UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

ROLE OF ppGpp IN REGULATING GLOBAL GENE EXPRESSION IN *Escherichia coli*

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

MATTHEW F. TRAXLER
Norman, Oklahoma
2009

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/318509

ROLE OF ppGpp IN REGULATING GLOBAL GENE EXPRESSION IN *Escherichia coli*

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF BOTANY AND MICROBIOLOGY

BY

_____
Dr. Tyrrell Conway, Chair

_____
Dr. Michael McInerney

_____
Dr. David Nagle

_____
Dr. Jimmy Ballard

_____
Dr. Randall Hewes

© Copyright by MATTHEW F. TRAXLER 2009
All Rights Reserved.

This dissertation is dedicated to those who have helped me the most:

My Mentor, Dr. Tyrrell Conway, who showed me the beauty of bacterial physiology, counseled me when answers seemed far away, and gave me the freedom to succeed.

And

My Students, Amanda, Frank, Dong-sul, Sean, Cecilia, Sarah, Stafford, and Vineetha, whose hands pipetted along with mine.

Acknowledgements

Research is a collaborative art form. I gratefully acknowledge these who have helped me along the way.

For help throughout my entire graduate career:

Tyrrell Conway, my mentor.

Joe Grissom, for critical bioinformatics support.

Chapter 2:

Dong-eun Chang, for technical teaching and experimental collaboration.

Chapter 3:

Sean Summers, **Cecilia Nguyen**, and **Vineetha Zacharia**, for experimental assisitance.

Joel Smith and **Arron Hightower**, for expert amino acid quantification.

Chapter 4:

Vineetha Zacharia, **Stafford Marquardt**, **Sean Summers**, **Cecilia Nguyen**, and **Sarah Stark**, for experimental assistance.

Table of Contents

Chapter 1: Literature review and thesis overview……...**1**

Chapter 2: Conceptual model of glucose-lactose diauxie in *Escherichia coli*: Evidence for regulation by guanosine 5',3'-bispyrophosphate (ppGpp)……**21**

Chapter 3: The global, ppGpp-mediated stringent response to amino acid starvation in *Escherichia coli*……..…....**49**

Chapter 4: Discretely calibrated regulatory loops controlled by ppGpp partition global gene expression across the 'feast to famine' gradient…...**112**

Chapter 5: Synthesis and Conclusions…….**161**

Note on appendices and high-thoughput data generated for this thesis…….**174**

Appendix 1: Supplementary data to accompany Chapter 2……...**175**

Appendix 2: Supplementary data to accompany Chapter 3……**181**

List of Tables

Chapter 3, Table 1……97

Chapter 4, Table 1……...144

List of Figures

Chapter 2, Fig 1……..41
Chapter 2, Fig 2……..42
Chapter 2, Fig 3……..43
Chapter 2, Fig 4……..44

Chapter 3, Fig 1…….....98
Chapter 3, Fig 2…….....99
Chapter 3, Fig 3……....100
Chapter 3, Fig 4……...…101
Chapter 3, Fig 5……....102
Chapter 3, Fig 6……...…103
Chapter 3, Fig 7……...…104
Chapter 3, Fig 8……105

Chapter 4, Fig 1……...…149
Chapter 4, Fig 2……150
Chapter 4, Fig 3……151
Chapter 4, Fig 4……152
Chapter 4, Fig 5……152
Chapter 4, Fig 6……...…153
Chapter 4, Fig 7……....154

Appendix 1, Fig 1……175
Appendix 1, Fig 2……177
Appendix 1, Fig 3……179

Appendix 2, Fig 1……193

Abstract

Bacteria grow when nutrient availability supports basic biochemical requirements and remain in stationary phase when basic needs go unmet. This deceptively simple phenomenon requires the orchestrated expression of thousands of genes. Free-living bacteria use a nucleotide second messenger, ppGpp, as a physiological signal and effector to appropriately coordinate global gene expression according to the nutritional quality of the environment. Over the last four decades, expression of many individual genes has been tied to the absence or presence of ppGpp, yet the full scope of gene expression mediated by ppGpp remained undefined. This dissertation defines the role of ppGpp in regulating global gene expression in a model bacterium, *Escherichia coli*.

Chapter 1: Literature review and thesis overview

Historical perspective

The work presented in this thesis is rooted in fundamental principles of bacterial physiology elucidated by such pioneers as François Jacob, Jacques Monod, Neils Kjeldgaard, Ole Maaloe, and their coworkers (19, 28). Bacteria grow when environmental conditions allow for their replication, and remain quiescent in environments that do not. When presented with several nutritional options, bacteria structure their metabolism to preferentially use the best resources first (19, 31). In rich medium, individual bacterial cells are large and maintain a high number of ribosomes (28, 31). Conversely, when grown in minimal medium or growth arrested, cells are small and contain a proportionately lower number of ribosomes (28, 31). Simple as these relationships might seem, we now know that each of these phenotypes requires the coordinated expression of hundreds of genes (11, 12, 27). How do bacteria, relatively 'simple' organisms, execute and balance these complex responses? In essence, the experiments done here are an effort to evaluate the regulation of these long-standing principles at a global scale by using high-throughput technologies. The results suggest that the regulatory molecule ppGpp lies at the apex of the systems that govern these fundamental aspects of prokaryotic physiology.

In the four decades since its discovery, the 'magic spot' aka ppGpp (guanosine-5', 3' -bispyrophosphate) has proven to be one of the most challenging bacterial genetic regulators to understand (36). Even now, central questions endure regarding its mechanism of action and role in regulating cell

physiology (36, 44). ppGpp was first discovered by Michael Cashel in 1969 as a mysterious hyperphosphorylated compound which accumulated after *Escherichia coli* cells were starved for amino acids (8). At that time, the so-called 'stringent response' to amino acid starvation was known to include the rapid down-regulation of stable RNA (tRNA and rRNA) production (28). Cashel also found that 'relaxed' mutants, which continued to make rRNA in the face of amino acid starvation, also failed to synthesize ppGpp (10). Thus, the connection was made: ppGpp is the arbiter of the stringent response. In the ensuing years, a great deal has been learned regarding ppGpp metabolism and its role in regulating a panoply of cellular processes (9).

Control of ppGpp metabolism

E. coli has two enzymes that govern the intracellular ppGpp concentration; RelA and SpoT (48). The ribosome-associated RelA monitors tRNA traffic at the ribosomal A site (47). Upon introduction of an unaminoacylated tRNA into the A site (as happens during amino acid starvation), RelA catalyzes the formation of one molecule of ppGpp from GDP and ATP, or one molecule of pppGpp from GTP and ATP (16). pppGpp is converted into the physiologically relevant ppGpp via the activity of GppA (15). At the time of catalysis, RelA, ppGpp, and the uncharged tRNA all dissociate from the ribosome, priming the system for another cycle and allowing RelA to potentially interact with other stalled ribosomes (47). Thus when intracellular amino acid concentrations fall below the threshold

required by the current translation rate, ribosomes stall and ppGpp accumulates rapidly.

SpoT is a bifunctional enzyme, capable of both ppGpp synthesis and degradation (48). In contrast to RelA, SpoT is important for determining the ppGpp level in response to a variety of starvation conditions, including starvation for carbon (48), phosphate (40), and iron (43). The ppGpp synthetic activity of SpoT is weaker than that of RelA, thus RelA is considered the primary source of ppGpp, while SpoT's role in setting the ppGpp level comes mainly from determining the rate of ppGpp degradation (36). Structural studies suggest that the two active sites in SpoT (one for ppGpp synthesis, the other for ppGpp degradation), may be active at different times, depending on the overall conformation of the enzyme (18). Thus, one conformation corresponds to synthase ON/hydrolase OFF, and the other conformation corresponds to synthase OFF/hydrolase ON (18, 36).

The signals governing SpoT activity are only now beginning to be understood. The location of SpoT in the cytoplasm and whether or not it interacts with the ribosome has been controversial. Most recently, at least one report has shown that SpoT can be isolated from immature ribosomes (a pre-50s particle), but not active ribosomes, depending on the preparation technique (21). SpoT has also been found to interact with at least one GTPase, CgtA (21, 37). A model put forth for *Vibrio cholera* suggests that CgtA functions to hold SpoT in the hydrolase ON state when nutrient availability is high (37). Even more recently, SpoT has been found to interact with the acyl carrier protein, ACP (5),

which functions as a carrier for carbon units in the synthesis of fatty acids (29). The current model is that the ratio of acylated ACP to unacylated ACP influences the activity of SpoT such that fatty acid starvation triggers SpoT-dependent ppGpp synthesis (5). This model is particularly attractive since it provides an explanation for how SpoT might sense carbon starvation (5).

Mechanism of action of ppGpp

Though much progress has been made regarding the mechanism of action of ppGpp, fundamental questions remain unresolved. Fairly early on, it was recognized that the target protein affected by ppGpp was RNA polymerase (RNAP) itself (42). This suggested that ppGpp might have the capacity to influence transcription at virtually all promoters. As the primary cellular reaction to amino acid starvation is the down-regulation of rRNA production, it was hypothesized that the binding of ppGpp to RNAP causes this down-regulation (38). ppGpp-dependent down-regulation of transcription from rRNA promoters has been demonstrated in vivo and in vitro (8, 32, 38, 42). It was also noted that amino acid biosynthetic genes were induced in response to amino acid starvation, concomitant with stringent down-regulation of rRNA genes (9, 38, 41). This induction of amino acid biosynthetic genes was also found to require ppGpp (41). The idea that amino acid biosynthetic genes require ppGpp for their induction was further bolstered by the observation that a strain lacking ppGpp (a *ΔrelA*, *ΔspoT* double knockout, also called $ppGpp^0$) is auxotrophic for multiple amino acids (48).

For many years, in vitro attempts to stimulate down-regulation of rRNA promoters by adding ppGpp to in vitro transcription systems yielded only a ~2-3 fold down-regulation, compared to the ~20 fold down-regulation observed in vivo (17). A major breakthrough was the discovery that a small RNAP binding protein, DksA, allowed for robust down-regulation of rRNA promoters with ppGpp and RNAP in vitro (32). The structure of DksA is similar to that of the Gre factors (GreA and GreB), which posses a long coiled-coil finger domain that inserts into the secondary channel of RNAP (the secondary channel serves as the entry portal for free rNTPs to reach the active site of RNAP) (35). Crystallographic studies also placed the binding site of ppGpp in the secondary channel of RNAP, near the active site (1). Thus, the proposed model was that ppGpp was held in place inside RNAP by DksA, which occupied the secondary channel (1). DksA is understood to potentiate the effect of ppGpp on RNAP. However, the binding site of ppGpp within RNAP has recently been called into question, as point mutations of the RNAP amino acid residues purported to interact with ppGpp did not affect regulation by ppGpp in an in vitro system (44).

ppGpp and DksA have been shown to regulate gene expression in several ways. This includes directly, passively, and indirectly via sigma factor competition. Each of these is considered in kind below.

Direct regulation by ppGpp

Current models for how ppGpp and DksA directly modify RNAP behavior require a detailed understanding of the normal steps that occur as RNAP initiates

transcription [for a review see (17)]. After the promoter sequence on DNA is recognized and bound by the sigma (σ) subunit of RNAP (an arrangement known as the closed complex), a series of conformational changes leads to the separation of the two DNA strands such that the transcriptional start site is positioned at the active site of RNAP. At this stage the RNAP/DNA complex is said to be in the 'open complex.' Ultimately, transition to the transcription elongation complex requires the addition of several rNTPs to the growing RNA chain and the dissociation of the σ-promoter DNA contacts (17). Accumulating evidence suggests that ppGpp and DksA work together to destabilize the open complex itself or some intermediate along the way to open complex formation (32). This destabilization results in the down regulation of promoters that form intrinsically unstable, short-lived open complexes, such as strong rRNA promoters (34). In contrast, promoters that form longer-lived open complexes, such as those found upstream of amino acid biosynthesis genes, are less prone to destabilization (2, 3). Direct induction by ppGpp and DksA of amino acid biosynthesis genes occurs in vitro via a mechanism that is not understood (33). The hypothesis advanced by the authors is that ppGpp/DksA may lower the free energy required for the conformational changes that accompany conversion from the closed to open complex (17, 33). Kinetic parameters determined by the sequence of a given promoter would then dictate whether the progression to open complex formation or open complex collapse is favored in the presence of ppGpp/DksA (17, 33).

Passive induction by ppGpp

During rapid growth, 60-80% of RNAP is devoted to transcription of rRNA genes (34). This high level of rRNA synthesis leads to production of the large number of ribosomes needed for rapid protein synthesis to support a maximal growth rate (28). Such a large sequestration of RNAP engaged in stable RNA synthesis allows only a relatively small amount of RNAP to undertake mRNA synthesis across the remainder of the genome (7, 45). When stress is encountered, such as during starvation, ppGpp accumulates rapidly, resulting in down-regulation of stable RNA synthesis and the liberation of the majority of sequestered RNAP (2, 7, 27). Thus ppGpp passively allows for induction of various promoters across the genome by liberating RNAP from stable RNA synthesis. This model makes intuitive sense given that the majority of RNAP is thought to be engaged in transcription at all times, leaving a relatively small pool of free RNAP (34, 45). Furthermore, many relatively weak promoters with longer-lived open-complexes require higher levels of RNAP when in competition with rRNA promoters in vitro (2).

Indirect regulation via sigma factor competition

In *E. coli* one housekeeping sigma factor, RpoD (σ^{70}), is responsible for mediating transcription from the overwhelming majority of promoters. However, *E. coli* has six other alternative sigma factors, each of which controls transcription of a discreet regulon [reviewed in (14)]. The alternative sigma factors include RpoS (σ^{38}), mediator of the general stress response, RpoN (σ^{54}), active during

nitrogen starvation, RpoE (σ^{24}), which responds to envelope stress, RpoH (σ^{32}), manager of the heat shock response, RpoF (σ^{28}), involved in flagella synthesis, and FecI (σ^{19}), regulator of ferric citrate uptake. Alternative sigma factors therefore compete with each other and σ^{70} to engage the RNAP core enzyme (E) and activate gene expression.

The first indication that this competition was not strictly influenced by the relative levels of each sigma factor came from an investigation with RpoS (24). Perhaps is not surprising that RpoS, which activates ~100 genes during almost any stressful situation (25, 46), is intimately linked to ppGpp (13, 26). RpoS is known to be controlled at the levels of transcription and protein stabilization by ppGpp (6, 13, 26). However, it was found that RpoS-dependent promoters also require ppGpp for their induction, even if the level of RpoS itself is normal (24). One possibility was that ppGpp somehow made RpoS more competitive for RNAP core enzyme. It was subsequently shown that ppGpp destabilizes the interaction between E and σ^{70}, thus affording more opportunity for RpoS to gain access to E (22). Such a scenario implies that all alternative sigma factors would be favored against σ^{70}, in the presence of ppGpp. This hypothesis seems to be correct as subsequent experiments showed that ppGpp enhanced the competitiveness of both RpoH and RpoE for E against σ^{70} (22, 30).

Outstanding questions

At the onset of the work presented in this thesis, there were many open questions regarding ppGpp. Moreover, ppGpp as a model physiological system

offered an opportunity to study several topics of keen interest to me. Fundamentally, I am interested in how bacteria make decisions based on environmental conditions. Bacterial responses are mostly limited to modifying expression of their genes at the transcriptional level.

Expression of each gene or operon is controlled through the interaction between regulatory proteins and DNA operator sites usually present in the upstream promoter region. In essence, the cells integrate information through these regulatory elements to 'compute' whether a given gene should be turned on or off (39). Examination of a wide range of bacterial genetic systems has shown that in some cases regulation is a relatively localized event, wherein a regulator controls expression of a single gene or operon in response to a single environmental attribute (i.e. sugar catabolism operons) (23). However, it is also clear that bacteria utilize global regulators to consider a single environmental condition at the promoters of a large number of genes (20, 23). Thus, bacteria utilize complex transcriptional networks to optimize their gene expression at a global scale (4, 20). As I became acquainted with the available literature, it became clear to me that ppGpp probably affected more systems than any other single regulatory system. Perhaps for this very reason, that is, ppGpp seemed control such a huge number of genes, the extent of gene expression mediated by ppGpp remained poorly defined. Endeavoring to answer this question became a major part of the work described here. Moreover, advances in microarray technology offered an exciting and unprecedented avenue for examining these questions. As the impact of ppGpp on global transcription is so large, ppGpp

itself probably rarely acts in isolation. Thus, another important question is: where does ppGpp fit in the context the global transcriptional network? The role of ppGpp in gene regulation *vis a vis* other global regulators is the other major question which underlies the experiments described here in three major chapters:

Preamble to Chapter Two: ppGpp coordinates global gene expression during glucose-lactose diauxie in Escherichia coli

When I entered Dr. Conway's lab as graduate student in the spring semester of 2003, three members of the lab had just published an important paper in Molecular Microbiology entitled: 'Gene expression profiling of *Escherichia coli* growth transitions: an expanded stringent response model' (11). The work presented in that paper was conducted mostly by a postdoctoral researcher, Dr. Dong-Eun Chang. Dr. Chang showed that two very different physiological conditions that caused temporary growth arrest (glucose-lactose diauxie and H_2O_2 treatment) both triggered the stringent response. In the first case, WT cells were grown in minimal medium containing glucose and lactose. Under these classic experimental conditions first examined by Jacques Monod and coworkers, *E. coli* selectively grows on the glucose first, undergoes a pause in growth (called diauxie), and then resumes growth on lactose (19). In his analysis of microarray data collected from multiple time points during transient growth arrest, Dr. Chang noted that gene expression within three major networks responded: genes of the general stress response, genes involved in carbon

scavenging, and genes involved in ribosome biosynthesis. The work described in the second chapter of this thesis is an effort to understand how the RpoS (stress) and Crp (carbon scavenging) neworks are integrated within the stringent response controlled by RelA/ppGpp. This work culminated in the development of a model which describes the timing and physiological components of the response to transient growth arrest. Results from this chapter are published (PMID: 16467149), and have been reformatted here for consistency.

Preamble to Chapter Three: The global, ppGpp-mediated stringent response to amino acid starvation in Escherichia coli

Having examined the role of ppGpp in response to carbon starvation, I wanted to consider the role of ppGpp in the cellular response to amino acid starvation, especially since amino acid starvation is the archetypal condition known to trigger the stringent response (9). I also felt that it was important to make sure that the experiments conducted could be applied equally to both the WT and the multiauxotrophic ppGpp^0 strain. To this end, I developed an experimental system based on starvation for isoleucine that met this criterion. As I began to gather array data for the WT, ppGpp^0 strain, and several other regulatory mutants, it became clear that before an analysis involving several regulatory players could be interpreted, the actual physiological extent of the ppGpp-dependent stringent response would need to be defined. This became the goal of the research presented in Chapter 3. We found that the ppGpp-mediated stringent response encompassed down-regulation of virtually all types

of macromolecular biosynthesis including genes involved in protein, nucleic acid, and fatty acid biosynthesis. We also found that metabolic gene expression was broadly restructured. In this experimental system, we also noted that the $ppGpp^0$ strain produced 50% more biomass than the WT, despite producing the same amount of total protein. Taken together these results suggested that ppGpp plays a fundamental role in calibrating virtually all macromolecular processes to the translational capacity of the cell. A model that integrates this information at the level of gene expression is presented. Results from this chapter are published (PMID: 18430135), and have been reformatted here for consistency.

Preamble to Chapter Four: Architecture of the stringent response

After having defined the physiological extent of gene expression mediated by ppGpp, I wanted to explore the regulatory hierarchy responsible for executing the global response to amino acid starvation. This is the focus of Chapter 4. Transcriptional patterns observed in the data collected for Chapter 3 suggested that the regulators RpoS and Lrp were active in developing the observed response to isoleucine starvation. Thus, the experimental system I designed in Chapter 3 served as a starting point for analyzing the contribution of these two regulators. After experimentally defining the RpoS and Lrp regulons, the activation time (response time) for each of these regulons was obtained using a large microarray time series of isoleucine starvation. This analysis, along with ppGpp measurements suggested a model in which low amounts of ppGpp are sufficient to allow activation of metabolic genes (such as amino acid biosynthetic

promoters), while genes controlled by RpoS (the general stress response) require a comparatively high amount of ppGpp for their activation. Experiments testing two predictions of this model offered an initial confirmation of the proposed regulatory framework. Implications of such a model for partitioning of metabolic and stress responses are discussed.

References

1. **Artsimovitch, I., V. Patlan, S. Sekine, M. N. Vassylyeva, T. Hosaka, K. Ochi, S. Yokoyama, and D. G. Vassylyev.** 2004. Structural basis for transcription regulation by alarmone ppGpp. Cell **117:**299-310.

2. **Barker, M. M., T. Gaal, and R. L. Gourse.** 2001. Mechanism of regulation of transcription initiation by ppGpp. II. Models for positive control based on properties of RNAP mutants and competition for RNAP. J Mol Biol **305:**689-702.

3. **Barker, M. M., T. Gaal, C. A. Josaitis, and R. L. Gourse.** 2001. Mechanism of regulation of transcription initiation by ppGpp. I. Effects of ppGpp on transcription initiation in vivo and in vitro. J Mol Biol **305:**673-88.

4. **Barrett, C. L., C. D. Herring, J. L. Reed, and B. O. Palsson.** 2005. The global transcriptional regulatory network for metabolism in Escherichia coli exhibits few dominant functional states. Proc Natl Acad Sci U S A **102:**19103-8.

5. **Battesti, A., and E. Bouveret.** 2006. Acyl carrier protein/SpoT interaction, the switch linking SpoT-dependent stress response to fatty acid metabolism. Mol Microbiol **62:**1048-63.

6. **Bougdour, A., and S. Gottesman.** 2007. ppGpp regulation of RpoS degradation via anti-adaptor protein IraP. Proc Natl Acad Sci U S A **104:**12896-901.

7. **Cabrera, J. E., and D. J. Jin.** 2003. The distribution of RNA polymerase in *Escherichia coli* is dynamic and sensitive to environmental cues. Mol Microbiol **50:**1493-505.

8. **Cashel, M., and J. Gallant.** 1969. Two compounds implicated in the function of the RC gene of *Escherichia coli*. Nature **221:**838-41.

9. **Cashel, M., D. R. Gentry, V. J. Hernandez, and D. Vinella.** 1996. The stringent response, p. 1458-1496. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.

10. **Cashel, M., and B. Kalbacher.** 1970. The control of ribonucleic acid synthesis in Escherichia coli. V. Characterization of a nucleotide associated with the stringent response. J Biol Chem **245:**2309-18.

11. **Chang, D. E., D. J. Smalley, and T. Conway.** 2002. Gene expression profiling of *Escherichia coli* growth transitions: an expanded stringent response model. Mol Microbiol **45:**289-306.

12. **Durfee, T., A. M. Hansen, H. Zhi, F. R. Blattner, and D. J. Jin.** 2008. Transcription profiling of the stringent response in Escherichia coli. J Bacteriol **190:**1084-96.

13. **Gentry, D. R., V. J. Hernandez, L. H. Nguyen, D. B. Jensen, and M. Cashel.** 1993. Synthesis of the stationary-phase sigma factor sigma s is positively regulated by ppGpp. J Bacteriol **175:**7982-9.

14. **Gruber, T. M., and C. A. Gross.** 2003. Multiple sigma subunits and the partitioning of bacterial transcription space. Annu Rev Microbiol **57:**441-66.

15. **Hara, A., and J. Sy.** 1983. Guanosine 5'-triphosphate, 3'-diphosphate 5'-phosphohydrolase. Purification and substrate specificity. J Biol Chem **258:**1678-83.

16. **Haseltine, W. A., and R. Block.** 1973. Synthesis of guanosine tetra- and pentaphosphate requires the presence of a codon-specific, uncharged transfer ribonucleic acid in the acceptor site of ribosomes. Proc Natl Acad Sci U S A **70:**1564-8.

17. **Haugen, S. P., W. Ross, and R. L. Gourse.** 2008. Advances in bacterial promoter recognition and its control by factors that do not bind DNA. Nat Rev Microbiol **6:**507-19.

18. **Hogg, T., U. Mechold, H. Malke, M. Cashel, and R. Hilgenfeld.** 2004. Conformational antagonism between opposing active sites in a bifunctional RelA/SpoT homolog modulates (p)ppGpp metabolism during the stringent response [corrected]. Cell **117:**57-68.

19. **Jacob, F., and J. Monod.** 1961. Genetic regulatory mechanisms in the synthesis of proteins. J. Mol Biol **3:**318-56.

20. **Janga, S. C., H. Salgado, A. Martinez-Antonio, and J. Collado-Vides.** 2007. Coordination logic of the sensing machinery in the transcriptional regulatory network of Escherichia coli. Nucleic Acids Res **35:**6963-72.

21. **Jiang, M., S. M. Sullivan, P. K. Wout, and J. R. Maddock.** 2007. G-protein control of the ribosome-associated stress response protein SpoT. J Bacteriol **189:**6140-7.

22. **Jishage, M., K. Kvint, V. Shingler, and T. Nystrom.** 2002. Regulation of sigma factor competition by the alarmone ppGpp. Genes Dev **16:**1260-70.

23. **Kaplan, S., A. Bren, A. Zaslaver, E. Dekel, and U. Alon.** 2008. Diverse two-dimensional input functions control bacterial sugar genes. Mol Cell **29:**786-92.

24. **Kvint, K., A. Farewell, and T. Nystrom.** 2000. RpoS-dependent promoters require guanosine tetraphosphate for induction even in the presence of high levels of sigma(s). J Biol Chem **275:**14795-8.

25. **Lacour, S., and P. Landini.** 2004. SigmaS-dependent gene expression at the onset of stationary phase in Escherichia coli: function of sigmaS-dependent genes and identification of their promoter sequences. J Bacteriol **186:**7186-95.

26. **Lange, R., D. Fischer, and R. Hengge-Aronis.** 1995. Identification of transcriptional start sites and the role of ppGpp in the expression of rpoS, the structural gene for the sigma S subunit of RNA polymerase in Escherichia coli. J Bacteriol **177:**4676-80.

27. **Liu, M., T. Durfee, J. E. Cabrera, K. Zhao, D. J. Jin, and F. R. Blattner.** 2005. Global transcriptional programs reveal a carbon source foraging strategy by Escherichia coli. J Biol Chem **280:**15921-7.

28. **Maaloe, O., and N. Kjeldgaard.** 1966. Control of Macromolecular Synthesis, vol. W. A. Benjamin, inc., New York.

29. **Magnuson, K., S. Jackowski, C. O. Rock, and J. E. Cronan, Jr.** 1993. Regulation of fatty acid biosynthesis in Escherichia coli. Microbiol Rev **57:**522-42.

30. **Magnusson, L. U., A. Farewell, and T. Nystrom.** 2005. ppGpp: a global regulator in Escherichia coli. Trends Microbiol **13:**236-42.

31. **Neidhardt, F. C., J. L. Ingraham, and M. Schaechter.** 1990. Physiology of the Bacterial Cell, 1st ed, vol. Sinauer Associates, Inc., Sunderland, MA.

32. **Paul, B. J., M. M. Barker, W. Ross, D. A. Schneider, C. Webb, J. W. Foster, and R. L. Gourse.** 2004. DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. Cell **118:**311-22.

33. **Paul, B. J., M. B. Berkmen, and R. L. Gourse.** 2005. DksA potentiates direct activation of amino acid promoters by ppGpp. Proc Natl Acad Sci U S A **102:**7823-8.

34. **Paul, B. J., W. Ross, T. Gaal, and R. L. Gourse.** 2004. rRNA transcription in *Escherichia coli*. Annu Rev Genet **38:**749-70.

35. **Perederina, A., V. Svetlov, M. N. Vassylyeva, T. H. Tahirov, S. Yokoyama, I. Artsimovitch, and D. G. Vassylyev.** 2004. Regulation through the secondary channel--structural framework for ppGpp-DksA synergism during transcription. Cell **118:**297-309.

36. **Potrykus, K., and M. Cashel.** 2008. (p)ppGpp: still magical? Annu Rev Microbiol **62:**35-51.

37. **Raskin, D. M., N. Judson, and J. J. Mekalanos.** 2007. Regulation of the stringent response is the essential function of the conserved bacterial G protein CgtA in Vibrio cholerae. Proc Natl Acad Sci U S A **104:**4636-41.

38. **Reiness, G., H. L. Yang, G. Zubay, and M. Cashel.** 1975. Effects of guanosine tetraphosphate on cell-free synthesis of Escherichia coli ribosomal RNA and other gene products. Proc Natl Acad Sci U S A **72:**2881-5.

39. **Setty, Y., A. E. Mayo, M. G. Surette, and U. Alon.** 2003. Detailed map of a cis-regulatory input function. Proc Natl Acad Sci U S A **100:**7702-7.

40. **Spira, B., N. Silberstein, and E. Yagil.** 1995. Guanosine 3',5'-bispyrophosphate (ppGpp) synthesis in cells of Escherichia coli starved for Pi. J Bacteriol **177:**4053-8.

41. **Stephens, J. C., S. W. Artz, and B. N. Ames.** 1975. Guanosine 5'-diphosphate 3'-diphosphate (ppGpp): positive effector for histidine operon transcription and general signal for amino-acid deficiency. Proc Natl Acad Sci U S A **72:**4389-93.

42. **Travers, A.** 1976. RNA polymerase specificity and the control of growth. Nature **263:**641-6.

43. **Vinella, D., C. Albrecht, M. Cashel, and R. D'Ari.** 2005. Iron limitation induces SpoT-dependent accumulation of ppGpp in Escherichia coli. Mol Microbiol **56:**958-70.

44. **Vrentas, C. E., T. Gaal, M. B. Berkmen, S. T. Rutherford, S. P. Haugen, D. G. Vassylyev, W. Ross, and R. L. Gourse.** 2008. Still looking for the magic spot: the crystallographically defined binding site for ppGpp on RNA polymerase is unlikely to be responsible for rRNA transcription regulation. J Mol Biol **377:**551-64.

45. **Wade, J. T., K. Struhl, S. J. Busby, and D. C. Grainger.** 2007. Genomic analysis of protein-DNA interactions in bacteria: insights into transcription and chromosome organization. Mol Microbiol **65:**21-6.

46. **Weber, H., T. Polen, J. Heuveling, V. F. Wendisch, and R. Hengge.** 2005. Genome-wide analysis of the general stress response network in Escherichia coli: sigmaS-dependent genes, promoters, and sigma factor selectivity. J Bacteriol **187:**1591-603.

47. **Wendrich, T. M., G. Blaha, D. N. Wilson, M. A. Marahiel, and K. H. Nierhaus.** 2002. Dissection of the mechanism for the stringent factor RelA. Mol Cell **10:**779-88.

48. **Xiao, H., M. Kalman, K. Ikehara, S. Zemel, G. Glaser, and M. Cashel.** 1991. Residual guanosine 3',5'-bispyrophosphate synthetic activity of *relA* null mutants can be eliminated by *spoT* null mutations. J Biol Chem **266:**5980-90.

Chapter 2: Conceptual model of glucose-lactose diauxie in *Escherichia coli*: Evidence for regulation by guanosine 5',3'-bispyrophosphate (ppGpp)

Abstract

Guanosine 5'3'-bispyrophosphate (ppGpp), also known as "magic spot", has been shown to bind prokaryotic RNA polymerase to down-regulate ribosome production and increase transcription of amino acid biosynthesis genes during the stringent response to amino acid starvation. Since many environmental growth perturbations cause ppGpp to accumulate, we hypothesize ppGpp to have an overarching role in regulating the genetic program that coordinates transitions between logarithmic growth (feast) and growth arrest (famine). We used the classic glucose-lactose diauxie as an experimental system to investigate the temporal changes in transcription that accompany growth arrest and recovery in wildtype *Escherichia coli* and in mutants that lack RelA (ppGpp synthetase) and other global regulators, i.e., RpoS and Crp. In particular, diauxie was delayed in the *relA* mutant and was accompanied by a 15% decrease in the number of carbon sources utilized and a 3-fold overall decrease in the induction of RpoS and Crp regulon genes. Thus the data significantly expand the previously known role of ppGpp and support a model wherein the ppGpp-dependent redistribution of RNA polymerase across the genome is the driving force behind control of the stringent response, general stress response, and starvation-induced carbon scavenging. Our conceptual model of diauxie describes these global control circuits as dynamic, interconnected, and dependent upon ppGpp for the efficient

temporal coordination of gene expression that programs the cell for transitions between feast and famine.

Introduction

The fitness of free-living organisms depends on their ability to withstand environmental insults and grow as rapidly as possible when conditions allow. Consequently, the coordination of growth control processes constitutes a fundamental level of regulation in prokaryotes. For this reason, the bacterial existence is often thought to be one of “feast and famine”(21). In the laboratory, nutritional conditions which cause biphasic growth provide a unique opportunity to investigate this most basic of bacterial behaviors. When cultured on a mixture of glucose and lactose, *E. coli* grows preferentially on glucose until the glucose is exhausted, resulting in growth arrest while the cells adjust to growth on lactose, i.e., diauxie. The genetic basis for biphasic sugar catabolism, elucidated by Jacob and Monod(18), is exemplified by *lac* operon induction, which is a textbook paradigm for illustrating genetic control. However, transcriptome analysis revealed that diauxie involves much more than induction of the *lac* operon, that diauxie is accompanied by a global response to growth arrest that apparently ensures recovery when conditions allow growth to resume(8). The purpose of this study is to dissect the regulatory networks that govern diauxie as a means for understanding how the cell integrates the response to growth arrest.

