J. Bacteriol.* **178,** 2926–2933.
48. Pedersen, S. (1984) *EMBO J.* **3,** 2895–2898.
49. Deana, A., Ehrlich, R. & Reiss, C. (1996) *J. Bacteriol.* **178,** 2718–2720.
50. Lee, K., Bernstein, J. A. & Cohen, S. N. (2002) *Mol. Microbiol.* **43,** 1445–1456.
51. Riley, M. (1997) *Nucleic Acids Res.* **25,** 51–52.
52. Ow, M. C., Liu, Q. & Kushner, S. R. (2000) *Mol. Microbiol.* **38,** 854–866.
53. O'Hara, E. B., Chekanova, J. A., Ingle, C. A., Kushner, Z. R., Peters, E. & Kushner, S. R. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 1807–1811.
54. Meyer, B. J., Bartman, A. E. & Schottel, J. L. (1996) *Gene* **179,** 263–270.
55. Patel, A. M. & Dunn, S. D. (1995) *J. Bacteriol.* **177,** 3917–3922.
56. Cam, K., Rome, G., Krisch, H. M. & Bouche, J. P. (1996) *Nucleic Acids Res.* **24,** 3065–3070.
57. Baga, M., Goransson, M., Normark, S. & Uhlin, B. E. (1988) *Cell* **52,** 197–206.
58. Vytvytska, O., Jakobsen, J. S., Balcunaite, G., Andersen, J. S., Baccarini, M. & von Gabain, A. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 14118–14123.
59. Zilhao, R., Cairrao, F., Regnier, P. & Arraiano, C. M. (1996) *Mol. Microbiol.* **20,** 1033–1042.
60. Granger, L. L., O'Hara, E. B., Wang, R. F., Meffen, F. V., Armstrong, K., Yancey, S. D., Babitzke, P. & Kushner, S. R. (1998) *J. Bacteriol.* **180,** 1920–1928.