We showed previously(8) that during steady-state logarithmic growth, gene expression in *E. coli* is quasi-steady state. In contrast, when glucose is exhausted and growth of the culture is arrested a major component of the transcriptome‘s adjustment to diauxie is the stringent response, which includes

down-regulation of a large number of transcription and translation apparatus genes, inhibition of ribosome synthesis, and induction of amino acid biosynthesis genes(6, 7). Also induced are general stress response and carbon scavenging genes, which apparently ensure survival during growth arrest and switching to alternative carbon sources. These genes are controlled primarily by the stationary phase sigma factor, RpoS(10, 14), and the cAMP receptor protein (Crp), which governs catabolite repression, a response to sugar limitation(20, 33). There is strong evidence for a connection between stringent control and the general stress response: ppGpp is required for RpoS accumulation(11) and ppGpp-bound RNA polymerase preferentially binds alternative sigma factors(22, 27, 29). Likewise, a connection between carbon scavenging and the general stress response is manifested as an RpoS-dependent tradeoff between induction of genes in the RpoS and Crp regulons(20). These published studies are indicative of a larger emerging theme in global gene regulation in prokaryotes: large scale regulatory circuits do not function independently of one another, but instead are finely calibrated to coordinate bacterial cell functions in response to environmental cues.

On the basis of the behavior of the transcriptome during diauxie and dependence of the general stress response on ppGpp, we(8) and others(27) hypothesized that ppGpp controls not only the stringent response, but also the regulatory networks that coordinate survival during stationary phase and resumption of growth following growth arrest. To dissect the roles of individual regulators of this process, we now compare the transcription profiles of mutants

lacking RelA, RpoS, and Crp across the diauxic time course. We show here that efficient induction of all genes that are significantly induced during diauxie, including primarily the Crp and RpoS regulons, is RelA-dependent, implying that ppGpp is at the apex of global regulation during times of carbon starvation. We incorporate these results into a conceptual model of glucose-lactose diauxie that places at the center of growth transitions the ppGpp-mediated balance between stringent-controlled repression of the translation apparatus and induction the general stress response and carbon scavenging regulons.

Results

Systematic regulatory mutant analysis during diauxie

K-means cluster analysis of the transcriptome dataset of wildtype *E. coli* during glucose-lactose diauxie(8) revealed three regulatory networks (RpoS, Crp, and RelA) that dominated the transcription profile (Supplementary Fig. 1). To further elucidate their role in diauxie, we cultured *rpoS*, *crp* and *relA* mutants on minimal medium containing a mixture of glucose and lactose as sole carbon sources. Total RNA was isolated during logarithmic growth in the glucose phase of diauxie and at approximately 10 min intervals during diauxie, and was analyzed in triplicate using whole-genome *E. coli* MG1655 oligonucleotide glass microarrays. The RNA control for all microarrays was from an early logarithmic phase culture of wildtype *E. coli* MG1655 on minimal glucose medium. The

datasets are available on the Internet (www.ou.edu/microarray). We consider the transcriptome analysis of these regulatory mutants, as follows.

Transcriptome of rpoS and crp mutants

Under diauxie conditions, the strain lacking RpoS grew normally, as shown previously(10). The strain lacking Crp was unable to resume growth on lactose, as expected(28). Figure 1 shows a transcriptome comparison for the mutant strains during the diauxic lag period with that of the wildtype strain. In the wildtype, exhaustion of glucose was accompanied by diauxie and a whole-genome expression profile characteristic of release from catabolite repression and induction of the general stress response; the Crp regulon was induced in the 10 min interval immediately preceding diauxie, while induction of the RpoS regulon occurred within the first 10 min following growth arrest. These results are consistent with the known diauxie-dependent kinetics of RpoS protein accumulation, which is slow(10), and cAMP accumulation, which is rapid(28) (Fig. 1).

To identify genes regulated by RpoS and/or Crp during diauxie we used K-means cluster analysis (K=22) of the entire dataset shown in Fig. 1a. This analysis revealed four clusters containing 97 highly regulated genes, which are shown in Fig 1b. The constituent genes of the Crp or RpoS regulons were not induced in the respective mutants (Fig 1b). Fourteen genes were not induced in the *rpoS* strain during the diauxic lag, including genes that are typically associated with the general stress response and known to be induced in an

RpoS-dependent manner, such as *bolA*, *dps, wrbA,* and *mscL*(37). Thirty-one genes were not induced in the *crp* strain when glucose was exhausted during diauxie, including genes such as *lacZA, mglBA, lamB, glpFK,* and *rbsD* that are known to depend on Crp for their expression(41). Finally, another 14 genes were not induced in either the *rpoS* or *crp* strains during diauxie, including *glgS* which is known to be regulated by both RpoS and Crp(15). The analysis also revealed repression of 37 genes, shown in Fig. 1b, that are known to be associated with the stringent response(8). Repression of these genes was not affected to a large extent by mutation of *rpoS* or *crp*. Taken together, these results are consistent with diauxie in the wildtype being accompanied by the stringent response and induction, simultaneously, of the RpoS-dependent general stress response and Crp-dependent scavenging for alternative carbon sources. The remaining genes that were not regulated to a large extent during diauxie are shown in grey in Fig. 1a.

Transcriptome of relA mutant

Since accumulation of RpoS and transcription of RpoS-dependent genes requires ppGpp(11), we wanted to determine whether a defect in ppGpp synthesis would affect expression of RpoS-dependent genes under diauxic conditions. There are two ppGpp synthetase enzymes to consider. The dogma is that ribosome-associated RelA synthesizes ppGpp in response to amino acid starvation and SpoT, which has both weak synthetase and ppGpp hydrolase activities, is responsible for ppGpp accumulation in response to carbon

starvation(7, 25). Upon growth arrest, regardless of whether it is synthesized by RelA or SpoT, ppGpp accumulates rapidly in the cell, binds to RNA polymerase (RNAP), and stimulates the down-regulation of the translation apparatus that characterizes the stringent response(1). The rapid accumulation of high levels of ppGpp during glucose lactose diauxie has been reported previously(13).

Ideally, this experiment should compare diauxie in the wildtype with that of a strain that is completely devoid of ppGpp. However, *spoT* mutants are only viable in a *relA* background and *relA spoT* strains are multiply auxotrophic for nine amino acids(34). Thus, we tested the wildtype and *relA spoT* strains with amino acids added to the growth medium and found they did not exhibit glucose-lactose diauxie (data not shown). Presumably this is because the amino acids served as carbon sources to support growth during the period of time when the *lac* operon was being induced. Regardless of the cause, it was not possible to culture the *relA spoT* strain under glucose-lactose diauxie conditions because of the amino acid requirement of this strain. However, the *relA* strain is able to grow on minimal medium without added amino acids. Therefore, we investigated the impact on the transcriptome of a *relA* mutation, which has been shown to extend diauxie in *E. coli* strains(7, 13, 16, 23). We observed large-scale differences in both the timing and extent of differential gene expression in the *relA* mutant during diauxie (Fig. 2). The down-regulation of the transcription and translation apparatus genes (listed in Fig. 1 and Supplementary Table 1) was delayed, reaching a minimum in the 26-36 min interval of diauxie, as opposed to the 0-10 min interval for the wildtype strain. Thus, the results are consistent with

the known role of ppGpp in stringent control of ribosome synthesis during growth arrest.

Altered induction of RpoS and Crp regulons in relA mutant

All members of the RpoS regulon exhibited delayed induction in the *relA* mutant, with the exception of one gene (*bolA*). The amplitude of the 'burst' of gene induction normally seen at the onset of diauxie in the wildtype was lessened 3-fold in the *rpoS* strain (compare Fig. 2b and Fig. 2d). Since *relA* mutants delay ppGpp accumulation during nutrient downshifts(2, 23) and the RpoS-dependent general stress response requires ppGpp(11, 22, 27), the data are consistent with a model which places ppGpp in control of the general stress response.

The *relA* mutant also exhibited diminished induction of the Crp regulon (Fig. 2b), with an average peak expression of Crp-activated genes that was 3-fold lower than that observed in the wildtype (Fig. 2d). In the wildtype strain induction of some Crp-dependent genes was immediate, occurring during the 10 min interval prior to diauxie, and constituted a first wave of gene induction in response to glucose starvation. Other Crp-dependent genes, including the *lac* operon, were not induced in the wildtype until the onset of diauxie. By contrast, in the *relA* mutant induction of the *lac* operon was delayed by 25 min (Supplementary Table 1). This likely is the ultimate cause of the lengthened diauxie of the *relA* strain. Thus, the *relA* mutation generally dampened expression of all CRP-dependent genes that normally are induced in the wildtype

during diauxie (Supplementary Table 1), as well as *rmf* (Supplementary Fig. 2), which is known to be outside control of both Crp and RpoS(17). Since *relA* strains are known to have higher than normal cAMP levels following nutrient downshifts(2) and normal cAMP levels following amino acid starvation(30), this argues that the RelA-dependent effect on Crp-dependent gene activation is not mediated by the intracellular cAMP concentration. Rather, in vitro transcription assays demonstrated that ppGpp is required for maximal induction of the *lac* operon and it has been suggested that this might also be the case for other catabolic genes and operons(31). The results shown in Fig 2. are consistent with the idea that maximal induction of Crp-activated genes depends on ppGpp.

In further support of the assertion that ppGpp is required for induction of carbon catabolism genes, phenotype arrays (Biolog GN2 microplates) showed a 15% decrease in the number of carbon sources used by the *relA* and *relA spoT* mutants (Supplementary Fig. 3). Specifically, we observed that alpha-hydroxybutyric acid, alpha-ketobutyric acid, propionate, D-saccharic acid, lactulose, Tween 40, and Tween 80 were not consumed by these mutants (Supplementary Table 2). These results are consistent with a model wherein RelA-dependent adjustment of intracellular ppGpp levels in response to starvation is required for normal induction of survival genes, including the Crp regulon.

Discussion

E. coli is a comparatively simple model system, yet a full understanding of the regulatory connections which shape prokaryotic physiology remains elusive. The strategy of iteratively examining the roles of several transcription factors in a single, complex physiological transition (i.e., growth arrest) provided a conceptual framework for integrating diverse cellular processes. This general strategy and the datasets generated here should be of value for systems biology.

To derive a conceptual model of diauxie that accounts for the global redistribution of gene expression in response to growth arrest, such as that caused by diauxie, we propose a simple RNAP switch model that is consistent with known biochemical parameters (regulatory mechanisms) of the stringent response (Fig. 3). To our knowledge, all growth perturbations result in rapid accumulation of ppGpp(7, 36), which binds to RNA polymerase (RNAP) and causes the down-regulation of the translation apparatus that characterizes the stringent response(7). Given that stable RNA synthesis constitutes up to ~80% of transcription in rapidly growing cells(3, 5), reduced transcription from these stringent promoters, which has been proposed to result from various mechanisms, including destabilization of the RNAP-promoter open complex or inactive dead-end promoter complexes(1, 24, 29), should greatly increase the availability of free RNAP(3, 7). The indirect or so called passive model postulates that the increased availability of RNAP caused by inhibition of rRNA transcription frees RNAP to bind to other promoters, i.e., amino acid biosynthetic genes and those shown in this study to be activated. Alternatively, the direct model of RelA-dependent activation proposes that ppGpp and DksA act directly

to stimulate RNAP at promoters of amino acid biosynthesis genes(7, 29). These models are not mutually exclusive and may in fact both contribute to the observed changes in gene expression caused by ppGpp. Since gene expression profiles reflect the distribution of RNAP at promoters across the genome(38), our data suggest that reprogramming of RNAP by binding ppGpp increases the transcription initiation frequency at many more promoters than had been recognized previously, including the CRP and RpoS regulons. Also, it was recently shown that growth arrest is accompanied by RelA-dependent, physical redistribution of RNAP(5). Thus, regardless of whether the ppGpp effect is direct or indirect, gene expression profiling of *E. coli* during diauxie supports the proposed switch model (Fig. 3), wherein ppGpp not only controls the down-regulation of stringent promoters, but also the activation of stress survival and carbon catabolism genes. We propose that ppGpp-dependent reprogramming of RNAP is the driving force behind differential gene expression during diauxie.

Our conceptual model of diauxie, based on the microarray data presented above and incorporating the RNAP switch model (Fig. 3), is given in Fig. 4. Diauxie involves much more than induction of the *lac* operon. Before the *lac* operon is induced, the general stress and stringent responses are induced and catabolite repression is released. Through the use of regulatory mutants, we show that the large-scale changes in the transcriptome during diauxie, including induction of the Crp and RpoS regulons and adaptation to metabolism of a less-preferred carbon source, requires RelA for efficient and timely control. By accounting for the ppGpp-dependent system that controls ribosome number in

bacteria(7, 12, 19, 26, 27, 35), our model incorporates one of the fundamental principles of bacterial physiology, i.e., that bacterial cell growth rate is determined by the number of ribosomes active in protein synthesis(19). Also, the model is consistent with the recently discovered roles of DksA in mediating physical interactions of ppGpp with RNAP(29) and ppGpp-dependent RpoS accumulation(4).

Our data bring to light a controversy regarding the roles of the two ppGpp synthetase enzymes in regulating the stringent response. The prevailing notion is that RelA responds to amino acid starvation, while SpoT governs ppGpp accumulation during carbon starvation(7). The data presented here indicate that starvation for glucose in the wildtype induces the stringent response. In addition, our results indicate that mutation of *relA* alters this response. However, our experiments do not distinguish whether this effect is mediated directly by RelA or indirectly, i.e., glucose starvation leads to amino acid starvation. We also note that the *relA* strain exhibits prolonged diauxie (Fig. 2) and diauxie is abolished in both the wildtype and mutant strains when amino acids are present (data not shown). Another possibility is that RelA, along with SpoT, can sense carbon starvation independently of amino acid pool fluctuations. However, this seems unlikely given the strong evidence for physical association of RelA with the ribosome, allowing it to monitor translational pausing and hence the amino acyl-tRNA pool(39). Thus, the model shown in Fig. 4 attributes ppGpp accumulation to amino acid starvation brought about indirectly from exhaustion of glucose, rather than directly by carbon starvation. Direct measurement of the amino acid

pool should provide insight into the physiological state caused by carbon starvation under diauxic conditions.

ppGpp is known to affect the overall physiological state of the cell through transcriptional regulation of a large number of promoters. Processes affected include stable RNA synthesis, amino acid biosynthesis, sigma factor competition, and induction of the σ^s-dependent stress response genes. The *lac* operon has been shown to require ppGpp for maximal expression(31). The results presented here extend the influence of ppGpp beyond control of *lac* to the larger Crp regulon and therefore to catabolite repression and carbon catabolism in general. As such, ppGpp signals the nutritional quality of the environment and coordinates adjustments to gene expression across a continuum that ranges from maximum growth and metabolism to complete growth arrest and damage control. Accordingly, our conceptual model (Fig. 4) places ppGpp at the apex of the stimulus-response pathways that allow *E. coli* to successfully negotiate growth arrest during diauxie. This regulatory network includes ppGpp-dependent control of the general stress response, carbon scavenging, and ribosome synthesis. The benefit of their coordinated regulation during growth transitions is critically important, as the energy that would otherwise have been spent on growth functions(12, 26) is now conserved while the cell diverts its attention to survival in stationary phase until conditions allow growth to resume. Thus, ppGpp controls the feast and famine existence and may therefore profoundly influence the activities of microbes in a host as well as survival between hosts.

Materials and Methods

Strains and growth conditions

E. coli MG1655 and isogenic mutants were cultured in a 2 l Biostat B fermentor (B. Braun Biotech International) containing 1 liter of Morpholinepropanesulfonic acid (MOPS) minimal medium with 0.5 g/l of glucose and 1.5 g/l of lactose, as described(8). The temperature was maintained at 37°C and pH was kept constant at 7.2 by the addition of 2 M NaOH. The dissolved oxygen level was maintained above 20% of saturation by adjusting the agitation speeds in the range of 270-500 rpm with fixed 1 l/min air flow. Growth was monitored as absorbance at 600nm. *E. coli* Δ*relA*251::kan^R was a gift from M. Cashel, constructed as described(40). The *E. coli* Δ*crp*::kan^R and Δ*rpoS*::kan^R strains were constructed by allelic replacement(9) of the entire genes. These mutant strains are isogenic with *E. coli* MG1655.

Microarray analysis

Microarray analysis was carried out essentially as described(38). Total RNA was extracted from cells, diluted (1:1) in ice-cold RNA*later* (Ambion) and purified using RNeasy columns (Qiagen), as described(8). RNA was labeled by first strand cDNA synthesis using reverse transcriptase, random primers, and aminoallyl-dUTP incorporation; Cy-3 and Cy-5 dyes were chemically coupled in vitro to the aminoallyl-derivatized cDNA. The oligonucleotide microarrays used in this study were printed on GAPS II slides (Corning) with a probe set containing

70 base oligonucleotide probes for all *E. coli* MG1655 genes (Operon Biotechnologies) using a Molecular Dynamics Gen III Array Spotter (Amersham Biosciences). Slides were hydrated and flash-dried, UV-cross-linked, and blocked with succinic anhydride, then equal amounts of the Cy-3 and Cy-5 labeled samples were hybridized in triplicate to microarrays using a Discovery system and ChipMap reagents (Ventana Medical Systems). For all microarrays, the experimental sample was labeled with Cy-5 and the control, from early logarithmic growth of *E. coli* MG1655 wildtype on minimal glucose medium, was labeled with Cy-3. Hybridized slides were scanned on a GenePix4000 scanner (Axon), the data collected using GenePix (ver. 5.0) software, and uploaded to our database for analysis (http://www.ou.edu/microarray). The data were normalized by a local Lowess algorithm(32) implemented on our database and the replicate arrays averaged for analysis. Clustering algorithms were implemented in Spotfire DecisionSite for Functional Genomics software.

Acknowledgements

We gratefully acknowledge the support of the National Institutes of Health RO1-AI48945-05. We thank Ming Yann (Frank) Wu and Marybeth Langer for technical assistance with microarrays and phenotype assays, and Mike Cashel and Moselio Schaechter for critical reading of the manuscript.

Figure Legends

Fig. 1. Transcriptome analysis of the diauxic lag in *E. coli* MG1655 wildtype, *rpoS*, and *crp* strains. (a) Log_2 ratio plot of time series microarray data. K-means cluster analysis (K=22) of the dataset revealed 4 clusters containing 97 significantly regulated genes (60 induced and 37 repressed genes). The induced genes belonged to two regulons: RpoS (cluster 2, red) and Crp (cluster 3, green), and cluster 4 contained genes regulated by both RpoS and Crp (purple). The repressed genes include the RelA/stringent response (cluster 1, yellow). All other genes are shown in the background (grey). (b) Hierarchical cluster analysis of significantly regulated genes. Ratios are displayed colorimetrically in a heatmap: blue indicates genes with >2.5 log_2-fold higher expression (~3 standard deviations) and yellow indicates <2.5 log_2-fold lower expression in the experimental condition compared to the control; the colors darken to black to indicate no change in expression. Strains and time points (min) are shown

above (a) below (b). Diauxie began at 0 min. The gene names are shown on the left, together with the corresponding K-means clusters.

Fig. 2. Transcriptome analysis of the diauxic lag in *E. coli* MG1655 wildtype and *relA* strains. Expression of RpoS and Crp regulons (identified as being significantly regulated in Fig. 1) for the wildtype (a and b) and *relA* (c and d) strains, shown as $\log_2$ ratio plot of time series microarray data for individual genes (a and c) and the average $\log_2$ expression ratios for regulons (b and d); RpoS (red), Crp (green), both RpoS and Crp (purple), transcription/translation apparatus genes (blue), and growth curve (black).

Fig. 3. RNAP switch model in *E. coli* (wildtype). Cells programmed for growth (left) have low intracellular levels of ppGpp and 80% of active RNAP is engaged in stable RNA synthesis, resulting in balanced growth; RNAP availability, cAMP, and RpoS levels are low, repressing transcription of Crp^- and RpoS-dependent carbon scavenging and stress response genes. Cells programmed for arrest (right) accumulate high intracellular levels of ppGpp, which binds to RNAP, perhaps synergistically with DksA, causing a genome-wide redistribution of transcription because stable RNA promoters are most sensitive to repression by ppGpp-reprogrammed RNAP; by a passive mechanism, RNAP availability increases significantly, as do cAMP and RpoS levels, leading to induction of Crp^- and RpoS-dependent genes, as well as other stringent response induced genes.

Fig. 4. Conceptual model of glucose-lactose diauxie in wildtype *E. coli* MG1655. During the -10 to 0 min interval, a cascade of responses stimulated by glucose exhaustion results in rapid accumulation of ppGpp, which binds to and reprograms RNAP, culminating in flipping of the RNAP switch, shown in Fig. 3 (see text). Concomitantly, the liberation of RNAP by repression of stable RNA synthesis, and accumulation of cAMP and the RpoS sigma factor, combine to activate transcription of the carbon scavenging and general stress response genes, respectively, which peaks at ~20 min. Transcription of the *lac* operon is first observed in the 0-10 min interval and remains high until lactose is exhausted. Metabolism of lactose during the 10 to 20 min interval leads to a return to pre-stimulus conditions, consistent with replenishment of the charged tRNA pool and degradation of ppGpp, which causes the RNAP switch to program the cells for growth. Growth resumes in the 20-30 min interval and transcript levels for the carbon scavenging and general stress response genes begin to fall, while those for the transcription and translation apparatus genes rise, reaching pre-diauxie levels by approximately 50 min.

Fig. 1

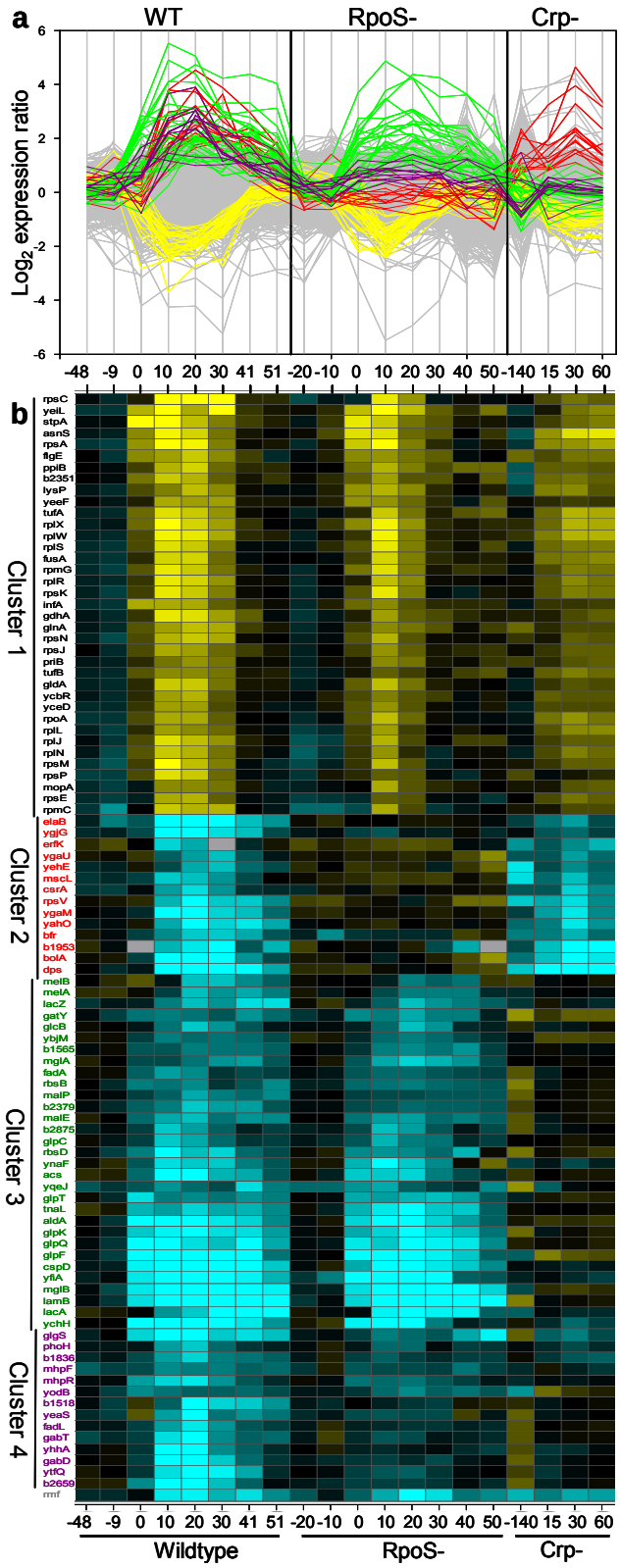

a
WT
RpoS-
Crp-
log2 expression ratio
6
4
2
0
-2
-4
-6
-48 -9 0 10 20 30 41 51 -20 -10 0 10 20 30 40 50 -140 15 30 60
b
Cluster 1
Cluster 2
Cluster 3
Cluster 4
rpsC
yeiL
stpA
asnS
rpsA
flgE
ppiB
b2351
lysP
yeeF
tufA
rplX
rplW
rplS
fusA
rpmG
rplR
rpsK
infA
gdhA
glnA
rpsN
rpsJ
priB
tufB
gldA
ycbR
yceD
rpoA
rplL
rplJ
rplN
rpsM
rpsP
mopA
rpsE
rpmC
elaB
yqjG
erfK
ygaU
yehE
mscL
csrA
rpsV
ygaM
yahO
btr
b1953
bolA
dps
melB
melA
lacZ
gatY
glcB
ybjM
b1565
mglA
fadA
rbsB
malP
b2379
malE
b2875
glpC
rbsD
ynaF
acs
yqeJ
glpT
tnaL
aldA
glpK
glpQ
glpF
cspD
yfiA
mglB
lamB
lacA
ychH
glgS
phoH
b1836
mhpF
mhpR
yodB
b1518
yeaS
fadL
gabT
yhhA
gabD
ytfQ
b2659
rmf
-48 -9 0 10 20 30 41 51 -20 -10 0 10 20 30 40 50 -140 15 30 60
Wildtype
RpoS-
Crp-

Fig. 2.

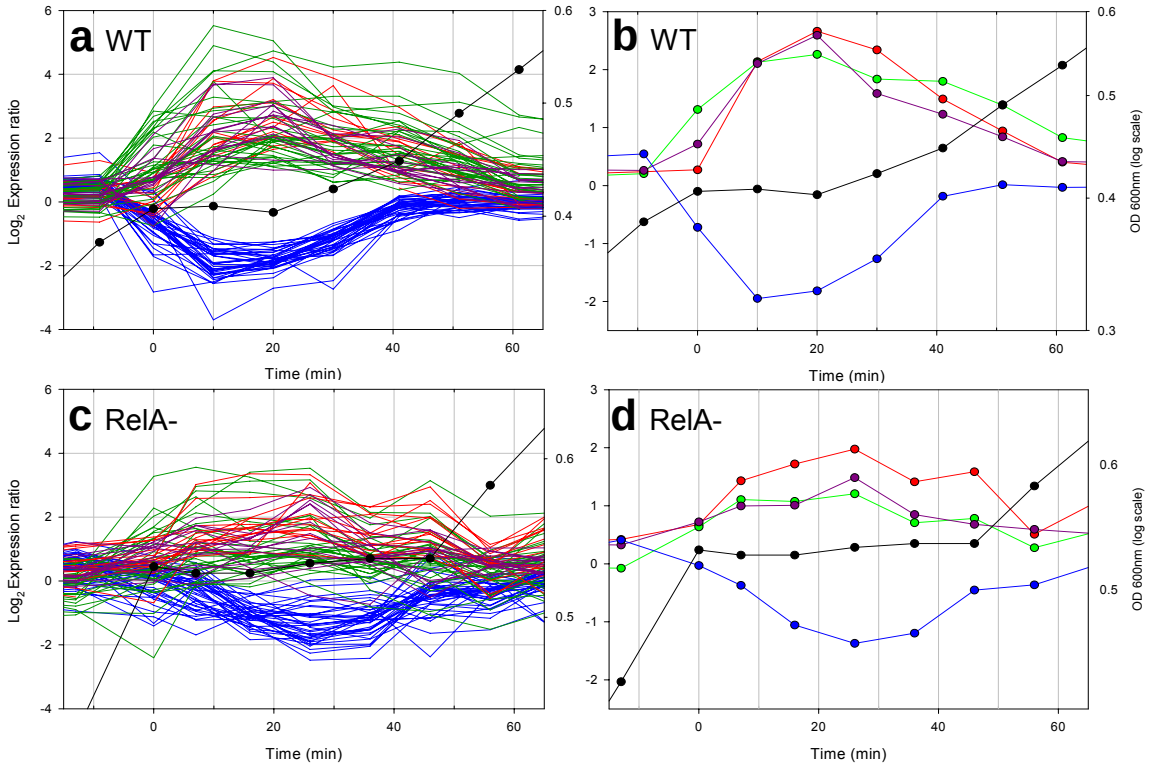

a WT
b WT
c RelA-
d RelA-
Log2 Expression ratio
OD 600nm (log scale)
Time (min)

Fig. 3.

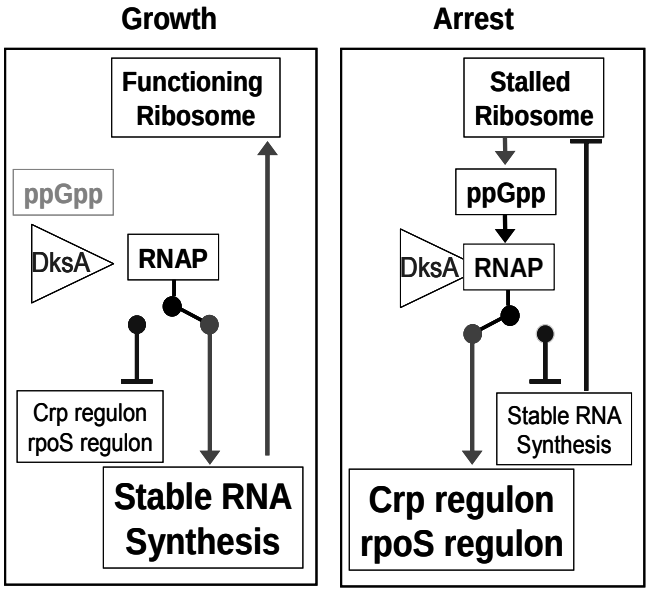
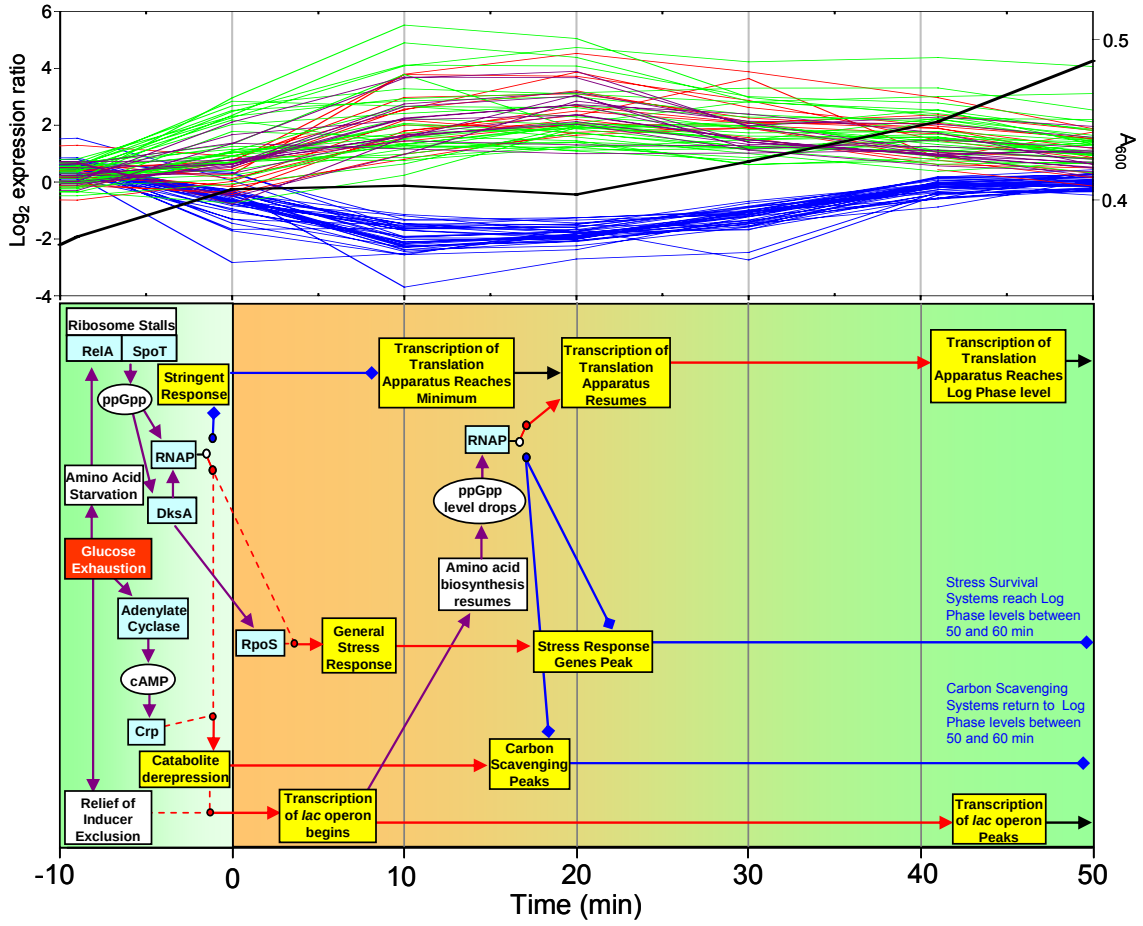

Growth
Arrest
Functioning Ribosome
ppGpp
DksA
RNAP
Crp regulon
rpoS regulon
Stable RNA Synthesis
Stalled Ribosome
ppGpp
DksA
RNAP
Stable RNA Synthesis
Crp regulon
rpoS regulon

Fig. 4.


Log2 expression ratio
A600
Time (min)
Ribosome Stalls
RelA
SpoT
Stringent Response
ppGpp
RNAP
Amino Acid Starvation
DksA
Glucose Exhaustion
Adenylate Cyclase
cAMP
Crp
Catabolite derepression
Relief of Inducer Exclusion
RpoS
General Stress Response
Transcription of Translation Apparatus Reaches Minimum
Transcription of Translation Apparatus Resumes
Transcription of Translation Apparatus Reaches Log Phase level
RNAP
ppGpp level drops
Amino acid biosynthesis resumes
Stress Response Genes Peak
Carbon Scavenging Peaks
Transcription of lac operon begins
Transcription of lac operon Peaks
Stress Survival Systems reach Log Phase levels between 50 and 60 min
Carbon Scavenging Systems return to Log Phase levels between 50 and 60 min

References

1. **Barker, M. M., T. Gaal, C. A. Josaitis, and R. L. Gourse.** 2001. Mechanism of regulation of transcription initiation by ppGpp. I. Effects of ppGpp on transcription initiation in vivo and in vitro. J Mol Biol **305:**673-88.

2. **Braedt, G., and J. Gallant.** 1977. Role of the *rel* gene product in the control of cyclic adenosine 3',5'-monophosphate accumulation. J Bacteriol **129:**564-6.

3. **Bremer, H., and P. P. Dennis.** 1996. Modulation of chemical composition and other parameters of the cell by growth rate, p. 1553-1569. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.

4. **Brown, L., D. Gentry, T. Elliott, and M. Cashel.** 2002. DksA affects ppGpp induction of RpoS at a translational level. J Bacteriol **184:**4455-65.

5. **Cabrera, J. E., and D. J. Jin.** 2003. The distribution of RNA polymerase in *Escherichia coli* is dynamic and sensitive to environmental cues. Mol Microbiol **50:**1493-505.

6. **Cashel, M., and J. Gallant.** 1969. Two compounds implicated in the function of the RC gene of *Escherichia coli*. Nature **221:**838-41.

7. **Cashel, M., D. R. Gentry, V. J. Hernandez, and D. Vinella.** 1996. The stringent response, p. 1458-1496. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.

8. **Chang, D. E., D. J. Smalley, and T. Conway.** 2002. Gene expression profiling of *Escherichia coli* growth transitions: an expanded stringent response model. Mol Microbiol **45:**289-306.

9. **Datsenko, K. A., and B. L. Wanner.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A **97:**6640-5.

10. **Fischer, D., A. Teich, P. Neubauer, and R. Hengge-Aronis.** 1998. The general stress sigma factor sigmaS of *Escherichia coli* is induced during diauxic shift from glucose to lactose. J Bacteriol **180:**6203-6.

11. **Gentry, D. R., V. J. Hernandez, L. H. Nguyen, D. B. Jensen, and M. Cashel.** 1993. Synthesis of the stationary-phase sigma factor sigma s is positively regulated by ppGpp. J Bacteriol **175:**7982-9.

12. **Gralla, J. D.** 2005. Escherichia coli ribosomal RNA transcription: regulatory roles for ppGpp, NTPs, architectural proteins and a polymerase-binding protein. Mol Microbiol **55:**973-7.

13. **Harshman, R. B., and H. Yamazaki.** 1971. Formation of ppGpp in a relaxed and stringent strain of *Escherichia coli* during diauxie lag. Biochemistry **10:**3980-2.

14. **Hengge-Aronis, R.** 1996. Regulation of gene expression during entry into stationary phase, p. 1497-1512. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.

15. **Hengge-Aronis, R., R. Lange, N. Henneberg, and D. Fischer.** 1993. Osmotic regulation of rpoS-dependent genes in Escherichia coli. J Bacteriol **175:**259-65.

16. **Ishiguro, E. E.** 1979. Regulation of peptidoglycan biosynthesis in *relA*+ and *relA*- strains of *Escherichia coli* during diauxic growth on glucose and lactose. Can J Microbiol **25:**1206-8.

17. **Izutsu, K., A. Wada, and C. Wada.** 2001. Expression of ribosome modulation factor (RMF) in Escherichia coli requires ppGpp. Genes Cells **6:**665-76.

18. **Jacob, F., and J. Monod.** 1961. Genetic regulatory mechanisms in the synthesis of proteins. J. Mol Biol **3:**318-56.

19. **Keener, J., and M. Nomura.** 1996. Regulation of ribosome synthesis, p. 14171431. *In* F. C. Neidhardt, R. Curtiss III, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.

20. **King, T., A. Ishihama, A. Kori, and T. Ferenci.** 2004. A regulatory trade-off as a source of strain variation in the species *Escherichia coli*. J Bacteriol **186:**5614-20.

21. **Koch, A. L.** 1971. The adaptive responses of *Escherichia coli* to a feast and famine existence. Adv Microb Physiol **6:**147-217.

22. **Kvint, K., A. Farewell, and T. Nystrom.** 2000. RpoS-dependent promoters require guanosine tetraphosphate for induction even in the presence of high levels of sigma(s). J Biol Chem **275:**14795-8.

23. **Lazzarini, R. A., M. Cashel, and J. Gallant.** 1971. On the regulation of guanosine tetraphosphate levels in stringent and relaxed strains of Escherichia coli. J Biol Chem **246:**4381-5.

24. **Maitra, A., I. Shulgina, and V. J. Hernandez.** 2005. Conversion of Active Promoter-RNA Polymerase Complexes into Inactive Promoter Bound Complexes in E. coli by the Transcription Effector, ppGpp. Mol Cell **17:**817-29.

25. **Metzger, S., G. Schreiber, E. Aizenman, M. Cashel, and G. Glaser.** 1989. Characterization of the relA1 mutation and a comparison of relA1 with new relA null alleles in Escherichia coli. J Biol Chem **264:**21146-52.

26. **Neidhardt, F. C., J. L. Ingraham, and M. Schaechter.** 1990. Physiology of the Bacterial Cell: a Molecular Approach, vol. Sinauer Associates, Sunderland, MA.

27. **Nystrom, T.** 2004. Growth versus maintenance: a trade-off dictated by RNA polymerase availability and sigma factor competition? Mol Microbiol **54:**855-62.

28. **Pastan, I., and S. Adhya.** 1976. Cyclic adenosine 5'-monophosphate in *Escherichia coli*. Bacteriol Rev **40:**527-51.

29. **Paul, B. J., W. Ross, T. Gaal, and R. L. Gourse.** 2004. rRNA transcription in *Escherichia coli*. Annu Rev Genet **38:**749-70.

30. **Primakoff, P.** 1981. In vivo role of the relA+ gene in regulation of the lac operon. J Bacteriol **145:**410-6.

31. **Primakoff, P., and S. W. Artz.** 1979. Positive control of lac operon expression in vitro by guanosine 5'-diphosphate 3'-diphosphate. Proc Natl Acad Sci U S A **76:**1726-30.

32. **Quackenbush, J.** 2002. Microarray data normalization and transformation. Nat Genet **32 Suppl:**496-501.

33. **Saier, M. H., T. M. Ramseier, and J. Reizer.** 1996. Regulation of carbon utilization, p. 1325-1343. *In* F. C. Neidhardt, R. Curtiss III, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D. C.

34. **Sarubbi, E., K. E. Rudd, and M. Cashel.** 1988. Basal ppGpp level adjustment shown by new *spoT* mutants affect steady state growth rates and *rrnA* ribosomal promoter regulation in *Escherichia coli*. Mol Gen Genet **213:**214-22.

35. **Travers, A.** 1974. The stringent response--21 years on. Basic Life Sci **3:**67-80.

36. **VanBogelen, R. A., P. M. Kelley, and F. C. Neidhardt.** 1987. Differential induction of heat shock, SOS, and oxidation stress regulons and accumulation of nucleotides in *Escherichia coli*. J Bacteriol **169:**26-32.

37. **Weber, H., T. Polen, J. Heuveling, V. F. Wendisch, and R. Hengge.** 2005. Genome-wide analysis of the general stress response network in Escherichia coli: sigmaS-dependent genes, promoters, and sigma factor selectivity. J Bacteriol **187:**1591-603.

38. **Wei, Y., J. M. Lee, C. Richmond, F. R. Blattner, J. A. Rafalski, and R. A. LaRossa.** 2001. High-density microarray-mediated gene expression profiling of *Escherichia coli*. J Bacteriol **183:**545-56.

39. **Wendrich, T. M., G. Blaha, D. N. Wilson, M. A. Marahiel, and K. H. Nierhaus.** 2002. Dissection of the mechanism for the stringent factor RelA. Mol Cell **10:**779-88.

40. **Xiao, H., M. Kalman, K. Ikehara, S. Zemel, G. Glaser, and M. Cashel.** 1991. Residual guanosine 3',5'-bispyrophosphate synthetic activity of *relA* null mutants can be eliminated by *spoT* null mutations. J Biol Chem **266:**5980-90.

41. **Zheng, D., C. Constantinidou, J. L. Hobman, and S. D. Minchin.** 2004. Identification of the CRP regulon using in vitro and in vivo transcriptional profiling. Nucleic Acids Res **32:**5874-93.

Chapter 3: The global, ppGpp-mediated stringent response to amino acid starvation in *Escherichia coli*

Abstract

The stringent response to amino acid starvation, whereby stable RNA synthesis is curtailed in favor of transcription of amino acid biosynthetic genes, is controlled by the alarmone ppGpp. To elucidate the extent of gene expression effected by ppGpp, we designed an experimental system based on starvation for isoleucine, which could be applied to both wild-type *Escherichia coli* and the multi-auxotrophic *relA spoT* mutant (ppGpp^0). We used microarrays to profile the response to amino acid starvation in both strains. The wildtype response included induction of the general stress response, down regulation of genes involved in production of macromolecular structures, and comprehensive restructuring of metabolic gene expression, but not induction of amino acid biosynthesis genes *en masse*. This restructuring of metabolism was confirmed using kinetic Biolog assays. These responses were profoundly altered in the ppGpp^0 strain. Furthermore, upon isoleucine starvation, the ppGpp^0 strain exhibited a larger cell size and continued growth, ultimately producing 50% more biomass than the wildtype, despite producing a similar amount of protein. This mutant phenotype correlated with aberrant gene expression in diverse processes including DNA replication, cell division, and fatty acid and membrane biosynthesis. We present a model that expands and functionally integrates the ppGpp-mediated stringent response to include

control of virtually all macromolecular synthesis and intermediary metabolism.

Introduction

When nutrients become limiting for growth, *E. coli* cells adjust their gene expression program from one that supports growth to one that allows for prolonged survival in stationary phase. In many bacteria, a key potentiator of this physiological switch is the accumulation of the alarmones guanosine 5',3' bispyrophosphate and guanosine pentaphosphate (ppGpp and pppGpp: collectively referred to here as ppGpp) (12). When amino acids become limiting, uncharged tRNAs bind to the ribosomal A site, signaling ribosome-associated RelA to synthesize ppGpp (61). Aided by DksA, ppGpp binds in the secondary channel of RNA polymerase (RNAP) near the active site (2); the immediate consequence is the cessation of transcription of stable RNAs (ribosomal and transfer RNAs), termed the stringent response(12, 34). ppGpp decreases the half-life of the open complex at most promoters tested thus far; the physiological result is the strong down regulation of promoters with intrinsically short half-lives, such as those of stable RNA genes (4). Since expression of ribosomal protein genes is controlled by rRNA levels, the stringent response includes a large-scale down regulation of the translation apparatus (43). As transcription of the translation apparatus genes and stable RNAs can account for a large percentage (60-80%) of the transcription occurring in rapidly growing cells, the liberation of RNAP from these genes is thought to passively allow up regulation of diverse promoters activated at the onset of stationary phase (3, 4). Also, ppGpp in concert with DksA has been shown to directly stimulate transcription from

promoters of several amino acid biosynthesis genes (42). In support of the 'passive mechanism', we recently showed that a *ΔrelA* mutant is constrained in its ability to up regulate genes in diverse regulatory networks during carbon starvation (52).

In addition to RelA, ppGpp is also produced by SpoT, apparently in response to diverse signals including carbon (64), iron (57), and fatty acid starvation (5). The synthase activity of SpoT is not as robust as that of RelA, and SpoT also contains ppGpp hydrolase activity, which mediates ppGpp turnover and thus is important for determining the intracellular ppGpp concentration (64). Mutants lacking *relA* and *spoT* are completely devoid of ppGpp ($ppGpp^0$), a state that results in a pleiotropic phenotype (64). Most notably, $ppGpp^0$ strains exhibit a 'relaxed' phenotype, i.e., stable RNA synthesis continues after exhaustion of amino acids (47). $ppGpp^0$ strains are also auxotrophic for eleven amino acids, apparently because ppGpp is required for effective transcription of amino acid biosynthetic genes (64). Additionally, relaxed strains exhibit a prolonged period of growth arrest after amino acid starvation has been relieved (55). Years of experimentation have linked ppGpp to a wide variety of physiological processes beyond translation and amino acid biosynthesis, including catabolite (de)repression (31, 52), DNA synthesis (14, 24), fatty acid metabolism (19, 23), general stress response (20), surface organelle production (fimbriae and flagella) (1, 36), and virulence (35). Such wide-ranging regulation suggests ppGpp is a critical element of the response network that allows cells to adapt their physiology to their surroundings; however, the manner in which these processes

are integrated within the stringent response remains unclear. Furthermore, the role of ppGpp in controlling intermediary metabolism beyond that of amino acid biosynthesis has not been experimentally defined.

To address these questions, we sought to examine the extent of the stringent response under conditions of amino acid starvation. To do so, we exploited a long-known metabolic anomaly characteristic of *E. coli* K12 strains, whereby starvation for isoleucine is caused by excess valine (32). To determine the extent of regulation by ppGpp we obtained transcription profiles of the wildtype (WT) and $ppGpp^0$ strains, using whole genome microarrays. In comparison to the WT, the strain lacking ppGpp was crippled in its ability to regulate genes involved in diverse areas of metabolism, including central metabolism, amino acid biosynthesis/degradation, and nucleotide biosynthesis. We also conducted a number of experiments that provide a context for interpreting these transcription profiles, including measurements of metabolites in the culture medium, changes in the metabolic proteome, viability assays, and measurements of total protein, RNA, and biomass. Based on these data, we present a model that expands and integrates our view of the metabolic and structural rearrangements that accompany the stringent response, which is at once more massive and finely tuned than previously appreciated.

Results

Experimental system for eliciting the stringent response to amino acid starvation

We were severely constrained in our choice of experimental systems because ppGpp^0 strains are multiply auxotrophic and because the stringent response is thought to broadly impact amino acid biosynthesis (12). To elicit the stringent response to amino acid starvation, many studies have utilized serine hydroxamate, which binds to and interferes with seryl-tRNA synthetases. While this strategy is excellent for inducing ppGpp accumulation, it falls short of modeling the concerted response to depletion of the intracellular amino acid pool because little or no new protein can be produced, post-treatment (51). Thus, after serine hydroxamate treatment, no reorganization of the proteome can occur.

Another widely used experimental system is based on the vulnerability of K-12 strains to valine toxicity (32). The first dedicated reaction in branched chain amino acid biosynthesis is catalyzed by acetohydroxy acid synthase (AHAS), which forms α-acetolactate from two molecules of pyruvate during the synthesis of valine or the formation of α-acetohydroxybutyrate from one molecule of pyruvate and one molecule of α-ketobutyrate during the synthesis of isoleucine, (for review, see (54). *E. coli* has three different AHAS enzymes: AHAS I (*ilvBN*) uses pyruvate exclusively for the production of valine, AHAS II (*ilvGM*) uses pyruvate and α-ketobutyrate for the production of isoleucine, and AHAS III (*ilvIH*) catalyzes both reactions but favors pyruvate and α-ketobutyrate as substrates. Both AHAS I and III are feedback inhibited by valine. K-12 strains of *E. coli* harbor a frame-shift mutation in the *ilvG* gene, which renders the AHAS II enzyme inactive. Thus, when isoleucine is limiting and valine is in excess, AHAS

I and III are inhibited, resulting in an inability to biosynthesize isoleucine. While starvation for isoleucine can be induced by dosing cells with valine in minimal medium (i.e., isoleucine absent) (32), this strategy was not available to us because of the limitations imposed by the multiple amino acid auxotrophy associated with lack of ppGpp in *E. coli* MG1655 Δ*relA* Δ*spoT* ($ppGpp^0$).

To circumvent this problem, we retroverted the valine-dosing strategy by imposing isoleucine starvation in the presence of the other 19 amino acids, including valine, which we reasoned would inhibit isoleucine biosynthesis, as described above. We grew the cells in MOPS medium containing glucose (0.2%) plus all 20 amino acids (60), except that isoleucine was provided at 60 μM instead of 400 μM, which allowed WT cells to reach an OD of 0.6-0.7 (Fig. 1A). At this OD, isoleucine was exhausted, but valine was still in excess (Fig. 1), and as expected, growth was arrested; this could be alleviated by re-addition of isoleucine (data not shown).

While we favored isoleucine starvation over serine hydroxamate treatment, this system is not without potential confounding factors. The growth arrest caused by excess valine has been attributed to metabolic consequences other than isoleucine starvation, namely α-ketobutyrate accumulation. This conclusion is based on valine inhibited cells grown on glucose minimal medium having a slightly above normal (144%) physiological intracellular level of isoleucine (25). Mathematical modeling predicted an accumulation of α-ketobutyrate (65), which is known to inhibit the glucose PTS transporter (16), leading to the idea that valine toxicity should be attributed to carbon starvation-

induced growth arrest (65). To test the possibility that glucose transport was inhibited in our experiments, we examined the array data, but could find no evidence for glucose starvation, i.e., catabolite derepression (discussed in detail below), as we previously observed for carbon starved cells (13, 52). Moreover, a microarray time course showed that the first genes induced in this experimental setup are those of the branched chain amino acid biosynthetic pathways (data not shown), implying that from a physiological stand point, the cells acutely sensed the depletion of isoleucine. A second potential complication was also considered: α-ketobutyrate accumulation, concomitant with high expression of the leucine biosynthetic pathway, leads to production of norleucine, a methionine analog (7). Norleucine can be incorporated into protein in place of methionine (15). However, the incorporation of norleucine into protein was found to be completely prevented by supplementation with exogenous methionine (7). Since our medium contained exogenous methionine, we expect any negative effects of norleucine accumulation to be minimized. Thus, we interpret our results in the context of isoleucine starvation.

Growth and pattern of ppGpp accumulation in isoleucine starved cultures

To maximize reproducibility, cultures were grown in 1 L volumes in a fermenter under steady pH and O_2 saturation levels. To examine the response to isoleucine starvation, isogenic *E. coli* MG1655 wild type (WT) and ppGpp^0 strains were grown in isoleucine-limited medium and samples were taken in log phase and following growth arrest. The WT and ppGpp^0 strains grew at similar

rates in logarithmic phase. However, the ppGpp^0 strain exhibited a prolonged lag phase before achieving robust growth; this accounts for the differences in the two growth curves (Fig 1). The response to limiting isoleucine was evident by an initial slowing of growth followed by growth arrest. Growth began to slow in the WT at OD of ~0.4 and growth arrested at OD of ~0.6. In the ppGpp^0 strain growth began to slow at OD ~0.5 and arrested at OD ~0.8 ppGpp began to accumulate in the WT before growth slowed, and was first detectable at OD ~0.3. The ppGpp level increased over the next 100 minutes, leveling off in growth arrested cells (OD ~0.6 at about 500 min in Fig. 1). ppGpp was undetectable in the Δ*relA* Δ*spoT* mutant. In our assays, the level of ppGpp that ultimately accumulated in the WT was ~800 pmoles/ml/OD, which equates to an intracellular concentration of ~0.9 mM. A similar intracellular concentration was observed after isoleucine starvation was provoked by addition of valine to cells growing in minimal media (56) .

Altered flux through amino acid degradative and biosynthetic pathways in response to isoleucine starvation is ppGpp-dependent

We sought to characterize the pattern of amino acid utilization under the experimental conditions described above to allow for better interpretation of the high-throughput data described in subsequent sections. During the course of growth and entry into growth arrest, culture samples were harvested, filtered, and the filtrates were immediately frozen. Samples were then analyzed by capillary electrophoresis-mass spectrometry (CE-MS) to measure the concentrations of 19

amino acids at each time point (cysteine could not be measured). The results for selected amino acids are shown in Fig 1. Concentrations of the 8 amino acids not shown in Fig. 1 did not change significantly over the course of the experiment. At an OD of ~0.4 in both cultures, isoleucine levels dropped below detectable limits, leading to growth arrest of the WT at OD ~0.6 and in the ppGpp^0 strain at OD ~0.8 (Fig. 1B). Starved WT cultures immediately resumed growth upon addition of isoleucine, accompanied by depletion of oxygen; these responses were significantly delayed in the ppGpp^0 strain (the period of delay depended on the length of growth arrest; data not shown).

The next three amino acids consumed from the medium were aspartate, followed by glutamine, followed by serine. Serine was included in the medium at a much higher level (10 mM) than the other amino acids since it also serves as a carbon source that is co-metabolized with glucose (60). In both strains, the usage of serine in the medium accelerated approximately 2-fold at the onset of growth arrest concomitant with a 2.8-fold decrease in the rate of glucose consumption in the WT (discussed below) (Fig. 1C-D). As the cells transitioned into growth arrest, we observed that the saturation of O_2 in the medium increased by approximately 20%, implying that the cells consumed less oxygen (Fig. 1C-D). Acetate accumulated steadily over the entire time-course. Taken together, these results imply that amino acid starvation causes aerobic metabolism of glucose to be curtailed in favor of fermentation of serine to acetate via pyruvate.

We observed a large difference in the pattern of glutamate production/consumption between the WT and ppGpp^0 strains (Fig. 1E-F). In the

WT, the level of glutamate in the medium increased across the time series, a trend that accelerated noticeably at the onset of growth arrest. In contrast, the level of glutamate in the ppGpp^0 culture decreased significantly over time, starting at the onset of growth arrest and dropping below detection by the conclusion of the experiment. Another observable trend was seen in the utilization of glycine, alanine, and leucine. In the WT culture, the levels of these three amino acids remained tightly associated, showing a modest decrease by the end of the experiment. In the ppGpp^0 culture, the levels of glycine, alanine, and leucine diverged at the onset of growth arrest, with the glycine level increasing to a much higher level than in the WT culture.

Overall, these results suggest that the metabolic response to isoleucine starvation is complex and involves the rerouting of flux through multiple pathways. Most notably these adjustments entailed the apparent conversion of serine to acetate as the primary energy-generating process and diminished metabolism of glucose. Differences in the patterns of amino acid utilization/formation between the WT and ppGpp^0 cultures suggest that the mutant strain was defective in its ability to restructure its metabolism. Specifically, the aberrant production of glycine by the ppGpp^0 strain suggests heteroclitic metabolism of serine, the precursor of glycine (46). Finally, the depletion of glutamate observed in the ppGpp^0 culture suggests a potential problem in maintaining the balance of glutamate and α-ketoglutarate, and thus has profound implications for the ability of the ppGpp^0 strain to fulfill the many

needs met by glutamate in redox homeostasis, as a metabolic precursor, and as an amine donor.

Overview of microarray datasets

To test the transcriptional response to isoleucine starvation, we extracted RNA from growth arrested WT and isogenic mutants (OD ~0.6 at 500 min for the WT and OD ~0.8 at 640 min for the ppGpp^0 strain). In the WT, this corresponded to the time of maximum ppGpp accumulation (Fig. 1A). The control RNA was extracted from exponentially growing WT cells in identical medium replete with isoleucine. Since the stringent response is known to inhibit stable RNA synthesis, it is likely that the proportion of mRNA to stable RNA is different in the WT compared to the ppGpp^0 strain. However, it is not possible to estimate these differences using the microarrays employed in this study because they do not contain probes for stable RNA transcripts. Moreover, the normalization strategy (RMA) used for the data processing offsets differences in total RNA and/or the relative proportion of mRNA and stable RNA. Hence, the data allow direct comparison of mRNA levels between strains and growth conditions, but do not compensate for gross differences in RNA content. The WT transcriptional response to isoleucine starvation was extensive, with 1024 genes differentially regulated >2-fold ($\log_2 = 1$). A list of these genes is shown in Supplemental table T1. Of the 532 genes that were induced >2-fold in the WT strain, about one-third are involved in metabolism (174 genes), 139 are involved in the RpoS-dependent general stress response, and >200 have unknown functions. The WT down

regulated 492 genes, including many genes associated with the translation apparatus (>40 ribosomal protein genes and 19 accessory translation genes), a hallmark of the stringent response.

Global comparison of the WT and isogenic ppGpp^0 mutant transcriptional responses to isoleucine starvation revealed profound differences between them (Fig. 2A). In fact, when the transcription profiles of the two strains were compared directly, 1427 genes (>30% of the genome) showed expression levels that deviated 2-fold or more in the ppGpp^0 strain. A list of these genes is available in Supplemental table T2. While both the WT and ppGpp^0 strain induced over 500 genes in response to isoleucine starvation, only 133 genes were commonly induced in both strains (Fig. 2B). Moreover, both strains down regulated over 450 genes >2-fold, but only 198 of these genes were down regulated in both. The comprehensive down regulation of translation apparatus genes observed in the WT was essentially absent in the ppGpp^0 strain, and induction of the RpoS-dependent general stress response was severely diminished (Fig. 2C). Overall, these results suggest that the response to isoleucine starvation is fundamentally altered in the ppGpp^0 strain.

We also obtained a transcription profile of an isogenic Δ*relA* strain starved for isoleucine and found that its response was strikingly similar to the ppGpp^0 strain, indicating a minimal role for SpoT in the stringent response to isoleucine starvation (Fig 2A lower panel). Genes which were expressed differently in the Δ*relA* and ppGpp^0 strain are listed in Supplementary table T3. Based on the similarity of the transcription profiles of the *ΔrelA* and ppGpp^0 strains, it is

reasonable to assume that the differences observed between the WT and ppGpp^0 strain qualitatively hold true also for the Δ*relA* strain. Thus, we focused our analysis on the ppGpp^0 strain, so the results could be interpreted in the complete absence of ppGpp.

The WT response to isoleucine starvation involves regulation of diverse metabolic pathways at the transcriptional level.

To interpret microarray data in a metabolic context, $\log_2$ gene expression ratios were overlaid onto metabolic maps (Figs. 3 and 4). This analysis shows a large number of genes in multiple pathways were differentially regulated in response to isoleucine starvation. We interpret these data with the caveat that the relationship between simple induction or repression of a pathway at the transcriptional level does not necessarily reflect the level of flux or active enzymes in these pathways. However, general correlations between gene expression and metabolic activity have been observed (40).

The WT induced genes in all branches of central metabolism (Fig. 3), including the pentose phosphate pathway, glycolysis, TCA cycle, and the glyoxylate shunt. Within the pentose phosphate pathway three genes were induced, including two in the non-oxidative branch, *tktB* and *talA*, which are known members of the RpoS regulon. Among glycolytic genes, three were induced, including two whose products catalyze the formation of metabolic intermediates known to exert feedback control of glycolytic flux (*fbaB* and *pykA*). Many genes encoding enzymes involved in pyruvate metabolism were induced in

the WT, suggesting that pyruvate is a critical nexus in the metabolic adjustment to isoleucine starvation. *sdaA*, which encodes the major serine deaminase, was induced 2.6-fold. This induction correlates well with increased serine uptake from the medium (Fig 1). When considered together, these observations are consistent with increased flux from glycolytic intermediates and serine to pyruvate. Expression of genes involved in the TCA cycle was complex with the induction of *acnA*, genes of the glyoxylate shunt, and the gene encoding malate synthase G, *maeB* and the down regulation of genes involved in the conversion of succinate to oxaloacetate. Together these alterations in gene expression suggest that carbon entering the TCA cycle from acetate or β-oxidation of fatty acids may be channeled to the production of pyruvate, as well as α-ketoglutarate, the precursor of glutamate, which accumulated in the medium under these conditions (Fig. 1).

Isoleucine starvation triggered induction of genes directly involved in the branched chain amino acid pathways as well as genes in pathways which generate precursors for branched chain amino acid biosynthesis. *E. coli* synthesizes isoleucine from two precursor metabolites: pyruvate and oxaloacetate, (for review, see (41)). Oxaloacetate from the TCA cycle is converted in one step to aspartate. Aspartate can also be made from asparagine in a single reaction. Aspartate is converted to threonine in five steps (Fig. 3). Threonine is then deaminated to form α-ketobutyrate, which is a substrate for AHAS II and III. Thus, from a physiological standpoint, the cell can convert asparagine→aspartate→threonine→isoleucine in 11 steps, 10 of which were up

regulated under the experimental conditions examined here. IlvGM (AHAS II) was the only one of the three AHAS enzymes whose genes were induced by isoleucine starvation in the WT (*ilvG* 22-fold and *ilvM* 25-fold). While not enzymatically effective, the up regulation of *ilvG* and *ilvM* likely represents the cells' attempt to induce valine insensitive AHAS II. In addition to *ilvG* and *ilvM*, all of the genes for enzymes involved in synthesis of the branched chain amino acids valine and isoleucine were strongly induced: *ilvC* (2.3-fold), *ilvE* (8.0-fold), and *ilvD* (11-fold). The genes of the *leuABCD* operon, which are responsible for leucine biosynthesis, were up regulated (8.5- to 26-fold). Collectively, this strong, comprehensive induction constitutes a direct, albeit impotent, response to isoleucine starvation.

The ppGpp^0 metabolic response to isoleucine starvation deviates from WT at the transcriptional level.

The transcriptional differences observed in the profiles of the WT and ppGpp^0 strains in response to isoleucine starvation are far-reaching. It is long known that ppGpp is required for stringent induction of genes involved in amino acid biosynthesis (12). This trend is observed in the data presented here, wherein the ppGpp^0 strain failed to induce genes associated with biosynthesis of the branched chain amino acids, as well as threonine and glutamate (Fig. 4). The transcription profiles show that induction of multiple genes involved in the metabolism of arginine, alanine, serine, and glutamate is contingent upon ppGpp. Most notably, the data presented here also establish that the expression of many

metabolic genes beyond the scope of amino acid biosynthesis is also ppGpp-dependent. These include genes of glycolysis, the pentose phosphate pathway, the TCA cycle and the glyoxylate shunt. The strong repression of numerous central metabolism genes observed in the $ppGpp^0$ strain implies a comprehensive down-shift in metabolic potential. Diminished carbon flux through glycolysis would be expected to impact production of precursor metabolites, and ultimately, the flux of carbon into other central metabolic pathways. Additionally, normal regulation of genes involved in pyruvate metabolism, a metabolic focal point, also required ppGpp. While the regulation of many of the metabolic genes discussed here may be indirectly regulated by ppGpp, taken together these results suggest that ppGpp plays a larger role in regulating intermediary metabolism than previously recognized.

Biolog analysis shows diversification of carbon source utilization in response to isoleucine limitation.

In light of the metabolic rearrangements shown in the transcription profiles, we sought a strategy to measure the response to isoleucine starvation with respect to changes in the overall metabolic capacity of the cells. To do this, we harvested cells for Biolog GN2 microplate analysis from the isoleucine-starved cultures and immediately added chloramphenicol to inhibit protein synthesis and preserve their metabolic capacity, as described elsewhere (26). Thus, the pattern of carbon sources utilized represents a metabolic snapshot of the enzymes present in the cells at the time of sampling. It should also be noted

that because the cells are washed and incubated in basal, minimal medium, the allosteric constraints applied by the components of the growth medium were relieved, theoretically allowing for an uninhibited display of metabolic capacity. Cells where harvested for Biolog analysis at times corresponding to active growth (OD 0.3), the onset of growth arrest, and 1.5 hours after the onset of growth arrest (Fig 5). The cells were incubated in Biolog plates for 24 hours in an Omnilog system, and the amount of reduced tetrazolium violet dye in each well was quantified every 15 min. The amount of dye reduced corresponds to the extent of carbon source oxidized. Dye reduction was plotted vs. time, and the area under each resulting kinetic curve was used as an overall measure of the cells' ability to utilize each carbon source.

Hierarchical cluster analysis of the WT Biolog data indicated that the number of carbon sources used increased as the cells progressed into growth arrest. Distinct groups of carbon sources were utilized at each of the three time points. The first group of carbon sources was utilized strongly at all time points and consisted mainly of carbohydrates that are assimilated directly into glycolysis. These included mannitol, gluconate, glycerol, fructose, mannose, N-acetyl-glucosamine, and glucose. Serine was the only amino acid that fell into this group. This is not surprising given the concomitant consumption of serine and glucose observed in Fig. 1C. The next cluster of carbon sources represents those that were utilized in the second and third time points, but not the first. Compounds in this category were diverse and included TCA cycle intermediates (α-ketoglutarate, succinate, and derivatives thereof), nucleotides (uridine,

inosine, and thymidine), and others (L-alanine, lactate, galactose, trehalose, psicose, etc.) Several intermediates of branched chain amino acid biosynthesis (L-threonine, α-ketobutyrate, α-hydroxybutyrate) were also utilized at the second time point, reflecting the specific induction of this pathway in response to isoleucine starvation. The third cluster of substrates, which was only utilized after 1.5 hours of growth arrest, was comprised of mainly amino acids and their derivatives including L-alaninamide, L-proline, D-alanine, L-asparagine, Glycyl-L-asparate, L-glutamate, Glycyl-L-glutamate, D-serine and also dextrin and glycogen. The induction of pathways indicated by these Biolog assays correlated well with the pattern of metabolic gene induction noted in the transcriptome profiles.

We considered whether the pattern of carbon source utilization in the WT reflected a foraging strategy, the induction of pathways to re-route internal flux, or both. The expanded metabolic capacity of the WT did not resemble Biolog results from carbon starved cells described elsewhere (26). Nor did we observe in our microarray experiments expression patterns of induced transporter genes consistent with known patterns of carbon or nitrogen foraging (13, 21, 33). The transporter expression pattern was not directly predictive of carbon sources readily utilized in the Biolog assays (data not shown). For these reasons, we think the WT Biolog assay results probably do not represent a foraging response alone, but are also indicative of a re-routing of intracellular flux through newly induced pathways. Thus, we interpret the expansion of the WT metabolic capacity as evidence for a large-scale reorientation of metabolism from

anabolism and macromolecular synthesis to reassimilation of carbon back into central metabolism and amino acid biosynthesis pathways. The observed metabolic rearrangement is a ppGpp-dependent process, as described below.

The ppGpp^0 strain was radically impacted, by comparison to the WT, in substrate utilization (Fig. 5). The only substrates readily metabolized by the ppGpp^0 strain were a subset of those found in the first cluster of carbon sources used by the WT, namely, L-serine, glucose, and a few other substrates that are directly assimilated into glycolysis. Furthermore, the number of carbon substrates utilized by the ppGpp^0 strain did not increase in response to growth arrest, rather, the ability of the cells to metabolize those substrates that were used during active growth actually declined. This result suggests that the metabolic capacity of the ppGpp^0 strain diminished in response to isoleucine starvation. The extreme nature of the ppGpp^0 Biolog phenotype prompted us to consider whether or not these results were artifactual. However, we think that the Biolog phenotype observed here is accurate because the transcriptome of the ppGpp^0 strain is consistent with a large-scale metabolic shut-down and the Δ*relA* strain had a nearly identical transcription profile and Biolog phenotype (data not shown). These data demonstrate that the transcriptional impairment observed in the ppGpp^0 strain leads to compromised ability to expand metabolic potential at the proteomic level.

The ppGpp^0 cells are viable and enlarged during isoleucine starvation

One possible explanation for the decreased metabolic activity exhibited by the ppGpp^0 strain in response to isoleucine starvation is simply that the cells had died. Plate counts of isoleucine starved ppGpp^0 strain indicated a ~1 log lower number of colony forming units across the entire time course (even during rapid growth) by comparison to the WT (data not shown). A several-fold lower plating efficiency (64) and filamentous morphology (36, 64) have been demonstrated previously for strains lacking ppGpp. To directly check whether or not the ppGpp^0 strain suffered a decrease in viability due to amino acid starvation, we stained culture samples with a live-dead stain and observed them using confocal microscopy (Fig. 6). This technique stains cells with compromised membrane integrity red (via propidium iodide) while intact (viable) cells are stained green (SYTO-9) (48). Samples were checked across the entire time-course. WT cells were large and rod-shaped during rapid growth and became coccoid as growth ceased (Fig. 6D). No increase in non-viable cells was observed during isoleucine starvation, with non-viable cells making up a negligible portion of the population. The ppGpp^0 cells were observed to be as long as or longer than the WT during rapid growth, with occasional filamentation (Fig. 6E). There was no decline in viability of the ppGpp^0 culture upon isoleucine starvation (< 1% die-off). However, we observed that the ppGpp^0 cells remained large and rod shaped, even 1.5 hours after growth had stopped. Stationary phase ppGpp^0 cells routinely ranged from 4- to 10-fold longer than WT stationary phase cells and in the most extreme cases, filaments up to ~200 μm long were observed (data not

shown). These results point to a larger role for ppGpp in modulating cell division, as has been suggested (45, 58).

The ppGpp^0 strain has altered macromolecular composition.

We also observed that the WT isoleucine-limited cultures routinely reached an OD of ~0.7, whereas the ppGpp^0 strain reproducibly reached a higher OD of ~0.9 (Fig. 6A vs.6B). Theoretically, the total amount of protein produced by both cultures is dictated by the amount of isoleucine included in the growth medium. However, precursors of other major cell components that contribute to biomass might still be formed from glucose, serine, etc., which were not limiting. Keeping in mind these possibilities, we measured total protein, biomass, and RNA production for both the WT and ppGpp^0 strain (Fig. 7). Both strains produced a comparable amount of protein (~150 µg/ml culture by 1.5 hours into stationary phase). In contrast, the ppGpp^0 strain produced an average of ~50% more biomass than the WT under identical conditions. From these findings, we conclude that ppGpp is required to maintain a normal protein:biomass ratio, and that during times of amino acid starvation, ppGpp plays a critical role in keeping the production of macromolecular components in line with the translational capacity of the cell. Since the defining phenotype of relaxed strains is continued stable RNA synthesis in times of starvation, we determined the RNA content of the WT and ppGpp^0 strains (Fig. 7). The WT RNA level did not increase significantly after an OD of 0.3, which correlates with the onset of ppGpp accumulation. As expected, the ppGpp^0 strain did not curtail

RNA synthesis, ultimately producing ~2.5-fold more RNA compared to the WT 1.5 hours after growth arrest. This increase in RNA accounts for ~45% of the mutant's extra biomass. Since only about one-half of the excess biomass produced by the ppGpp^0 strain was RNA, we conclude that ppGpp accumulation in the WT also limits production of other macromolecular cell components, i.e., cell membranes, cell wall, and DNA.

The ppGpp^0 strain shows deviations in gene expression for major physiological processes

Prompted by observations that cells lacking ppGpp are much larger than WT stationary phase cells, and that this likely contributes to a larger biomass yield, we interrogated the microarray data to look for abnormal expression of genes that could account for this phenotype. Accordingly, we considered genes in various functional categories whose difference in expression between the WT and the ppGpp^0 strain was ≥2-fold. These differences are summarized in Table 1. Heat maps and comparative expression values are available for these functional groups in supplementary Fig 1 with an accompanying note.

In general, the data presented in Table 1 show a trend in which genes involved in macromolecular synthesis/biomass production were consistently expressed at higher levels in the ppGpp^0 strain than the WT. This included some 92 genes involved in cell division, DNA replication, and the biosynthesis of nucleotides, fatty acids, cell wall, and LPS/outer membrane. Interestingly, the WT strain did not exhibit increased expression of genes associated with the SOS

response to DNA damage, however, 20 genes involved in DNA repair were expressed higher in the $ppGpp^0$ strain. Taken together with the higher expression of genes involved in DNA replication, this result suggests that chromosome replication continued abnormally in the $ppGpp^0$ strain, ultimately resulting in DNA damage. Another trend evident in Table 1 is that whereas the WT induced genes involved in catabolism of macromolecular precursors, the $ppGpp^0$ strain did not. This included the lower expression in $ppGpp^0$ strain of 11 genes involved in nucleotide catabolism and fatty acid β-oxidation. The $ppGpp^0$ strain also expressed six genes more highly that are involved in the salvage of endogenous nucleotide precursors for the production of new nucleotides. This trend suggests that the $ppGpp^0$ strain actively attempted to salvage nucleotides in keeping with continued synthesis of nucleic acid. Finally, the $ppGpp^0$ strain failed to induce genes involved in glycogen metabolism as observed in the WT, suggesting that glycogen probably does not contribute to the higher biomass produced by the $ppGpp^0$ strain.

Discussion

Summary of results

We sought to examine the extent of the stringent response to amino acid starvation. To do so, we used isoleucine starvation as a model system. Global transcriptome profiling showed that isoleucine starvation of the WT resulted in changes in expression of genes in many metabolic pathways, curtailed

expression of genes involved in macromolecular synthesis, and initiated the general stress response. In stark contrast, the ppGpp^0 strain failed to make these changes. To further examine the transcriptional response to isoleucine starvation, snapshots of the metabolic capacities of the WT and ppGpp^0 mutant were determined using Biolog GN2 microplates in a protocol that prevented changes to the functional metabolic proteome. The results showed that the WT greatly expanded its repertoire of active metabolic pathways, while the ppGpp^0 strain failed to diversify its metabolic capacity and showed diminished ability to utilize pathways that were active before the onset of isoleucine starvation.

Noting that the ppGpp^0 strain showed greatly diminished metabolic activity, we also checked for a possible decrease in viability as a result of amino acid starvation through differential staining/confocal microscopy. Although we noted no decrease in cell viability based on membrane integrity, we did observe that ppGpp^0 mutant cells were considerably longer than WT isoleucine-starved cells at all time points and ppGpp^0 cultures routinely reached a higher density. We found that although the two strains produced similar amounts of protein, the ppGpp^0 strain continued to grow unchecked, producing 50% more biomass than the WT. Nearly one-half of this extra biomass was composed of RNA. Furthermore, we found that this mutant phenotype correlates with aberrant gene expression in diverse cellular processes including cell division, DNA replication, and nucleotide, fatty acid, cell wall, and LPS/outer membrane biosynthesis. The comprehensive deficiencies of the ppGpp^0 strain, as evidenced by the global measurements presented here, imply that ppGpp is the pivotal signal required to

successfully develop virtually all physiological responses to amino acid starvation.

A recent study described the transcriptional changes resulting from serine hydroxamate treatment (18). This report compared the WT transcriptional response to that of a *relA*Δ*251* strain. Though the strains and conditions differ from ours, several overlapping trends are evident. These include down regulation of genes involved in translation and induction of the RpoS-dependent general stress response. However, we note that in these cases the response to isoleucine starvation was more robust, with >40 ribosomal protein genes down regulated compared to only 5 genes down regulated in response to serine hydroxamate treatment. Also, >130 RpoS regulon members were induced in isoleucine starved cells compared to ~20 RpoS regulon members induced by serine hydroxamate. The latter difference may be attributable to inhibited translation caused by serine hydroxamate and therefore decreased accumulation of RpoS. Global defects in regulation were observed for the *relA*Δ*251* strain in response to serine hydroxamate treatment, but the defects were far less comprehensive than those observed here for the ppGpp^0 and Δ*relA* strains starved for isoleucine. Differences in the transcriptomes of isoleucine starved ppGpp^0 cultures and the serine hydroxamate treated *relA*Δ*251* strain likely result from the residual ppGpp present as result of having an intact *spoT* allele in the latter strain, as well as physiological differences between isoleucine starvation and serine hydroxamate induced translation inhibition.

Metabolic restructuring during the stringent response to isoleucine starvation

Often the stringent response to amino acid starvation is thought of as a general reorganization of transcription involving down regulation of stable RNA synthesis and large-scale induction of amino acid biosynthetic operons. While the basic elements of this paradigm are clearly correct, the data presented in this report suggest that the response to starvation for a given amino acid leads to a more complex pattern of transcription than previously thought, especially with respect to metabolic genes. Instead of a general stimulation of amino acid biosynthetic genes, we observed induction of a range of amino acid biosynthetic and catabolic pathways, which, given the composition of the growth medium, would best allow the cell to route metabolic flux into formation of the limiting amino acid. The transcriptome data obtained for the mutant lacking ppGpp showed radically impacted ability to appropriately regulate expression of genes involved in almost all areas of metabolism, from central metabolism to disparate biosynthetic pathways. We therefore suggest that the transcriptional program initiated by ppGpp accumulation is a larger framework within which the pathways specific for the limiting amino acid are readily induced, accompanied by changes in expression of central metabolic genes as necessary to maximize the production of the required precursor metabolites. While it remains to be tested whether alternative metabolic rearrangements occur when cultures are starved for other amino acids, we envision the stringent response to be a global response to nutritional stress that halts growth processes and makes those resources

available for the cell to specifically remediate the offending stress, as described in the model outlined below.

The temporal unfolding of a complex global response likely requires continued feedback between the changing transcriptome and the resulting proteome for its appropriate development. Under the conditions employed here, since isoleucine is not replenished once it is depleted, translation which occurs after the exhaustion of exogenous isoleucine is likely made possible through the recycling of amino acids liberated from turnover of existing cellular proteins. Indeed, significant reorganization of the proteome was found to occur in the WT after the exhaustion of isoleucine, as evidenced by the cells' changing metabolic capacities measured by kinetic Biolog assays. This did not happen in the $ppGpp^0$ strain. In keeping with the constrained expression of metabolic genes observed in our array experiments, the $ppGpp^0$ strain displayed no ability to diversify its repertoire of readily metabolized carbon sources in our Biolog assays, further implicating ppGpp in the large-scale restructuring of metabolism in response to amino acid starvation.

Functional integration of the stringent response

The array data presented here must be considered in the larger context of metabolic and structural trends evident in stationary phase cells. The medium used in our experiments contains both glucose and serine as carbon sources as well as all amino acids necessary for translation. During growth under these conditions, glucose may be used primarily for production of precursor metabolites

that originate in the pentose phosphate pathway and glycolysis, i.e. nucleotides and membrane/cell wall components, while serine, in addition to being integrated into protein and providing one-carbon units for nucleotide biosynthesis, might be used primarily to drive energy generation via acetate overflow metabolism. Together, nucleic acids and membrane/cell wall components comprise about 38% of the dry weight of growing cells, while protein makes up about 50% (39). Upon isoleucine starvation, the WT efficiently stops growing, shuts down ribosome synthesis, and assumes a coccoid morphology. In addition, the microarray data presented here and elsewhere (13) show a concomitant inhibition of DNA replication, and macromolecule biosynthesis, i.e., membrane/cell wall components. The starved cells switch from a primarily anabolic mode devoted to generating building blocks for biomass production, to a catabolic mode which re-assimilates unused nucleotides and fatty acids back into central metabolism (as evidenced by gene expression and Biolog assays). Accordingly, glucose consumption slows as demand for these components is reduced, while serine consumption accelerates to allow for continued energy production via fermentation to acetate. The observed changes in the expression of central metabolic genes may both reflect and effect this metabolic reorientation. Moreover, virtually all of these metabolic changes require ppGpp for their manifestation.

The data presented in this report substantiate a wide range of observations made over many years implicating ppGpp in diverse processes including peptidoglycan synthesis (28), cell division (45, 64), DNA replication (24,

59), fatty acid and phospholipid biosynthesis (23, 44, 49), nucleotide biosynthesis (53), and glycogen metabolism (50). In most cases, the connection of one of these processes to the stringent response was established through the activity of a given enzyme (peptidoglycan and DNA replication), synthesis of a given enzyme (nucleotide and glycogen biosynthesis), or indirectly (cell division). Our transcriptome profiles show that all of these previous observations are parts of larger trends, which are evident at the transcriptional level.

A data-driven model of the constituent processes encompassed by the ppGpp-dependent stringent response is presented in Fig. 8. We assume the changes in gene expression observed here result from two basic signal inputs: i) ribosome stalling which leads to ppGpp accumulation and ii) starvation for isoleucine, which results in allosteric and transcriptional control of branched chain amino acid biosynthetic pathways. In this model, ppGpp accumulation in response to isoleucine starvation leads to stringent down regulation of several processes including repression of DNA replication, ribosome synthesis, nucleotide biosynthesis, phospholipid biosynthesis, cell envelope synthesis, and cell division. Processes requiring ppGpp for their induction include the general stress response, central metabolism, nucleotide catabolism, fatty acid β-oxidation, and amino acid biosynthesis/catabolism genes. Down regulation of macromolecular synthesis, and induction of stationary phase morphogenes, cause stationary phase cells to take on their classical coccoid morphology. Metabolic restructuring leads to a reversal of anabolic flux (i.e., into biomass) to re-assimilation of carbon previously allocated to nucleotide and fatty acid

synthesis. Flux into central metabolism combined with changes in expression of central metabolic genes serves to accommodate the redistribution of metabolites into amino acid biosynthetic pathways. We also note that this extensive metabolic restructuring is accompanied by the induction of the general stress response, which prepares the cells for long-term survival in stationary phase. We expect that other regulators, especially cAMP/Crp and Lrp likely play a role in controlling the transcriptional response to isoleucine starvation. However, since the manner in which isoleucine starvation may trigger signaling through these pathways is unclear, we have not attempted to integrate them into the model presented here.

Global transcription patterns and the regulatory mechanism of ppGpp

With the data presented in this report, we now have transcriptome profiles for cells impacted in their ability to accumulate ppGpp under two very different starvation conditions, i.e., carbon starvation (52) and amino acid starvation (this study). A qualitative comparison reveals two trends present in both data sets: down regulation of translation apparatus genes typical of the stringent response and induction of the RpoS-dependent general stress response. Beyond these shared responses, starvation for carbon or amino acids elicited induction of different condition-specific stimulons (i.e., carbon foraging or isoleucine biosynthesis genes, respectively). Moreover, under both starvation conditions, the induction of the corresponding stimulon was ppGpp-dependent. Based on

these observations, we hypothesize that efficient global induction of all starvation-specific stimulons requires ppGpp.

ppGpp, along with DksA, is thought to exert control over global gene expression via multiple mechanisms. These include alterations in sigma factor-core RNAP interactions (35), direct down regulation (37, 43), passive induction by increased RNAP availability (3), and direct activation (42). Here, we consider each of these mechanisms in light of the genome-wide expression data.

Induction of the general stress response involves ppGpp accumulation in several ways, and the data presented here (i.e., Fig. 2C) are compatible with each of them. First, ppGpp enhances the competitiveness of core RNAP for alternative sigma factors (30), resulting in poor expression of genes controlled by alternative sigma factors in strains lacking ppGpp. Second, expression of the *rpoS* gene was induced 6.9-fold in the WT compared to 2.3-fold in the ppGpp^0 strain (data not shown). Third, *yaiB* (renamed *iraP*) was recently found to encode a protein that stabilizes RpoS via binding to the anti-sigma factor RssB in response to carbon and phosphate starvation (9). Moreover, induction of *iraP* transcription was shown to be ppGpp-dependent (8). Our array data confirm that ppGpp is required for *iraP* induction, i.e., *iraP* was induced 10.7-fold in the WT compared to 2.2-fold in the ppGpp^0 strain (Supplementary Table T1). Thus, poor induction of the RpoS regulon in the ppGpp^0 strain likely results from lower induction of the *rpoS* gene, weak stabilization of RpoS by IraP, and poor competition of core RNAP for RpoS.

Perhaps the best understood mechanism of gene regulation by ppGpp/DksA applies to the direct down regulation of rRNA transcription (43). Extensive work has demonstrated that rRNA promoters form open complexes that are intrinsically unstable, which makes them particularly prone to open complex collapse, as occurs with ppGpp/DksA-bound RNAP (22, 43). Thus, when ppGpp accumulates, there is a rapid and robust cessation of stable RNA synthesis (43). An alternative model for direct down regulation was recently suggested, whereby ppGpp-bound RNAP forms dead-end complexes at stringently-controlled promoters, leading to occluded transcription (37). Since transcription of ribosomal protein genes is regulated directly by rRNA levels, we take the comprehensive down regulation of the translation apparatus as evidence of the cessation of stable RNA synthesis. As expected, this down regulation was ppGpp-dependent (Fig. 2C). The down regulation of translation apparatus genes observed here is compatible with both proposed mechanisms of direct down regulation of rRNA synthesis; we cannot distinguish between the two models based on transcriptional output alone.

One way ppGpp might contribute to induction of the many genes observed here is by mediating the balance between RNAP engaged in transcription of the translation apparatus versus transcription of all other genes (3). According to this passive model, the increased availability of RNAP for transcription of stringently induced promoters results from the liberation of RNAP previously sequestered in stable-RNA synthesis. We previously showed that global expression patterns observed during carbon starvation were consistent with this model (52).

However, the WT and ppGpp^0 strains induced transcription of a similar number of genes in response to isoleucine starvation (Fig. 2B), despite the absence of down regulation of translation apparatus genes in the strain lacking ppGpp. Thus, we hypothesize that the slower onset of growth arrest during isoleucine starvation (about 100 min compared to <10 min for carbon starvation) relaxes the tight relationship between transcription of the translation apparatus and other genes across the genome. While the passive mechanism allows for global transcriptional flexibility during times of sudden starvation, this constraint may be loosened when starvation is encountered over longer time scales (as occurs in complex nutrient mixtures). If the passive mechanism was the driving force behind the WT transcriptional response to isoleucine starvation, the ppGpp^0 strain would exhibit a similar transcriptional response if given enough time, possibly through reduced transcription initiation at rRNA gene promoters due to depletion of transcription initiating NTP pools (43). However, the transcriptional response of the ppGpp^0 strain did not approximate that of the WT (Fig. 2A), suggesting an active role for ppGpp in gene induction in response to isoleucine starvation, as discussed below.

ppGpp, in concert with DksA, is known to be sufficient for direct induction of transcription of several amino acid biosynthetic operons *in vitro*, (42). However, we did not observe stimulation of amino acid biosynthesis genes *en masse* during isoleucine starvation, suggesting that the specific induction of amino acid biosynthesis genes requires ppGpp *and* the action of specific regulators (i.e., Lrp, LeuO, and/or IlvY, etc.). Thus, we favor a synergistic model

that requires ppGpp-bound RNAP acting together with other transcription factors at a wide range of promoters to produce the stressor-specific response observed in the WT. Alternatively, the patterns of transcription observed in the WT may be the product of a regulatory cascade initiated by ppGpp accumulation and accomplished ultimately (or in part) by other transcription factors. These mechanisms are not mutually exclusive, and experiments to delineate the regulatory architecture of the stringent response are ongoing.

Concluding remarks

In this report, we examined the global response to amino acid starvation, including changes in the transcriptome, metabolic proteome, and composition of the growth medium. Our results indicated that the adjustments made by the WT in response to isoleucine starvation entail a large-scale restructuring of cellular metabolism that includes differential expression of genes involved in, and flux through, central metabolism, amino acid catabolism/anabolism, nucleotide biosynthesis/catabolism, and fatty acid metabolism. These responses were totally dependent on the alarmone ppGpp and profoundly absent in the $ppGpp^0$ strain. We observed that cells lacking ppGpp produced significantly more RNA and biomass (but not more protein) than the WT during amino acid starvation. This observation, considered in light of the extremely aberrant patterns of transcription observed in the $ppGpp^0$ strain, highlights the vital role of ppGpp in relaying information about the translational status of the cell not only to genes involved in translation and amino acid biosynthesis, but also to genes involved in

intermediary metabolism and macromolecule synthesis. The comprehensive deficiencies of the ppGpp^0 strain, as evidenced by the global measurements presented here, suggest that ppGpp is the primary signal used by *E. coli* cells to adjust their reproductive potential to that defined by their nutritional environment.

Materials and Methods

Bacterial strains and growth conditions

All strains used in this study were derivatives of *E. coli K-12* strain MG1655. The Δ*relA,* and Δ*relA* Δ*spoT* (ppGpp^0) strains were constructed for this study using a modified version of the method described by Datsenko and Wanner (17). The Δ*relA,* and ppGpp^0 strains were made marker-less by removal of antibiotic cassettes using surrounding FRT sites and confirmed by sequencing and PCR. The WT and isogenic mutants were cultured in a 2-liter Biostat B fermentor (Braun Biotech) containing 1 liter of morpholinepropanesulfonic acid (MOPS) medium (38) with 2.0 g/liter glucose and amino acids at the concentrations described in (60), with the exception that isoleucine was included at 60 μM instead of the usual 400 μM. The growth medium did not contain uracil, which has been shown to stimulate growth of *E. coli* MG1655, which has an *rph* frameshift mutation (29). However, inclusion of uracil had no effect on logarithmic growth, growth arrest caused by isoleucine starvation, or rescue of growth by addition of isoleucine (data not shown). The temperature was maintained at 37°C, and pH was kept constant at 7.4 by the addition of 1 M

NaOH. The dissolved oxygen level was maintained above 40% of saturation by adjusting the agitation speeds in the range of 270–500 rpm with fixed 1.5 liter/min air flow. Growth was monitored as absorbance at 600 nm with a Beckman-Coulter DU 800 spectrophotometer.

Amino acid analysis

Media samples were withdrawn from the bioreactor directly into a 10 ml syringe and then filtered through 0.4 μm filters. Samples were then frozen immediately. Amino acid concentrations in the filtered culture media were determined using capillary electrophoresis-mass spectrometry (CE-MS) with a sheath-fluid electrospray interface. This method was modified from a similar method described previously (62). Briefly, CE-MS was performed using an Agilent G1600 capillary electrophoresis connected to an Agilent 1946A single quadrupole mass spectrometer with an electrospray ionization source using the Agilent capillary electrophoresis spray needle adapter (Palo Alto, CA, USA). Fused-silica capillaries were used for the separation (Polymicro Technologies, Phoenix, AZ, USA). All capillaries used in the CE-MS analysis were 50 mm I.D. and had a total length of 70 cm. Sheath liquid composed of a 1:1 (v/v) mix of 2-propanol and HPLC-grade water with 5 mM formic acid was supplied to the co-axial sheath-fluid interface at a flow rate of 4 μl/min. The nebulizer pressure was maintained at 3 psi. The separation was performed using a +25 kV potential, which produced a current of 47 μA. Samples were injected hydrostatically, at 250 mbar*s (6.1 nl). Mass spectral data was collected in the selected-ion

monitoring (SIM) mode using the [M+H]+ m/z for each amino acid. The drying gas temperature was set to 70°C with a drying gas flow rate of 12.0 l/min. The potential on the MS capillary was maintained at 3800 V and the fragmentation voltage was set to 70 V. The running electrolyte was 1.0 M formic acid prepared in HPLC-grade water with no pH adjustment.

All samples and standards were diluted 1:1 in 10 mM HCl containing 200 μM ethionine (internal standard). The method of internal standards calibration was applied to each amino acid. Standards containing 10, 100, 250, and 1000 μM of each amino acid were analyzed in triplicate. All standards were commercial amino acids (Sigma Chemical, St. Louis, MO, USA). All amino acid calibration curves displayed a linearity of 0.995 or greater.

Nucleotide extraction and ppGpp quantification

Nucleotides were extracted as described, with minor modifications (6). Five ml of culture was sampled directly into a 15 ml falcon tube containing 0.5 ml of 11M formic acid. The sample was vigorously mixed and chilled on ice. One ml aliquots of this mixture were incubated at 0° C on ice for 30 min. These one ml samples were centrifuged at 4° at 10000 RPM for 5 min. The supernatant was then filtered through 0.2 μm filters and stored at -20° until HPLC analysis.

ppGpp was quantified by anion exchange HPLC using a Mono Q 5/50 GL column (GE Healthcare). 250 μl of supernatant was injected under initial conditions of 95% 20mM Tris (pH 8.0) and 5% 20mM Tris + 1.5 M sodium formate (pH 8.0). This initial condition was maintained for 5 min. Absorbance at

260nm was used to detect eluted nucleotides. Over a period of 30 min, the level of sodium formate buffer was ramped up to 60%. ppGpp was identified as a peak which eluted at ~30.9 min. Samples were run in duplicate for three separate WT time course experiments. Representative results are shown in Fig 1A. ppGpp standard was purchased from TriLink Biosciences. Standard curves established that the linear range of detection of ppGpp is 50nM to 100μM.

Microarray analysis

Cells were sampled directly from the fermenter into an equal volume of ice-cold RNA*later* (Ambion) and total RNA was extracted using Qiagen RNeasy Minikits with optional DNAse treatment steps. RNA was checked for integrity by gel electrophoresis and maintained in a 2:1 dilution of EtOH at -80°C until labeling. RNA was converted to cDNA by first strand synthesis using Superscript II (Invitrogen) and random hexamers, according to the manufacturer's specifications. The cDNA was fragmented and biotinylated (Enzo Kit, Roche Diagnostics) according to the Affymetrix prokaryotic labeling protocol. The microarrays used in this study were custom built Affymetrix GeneChips containing probes for several prokaryotic genomes including *E. coli* K12 MG1655, *E. coli* O157:H7 EDL933, *Bacteriodes thetaiotaomicron* VPI-5482*, Enterococcus faecalis* V583*, Salmonella typhimurium* LT2*,* and *Bacillus anthracis*. Biotinylated samples were prepared according to the manufacturer's instructions and hybridized for 16 hours at 60°C. Hybridized arrays were stained using Affymetrix protocol ProkGE_WS2v2_450. Stained microarrays were

scanned and the raw data files (.cel) were further analyzed using RMA processing with quartile normalization (27). All samples were duplicated biologically and technically; r^2 was >0.95 for all WT and $ppGpp^0$ replicates. We considered genes to be significantly induced or repressed if the absolute value of the expression ratio was >2-fold (63). Hierarchical clustering algorithms were implemented in DecisionSite for Functional Genomics (Spotfire). The microarray data were deposited at Array Express (http://www.ebi.ac.uk/miamexpress/), accession # E-MEXP-1370.

Kinetic Biolog assays

Biolog assays were carried out essentially as described (26). Briefly, a volume of cells sufficient to inoculate 2 Biolog plates was harvested from the fermenter and placed in 50 ml falcon tubes containing chloramphenicol (final concentration 25 µg/ml). Samples were kept warm (37°C) through all manipulations. Cells were washed 3 times with 15 ml MOPS medium (without carbon sources) containing chloramphenicol (25 µg/ml), by centrifugation at ~3700 rpm for 8 min. After washing, cells were resuspended in the same medium to an OD 600nm of ~0.3. Biolog GN2 microplates were then loaded (150 µl in each well) and incubated for 24 hours in an Omnilog system. Dye reduction was monitored every 15 minutes. The areas under the resulting curves were then averaged as an overall indicator of substrate utilization. All Biolog measurements reported are the averages of 2 biological, and 4 technical

replicates. Hierarchical clustering algorithms were implemented in DecisionSite (Spotfire).

Confocal microscopy

WT and ppGpp^0 strains were grown under isoleucine starvation conditions as described above. At the indicated time points, 1 ml of culture medium containing living cells was harvested and stained by adding 20 µl of 1.0 mg/ml propidium iodide and 1.2 µl of 5mM SYTO 9. Samples were incubated for 15 minutes in darkness. Five µl of stained sample was placed in a wet mount and visualized using a Olympus FluoView 500 confocal microscope with Blue Argon (488 nm) and Green Helium Neon (543 nm) lasers for excitation of SYTO 9 and propidium iodide fluorophores, respectively. FITC and Propidium Iodide filter settings were used for fluorescence detection.

Biomass, Protein, and RNA quantification

For biomass quantification, cultures were grown as described above. Two 25 ml samples were taken at an OD of ~0.33, three 15 ml samples at the onset of stationary phase (corresponding to time point C in Fig 6 for each strain), and three 15 ml samples were taken 1.5 hours into growth arrest (corresponding to time point E in Fig 6 for each strain). Cells were chilled on ice immediately after sampling and then centrifuged for 20 min at 3700 rpm at 4° C. Supernatant was removed and the remaining biomass was resuspended in 1 ml of deionized water. The samples were transferred to pre-weighed aluminum weigh-boats.

Samples were desiccated at 105° C overnight, and dry weights were obtained. The biomass values shown in Figure 6 are averages of three independent cultures for each of the WT and the ppGpp^0 strains.

For protein quantification, 5 ml of living cells were harvested directly into 15 ml falcon tubes containing chloramphenicol (final concentration was 25 µg/ml) at the same time points described for biomass quantification. One ml of cells was pelleted by centrifugation for 10 min at 14000 rpm and supernatant was removed. Cells were resuspended in 1 ml of sonication buffer (10mM Tris-HCl, 0.1mM EDTA, 5% glycerol, 150mM NaCl, pH 7.65), and frozen at -20° C until further use. Thawed samples were intermittently sonicated and cooled on ice at 5 sec intervals, 3 times. Sonicated samples were pelleted by centrifugation and supernatant containing protein was assayed according to the Bradford method of protein quantification (10). Standard curves for the Bradford assay were done using BSA in sonication buffer. The protein values shown in Figure 6 are averages of three independent cultures for each of the WT and the ppGpp^0 strains. The assay was replicated in quadruplicate at each time point for each culture.

Total RNA quantification was carried out as described previously, with minor modifications (11). Briefly, 15 ml of culture was sampled into a 50 ml falcon tube containing 3 ml of 3.0 M ice-cold trichloroacetic acid. The solution was mixed by inversion and cooled briefly on ice. Three aliquots of the sample were processed separately. Six ml of sample was pipetted onto a glass fiber filter atop a vacuum flask. The filter was then rinsed 3 times with 5 ml volumes of

nanopure water. The dried filter was placed in a screw-top glass vial with 2 ml of 0.2 M NaOH. The vials were incubated horizontally at RT with gentle rocking for 16-18 hours. Two ml of cold, 0.5 M perchloric acid was added to each vial and mixed by inversion. At least one ml of the resulting mixture was filtered through a 0.2 µm filter. The RNA concentration was measured spectrophotometerically on a Beckman-Coulter DU 800 using the 'RNA' setting. The total RNA values reported in Fig. 7 are the averages of 3 replicates from 3 independent WT cultures (9 total replicates for each timepoint) and two independent ppGpp^0 strain cultures (6 total replicates for each timepoint).

Acknowledgements

The authors wish to thank Stafford Marquardt for writing and implementing the Gene Mapper program, which overlaid array data values onto the pathway diagrams in Figs. 3 and 4. Joe Grissom also provided critical bioinformatics support. Sarah Stark helped with total RNA measurement. We thank Dr. Lawrence Reitzer for critical reading of parts of the manuscript. We gratefully acknowledge Dr. Michael Cashel for theoretical advice regarding quantification of ppGpp by anion exchange and Neil Wofford for practical HPLC advice. Finally, we thank Dr. Randall Hewes for expert guidance regarding confocal microscopy. This research was supported by NIH awards RO1AI48945, P20RR016478, U24GM077905.

Figure Legends

Fig. 1. Growth curves and selected medium component concentrations for WT and ppGpp^0 strains.

A. and B. Growth curves and ppGpp accumulation for WT and ppGpp^0 strains, respectively, grown at 37° C in MOPS medium with glucose plus all 20 amino acids, with a limiting amount of isoleucine. ppGpp was undetectable in the mutant strain. Time points marked in green indicate times of sampling for medium component analysis presented in C-F. Black arrows indicate sampling times for microarray analysis. Note the higher OD reached by the ppGpp^0 strain despite being grown in identical medium.

C. and D. Concentrations of carbon sources and acetate, and % O_2 saturation for WT and ppGpp^0 strains, respectively.

E. and F. Concentrations of selected amino acids. Concentrations of amino acids not shown did not change appreciably over the course of the experiments.

Fig. 2. Comparisons of microarray results for WT and ppGpp^0 strains. Control RNA was extracted from log phase WT cells at an OD of ~0.4 grown in medium with replete isoleucine. Test RNA was harvested from WT, Δ*relA* and ppGpp^0 strains ~40 min after onset of growth arrest. All test array data were normalized to control array data before comparative analysis of strain-specific responses to isoleucine starvation. Array data presented in all figures are $\log_2$ expression ratios (test:control).

A. Upper plot: comparison of isoleucine-starved WT and ppGpp^0 strain transcriptome profiles. Lower plot: comparison of isoleucine-starved ΔrelA and ppGpp^0 strain transcriptome profiles.

B. Upper diagram: Venn diagram comparing up regulated genes between the WT and ppGpp^0 strain. Lower diagram: Venn diagram comparing down regulated genes between the WT and ppGpp^0 strain.

C. Heat maps of $\log_2$ expression ratios for the WT and ppGpp^0 strain for ribosomal protein genes, other genes involved in translation, and the general stress response. All genes shown in C differed in their expression >2-fold between the WT and ppGpp^0 strain. Ribosomal protein and translation genes shown were all down regulated >2-fold in the WT, while all the general stress response genes shown were induced >2-fold in the WT.

Fig. 3. WT transcriptome data overlaid on selected metabolic pathways. Genes up regulated >2-fold are shown in red, while genes down regulated >2-fold are shown in green. Genes whose expression did not change >2-fold are shown in black. Where multiple gene products are required for a single conversion, the corresponding arrow is colored according to the average expression value of the corresponding genes. Amino acids are in blue font. Background colors are intended to delineate various pathways as follows: glycolysis: green (upper center), pentose phosphate pathway: dark blue (upper left), TCA cycle: yellow (lower right), threonine biosynthesis: light orange (middle left), branched chain amino acid biosynthesis: purple (lower left), serine metabolism: dark orange

(middle right), acetate metabolism: light blue (middle right), glutamate biosynthesis: red (lower center), arginine degradation: light green (bottom right).

Fig. 4. ppGpp^0 strain transcriptome data overlaid on selected metabolic pathways. Color coding is identical to that for Fig. 3.

Fig. 5. Cluster analysis of carbon sources utilized in kinetic Biolog assays for WT and ppGpp^0 strains. White represents negligible utilization, dark blue represents maximal utilization, and bright red represents half-maximal utilization, as shown by key. Units are arbitrary. Column 1 contains data for both strains at OD ~0.3, during rapid growth before isoleucine starvation. Column 2 contains data for cells harvested at the onset of isoleucine starvation. Column 3 contains data for cells harvested 1.5 hours into growth arrest. These sampling times correspond to time points T2, T3 and T5 in Fig 6A and 6B. Clusters A, B, and C are described in the text.

Fig. 6. Growth curves and viability staining of WT and ppGpp^0 strains.

A. Growth curve of WT for viability staining. Time points where viability was tested are marked in green and labeled T1-T5.

B. Growth curve of ppGpp^0 strain for viability staining. Labels are as in (A).

C. Micrographs of differentially stained WT cells harvested at time points T1-T5 as labeled in (A). Cells stained green are viable, while cells stained red have compromised membranes (non-viable). Gold scale bar = 50 μm.

D. Micrographs of differentially stained ppGpp^0 cells harvested at time points T1-T5 as labeled in (B). Staining and scale bars are identical to (C).

Fig. 7. Total protein, biomass, and RNA produced by WT (black bars) and ppGpp^0 strains (grey bars) during isoleucine starvation. The first time point (rapid growth) corresponds to T2 in (Fig 6A and 6B). Onset of growth arrest corresponds to (T3) in (Fig 6A and 6B). Final time point (1.5 hrs into growth arrest) corresponds to (T5) in (Fig 6A and 6B). Error bars indicate standard deviations. Both strains produce a similar amount of protein, however, the ppGpp^0 strain produces 50% more biomass and >150% more RNA than the WT.

Fig. 8. Physiological model of the ppGpp-mediated response to isoleucine starvation. Solid arrows represent positive regulation while solid lines with flat ends indicate negative regulation. Solid lines ending with both lines and arrows denote complex regulation. Dashed lines indicate functional effects (eg. Nucleotide catabolism generates intermediates which are funneled into central metabolism). Regulatory relationships are not necessarily direct.

Table 1. Number of genes in various functional groups differentially expressed in relA spoT- strain

Functional group	Total number differentially expressed	Genes expressed >2 fold higher	Genes expressed >2 fold lower
Cell Division	15	9	6
DNA replication	16	15	1
DNA repair	23	20	3
Nucleotide Biosynthesis	27	24	3
Nucleotide Catabolism	6	0	6
Nucleotide Salvage	6	6	0
Fatty acid/Phospholipid Biosynthesis	17	12	5
Fatty acid B-oxidation	6	1	5
Peptidoglycan Biosynthesis	13	9	4
LPS/Outer membrane Biosynthesis	24	23	1
Glycogen metabolism	7	0	7

Fig. 1.

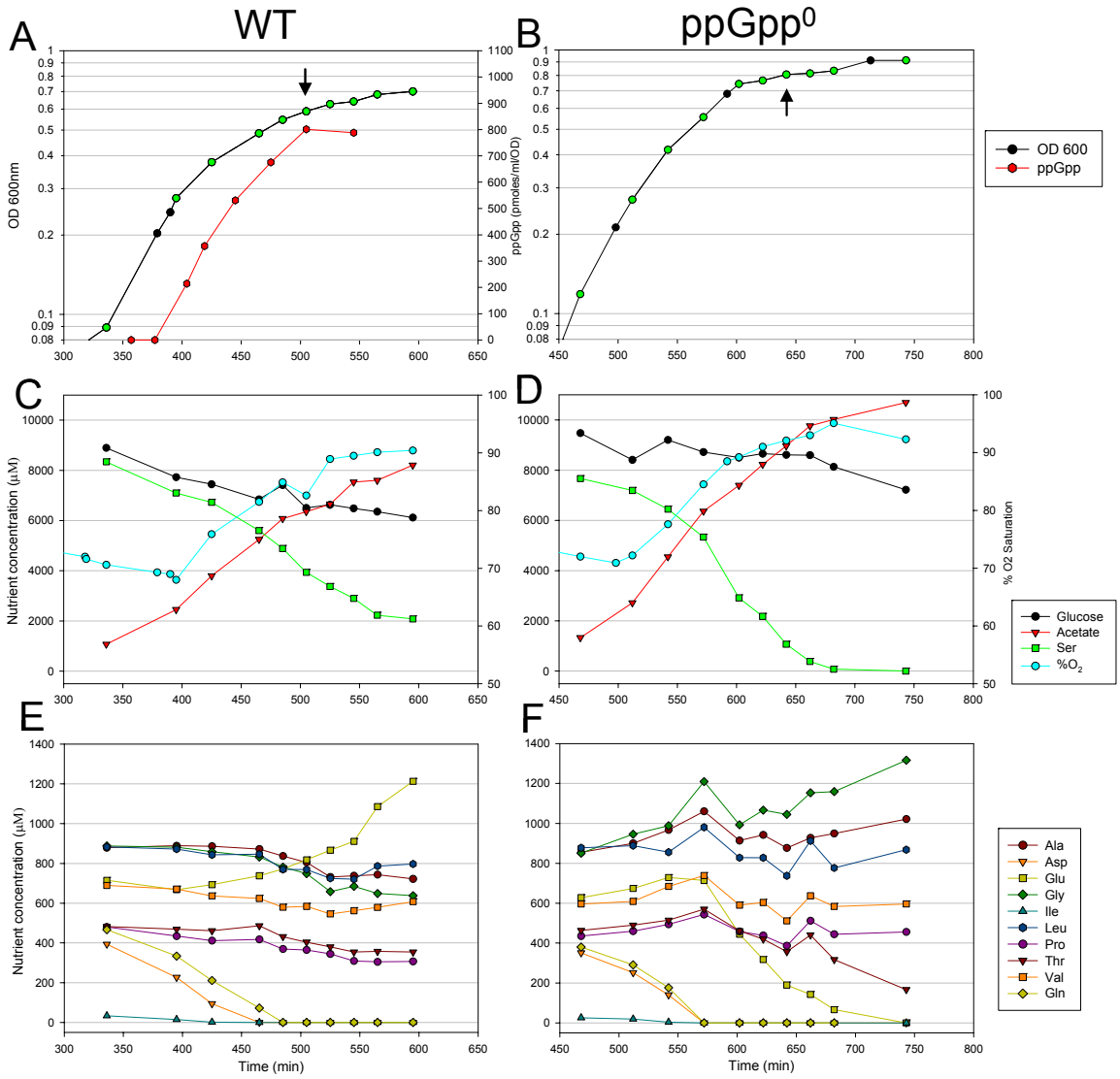

A
WT
B
ppGpp0
OD 600nm
ppGpp (pmoles/mlOD)
OD 600
ppGpp
C
D
Nutrient concentration (µM)
% O2 Saturation
Glucose
Acetate
Ser
%O2
E
F
Nutrient concentration (µM)
Time (min)
Time (min)
Ala
Asp
Glu
Gly
Ile
Leu
Pro
Thr
Val
Gln

Fig. 2.

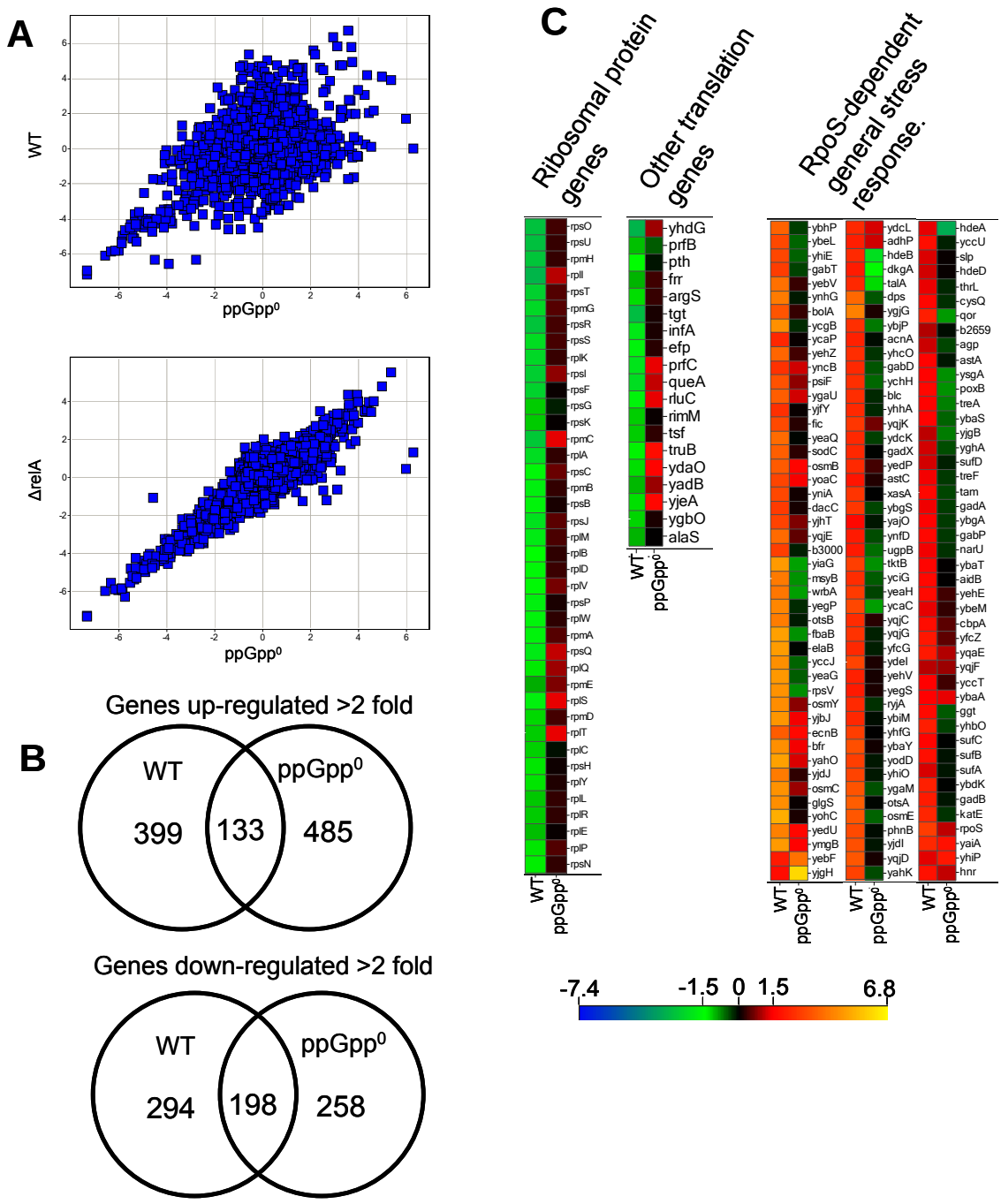

A
WT
ppGpp0
ΔrelA
ppGpp0
B
Genes up-regulated >2 fold
WT
399
133
ppGpp0
485
Genes down-regulated >2 fold
WT
294
198
ppGpp0
258
C
Ribosomal protein genes
Other translation genes
RpoS-dependent general stress response.
WT
ppGpp0
-7.4
-1.5
0
1.5
6.8

Fig. 3.

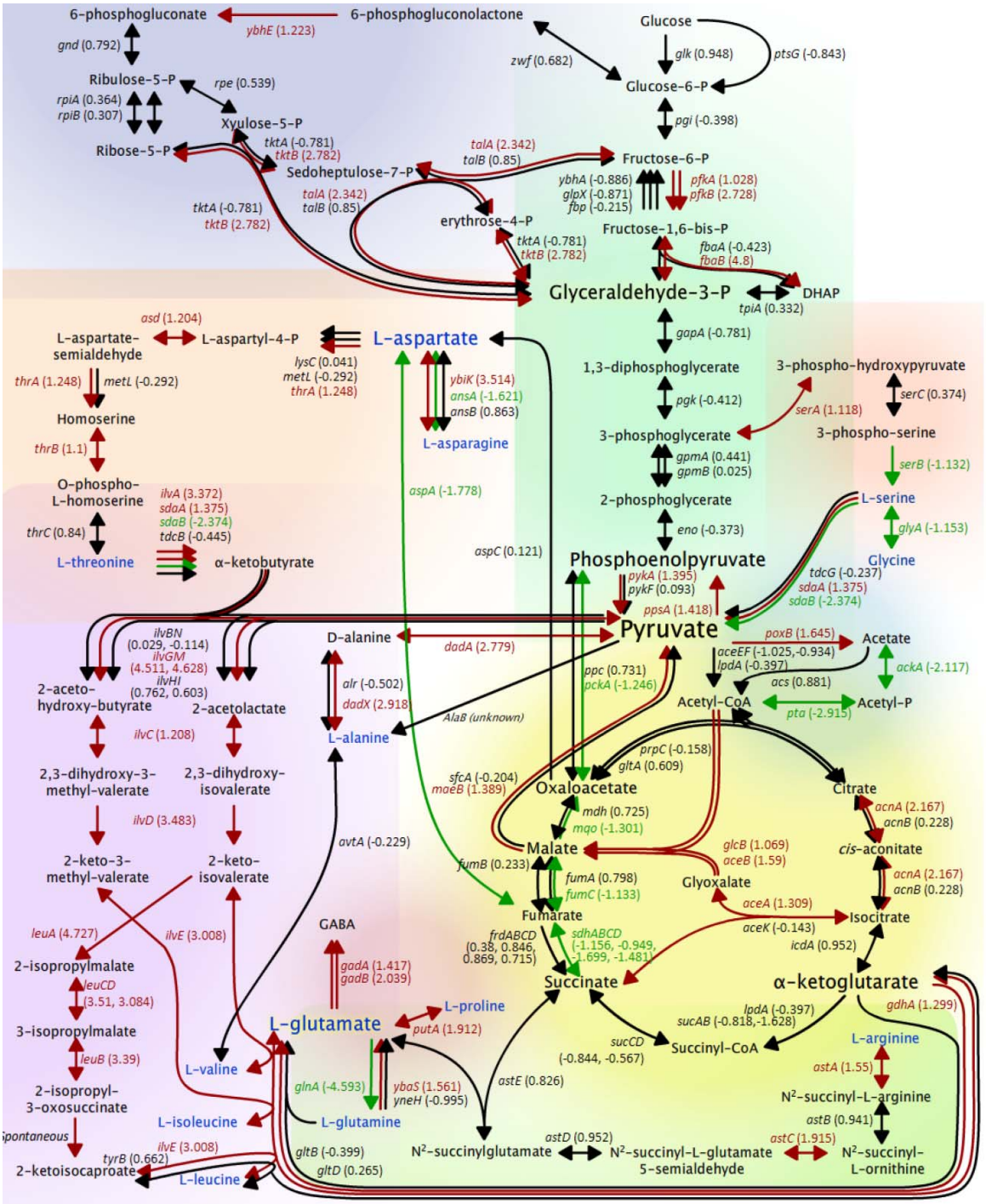

6-phosphogluconate
6-phosphogluconolactone
ybhE (1.223)
gnd (0.792)
Ribulose-5-P
rpe (0.539)
rpiA (0.364)
rpiB (0.307)
Xylulose-5-P
Ribose-5-P
tktA (-0.781)
tktB (2.782)
Sedoheptulose-7-P
talA (2.342)
talB (0.85)
Glucose
zwf (0.682)
glk (0.948)
ptsG (-0.843)
Glucose-6-P
pgi (-0.398)
Fructose-6-P
ybhA (-0.886)
glpX (-0.871)
fbp (-0.215)
pfkA (1.028)
pfkB (2.728)
erythrose-4-P
Fructose-1,6-bis-P
fbaA (-0.423)
fbaB (4.8)
Glyceraldehyde-3-P
DHAP
tpiA (0.332)
gapA (-0.781)
1,3-diphosphoglycerate
pgk (-0.412)
3-phosphoglycerate
gpmA (0.441)
gpmB (0.025)
2-phosphoglycerate
eno (-0.373)
Phosphoenolpyruvate
pykA (1.395)
pykF (0.093)
ppsA (1.418)
Pyruvate
3-phospho-hydroxypyruvate
serC (0.374)
serA (1.118)
3-phospho-serine
serB (-1.132)
L-serine
glyA (-1.153)
Glycine
tdcG (-0.237)
sdaA (1.375)
sdaB (-2.374)
asd (1.204)
L-aspartate-semialdehyde
L-aspartyl-4-P
L-aspartate
lysC (0.041)
metL (-0.292)
thrA (1.248)
ybiK (3.514)
ansA (-1.621)
ansB (0.863)
L-asparagine
aspA (-1.778)
aspC (0.121)
thrA (1.248)
metL (-0.292)
Homoserine
thrB (1.1)
O-phospho-L-homoserine
thrC (0.84)
L-threonine
ilvA (3.372)
sdaA (1.375)
sdaB (-2.374)
tdcB (-0.445)
α-ketobutyrate
ilvBN (0.029, -0.114)
ilvGM (4.511, 4.628)
ilvHI (0.762, 0.603)
2-aceto-hydroxy-butyrate
2-acetolactate
ilvC (1.208)
2,3-dihydroxy-3-methyl-valerate
2,3-dihydroxy-isovalerate
ilvD (3.483)
2-keto-3-methyl-valerate
2-keto-isovalerate
leuA (4.727)
ilvE (3.008)
2-isopropylmalate
leuCD (3.51, 3.084)
3-isopropylmalate
leuB (3.39)
2-isopropyl-3-oxosuccinate
spontaneous
2-ketoisocaproate
tyrB (0.662)
ilvE (3.008)
L-isoleucine
L-valine
L-leucine
avtA (-0.229)
D-alanine
dadA (2.779)
alr (-0.502)
dadX (2.918)
L-alanine
AlaB (unknown)
poxB (1.645)
Acetate
aceEF (-1.025,-0.934)
lpdA (-0.397)
ackA (-2.117)
acs (0.881)
Acetyl-CoA
pta (-2.915)
Acetyl-P
ppc (0.731)
pckA (-1.246)
prpC (-0.158)
gltA (0.609)
sfcA (-0.204)
maeB (1.389)
Oxaloacetate
Citrate
mdh (0.725)
mqo (-1.301)
Malate
acnA (2.167)
acnB (0.228)
cis-aconitate
glcB (1.069)
aceB (1.59)
Glyoxalate
fumB (0.233)
fumA (0.798)
fumC (-1.133)
Fumarate
aceA (1.309)
Isocitrate
aceK (-0.143)
icdA (0.952)
frdABCD (0.38, 0.846, 0.869, 0.715)
sdhABCD (-1.156, -0.949, -1.699, -1.481)
Succinate
α-ketoglutarate
gdhA (1.299)
lpdA (-0.397)
sucAB (-0.818,-1.628)
sucCD (-0.844, -0.567)
Succinyl-CoA
GABA
gadA (1.417)
gadB (2.039)
L-glutamate
L-proline
putA (1.912)
L-arginine
astA (1.55)
glnA (-4.593)
ybaS (1.561)
yneH (-0.995)
L-glutamine
astE (0.826)
N2-succinyl-L-arginine
astB (0.941)
gltB (-0.399)
gltD (0.265)
astD (0.952)
N2-succinylglutamate
N2-succinyl-L-glutamate 5-semialdehyde
astC (1.915)
N2-succinyl-L-ornithine

Fig. 4.

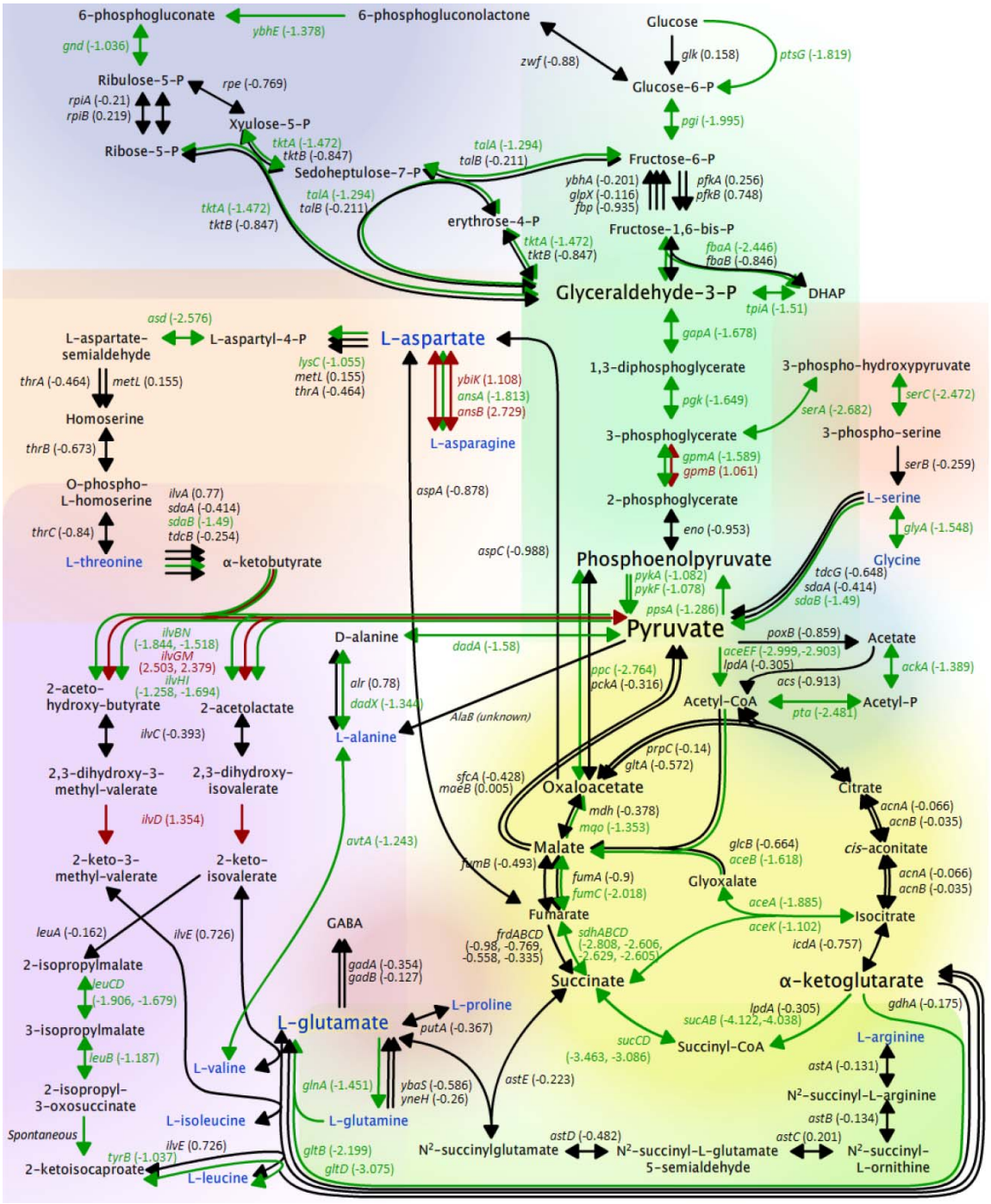

6-phosphogluconate
6-phosphogluconolactone
ybhE (-1.378)
gnd (-1.036)
Glucose
zwf (-0.88)
glk (0.158)
ptsG (-1.819)
Glucose-6-P
Ribulose-5-P
rpe (-0.769)
rpiA (-0.21)
rpiB (0.219)
Xylulose-5-P
pgi (-1.995)
Ribose-5-P
tktA (-1.472)
tktB (-0.847)
Sedoheptulose-7-P
talA (-1.294)
talB (-0.211)
Fructose-6-P
ybhA (-0.201)
glpX (-0.116)
fbp (-0.935)
pfkA (0.256)
pfkB (0.748)
erythrose-4-P
Fructose-1,6-bis-P
fbaA (-2.446)
fbaB (-0.846)
Glyceraldehyde-3-P
DHAP
tpiA (-1.51)
gapA (-1.678)
asd (-2.576)
L-aspartate-semialdehyde
L-aspartyl-4-P
L-aspartate
lysC (-1.055)
metL (0.155)
thrA (-0.464)
ybiK (1.108)
ansA (-1.813)
ansB (2.729)
L-asparagine
1,3-diphosphoglycerate
3-phospho-hydroxypyruvate
serC (-2.472)
serA (-2.682)
pgk (-1.649)
3-phosphoglycerate
3-phospho-serine
serB (-0.259)
gpmA (-1.589)
gpmB (1.061)
Homoserine
thrB (-0.673)
O-phospho-L-homoserine
ilvA (0.77)
sdaA (-0.414)
sdaB (-1.49)
tdcB (-0.254)
aspA (-0.878)
2-phosphoglycerate
L-serine
glyA (-1.548)
thrC (-0.84)
L-threonine
α-ketobutyrate
aspC (-0.988)
eno (-0.953)
Phosphoenolpyruvate
Glycine
tdcG (-0.648)
sdaA (-0.414)
sdaB (-1.49)
pykA (-1.082)
pykF (-1.078)
ppsA (-1.286)
Pyruvate
ilvBN (-1.844, -1.518)
ilvGM (2.503, 2.379)
ilvHI (-1.258, -1.694)
D-alanine
dadA (-1.58)
poxB (-0.859)
Acetate
aceEF (-2.999,-2.903)
lpdA (-0.305)
ackA (-1.389)
2-aceto-hydroxy-butyrate
2-acetolactate
alr (0.78)
dadX (-1.344)
ppc (-2.764)
pckA (-0.316)
acs (-0.913)
Acetyl-CoA
pta (-2.481)
Acetyl-P
AlaB (unknown)
ilvC (-0.393)
L-alanine
2,3-dihydroxy-3-methyl-valerate
2,3-dihydroxy-isovalerate
prpC (-0.14)
gltA (-0.572)
sfcA (-0.428)
maeB (0.005)
Oxaloacetate
Citrate
acnA (-0.066)
acnB (-0.035)
ilvD (1.354)
mdh (-0.378)
mqo (-1.353)
Malate
cis-aconitate
glcB (-0.664)
aceB (-1.618)
avtA (-1.243)
fumB (-0.493)
fumA (-0.9)
fumC (-2.018)
Glyoxalate
aceA (-1.885)
2-keto-3-methyl-valerate
2-keto-isovalerate
Fumarate
aceK (-1.102)
Isocitrate
icdA (-0.757)
leuA (-0.162)
ilvE (0.726)
GABA
frdABCD (-0.98, -0.769, -0.558, -0.335)
sdhABCD (-2.808, -2.606, -2.629, -2.605)
gadA (-0.354)
gadB (-0.127)
Succinate
α-ketoglutarate
2-isopropylmalate
leuCD (-1.906, -1.679)
L-proline
putA (-0.367)
gdhA (-0.175)
3-isopropylmalate
leuB (-1.187)
L-glutamate
lpdA (-0.305)
sucAB (-4.122,-4.038)
sucCD (-3.463, -3.086)
Succinyl-CoA
L-arginine
L-valine
astA (-0.131)
glnA (-1.451)
ybaS (-0.586)
yneH (-0.26)
astE (-0.223)
N²-succinyl-L-arginine
2-isopropyl-3-oxosuccinate
L-isoleucine
L-glutamine
astB (-0.134)
Spontaneous
ilvE (0.726)
astD (-0.482)
astC (0.201)
gltB (-2.199)
gltD (-3.075)
N²-succinylglutamate
N²-succinyl-L-glutamate 5-semialdehyde
N²-succinyl-L-ornithine
tyrB (-1.037)
2-ketoisocaproate
L-leucine

Fig. 5.

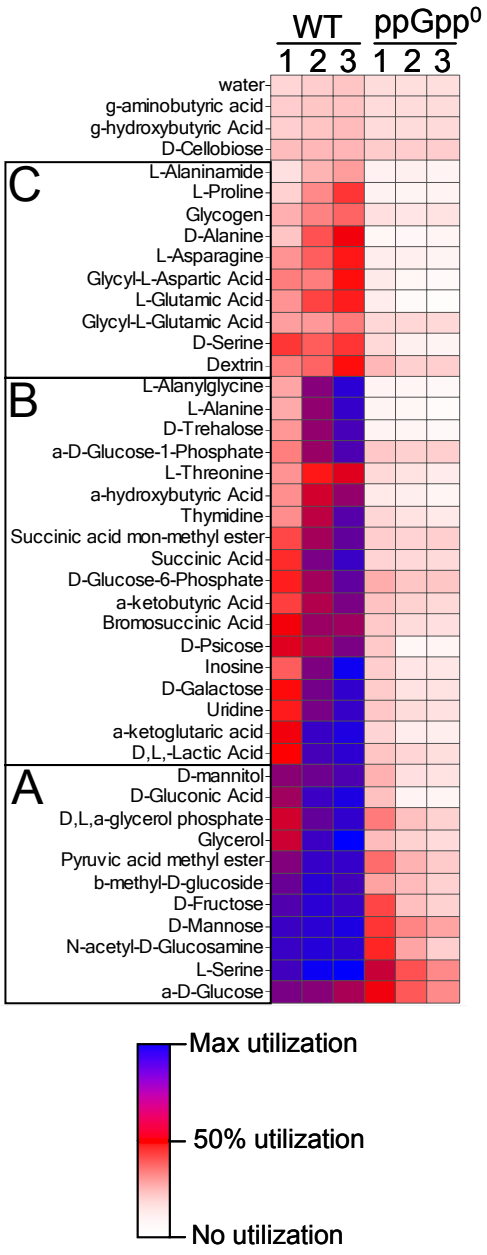

WT
ppGpp0
1 2 3 1 2 3
water
g-aminobutyric acid
g-hydroxybutyric Acid
D-Cellobiose
C
L-Alaninamide
L-Proline
Glycogen
D-Alanine
L-Asparagine
Glycyl-L-Aspartic Acid
L-Glutamic Acid
Glycyl-L-Glutamic Acid
D-Serine
Dextrin
B
L-Alanylglycine
L-Alanine
D-Trehalose
a-D-Glucose-1-Phosphate
L-Threonine
a-hydroxybutyric Acid
Thymidine
Succinic acid mon-methyl ester
Succinic Acid
D-Glucose-6-Phosphate
a-ketobutyric Acid
Bromosuccinic Acid
D-Psicose
Inosine
D-Galactose
Uridine
a-ketoglutaric acid
D,L,-Lactic Acid
A
D-mannitol
D-Gluconic Acid
D,L,a-glycerol phosphate
Glycerol
Pyruvic acid methyl ester
b-methyl-D-glucoside
D-Fructose
D-Mannose
N-acetyl-D-Glucosamine
L-Serine
a-D-Glucose
Max utilization
50% utilization
No utilization

Fig. 6.

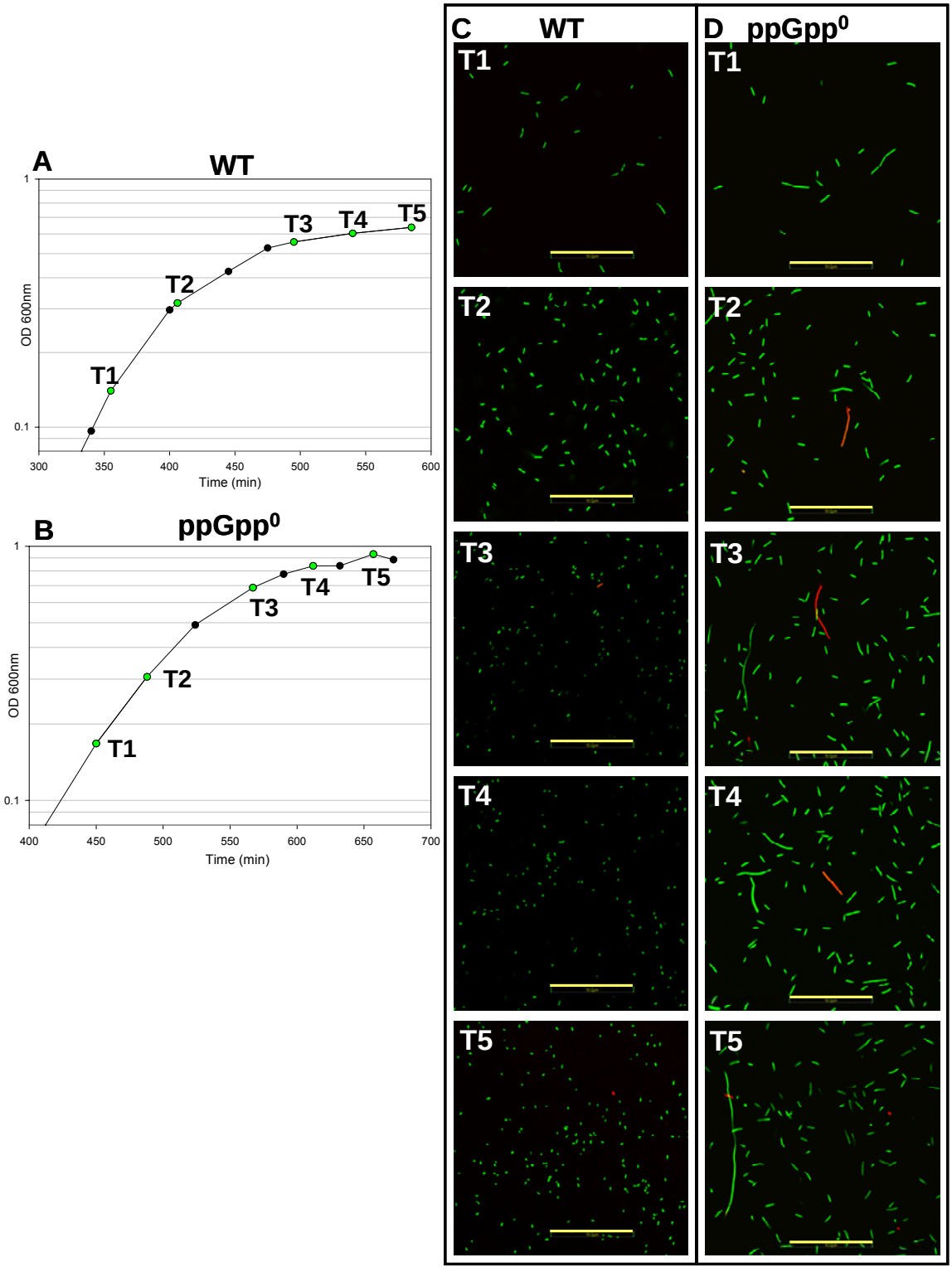

A
WT
OD 600nm
Time (min)
T1
T2
T3
T4
T5
B
ppGpp0
OD 600nm
Time (min)
T1
T2
T3
T4
T5
C
WT
D
ppGpp0
T1
T2
T3
T4
T5

Fig. 7.

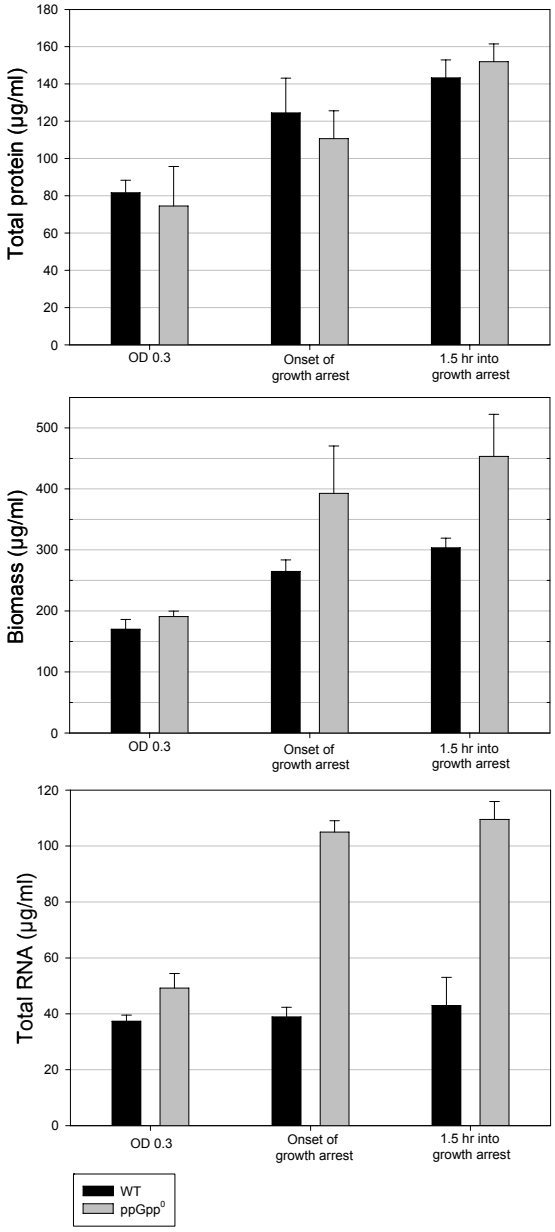
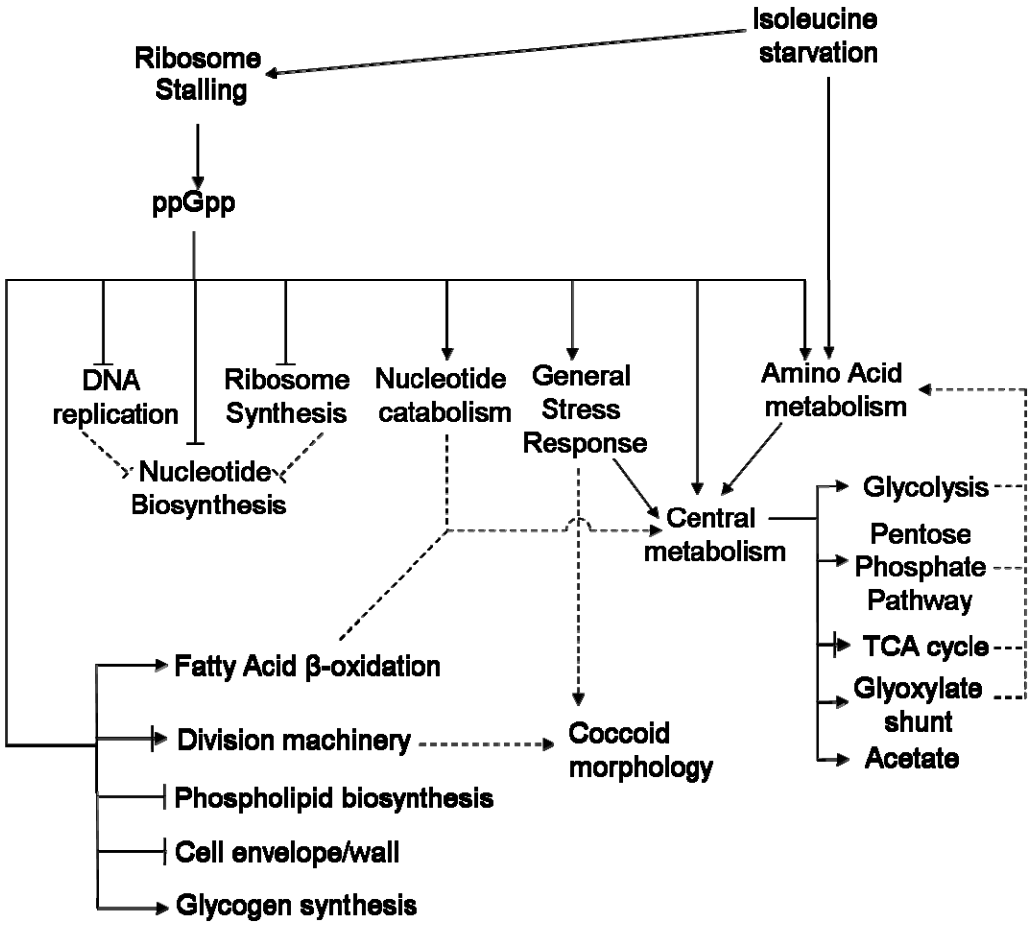

Total protein (µg/ml)
180
160
140
120
100
80
60
40
20
0
OD 0.3
Onset of growth arrest
1.5 hr into growth arrest
Biomass (µg/ml)
500
400
300
200
100
0
OD 0.3
Onset of growth arrest
1.5 hr into growth arrest
Total RNA (µg/ml)
120
100
80
60
40
20
0
OD 0.3
Onset of growth arrest
1.5 hr into growth arrest
WT
ppGpp0

Fig. 8.


Isoleucine starvation
Ribosome Stalling
ppGpp
DNA replication
Ribosome Synthesis
Nucleotide catabolism
General Stress Response
Amino Acid metabolism
Nucleotide Biosynthesis
Central metabolism
Glycolysis
Pentose Phosphate Pathway
TCA cycle
Glyoxylate shunt
Acetate
Fatty Acid β-oxidation
Division machinery
Coccoid morphology
Phospholipid biosynthesis
Cell envelope/wall
Glycogen synthesis

References

1. **Aberg, A., V. Shingler, and C. Balsalobre.** 2006. (p)ppGpp regulates type 1 fimbriation of Escherichia coli by modulating the expression of the site-specific recombinase FimB. Mol Microbiol **60:**1520-33.

2. **Artsimovitch, I., V. Patlan, S. Sekine, M. N. Vassylyeva, T. Hosaka, K. Ochi, S. Yokoyama, and D. G. Vassylyev.** 2004. Structural basis for transcription regulation by alarmone ppGpp. Cell **117:**299-310.

3. **Barker, M. M., T. Gaal, and R. L. Gourse.** 2001. Mechanism of regulation of transcription initiation by ppGpp. II. Models for positive control based on properties of RNAP mutants and competition for RNAP. J Mol Biol **305:**689-702.

4. **Barker, M. M., T. Gaal, C. A. Josaitis, and R. L. Gourse.** 2001. Mechanism of regulation of transcription initiation by ppGpp. I. Effects of ppGpp on transcription initiation in vivo and in vitro. J Mol Biol **305:**673-88.

5. **Battesti, A., and E. Bouveret.** 2006. Acyl carrier protein/SpoT interaction, the switch linking SpoT-dependent stress response to fatty acid metabolism. Mol Microbiol **62:**1048-63.

6. **Bochner, B. R., and B. N. Ames.** 1982. Complete analysis of cellular nucleotides by two-dimensional thin layer chromatography. J Biol Chem **257:**9759-69.

7. **Bogosian, G., B. N. Violand, E. J. Dorward-King, W. E. Workman, P. E. Jung, and J. F. Kane.** 1989. Biosynthesis and incorporation into protein of norleucine by Escherichia coli. J Biol Chem **264:**531-9.

8. **Bougdour, A., and S. Gottesman.** 2007. ppGpp regulation of RpoS degradation via anti-adaptor protein IraP. Proc Natl Acad Sci U S A **104:**12896-901.

9. **Bougdour, A., S. Wickner, and S. Gottesman.** 2006. Modulating RssB activity: IraP, a novel regulator of sigma(S) stability in Escherichia coli. Genes Dev **20:**884-97.

10. **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:**248-54.

11. **Brunschede, H., T. L. Dove, and H. Bremer.** 1977. Establishment of exponential growth after a nutritional shift-up in Escherichia coli B/r:

accumulation of deoxyribonucleic acid, ribonucleic acid, and protein. J Bacteriol **129:**1020-33.

12. **Cashel, M., D. R. Gentry, V. J. Hernandez, and D. Vinella.** 1996. The stringent response, p. 1458-1496. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.

13. **Chang, D. E., D. J. Smalley, and T. Conway.** 2002. Gene expression profiling of *Escherichia coli* growth transitions: an expanded stringent response model. Mol Microbiol **45:**289-306.

14. **Chiaramello, A. E., and J. W. Zyskind.** 1990. Coupling of DNA replication to growth rate in Escherichia coli: a possible role for guanosine tetraphosphate. J Bacteriol **172:**2013-9.

15. **Cohen, G. N., and R. Munier.** 1956. [Incorporation of structural analogues of amino acids in bacterial proteins.]. Biochim Biophys Acta **21:**592-3.

16. **Danchin, A., L. Dondon, and J. Daniel.** 1984. Metabolic alterations mediated by 2-ketobutyrate in Escherichia coli K12. Mol Gen Genet **193:**473-8.

17. **Datsenko, K. A., and B. L. Wanner.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A **97:**6640-5.

18. **Durfee, T., A. M. Hansen, H. Zhi, F. R. Blattner, and D. J. Jin.** 2008. Transcription profiling of the stringent response in Escherichia coli. J Bacteriol **190:**1084-96.

19. **Eichel, J., Y. Y. Chang, D. Riesenberg, and J. E. Cronan, Jr.** 1999. Effect of ppGpp on Escherichia coli cyclopropane fatty acid synthesis is mediated through the RpoS sigma factor (sigmaS). J Bacteriol **181:**572-6.

20. **Gentry, D. R., V. J. Hernandez, L. H. Nguyen, D. B. Jensen, and M. Cashel.** 1993. Synthesis of the stationary-phase sigma factor sigma s is positively regulated by ppGpp. J Bacteriol **175:**7982-9.

21. **Gyaneshwar, P., O. Paliy, J. McAuliffe, D. L. Popham, M. I. Jordan, and S. Kustu.** 2005. Sulfur and nitrogen limitation in Escherichia coli K-12: specific homeostatic responses. J Bacteriol **187:**1074-90.

22. **Haugen, S. P., M. B. Berkmen, W. Ross, T. Gaal, C. Ward, and R. L. Gourse.** 2006. rRNA promoter regulation by nonoptimal binding of sigma

region 1.2: an additional recognition element for RNA polymerase. Cell **125:**1069-82.

23. **Heath, R. J., S. Jackowski, and C. O. Rock.** 1994. Guanosine tetraphosphate inhibition of fatty acid and phospholipid synthesis in Escherichia coli is relieved by overexpression of glycerol-3-phosphate acyltransferase (plsB). J Biol Chem **269:**26584-90.

24. **Hernandez, V. J., and H. Bremer.** 1993. Characterization of RNA and DNA synthesis in Escherichia coli strains devoid of ppGpp. J Biol Chem **268:**10851-62.

25. **Herring, P. A., B. L. McKnight, and J. H. Jackson.** 1995. Channeling behavior and activity models for Escherichia coli K-12 acetohydroxy acid synthases at physiological substrate levels. Biochem Biophys Res Commun **207:**48-54.

26. **Ihssen, J., and T. Egli.** 2005. Global physiological analysis of carbon- and energy-limited growing Escherichia coli confirms a high degree of catabolic flexibility and preparedness for mixed substrate utilization. Environ Microbiol **7:**1568-81.

27. **Irizarry, R. A., B. Hobbs, F. Collin, Y. D. Beazer-Barclay, K. J. Antonellis, U. Scherf, and T. P. Speed.** 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics **4:**249-64.

28. **Ishiguro, E. E., and W. D. Ramey.** 1978. Involvement of the relA gene product and feedback inhibition in the regulation of DUP-N-acetylmuramyl-peptide synthesis in Escherichia coli. J Bacteriol **135:**766-74.

29. **Jensen, K. F.** 1993. The Escherichia coli K-12 "wild types" W3110 and MG1655 have an rph frameshift mutation that leads to pyrimidine starvation due to low pyrE expression levels. J Bacteriol **175:**3401-7.

30. **Jishage, M., K. Kvint, V. Shingler, and T. Nystrom.** 2002. Regulation of sigma factor competition by the alarmone ppGpp. Genes Dev **16:**1260-70.

31. **Johansson, J., C. Balsalobre, S. Y. Wang, J. Urbonaviciene, D. J. Jin, B. Sonden, and B. E. Uhlin.** 2000. Nucleoid proteins stimulate stringently controlled bacterial promoters: a link between the cAMP-CRP and the (p)ppGpp regulons in Escherichia coli. Cell **102:**475-85.

32. **Leavitt, R. I., and H. E. Umbarger.** 1962. Isoleucine and valine metabolism in Escherichia coli. XI. Valine inhibition of the growth of Escherichia coli strain K-12. J Bacteriol **83:**624-30.

33. **Liu, M., T. Durfee, J. E. Cabrera, K. Zhao, D. J. Jin, and F. R. Blattner.** 2005. Global transcriptional programs reveal a carbon source foraging strategy by Escherichia coli. J Biol Chem **280:**15921-7.

34. **Maaloe, O., and N. Kjeldgaard.** 1966. Control of Macromolecular Synthesis, vol. W. A. Benjamin, inc., New York.

35. **Magnusson, L. U., A. Farewell, and T. Nystrom.** 2005. ppGpp: a global regulator in Escherichia coli. Trends Microbiol **13:**236-42.

36. **Magnusson, L. U., B. Gummesson, P. Joksimovic, A. Farewell, and T. Nystrom.** 2007. Identical, independent, and opposing roles of ppGpp and DksA in Escherichia coli. J Bacteriol **189:**5193-202.

37. **Maitra, A., I. Shulgina, and V. J. Hernandez.** 2005. Conversion of Active Promoter-RNA Polymerase Complexes into Inactive Promoter Bound Complexes in E. coli by the Transcription Effector, ppGpp. Mol Cell **17:**817-29.

38. **Neidhardt, F. C., P. L. Bloch, and D. F. Smith.** 1974. Culture medium for enterobacteria. J Bacteriol **119:**736-47.

39. **Neidhardt, F. C., J. L. Ingraham, and M. Schaechter.** 1990. Physiology of the Bacterial Cell, 1st ed, vol. Sinauer Associates, Inc., Sunderland, MA.

40. **Oh, M. K., L. Rohlin, K. C. Kao, and J. C. Liao.** 2002. Global expression profiling of acetate-grown Escherichia coli. J Biol Chem **277:**13175-83.

41. **Patte, J.-C.** 1996. Biosynthesis of Threonine and Lysine. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology., 2nd ed. ASM Press, Washington, D.C.

42. **Paul, B. J., M. B. Berkmen, and R. L. Gourse.** 2005. DksA potentiates direct activation of amino acid promoters by ppGpp. Proc Natl Acad Sci U S A **102:**7823-8.

43. **Paul, B. J., W. Ross, T. Gaal, and R. L. Gourse.** 2004. rRNA transcription in *Escherichia coli*. Annu Rev Genet **38:**749-70.

44. **Podkovyrov, S. M., and T. J. Larson.** 1996. Identification of promoter and stringent regulation of transcription of the fabH, fabD and fabG genes encoding fatty acid biosynthetic enzymes of Escherichia coli. Nucleic Acids Res **24:**1747-52.

45. **Schreiber, G., E. Z. Ron, and G. Glaser.** 1995. ppGpp-mediated regulation of DNA replication and cell division in Escherichia coli. Curr Microbiol **30:**27-32.

46. **Stauffer, G. V., and J. E. Brenchley.** 1974. Evidence for the involvement of serine transhydroxymethylase in serine and glycine interconversions in Salmonella typhimurium. Genetics **77:**185-98.

47. **Stent, G. S., and S. Brenner.** 1961. A genetic locus for the regulation of ribonucleic acid synthesis. Proc Natl Acad Sci U S A **47:**2005-14.

48. **Stocks, S. M.** 2004. Mechanism and use of the commercially available viability stain, BacLight. Cytometry A **61:**189-95.

49. **Taguchi, M., K. Izui, and H. Katsuki.** 1980. Augmentation of cyclopropane fatty acid synthesis under stringent control in Escherichia coli. J Biochem (Tokyo) **88:**1879-82.

50. **Taguchi, M., K. Izui, and H. Katsuki.** 1980. Augmentation of glycogen synthesis under stringent control in Escherichia coli. J Biochem (Tokyo) **88:**379-87.

51. **Tosa, T., and L. I. Pizer.** 1971. Biochemical bases for the antimetabolite action of L-serine hydroxamate. J Bacteriol **106:**972-82.

52. **Traxler, M. F., D. E. Chang, and T. Conway.** 2006. Guanosine 3',5'-bispyrophosphate coordinates global gene expression during glucose-lactose diauxie in Escherichia coli. Proc Natl Acad Sci U S A **103:**2374-9.

53. **Turnbough, C. L., Jr.** 1983. Regulation of Escherichia coli aspartate transcarbamylase synthesis by guanosine tetraphosphate and pyrimidine ribonucleoside triphosphates. J Bacteriol **153:**998-1007.

54. **Umbarger, H. E.** 1996. Biosynthesis of the Branched-Chain Amino Acids. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology., 2nd ed. ASM Press, Washington, D.C.

55. **Uzan, M., and A. Danchin.** 1978. Correlation between the serine sensitivity and the derepressibility of the ilv genes in Escherichia coli relA-mutants. Mol Gen Genet **165:**21-30.

56. **VanBogelen, R. A., P. M. Kelley, and F. C. Neidhardt.** 1987. Differential induction of heat shock, SOS, and oxidation stress regulons and accumulation of nucleotides in *Escherichia coli*. J Bacteriol **169:**26-32.

57. **Vinella, D., C. Albrecht, M. Cashel, and R. D'Ari.** 2005. Iron limitation induces SpoT-dependent accumulation of ppGpp in Escherichia coli. Mol Microbiol **56:**958-70.

58. **Vinella, D., and R. D'Ari.** 1995. Overview of controls in the Escherichia coli cell cycle. Bioessays **17:**527-36.

59. **Wang, J. D., G. M. Sanders, and A. D. Grossman.** 2007. Nutritional control of elongation of DNA replication by (p)ppGpp. Cell **128:**865-75.

60. **Wanner, B. L., R. Kodaira, and F. C. Neidhart.** 1977. Physiological regulation of a decontrolled lac operon. J Bacteriol **130:**212-22.

61. **Wendrich, T. M., G. Blaha, D. N. Wilson, M. A. Marahiel, and K. H. Nierhaus.** 2002. Dissection of the mechanism for the stringent factor RelA. Mol Cell **10:**779-88.

62. **Williams, B. J., C. J. Cameron, R. Workman, C. D. Broeckling, L. W. Sumner, and J. T. Smith.** 2007. Amino acid profiling in plant cell cultures: an inter-laboratory comparison of CE-MS and GC-MS. Electrophoresis **28:**1371-9.

63. **Wren, J. D., and T. Conway.** 2006. Meta-analysis of published transcriptional and translational fold changes reveals a preference for low-fold inductions. Omics **10:**15-27.

64. **Xiao, H., M. Kalman, K. Ikehara, S. Zemel, G. Glaser, and M. Cashel.** 1991. Residual guanosine 3',5'-bispyrophosphate synthetic activity of *relA* null mutants can be eliminated by *spoT* null mutations. J Biol Chem **266:**5980-90.

65. **Yang, C. R., B. E. Shapiro, S. P. Hung, E. D. Mjolsness, and G. W. Hatfield.** 2005. A mathematical model for the branched chain amino acid biosynthetic pathways of Escherichia coli K12. J Biol Chem **280:**11224-32.

Chapter 4: Discretely calibrated regulatory loops controlled by ppGpp partition global gene expression across the ‘feast to famine’ gradient

Abstract

Bacteria profoundly reorganize their global gene expression when faced with nutrient exhaustion. In *Escherichia coli* and almost all other free-living bacteria, the alarmone ppGpp facilitates this massive response by directly or indirectly coordinating the down-regulation of genes of the translation apparatus, and the induction of amino acid biosynthetic genes and the general stress response. Such a large reorientation likely requires the cooperative activities of many different genetic regulators, yet the structure of the transcription network that functions below ppGpp remains poorly defined. Using isoleucine starvation as an experimental model system for amino acid starvation, we identified genes that required ppGpp, Lrp, and RpoS for their induction. Surprisingly, despite the fact that the overwhelming majority of genes controlled by Lrp and RpoS required ppGpp for their activation, we found that these two regulons were not induced simultaneously. The data reported here suggest that metabolic genes, such as those of the Lrp regulon, require only a low basal level of ppGpp for their efficient induction. In contrast, the RpoS-dependent general stress response is not robustly induced until relatively high levels of ppGpp accumulate. We propose a data-driven model that explains how bacterial cells allocate transcriptional resources between metabolic and stress survival processes by discretely tuning regulatory activities to a central indicator of cellular physiology. The regulatory structure that emerges is consistent with a rheostatic model of the stringent response

that allows cells to efficiently adapt to a wide range of nutritional environments.

Introduction

Across a continuum of possible nutritional environments ranging from 'feast' to 'famine,' *Escherichia coli* cells attempt to compensate for environmental deficiencies by activating endogenous biosynthetic pathways (12). Should conditions deteriorate to the point that these biosynthetic pathways can no longer remedy the situation, the cells transition into stationary phase, a physiological state oriented toward protection of cellular structures and long-term survival (21). Thus, as the quality of the environment diminishes, the cells must properly allocate resources between biosynthetic and stress/survival functions. The structure of the transcription network used by cells to balance these processes across this gradient remains incompletely understood.

In almost all bacteria, growth arrest prompts a restructuring of global transcription patterns known as the stringent response (15, 45). The alarmone ppGpp is the arbiter of the stringent response and lies at the apex of the network that governs global gene expression in response to nutrient limitation in *E. coli* (15). This assertion rests on the observation that cells lacking ppGpp exhibit profoundly altered global gene expression patterns during carbon and amino acid starvation (19, 52, 53). While these studies suggest that ppGpp controls one of the largest transcription networks in the bacterial cell, lingering questions regarding the mechanisms underlying regulation by ppGpp and the sheer size of the ppGpp regulon have led to the omission of ppGpp from all large-scale, computational transcriptional network analyses to date. Most experiments

designed to examine the stringent response/stationary phase physiology utilize conditions of 'feast or famine.' However, the ppGpp level is inversely proportional to the steady-state growth rate, i.e., a condition which supports a sub-optimal growth rate leads to an elevated, but not maximal level of ppGpp (32, 48-50). This suggests that the stringent response is rheostatic rather than 'all or nothing.' As ppGpp level serves as a general indicator of the nutritional state of the cell, it is a logical signal to which other regulators might key their activity. Indeed, global expression profiles from cells during the stringent response suggest that a multitude of other regulators are involved, depending on the type and severity of stress encountered (19, 52, 53). In our previous investigations designed to examine the physiological extent of the ppGpp-mediated stringent response, we observed that many genes known to be regulated by the alternative sigma factor RpoS and the DNA binding-protein Lrp were induced (53). Thus, as a step to understanding the larger architecture of the stringent response, here we examine the Lrp and RpoS networks as components of the ppGpp regulon, and consider how these two networks might influence each other as well.

The primary synthase of ppGpp is RelA, which catalyzes the production of ppGpp in response to amino acid starvation (15, 59). A secondary ppGpp synthase, SpoT, produces ppGpp in response to diverse stresses including carbon (61), iron (55), and fatty acid starvation (8). SpoT also contains ppGpp hydrolase activity, and thus plays a crucial role in regulating the overall level of ppGpp (49, 61). ppGpp binds directly to RNA polymerase (RNAP) with the help

of the RNAP-binding protein DksA (3, 20). ppGpp and DksA compromise the ability of RNAP to form a productive open complex at intrinsically unstable promoters (e.g. ribosomal RNA promoters), via a mechanism that is not yet completely understood (20, 47, 56). Conversely, ppGpp and DksA have been shown to directly stimulate transcription of amino acid biosynthesis genes, which have a much longer open complex half-life. Thus, when ppGpp accumulates, RNAP is liberated from rRNA promoters and becomes available for transcription of diverse promoters across the genome (5). At the level of transcription, we and others have found the stringent response to include the down regulation of diverse types of macromolecular synthesis (Protein, DNA, RNA, fatty acids, etc.), a broad-scale restructuring of intermediary metabolism (including amino biosynthesis), and the induction of the RpoS-mediated general stress response (1, 19, 53).

Lrp (leucine responsive protein) is a global transcription factor known to regulate a large number of genes involved in amino acid biosynthesis, uptake, and degradation (16, 51). Current structural models suggest that Lrp binds DNA as an octomer, wrapping the DNA around itself (18), however, the molecular mechanism underlying the regulatory activity of Lrp is not completely understood (13). A portion of the Lrp regulon displays induction or down-regulation in response to the availability of exogenous leucine. However, other genes are regulated by Lrp independently of leucine availability (14, 33, 42). A recent systems-level of analysis of the Lrp network suggested that the different outputs regulated by Lrp encompass several coherent physiological states that balance

between amino acid uptake, degradation and biosynthesis (16). Thus, to the extent that Lrp regulates induction of amino acid biosynthesis/metabolism genes, its role is complementary to that of ppGpp during the stringent response to amino acid starvation.

RpoS is the mediator of the general stress response in *E. coli* (35). It is known to control >100 genes in response to diverse conditions including starvation, oxidative, and osmotic stresses (28, 58). During times of low stress (e.g. nutrient excess), RpoS levels are kept low by submaximal transcription of *rpoS* and by high levels of degradation [reviewed in (22)]. Interplay between a suite of proteins (RssB, IraP, IraD, IraM, etc) stabilizes the level of RpoS according to the physiological status of the cell by promoting or inhibiting the degradation RpoS by ClpXP protease (10, 11, 39). The RpoS-mediated general stress response requires ppGpp for its development (27). The connections between ppGpp and RpoS are several-fold: i) ppGpp is required for increased transcription of *rpoS* during entry into stationary phase (30), ii) ppGpp facilitates competition of alternative sigma factors (including RpoS) with the housekeeping sigma (σ^{70}) for core RNAP binding (26), and iii) ppGpp is required for increased transcription of the anti-adaptor protein gene *iraP*, which results in inhibited proteolysis of RpoS (11).

The observation that genes of both the Lrp and RpoS regulons (which control nutritional and stress responses, respectively) appeared to require ppGpp for their induction prompted us to consider how these networks/processes were integrated as components of the stringent response. We used whole-genome

microarrays to experimentally determine the extent and timing of gene expression controlled by ppGpp, Lrp, and RpoS. Our results suggest a model in which genes of the Lrp regulon require a lower threshold level of ppGpp for their induction, while RpoS-dependent genes require a relatively high level of ppGpp for their induction. Further experimentation correlating regulon-representative promoter activity and ppGpp accumulation patterns validated several aspects of our proposed model. We interpret these results to mean that at a systems level, nutritional genes can readily respond when ppGpp is at basal or only slightly elevated levels (e.g. sub-optimal conditions which still allow for growth) while the general stress response is only fully developed under sustained growth limiting conditions (which lead to high levels of ppGpp accumulation). Our model illustrates how a single signal molecule, ppGpp, can drive an independent feed-back loop coupled in parallel with a feed-forward loop to appropriately balance basic biosynthetic and survival processes across the feast to famine gradient.

Results

Overall strategy

To investigate the regulatory architecture of the stringent response, we developed an experimental system based on starvation for isoleucine which could be equally applied to all strains used here, including the multi-auxotrophic ppGpp^0 strain. This system has been described in detail elsewhere (53). Briefly, *E. coli* K-12 strains cannot grow in the absence of isoleucine when valine is

present due to frame-shift mutation in the *ilvG* gene which renders the encoded protein (valine insensitive acetohydroxy-acid synthase II) inactive (31). Thus, isoleucine depletion in the presence of excess valine serves as an experimental model system for amino acid starvation and the stringent response.

Relative contributions of Lrp and RpoS in response to isoleucine starvation

An analysis of transcription profiles of WT and ppGpp^0 strains obtained during isoleucine starvation suggested that the regulators Lrp and RpoS were also activated as many genes known to be regulated by these proteins were induced compared to WT under non-growth limiting conditions (53). To elucidate the role of Lrp and RpoS during the stringent response, we grew strains lacking these regulators in isoleucine-limiting medium (Fig. 1). This MOPS medium includes glucose (0.2%) as a carbon source and all twenty amino acids in the amounts described in (57), with the exception that isoleucine was only present at a starting concentration of 60 μM. Under these conditions, the cultures exited logarithmic growth around OD 0.3, ultimately achieving a final OD of 0.6-0.7. RNA was harvested after the cells had transitioned into stationary phase, an OD of ~0.6 (arrows, Fig. 1). The transcriptomes of the mutant strains were analyzed using Affymetrix Genechips.

For our analysis, we considered expression profiles from all four strains: WT, ppGpp^0, Lrp-, and RpoS-. Transcriptomes from each of these four strains starved for isoleucine were compared to the transcriptome of the WT strain during logarithmic growth in identical medium, except that isoleucine was replete.

Using a simple 2-fold criterion (60), we categorized genes induced in the WT in response to isoleucine starvation according to dependence on ppGpp, Lrp, or RpoS for their induction (Fig. 2). This analysis leads to several important conclusions: i) the majority of genes induced in response to isoleucine starvation require ppGpp for their normal induction (365 required it, vs. 153 that did not), ii) Lrp controls a smaller subset of genes than RpoS (39 controlled by Lrp vs 133 controlled by RpoS, 11 were controlled by both), and iii) the great majority of genes in both the Lrp and RpoS regulons also require ppGpp for their full induction (153 required ppGpp, 8 did not). These results suggest a regulatory scheme in which ppGpp serves as the apex regulator, with Lrp and RpoS functioning to control discrete subsets of the larger ppGpp regulon.

Physiological roles of Lrp and RpoS

Our data suggest clear physiological roles for Lrp and RpoS based on the content of their respective regulons, and these data correlate well with other studies which have examined global regulation by these regulators (16, 51, 58). Under isoleucine starvation, Lrp was required for induction of a range of metabolic genes (Fig. 2B). This included genes involved amino acid metabolism such as leucine biosynthesis (*leu* genes), threonine biosynthesis (*thrA* and *thrB*), alanine metabolism (*dadAX*), and serine metabolism (*serA*). Lrp was also required for induction of the glyoxylate shunt genes *aceA* and *aceB*, as well as the malic enzyme *maeB*. Taken together, these data suggest that Lrp, along with

ppGpp, plays a role in coordinating the metabolic response to isoleucine starvation.

RpoS was required for the induction of a range of genes with various roles in preparing for long-term survival in stationary phase (Fig. 2E). These included systems for protecting the cell against oxidative stress (*sodC*, *dps*, *wrbA*), osmotic stress (several *osm* genes, *treA*, *otsA*, *otsB*), and metabolic genes known to be RpoS-dependent (*poxB*, *talA*, and *fbaB*). Thus the RpoS-dependent general stress response is initiated in response to isoleucine starvation. The Lrp and RpoS regulons are known to overlap. Eleven genes required both Lrp and Rpos for their induction (Fig. 2C). This group was mostly composed of genes involved in acid tolerance (*gadA*, *gadB*, *hdeA*, *hdeB*, *hdeD*, *xasA*, *gabP*). From these results, we conclude that the Lrp and RpoS networks serve specific functions within the larger stringent response to isoleucine starvation: Lrp works in conjunction with ppGpp to induce genes involved in amino acid metabolism (with the presumed function of initiating amino acid biosynthesis) while RpoS, together with ppGpp, ready the cell for survival under prolonged starvation conditions.

One possible way that ppGpp might influence the Lrp and RpoS regulons is by directly regulating the expression of the *lrp* and *rpoS* genes. There are conflicting reports regarding the role of ppGpp in affecting transcription of *lrp*. ppGpp was required for a ~5 fold induction of *lrp* upon transition to stationary phase in a rich defined medium (29). However, *lrp* transcription was found to unaffected in both stringent and relaxed strains in response to arginine starvation

(these experiments were done in an arginine auxotrophic genetic background) (4). Under isoleucine starvation conditions tested here, the WT did not significantly induce *lrp* expression and the difference in *lrp* expression between the WT and ppGpp^0 strain was ~1.8 fold (Fig. 2D). Thus, under these conditions, we conclude that the failure of the ppGpp strain to induce the Lrp regulon was not likely due to impaired transcription of the *lrp* gene. In contrast, while both the WT and ppGpp^0 strain induced *rpoS* >2-fold, *rpoS* transcription was much lower in the ppGpp^0 strain (2.2 induced fold vs. 6.7 fold in the WT). Thus, impaired induction of the RpoS regulon in the ppGpp^0 strain can be accounted for, at least in part, by poor induction of *rpoS* at the transcriptional level.

Differential induction times of the Lrp and RpoS regulons

Having delineated the Lrp and RpoS regulons induced during isoleucine starvation, we sought to determine the timing of these components of the stringent response. To address this question, we conducted a twelve-point microarray time-course of WT cells starting during logarithmic growth (OD 0.14) and subsequently as they starved for isoleucine. We then examined the behavior of the Lrp and RpoS regulons in the time-course data set (Fig. 3). Several trends were readily apparent when the regulon data was overlaid onto to the timecourse data. First, many genes of the Lrp regulon, including those involved in amino acid biosynthesis/metabolism were induced very early on, before the culture reached an OD of 0.3. This finding was somewhat unexpected given that the cells were still in logarithmic growth phase at this time. In contrast,

we found that the RpoS regulon was induced later than the Lrp regulon. To quantify these potential differences in timing, we undertook a systems-level analysis.

The response time of a gene (defined as the time at which expression reaches its half-maximal level) offers a quantitative measure of when a gene can be said to be induced (2, 38). To compare the relative response times of the Lrp and RpoS regulons, we averaged the expression levels of all the genes in each regulon at each time point. We then plotted a sigmoid regression curve for each set of averaged regulon values vs. time (Fig. 4). We found that the response time for the Lrp regulon was ~45 min before the RpoS regulon. Genes that required both Lrp and RpoS for their induction had an intermediate induction time, ~30 min after the Lrp regulon, and ~15 min before the RpoS regulon. Taken together these data suggest a regulatory and physiological hierarchy under the control of ppGpp. Specifically, the initial reaction to isoleucine starvation entails the induction of genes involved in amino acid biosynthesis mediated by Lrp; this induction occurs before the cells begin to transition into stationary phase. In contrast, the RpoS-mediated general stress response is not induced until later, as the growth rate slows.

These observations lead to a central question: How can a single signal molecule, ppGpp (which is required for expression of both the Lrp and RpoS regulons), differentially dictate the output timing of these two genetic networks? We suggest that this question can be most simply and productively considered in terms of ppGpp level and regulon output. We observed previously that ppGpp

accumulated over ~100 min in response to isoleucine starvation (53), a relatively long time compared to most experimental conditions that provoke robust ppGpp accumulation, usually within 10-15-min (32, 48, 54). The slow ppGpp accumulation associated with isoleucine starvation suggests that the cells experienced a gradual decline in environmental quality, a situation that might approximate a continuum from feast to famine. We considered the accumulation of ppGpp relative to the induction times of the Lrp and RpoS regulons (Fig. 4), and noted that the induction time of the Lrp regulon corresponded to a lower level of ppGpp (< 100 pmol/ml/OD), while the RpoS regulon was induced when the level of ppGpp was higher (~400 pmol/ml/OD). This correlation suggests a model in which the Lrp and RpoS regulons require different amounts of ppGpp for their induction (Fig. 5).

Altering the ppGpp accumulation rate predictably impacts the response times of the Lrp and RpoS regulons

We sought a way to test if the Lrp and RpoS regulons required different threshold levels for their induction. A key prediction that follows from this model is that if the two regulons require different ppGpp levels for their induction, then altering the rate of ppGpp accumulation should impact the relative response times of the two regulons. For example, we observed that when ppGpp accumulated over ~100 min, the difference in response times between the Lrp and RpoS regulons was ~45 min. We reasoned that if we could accelerate the

rate of ppGpp accumulation, then the difference in the response times between the Lrp and RpoS networks would be reduced.

To test this hypothesis, we needed an experimental system that would allow for relatively easy measurement of key parameters (ie gene induction and ppGpp accumulation) and would allow for rapid culture manipulation. To monitor gene induction of the Lrp and RpoS regulons, we chose two representative genes from each whose promoter activities could serve as behavioral indicators for the larger regulons. For the Lrp regulon we chose the *leuL* and *dadA* promoters; genes under control of both of these promoters were influenced by *lrp* mutation in our data sets, and are previously known to be controlled by Lrp (33, 65). To represent the RpoS regulon, we selected the promoters for the *yahO* and *wrbA* genes. Both of these genes responded robustly and were RpoS-dependent in our array experiments, and both are previously known to be controlled by RpoS (23, 28, 62). These four promoter regions were cloned into pUA66, a very low copy plasmid that has been used extensively for similar genetic analyses, upstream of a fast-folding GFP allele (46, 63, 64). Thus, we monitored GFP fluorescence as an indicator of the respective promoter activities (and Lrp and RpoS regulon induction). This setup allowed us to grow the four reporter strains in parallel in 50 ml cultures that facilitated experimental manipulation.

We first confirmed that the Lrp regulon was induced before the RpoS regulon in our four-flask, GFP system (Fig. 6A). We found that the Lrp-dependent promoters were induced an average of 26 min before the RpoS-

dependent promoters. While this 26-min difference is shorter than the 45 min whole-regulon measurements in our array data, it represents a statistically robust difference in timing between the two regulons ($p < 0.001$). In our normal isoleucine starvation regime, the cells deplete isoleucine from the medium until growth can no longer be supported. To abruptly trigger isoleucine starvation (and hence accelerate ppGpp accumulation), we grew the four reporter strains in medium with glucose and all twenty amino acids in excess. When the cultures reached an OD of ~0.3-0.4 we rapidly collected the cells on glass fiber filters and immediately resuspended them in medium containing glucose, 19 amino acids, and no isoleucine. Under these conditions we found that ppGpp accumulates to ~800 pmol/ml/OD in ~10 min (Fig. 6B). Under these abrupt isoleucine starvation conditions the average difference in the response times between Lrp-dependent and RpoS-dependent genes was reduced to a statistically insignificant 6.6 minutes ($p = 0.224$). Thus, one prediction of our proposed model appears to hold true: response times of Lrp and RpoS-dependent gene expression vary in accordance with the ppGpp accumulation rate.

Integration of the Lrp and RpoS networks within the stringent response

In considering other predictions implied by our proposed model (Fig. 5), we noted other features which may serve to partition expression of the Lrp and RpoS regulons under a range of physiological conditions. The Lrp and RpoS networks triggered by ppGpp constitute different types of physiological/regulatory motifs. The Lrp side of the model contains a physiological feedback loop. Under

amino acid limiting conditions, the series of events in this feedback loop begins with ppGpp accumulation. A relatively low level of ppGpp and Lrp (activated by an unknown mechanism) together allow for induction of the Lrp regulon. The physiological result of Lrp regulon induction is biosynthesis of the missing amino acids. Newly synthesized amino acids in turn modulate the ppGpp level. The RpoS side of the model contains a coherent feed-forward loop in which increased ppGpp prompts RpoS induction (at the transcriptional level) and stabilization (via IraP). RpoS and ppGpp then work together (via enhanced alternative sigma factor competition) to induce transcription of the RpoS regulon. A critical element of this model is that since the Lrp regulon might indirectly influence ppGpp levels, Lrp activity has the potential to indirectly modulate the induction of the RpoS regulon. This important connection within the stringent response offers a directly testable prediction which we sought to investigate experimentally.

If our model is correct, then Lrp activity should play a role in setting the level of ppGpp under growth conditions that require de novo synthesis of certain amino acids. However, Lrp is a pleitropic regulator and efforts to study it have long been frustrated by the lack of a straightforward phenotype (41). For example, one of the few known growth phenotypes of Lrp- mutants is marginally slower growth in glucose minimal medium (41, 51). Thus, while Lrp is clearly implicated in regulation of many amino acid biosynthetic genes, Δ*lrp* strains are not strict amino acid auxotrophs. These considerations led us to design an experimental regime in which starvation for all amino acids could be abruptly triggered. The experimental system we devised begins with cells in balanced

growth with all amino acids in abundance. Rapid collection and resuspension of the cells in glucose minimal medium disturbs the homeostasis previously maintained by the cells (i.e. the needs met by the originally available exogenous amino acids are suddenly unmet). The onset of amino acid starvation would be expected to stimulate robust, transient ppGpp accumulation. Proper induction of the Lrp regulon (along with other metabolic pathways) should eventually lead to endogenous biosynthesis of amino acids. In this framework, Lrp allows the cell to efficiently achieve a new homeostatic balance, resulting in the lowering of the ppGpp level and the resumption of growth. Accordingly, the RpoS regulon, which would be induced as ppGpp reaches high levels, should also return to a low level of expression as growth resumes. The model also suggests that under these amino acid starvation conditions, a strain lacking Lrp would also rapidly accumulate ppGpp. However, Δ*lrp* cells would be impacted in their ability to biosynthesize some amino acids and would thus continue to maintain a high level of ppGpp.

To examine the role of Lrp in influencing the ppGpp level in response to abrupt amino acid starvation, we grew 50 ml cultures in minimal glucose medium containing double the concentration of 18 amino acids (tyrosine and isoleucine were also included at normal levels). At an OD of ~0.3-0.4 we rapidly collected the cells on glass fiber filters and resuspended them in glucose minimal medium (Fig. 7). In the WT, this treatment triggered a diauxie of ~160 min. In contrast, growth was arrested in the Lrp mutant (outgrowth occurred only after many hours, i.e. overnight). Under these conditions, downshift resulted in a very high

level of ppGpp accumulation in both strains, peaking at 30 min post-filtration at ~2,400 pmol/ml/OD. This high level presumably reflects starvation for all 20 amino acids, as opposed to a single amino acid in our previously described experiments (Fig 4 and 6). In the WT, growth resumed as the ppGpp dropped to a level below ~400, ultimately falling to a level below 100 pmol/ml/OD. In contrast, the Lrp- mutant ppGpp level declined only to ~700 pmol/ml/OD; a level comparable to that reached by cells starved for isoleucine (Fig 4). This level was maintained for at least 270 min after amino acid starvation was induced. Thus, in accordance with the proposed model, we observe that under some conditions which lead to amino acid starvation, Lrp activity can influence the ppGpp level. Moreover, the proposed model not only validated the role of Lrp in modulating ppGpp level, but also led to discovery of a new phenotype for Lrp- cells.

Discussion

In this report, we examined the Lrp and RpoS transcription networks as components of the stringent response. We experimentally determined the Lrp and RpoS regulons induced in response to isoleucine starvation and found broad regulatory overlap between these gene systems and the ppGpp regulon. Next we elucidated the timing of induction of the Lrp and RpoS networks in a microarray timecourse obtained as cells starved for isoleucine. We found that the Lrp regulon was induced before the RpoS regulon. The response times of the two networks corresponded to different threshold levels of ppGpp

accumulation, suggesting a model in which the Lrp regulon requires a relatively low level of ppGpp for its induction compared to the RpoS regulon. Based on this model, we predicted that accelerating the ppGpp accumulation rate would reduce the difference in the response times of these two regulons. We devised an experimental system based on GFP production from Lrp and RpoS dependent promoters which allowed for ready testing of this hypothesis. We found that when ppGpp accumulates rapidly (in <15 min as opposed to 100 min), the difference in response times of the two regulons was irresolvable. Finally, our proposed model also postulates a feedback loop wherein Lrp activity indirectly regulates the ppGpp level by controlling expression of amino acid biosynthesis genes. In support of this hypothesis, we observed that an Lrp mutant maintained an abnormally high level of ppGpp after a harsh downshift from medium replete with amino acids to glucose minimal medium. Taken together, these data suggest a framework for understanding the regulatory structure governing gene induction during the stringent response.

As cells encounter declining environmental quality they must judiciously allocate cellular resources. The model presented in Fig. 5 considers how cells might partition their gene expression (and hence manage major cellular processes) under three conditions along the feast to famine gradient. In rich environments the ppGpp level is low and both biosynthetic (e.g. amino acid synthetic pathways) and stress survival mechanisms are kept at a minimal level of activity. Under intermediate, or 'hunger' conditions (such as those found in minimal medium) a slightly elevated level of ppGpp is maintained. The data

presented here suggests that biosynthetic processes require only a relatively low level of ppGpp for their induction; with this sensitivity being theoretically mediated through promoter structure and/or signaling through other regulatory proteins (e.g. Lrp). Under these conditions, biosynthesis of amino acids sets ppGpp at a relatively low level, a new homeostatic balance is achieved, and growth can continue. Should a critical nutrient become limiting, starvation occurs, resulting in high levels of ppGpp. Only when ppGpp reaches a high level is the general stress response initiated and the cell transitions into survival mode. Thus, biosynthetic and stress responses are tuned to require different levels of ppGpp, which in turn signify the overall physiological state of the cell.

What threshold levels are required for different components of the stringent response to function? This question must be considered in light of ppGpp levels associated with different physiological conditions. In quickly growing cells (~30 min doubling time, with glucose and casamino acids) the basal level of ppGpp has been measured at ~20-40 pmols/OD (32, 48, 50). This basal level of ppGpp ranges up to ~80 pmols/OD for slower growing cells (~200 min/doubling, with alanine as the sole carbon/energy source) (32, 48). In contrast, conditions which lead to growth arrest often prompt ppGpp levels in excess of ~800 pmols/OD [(32, 48, 54), and this report]. Our measurements for ppGpp under steady state growth conditions are in broad agreement with these levels (WT growth in minimal medium, post-diauxie Fig. 7 and data not shown), however, our HPLC method is not sensitive enough to reliably distinguish small differences in ppGpp associated with different rates of steady state growth. The

overarching conclusion drawn from these measurements is that the range of basal ppGpp levels found during growth, even slow growth, is small compared to ppGpp levels found during growth arrest. This distinction is important as it implies that even the relatively low level of ppGpp found during slow growth (eg. <80 pmols/OD) is sufficient to allow for induction of biosynthetic genes (which must be induced in order to grow in minimal medium). We do not contend that the observed ppGpp levels at the times of Lrp and RpoS regulon induction represent the exact ppGpp levels required for induction of the respective regulons. Rather we suggest that low basal levels of ppGpp present during growth (20-80 pmols/OD) are sufficient to allow for induction of the Lrp regulon. Accordingly, we observed that the response time of the Lrp regulon occurred while the cells were still actively growing, not during transition to stationary phase (Fig. 4). Thus, we expect that the timing of induction of the Lrp regulon observed here is determined primarily by signaling through Lrp.

From our measurements we conclude that ppGpp levels of > ~400 pmols/ml/OD correspond to growth arrest. Specifically, this is based on our observations that cells starving for isoleucine exhibit little growth after this level of ppGpp reached (Fig. 4), and that growth resumed after ppGpp dropped below this level post-diauxie (Fig. 7). This level of ppGpp also correlated with the response time of the RpoS regulon (Figs. 4 and 6). Thus, the 'relatively high' level of ppGpp required for full RpoS regulon induction is likely ~400 pmol/ml/OD. The experiments described here do not imply a mechanism for how the requirement for a higher amount of ppGpp might be built into the RpoS signaling

pathway. We hypothesize that the manifold levels of control exerted by ppGpp on RpoS, including inhibition of proteolysis via IraP (11), and/or at the level of sigma factor competition (26), may play a role in defining the threshold level of ppGpp required for RpoS activity (considered in detail below). Since ppGpp is also implicated in control of RpoS at a transcriptional level (30), the ppGpp/RpoS signaling pathway can be considered a coherent feed-forward loop (ie. ppGpp is required for *rpoS* transcription and for induction of the RpoS regulon). Such feed-forward loops can introduce a time delay between stimulus and response and insulate systems from short-lived spikes in stimulus (noise) (38).

Our observation that an Lrp mutant had higher levels of ppGpp following an amino acid downshift illustrates another critical connection within the architecture of the stringent response (Fig. 7). This finding implies that ppGpp accumulation is subject to feedback control resulting from the activity of biosynthetic enzymes via the endogenous amino acid pool. In this scenario, when exogenous amino acids first become limiting, ppGpp accumulates and works in conjunction with other regulators such as Lrp to activate transcription of biosynthetic genes. The resulting increase in endogenous amino acids lowers the ppGpp level until growth can resume. Since ppGpp level also influences induction of the RpoS regulon, the feedback loop proposed above describes a mechanism for homeostatic (metabolic) control of the general stress response.

Several advantages of the proposed regulatory network structure are apparent. First, the low level of ppGpp required for induction of biosynthetic genes allows for rapid and flexible response to metabolic perturbations. Second,

the high threshold level of ppGpp required for induction of the general stress response could work to buffer against erroneous activation of a large number of stationary phase genes in situations of transient or easily remediated nutritional stresses. Thus, only cells experiencing acute interruption of metabolic homeostasis develop a full-fledged general stress response. The discrete tuning of each of the proposed regulatory loops allows for multiple response systems to be tethered to a single indicator of cellular physiology. Moreover, as the threshold ppGpp level of the general response is set appropriately high, a range of signaling through RpoS is likely possible across levels of ppGpp that still allow for active growth. Indeed such a range of induction of the RpoS regulon is observed in the excellent published transcriptome data sets of cells growing at several different rates (34). The evidence presented in this report supplies a rationale for how bacterial cells utilize ppGpp to partition global gene expression for rapid growth, biosynthetic processes, and stress response across environments ranging from feast to famine. These findings also underscore the idea that the stringent response is not an all-or-nothing phenomenon, but a rheostatic system that can be dialed up or down depending on the richness of the nutritional environment.

How does the proposed regulatory framework function in light of the demonstrated mechanisms of ppGpp action? We enter this discussion with several key assumptions. Based on the data presented here, and others' observations (32, 48, 50), rheostat-like function of the stringent response occurs with ppGpp concentrations between 20-400 pmol/ml/OD, with doubling times

>200 min occurring at [ppGpp] >80 pmol/ml/OD and little or no growth occurring at ppGpp concentrations >400 pmol/ml/OD. If it is assumed that there is a single ppGpp binding site on RNAP, then, at the level of RNAP interaction, the ppGpp signal is binary (either ppGpp is bound or not). We expect that rheostatic function corresponds to varying the fraction of the total RNAP pool that is bound by ppGpp. In scenarios that lead to ppGpp accumulation of >400 pmol/ml/OD, RNAP is probably saturated with ppGpp. We consider how passive induction, direct induction, and sigma factor competition might each function within the proposed architecture of the stringent response.

The passive model of induction suggests that ppGpp is required for liberation of RNAP from transcription of stable RNAs; only RNAP freed from stable RNA synthesis is available for transcription of other mRNAs [for reviews see (37, 44, 45)]. An attractive feature of this model is that the relative proportions of these two populations of RNAP could be easily controlled by raising or lowering the ppGpp concentration. ppGpp binding to RNAP is sufficient to mediate direct induction of certain promoters, including some promoters for amino acid biosynthesis (6, 43). Both passive and direct control by ppGpp is thought to result from altered RNAP-promoter interactions, with the overall effect being dictated by the promoter nucleotide sequence (20, 44). Thus, under conditions which allow for growth, a low level of ppGpp is present and a correspondingly small pool of free RNAP is available for transcription of condition-specific mRNAs. In this scenario, we suggest that little or no increase in ppGpp is required to allow for induction of the small set of promoter that need

activation when a single amino acid is exhausted from the environment. Thus, the cumulative effect of both the passive and direct mechanisms might render amino acid biosynthesis gene promoters exquisitely sensitive to small increases (or simply low levels) of ppGpp. The timing of this induction would then be controlled by signaling through accessory regulators such as Lrp.

The data presented here suggest that the RpoS regulon is not maximally induced until a relatively high concentration of ppGpp accumulates. Competition for RNAP core enzyme (E) between the housekeeping sigma factor (σ^{70}) and alternative sigma factors, such as RpoS, has been shown to be influenced by ppGpp (26, 27, 37). Specifically, ppGpp-bound RNAP is more readily bound by alternative sigma factors (26). ppGpp is proposed to divide the RNAP population into two populations (bound and unbound) in the models described above. The sigma competition model suggests that only the ppGpp-bound RNAP population would be accessible to RpoS, as dictated by the affinity of RpoS for E. Indeed, even in competition with σ^{70} for ppGpp-bound RNAP, RpoS would be expected to successfully engage only some fraction of ppGpp-bound E. Thus, at low levels of ppGpp, the amount of E bound by RpoS would be expected to be extremely small (approaching zero). Superimposed upon this framework is the fact that RpoS levels are also subject to control by ppGpp [through transcriptional control and indirect control over RpoS proteolysis via IraP (11, 30)]. We suggest that the dual control of RpoS level and sigma factor competition by ppGpp may play a role in setting the high threshold of ppGpp required for robust induction of the general stress response.

Transcription network analyses have begun to illuminate how the topography large genetic networks can change in response to environmental cues (7, 16, 36). Systems analyses have shown that transcription factors are organized into appropriately structured regulatory motifs (2). The challenge going forward is to integrate the two into a framework that allows for ready testing of model predictions, and to likewise use data generated by high-throughput studies to inform model construction. The data presented in this report outline basic parameters that will hopefully allow for integration of ppGpp-mediated transcription control into *in silico* models of bacterial transcription networks. Moreover, the data-driven models we describe for partitioning gene expression across the continuum of the stringent response suggested hypotheses which we were able to test experimentally. At a fundamental level, the results illustrate how bacterial cells can utilize a single indicator of cellular physiological state in combination with discretely calibrated regulatory systems to: a) establish a new homeostatic balance, or b) protect the cell in the event that homeostasis cannot be maintained (e.g. starvation).

Materials and Methods

Bacterial strains and growth conditions

All strains used in this study were derivatives of *E. coli K-12* strain MG1655. A list of strains and plasmids used appears in Table 1. All mutant strains were constructed for this study using a modified version of the method described by

Datsenko and Wanner (17). The Δ*relA* Δ*spoT* (ppGpp^0) and Δ*rpoS* strains have been used in previous studies (52, 53). Marker-less mutants were made by removal of antibiotic cassettes using surrounding FRT sites and confirmed by sequencing and PCR.

For array studies the WT and isogenic mutants were cultured in a 2-liter Biostat B fermentor (Braun Biotech) containing 1 liter of morpholinepropanesulfonic acid (MOPS) medium (40) with 2.0 g/liter glucose and amino acids at the concentrations described in (57), with the exception that isoleucine was included at 60 μM instead of the usual 400 μM. The growth medium did not contain uracil, which has been shown to stimulate growth of *E. coli* MG1655, which has an *rph* frameshift mutation (25). However, inclusion of uracil had no effect on logarithmic growth, growth arrest caused by isoleucine starvation, or rescue of growth by addition of isoleucine (data not shown). The temperature was maintained at 37°C, and pH was kept constant at 7.4 by the addition of 1 M NaOH. The dissolved oxygen level was maintained above 40% of saturation by adjusting the agitation speeds in the range of 270–500 rpm with fixed 1.5 liter/min air flow. Growth was monitored as absorbance at 600 nm with a Beckman-Coulter DU 800 spectrophotometer.

For GFP isoleucine starvation experiments, strains were grown in 50 ml cultures in 500 ml flasks in medium as described above. Flasks were incubated shaking at 250 rpm at 37° C. Kanamycin was included at 25 μg/ml to maintain pUA66 derivatives. Fifty ml cultures were inoculated to a calculated starting OD of 0.0015 from overnight 5-ml seed cultures (which were not isoleucine limited).

To induce abrupt isoleucine starvation, 50 ml cultures were started in medium replete with isoleucine (400 μm). During log growth (OD 0.3-0.4), cells were rapidly harvested on a glass fiber filter atop a vacuum tower. Filters were immediately dropped into identical flasks with pre-warmed containing MOPS medium with all amino acids except zero isoleucine, and placed into the shaking incubator. Filters were removed after 10 min of shaking incubation. For quantification of ppGpp in the Lrp- mutant (Fig 7), strains were grown in MOPS medium with double the normal amount of all amino acids except tyrosine and isoleucine (which were provided at normal levels).

Nucleotide extraction and ppGpp quantification

Nucleotides were extracted as described, with minor modifications (9). For fermentor cultures, 5 ml of culture was sampled directly into a 15 ml round-bottom tube containing 0.5 ml of 11M formic acid. For 50 ml flask cultures, 1 ml of culture was sampled into 1.5 ml eppendorf tubes containing 0.1 ml of 11M formic acid. The sample was vigorously mixed and chilled on ice. One ml aliquots of this mixture were incubated at 0° C in an ice water bath for 45 min with periodic vortexing. These one ml samples were centrifuged at 4° at 6000 RPM for 5 min. The supernatant was then filtered through 0.2 μm filters and stored at -20° until HPLC analysis.

ppGpp was quantified by anion exchange HPLC using a Mono Q 5/50 GL column (GE Healthcare). 250 μl of supernatant was injected under initial conditions of 95% 20mM Tris (pH 8.0) and 5% 20mM Tris + 1.5 M sodium

formate (pH 8.0). This initial condition was maintained for 5 min. Absorbance at 254 nm was used to detect eluted nucleotides. Over a period of 30 min, the level of sodium formate buffer was ramped up to 65%. ppGpp was identified as a peak which eluted at ~28min (or ~52% 1.5 M sodium formate buffer). Samples were run in duplicate for at least two separate time course experiments. Combined results for at least two experiments are shown in Figs. 6 and 7. Representative results are shown in Fig 4. ppGpp standard was purchased from TriLink Biosciences. Standard curves established that the linear range of detection of ppGpp was 10nM to 100μM.

Microarray analysis

Cells were sampled directly from the fermentor into an equal volume of ice-cold RNA*later* (Ambion) and total RNA was extracted using Qiagen RNeasy Minikits with optional DNAse treatment steps. RNA was checked for integrity by gel electrophoresis and maintained in a 2:1 dilution of EtOH at -80°C until labeling. RNA was converted to cDNA by first strand synthesis using Superscript II (Invitrogen) and random hexamers, according to the manufacturer's specifications. The cDNA was fragmented and biotinylated (Enzo Kit, Roche Diagnostics) according to the Affymetrix prokaryotic labeling protocol.

The microarrays used for single time-point mutant analysis were custom built Affymetrix GeneChips containing probes for several prokaryotic genomes including *E. coli* K12 MG1655, *E. coli* O157:H7 EDL933, *Bacteriodes thetaiotaomicron* VPI-5482*, Enterococcus faecalis* V583*, Salmonella typhimurium*

LT2, and *Bacillus anthracis*. Biotinylated samples were prepared according to the manufacturer's instructions and hybridized for 16 hours at 60°C. Hybridized arrays were stained using Affymetrix protocol ProkGE_WS2v2_450. Stained microarrays were scanned and the raw data files (.cel) were further analyzed using RMA processing with quartile normalization (24). WT and mutant samples were duplicated biologically and technically; r^2 was >0.95 for all replicates.

For the 12-point WT isoleucine starvation timecourse (Fig. 3-4), Affymetrix E. coli 2.0 genome arrays were used as directed by the manufacturer. Data shown is based on a single timecourse. Array data from two biological replicates sampled from rapidly growing WT cells grown in medium replete all amino acids were averaged to serve as the control values for comparison to the timecourse arrays (r^2 = 0.998 for control arrays). Three-parameter sigmoidal regressions were plotted through averaged data points collected for each strain to obtain the response times. Statistical analysis was done using SigmaPlot 8.0. We considered genes to be significantly induced or repressed if the absolute value of the expression ratio was >2-fold (60). Hierarchical clustering algorithms were implemented in DecisionSite for Functional Genomics (Spotfire). The microarray data were deposited at Array Express (http://www.ebi.ac.uk/miamexpress/), accession # E-MEXP-1370.

GFP experiments

Growth conditions for GFP strains are described under 'Bacterial strains and growth conditions.' One hundred µl samples for fluorescence reading were

pipetted into transparent 96-well microplates (Corning) in triplicate. Wells contained 1 µl of x 25 mg/ml chloramphenicol to quench protein synthesis. Fluorescence (485nm excitation and 520nm emission) and optical densities (600 nm) were read on a FLUOstar Optima fluorimeter (BMG Labtech). For each fluorescent strain, uninduced control values were obtained as the average fluorescence across the time of log growth (OD ~0.2 to ~1.0) in medium with all 20 amino acids available at non-limiting levels. Control values were based on two biologically replicated growth curves for each strain. Experimental fluorescence values were compared to these control values to yield Log_2 expression ratios. Plots shown in Fig 6 contain data from three experimental runs for each strain. Three-parameter sigmoidal regressions were plotted through all data points collected for each strain to obtain the response times. Statistical analysis was done using SigmaPlot 8.0.

Table 1: Strains used in this study

Strain	Relevant Genotype	Source or Derivation
MG1655 seq	WT	Sequenced strain
MFT704	Δ*relA*::FRT, Δ*spoT*::FRT	ref 52
MFT	Δ*rpoS*::kan	ref 53
MFT760	Δ*lrp*::FRT	This study
MFT720	WT MG1655 x pUA66::P_{leuL}	This study
MFT721	WT MG1655 x pUA66::P_{dadA}	This study
MFT722	WT MG1655 x pUA66::P_{yahO}	This study
MFT723	WT MG1655 x pUA66::P_{wrbA}	This study

Figure Legends

Fig. 1. Growth curves for isoleucine starved cultures. Cultures were grown as described in materials and methods. Arrows indicate time points when RNA was harvested for microarray analysis.

Fig. 2. Contributions of Lrp and RpoS-mediated gene induction to the ppGpp-dependent stringent response. Only genes which were induced >2-fold in the WT were considered for the analysis. Genes that were expressed at significantly higher levels in any of the mutants were also omitted (14 genes). Thus, the genes shown require a given regulator for their full induction. (A) Venn diagram of overlapping regulons. Numbers of genes in each field are shown in red. Each field is labeled 1.1-3.1 for identification in B-E. 'Unaffected' indicates genes that were induced in all strains. (B) Heat map of genes that require ppGpp and Lrp for their induction. Gene names are shown to the left and strains are shown at bottom. Color legend for Log_2 expression is shown below. (C) Genes that require ppGpp, RpoS, and Lrp for induction. (D) Expression of the *lrp* and *rpoS* genes in array experiments. (E) Genes that require ppGpp and RpoS for their induction.

Fig. 3. Timecourse (WT) heat maps of genes in the regulons identified in Fig 2 as cells starved for isoleucine. Number 2.1, 2.3, and 3.1 refer to Venn fields in Fig 2. Numbers below each heat maps refer to the timepoints identified in Fig 4.

Fig. 4. WT growth curve, average regulon response times, and ppGpp accumulation as cells starved for isoleucine. Numbered timepoints indicate times of RNA sampling for array data shown in Fig 3. Expression of each regulon shown in Fig 3 was averaged at each time point and plotted. Sigmoidal regression curves were plotted for each regulon and the response time (time of ½ maximal induction) is marked as a solid vertical line.

Fig. 5. Model for regulatory architecture within the stringent response under three different environmental conditions. Arrows indicate positive relationships, flat ends represent negative relationships. Green lines indicate active vs. inactive (black) pathways. (A) When nutrients are plentiful, both the metabolic Lrp regulon and the stress survival RpoS regulon are inactive. ppGpp accumulation is prevented by exogenously available nutrients (ie amino acids). (B) When an amino acid becomes limiting, ppGpp accumulates. Signaling through a regulatory/sensory protein such as Lrp works in conjunction with a low level of ppGpp (x) to activate transcription of amino acid biosynthetic genes. Endogenously produced amino acids cause the ppGpp level to decline, inactivating the general stress response [which requires a high level of ppGpp (y)] and allowing growth to continue. (C) When the environment no longer supports growth biosynthetic pathways are incapable of yielding necessary metabolites. ppGpp accumulates to a high level, fully activating the general stress response.

Fig. 6. Differences in response times of the Lrp and RpoS regulons are reduced when ppGpp accumulates rapidly. (A) Fluorescence directed by RpoS-dependent (P_{yahO} and P_{wrbA}) and Lrp-dependent promoters (P_{leuL} and P_{dadA}) under conditions wherein culture growth results in exhaustion of exogenous isoleucine. ppGpp accumulates slowly (in > 100 min) under such a scenario. Response times are marked as vertical colored lines. T_0 = time of inoculation (B) Fluorescence from the same promoters as in A except that cells were grown in media replete with all 20 amino acids, collected on filters at T_0, and resuspended in medium with all amino acids except isoleucine. ppGpp accumulates rapidly and the differences in response times between promoters are reduced. (C) Response times of each promoter from panel A. Blue lines indicate regulon the average time of promoters from each regulon. (D) Response times from promoters from panel B. Dotted blue lines indicate that the difference in regulon response times is statistically insignificant.

Fig. 7. Lrp activity influences ppGpp level. (A) Growth curves of WT (black) and Δ*lrp* (blue) strains grown in MOPS medium with 2x the normal amount of 18 amino acids (tyrosine and isoleucine were included at normal levels). Cells were collected on filters and resuspended in MOPS minimal medium at T_0. Lrp cells adhered more readily to glass fiber filters than the WT. (B) ppGpp kinetics of WT and Δ*lrp* strains. In the WT, ppGpp accumulated rapidly after filtration and

remained high until growth resumed around 150 min. ppGpp only declined to ~700 pmol/ml/OD in the Δ*lrp* strain, which did not resume growth for many hours.

Fig. 1.

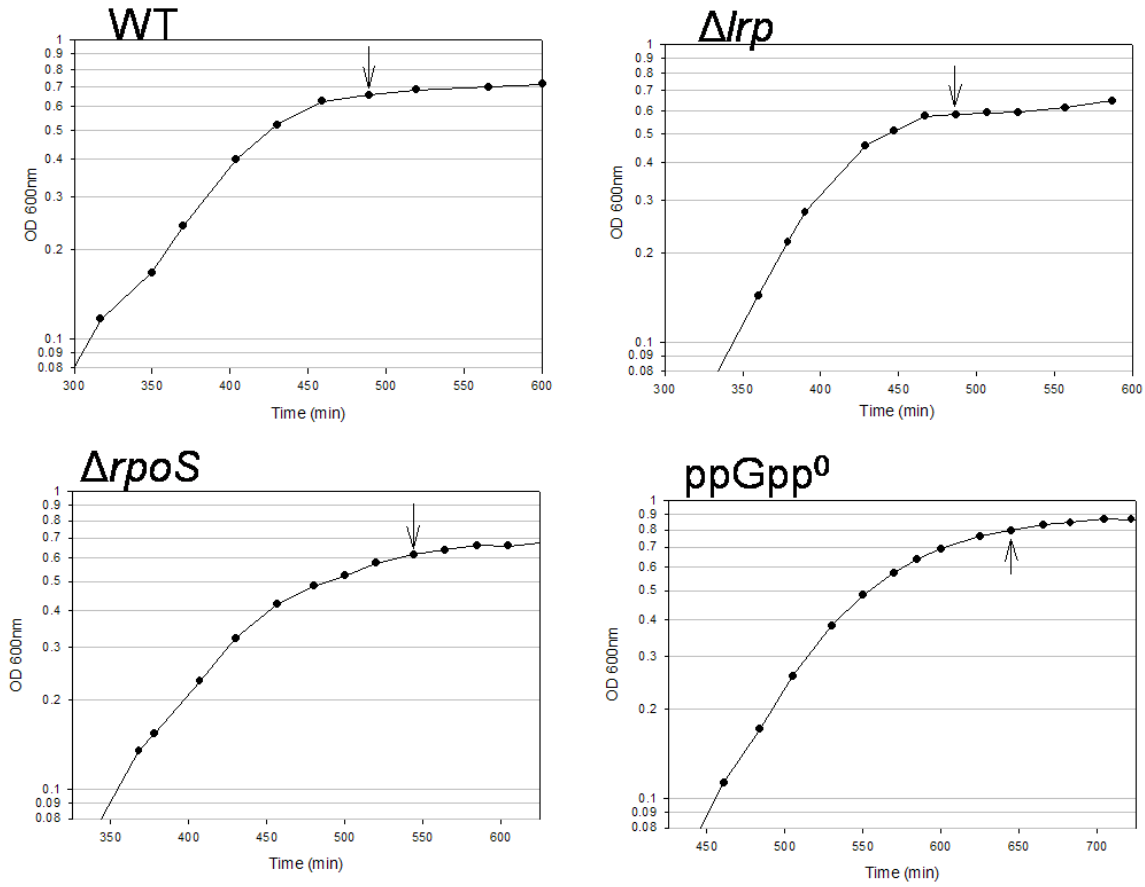

WT
OD 600nm
Time (min)
Δlrp
OD 600nm
Time (min)
ΔrpoS
OD 600nm
Time (min)
ppGpp⁰
OD 600nm
Time (min)

Fig. 2.

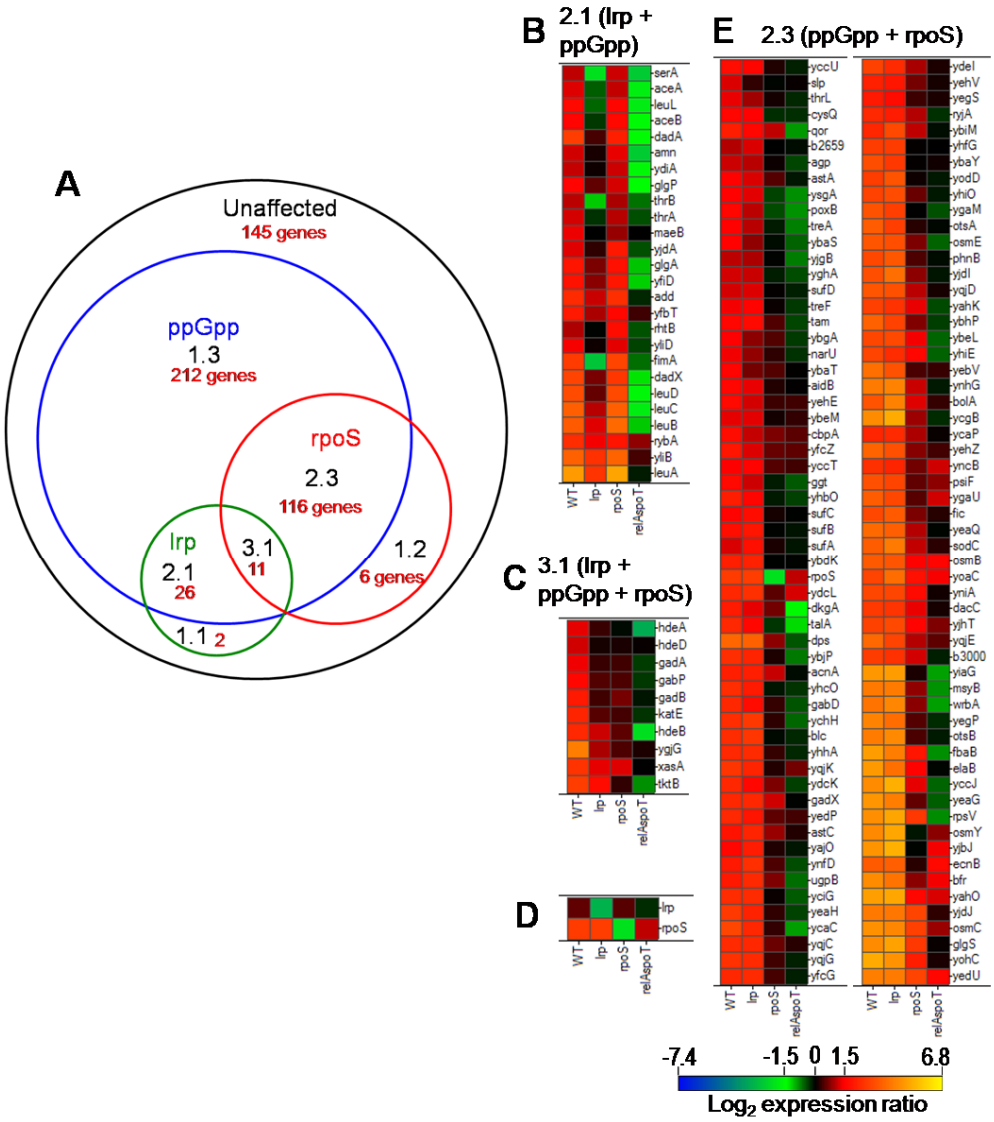

A
Unaffected
145 genes
ppGpp
1.3
212 genes
rpoS
2.3
116 genes
lrp
2.1
26
3.1
11
1.2
6 genes
1.1 2
B
2.1 (lrp + ppGpp)
C
3.1 (lrp + ppGpp + rpoS)
D
E
2.3 (ppGpp + rpoS)
WT
lrp
rpoS
relAspoT
-7.4
-1.5
0
1.5
6.8
Log2 expression ratio

Fig. 3.

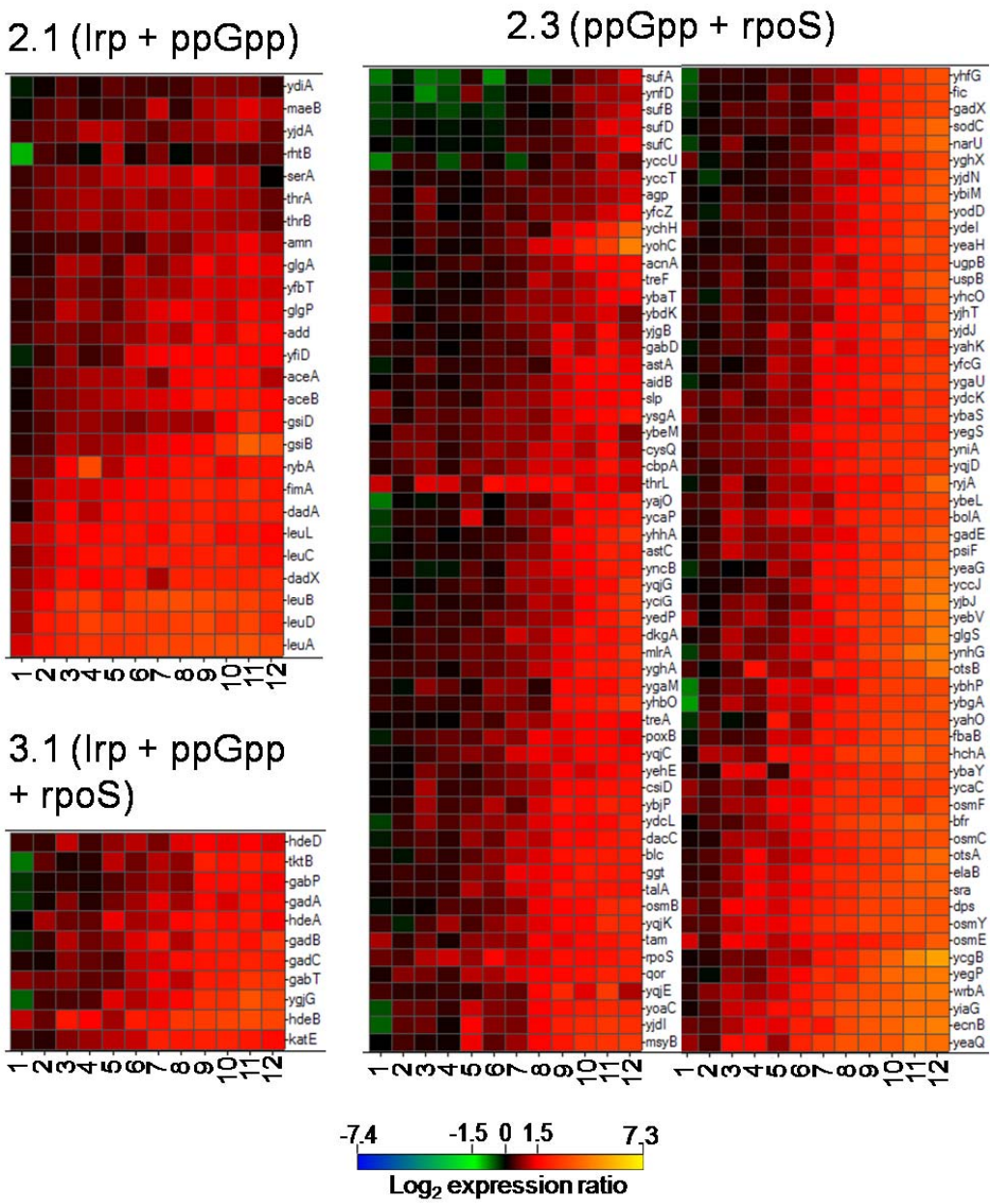

2.1 (lrp + ppGpp)
ydiA
maeB
yjdA
rhtB
serA
thrA
thrB
amn
glgA
yfbT
glgP
add
yfiD
aceA
aceB
gsiD
gsiB
rybA
fimA
dadA
leuL
leuC
dadX
leuB
leuD
leuA
1 2 3 4 5 6 7 8 9 10 11 12
2.3 (ppGpp + rpoS)
sufA
ynfD
sufB
sufD
sufC
yccU
yccT
agp
yfcZ
ychH
yohC
acnA
treF
ybaT
ybdK
yjgB
gabD
astA
aidB
slp
ysgA
ybeM
cysQ
cbpA
thrL
yajO
ycaP
yhhA
astC
yncB
yqjG
yciG
yedP
dkgA
mlrA
yghA
ygaM
yhbO
treA
poxB
yqjC
yehE
csiD
ybjP
ydcL
dacC
blc
ggt
talA
osmB
yqjK
tam
rpoS
qor
yqjE
yoaC
ydjI
msyB
yhfG
fic
gadX
sodC
narU
yghX
yjdN
ybiM
yodD
ydeI
yeaH
ugpB
uspB
yhcO
yjhT
yjdJ
yahK
yfcG
ygaU
ydcK
ybaS
yegS
yniA
yqjD
ryjA
ybeL
bolA
gadE
psiF
yeaG
yccJ
yjbJ
yebV
glgS
ynhG
otsB
ybhP
ybgA
yahO
fbaB
hchA
ybaY
ycaC
osmF
bfr
osmC
otsA
elaB
sra
dps
osmY
osmE
ycgB
yegP
wrbA
yiaG
ecnB
yeaQ
1 2 3 4 5 6 7 8 9 10 11 12
3.1 (lrp + ppGpp + rpoS)
hdeD
tktB
gabP
gadA
hdeA
gadB
gadC
gabT
ygjG
hdeB
katE
1 2 3 4 5 6 7 8 9 10 11 12
-7.4
-1.5 0 1.5
7.3
Log2 expression ratio

Fig. 4.

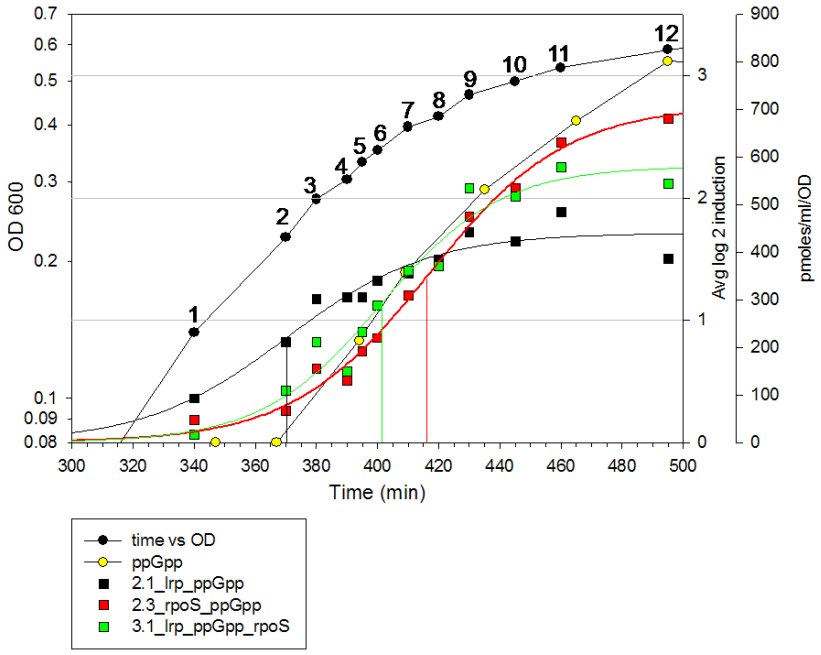

OD 600
Avg log 2 induction
pmoles/ml/OD
Time (min)
time vs OD
ppGpp
2.1_lrp_ppGpp
2.3_rpoS_ppGpp
3.1_lrp_ppGpp_rpoS

Fig. 5.

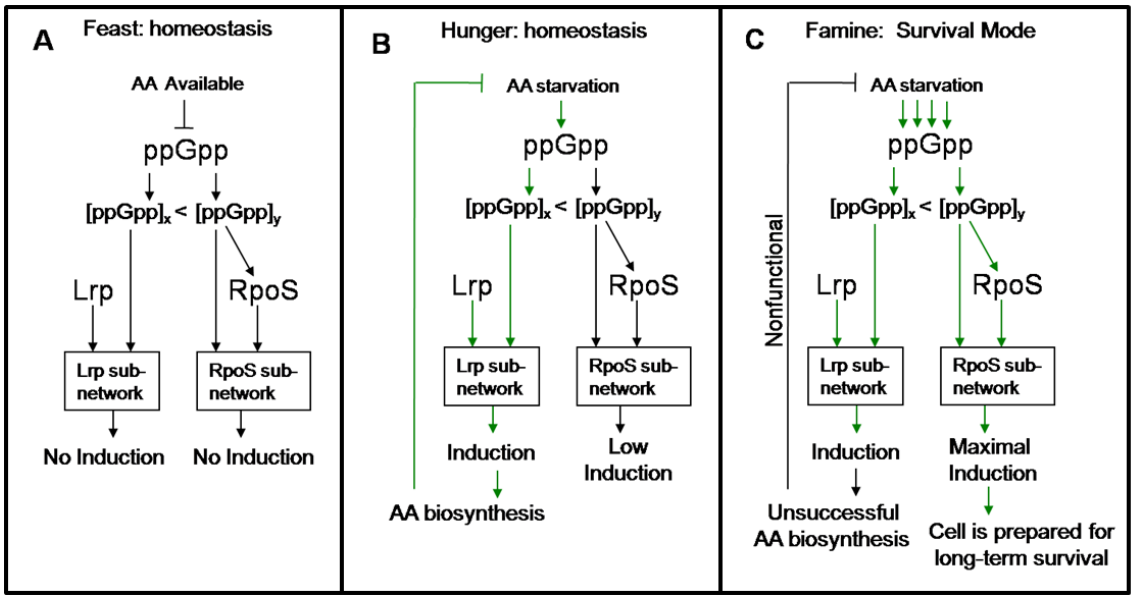

A
Feast: homeostasis
AA Available
ppGpp
[ppGpp]x < [ppGpp]y
Lrp
RpoS
Lrp sub-network
RpoS sub-network
No Induction
No Induction
B
Hunger: homeostasis
AA starvation
ppGpp
[ppGpp]x < [ppGpp]y
Lrp
RpoS
Lrp sub-network
RpoS sub-network
Induction
Low Induction
AA biosynthesis
C
Famine: Survival Mode
AA starvation
ppGpp
[ppGpp]x < [ppGpp]y
Nonfunctional
Lrp
RpoS
Lrp sub-network
RpoS sub-network
Induction
Maximal Induction
Unsuccessful AA biosynthesis
Cell is prepared for long-term survival

Fig. 6.

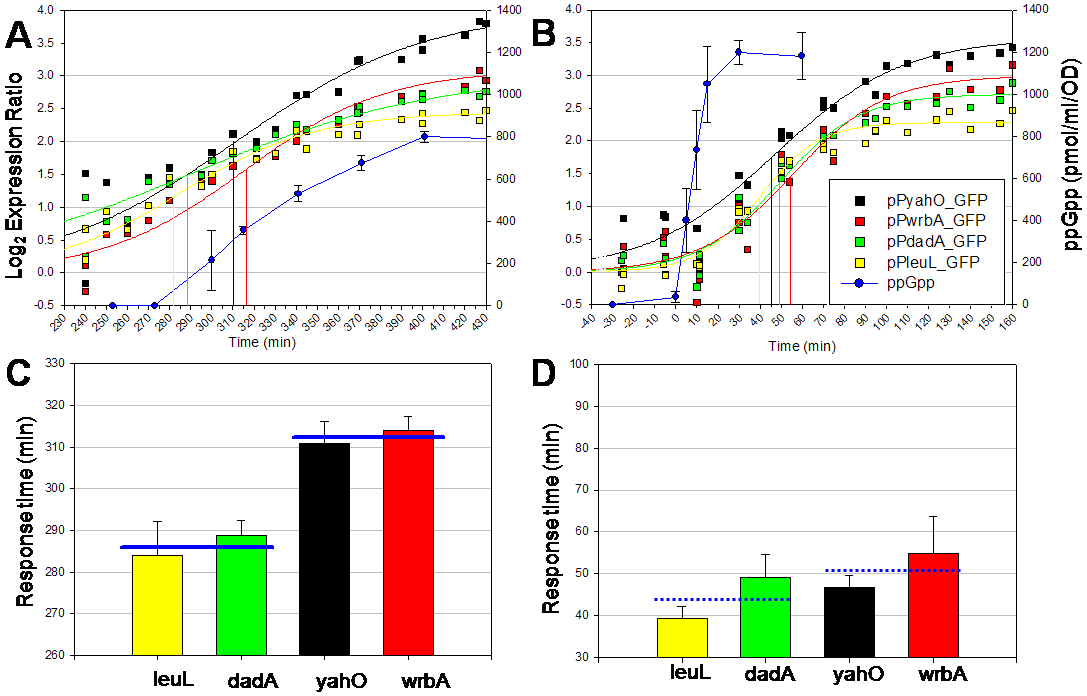
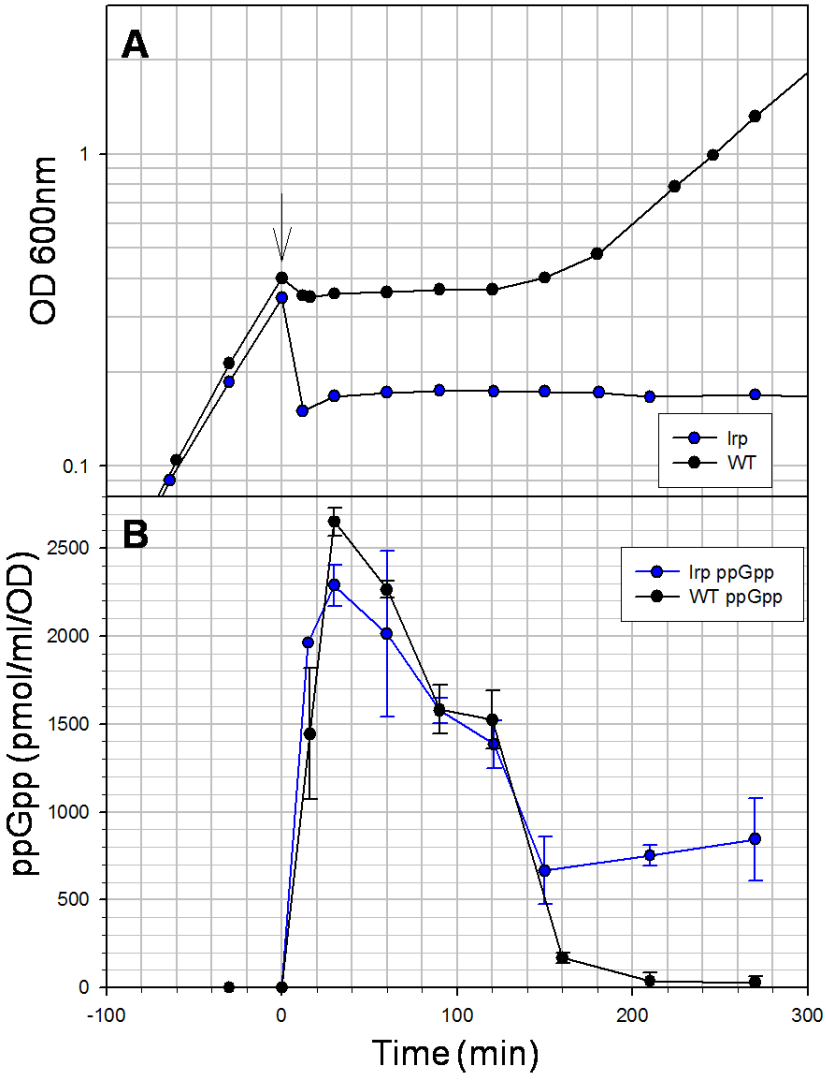

A
B
C
D
Log2 Expression Ratio
Time (min)
ppGpp (pmol/ml/OD)
pPyahO_GFP
pPwrbA_GFP
pPdadA_GFP
pPleuL_GFP
ppGpp
Responsetime (min)
leuL
dadA
yahO
wrbA

Fig. 7.


A
B
OD 600nm
1
0.1
lrp
WT
lrp ppGpp
WT ppGpp
ppGpp (pmol/ml/OD)
2500
2000
1500
1000
500
0
-100
0
100
200
300
Time (min)

References

1. **Aberg, A., J. Fernandez-Vazquez, J. D. Cabrer-Panes, A. Sanchez, and C. Balsalobre.** 2009. Similar and divergent outcomes of ppGpp and DksA deficiencies on transcription in Escherichia coli. J Bacteriol.

2. **Alon, U.** 2007. Network motifs: theory and experimental approaches. Nat Rev Genet **8:**450-61.

3. **Artsimovitch, I., V. Patlan, S. Sekine, M. N. Vassylyeva, T. Hosaka, K. Ochi, S. Yokoyama, and D. G. Vassylyev.** 2004. Structural basis for transcription regulation by alarmone ppGpp. Cell **117:**299-310.

4. **Baccigalupi, L., R. Marasco, E. Ricca, M. De Felice, and M. Sacco.** 1995. Control of ilvIH transcription during amino acid downshift in stringent and relaxed strains of Escherichia coli. FEMS Microbiol Lett **131:**95-8.

5. **Barker, M. M., T. Gaal, and R. L. Gourse.** 2001. Mechanism of regulation of transcription initiation by ppGpp. II. Models for positive control based on properties of RNAP mutants and competition for RNAP. J Mol Biol **305:**689-702.

6. **Barker, M. M., T. Gaal, C. A. Josaitis, and R. L. Gourse.** 2001. Mechanism of regulation of transcription initiation by ppGpp. I. Effects of ppGpp on transcription initiation in vivo and in vitro. J Mol Biol **305:**673-88.

7. **Barrett, C. L., C. D. Herring, J. L. Reed, and B. O. Palsson.** 2005. The global transcriptional regulatory network for metabolism in Escherichia coli exhibits few dominant functional states. Proc Natl Acad Sci U S A **102:**19103-8.

8. **Battesti, A., and E. Bouveret.** 2006. Acyl carrier protein/SpoT interaction, the switch linking SpoT-dependent stress response to fatty acid metabolism. Mol Microbiol **62:**1048-63.

9. **Bochner, B. R., and B. N. Ames.** 1982. Complete analysis of cellular nucleotides by two-dimensional thin layer chromatography. J Biol Chem **257:**9759-69.

10. **Bougdour, A., C. Cunning, P. J. Baptiste, T. Elliott, and S. Gottesman.** 2008. Multiple pathways for regulation of sigmaS (RpoS) stability in Escherichia coli via the action of multiple anti-adaptors. Mol Microbiol **68:**298-313.

11. **Bougdour, A., and S. Gottesman.** 2007. ppGpp regulation of RpoS degradation via anti-adaptor protein IraP. Proc Natl Acad Sci U S A **104:**12896-901.

12. **Bremer, H., and P. P. Dennis.** 1996. Modulation of chemical composition and other parameters of the cell by growth rate, p. 1553-1569. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.

13. **Brinkman, A. B., T. J. Ettema, W. M. de Vos, and J. van der Oost.** 2003. The Lrp family of transcriptional regulators. Mol Microbiol **48:**287-94.

14. **Calvo, J. M., and R. G. Matthews.** 1994. The leucine-responsive regulatory protein, a global regulator of metabolism in Escherichia coli. Microbiol Rev **58:**466-90.

15. **Cashel, M., D. R. Gentry, V. J. Hernandez, and D. Vinella.** 1996. The stringent response, p. 1458-1496. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.

16. **Cho, B. K., C. L. Barrett, E. M. Knight, Y. S. Park, and B. O. Palsson.** 2008. Genome-scale reconstruction of the Lrp regulatory network in Escherichia coli. Proc Natl Acad Sci U S A **105:**19462-7.

17. **Datsenko, K. A., and B. L. Wanner.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A **97:**6640-5.

18. **de los Rios, S., and J. J. Perona.** 2007. Structure of the Escherichia coli leucine-responsive regulatory protein Lrp reveals a novel octameric assembly. J Mol Biol **366:**1589-602.

19. **Durfee, T., A. M. Hansen, H. Zhi, F. R. Blattner, and D. J. Jin.** 2008. Transcription profiling of the stringent response in Escherichia coli. J Bacteriol **190:**1084-96.

20. **Haugen, S. P., W. Ross, and R. L. Gourse.** 2008. Advances in bacterial promoter recognition and its control by factors that do not bind DNA. Nat Rev Microbiol **6:**507-19.

21. **Hengge-Aronis, R.** 1996. Regulation of gene expression during entry into stationary phase, p. 1497-1512. *In* F. C. Neidhardt, R. Curtiss III, J. L.

Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.

22. **Hengge-Aronis, R.** 2002. Signal transduction and regulatory mechanisms involved in control of the sigma(S) (RpoS) subunit of RNA polymerase. Microbiol Mol Biol Rev **66:**373-95, table of contents.

23. **Ibanez-Ruiz, M., V. Robbe-Saule, D. Hermant, S. Labrude, and F. Norel.** 2000. Identification of RpoS (sigma(S))-regulated genes in Salmonella enterica serovar typhimurium. J Bacteriol **182:**5749-56.

24. **Irizarry, R. A., B. Hobbs, F. Collin, Y. D. Beazer-Barclay, K. J. Antonellis, U. Scherf, and T. P. Speed.** 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics **4:**249-64.

25. **Jensen, K. F.** 1993. The Escherichia coli K-12 "wild types" W3110 and MG1655 have an rph frameshift mutation that leads to pyrimidine starvation due to low pyrE expression levels. J Bacteriol **175:**3401-7.

26. **Jishage, M., K. Kvint, V. Shingler, and T. Nystrom.** 2002. Regulation of sigma factor competition by the alarmone ppGpp. Genes Dev **16:**1260-70.

27. **Kvint, K., A. Farewell, and T. Nystrom.** 2000. RpoS-dependent promoters require guanosine tetraphosphate for induction even in the presence of high levels of sigma(s). J Biol Chem **275:**14795-8.

28. **Lacour, S., and P. Landini.** 2004. SigmaS-dependent gene expression at the onset of stationary phase in Escherichia coli: function of sigmaS-dependent genes and identification of their promoter sequences. J Bacteriol **186:**7186-95.

29. **Landgraf, J. R., J. Wu, and J. M. Calvo.** 1996. Effects of nutrition and growth rate on Lrp levels in Escherichia coli. J Bacteriol **178:**6930-6.

30. **Lange, R., D. Fischer, and R. Hengge-Aronis.** 1995. Identification of transcriptional start sites and the role of ppGpp in the expression of rpoS, the structural gene for the sigma S subunit of RNA polymerase in Escherichia coli. J Bacteriol **177:**4676-80.

31. **Lawther, R. P., D. H. Calhoun, C. W. Adams, C. A. Hauser, J. Gray, and G. W. Hatfield.** 1981. Molecular basis of valine resistance in Escherichia coli K-12. Proc Natl Acad Sci U S A **78:**922-5.

32. **Lazzarini, R. A., M. Cashel, and J. Gallant.** 1971. On the regulation of guanosine tetraphosphate levels in stringent and relaxed strains of Escherichia coli. J Biol Chem **246:**4381-5.

33. **Lin, R., R. D'Ari, and E. B. Newman.** 1992. Lambda placMu insertions in genes of the leucine regulon: extension of the regulon to genes not regulated by leucine. J Bacteriol **174:**1948-55.

34. **Liu, M., T. Durfee, J. E. Cabrera, K. Zhao, D. J. Jin, and F. R. Blattner.** 2005. Global transcriptional programs reveal a carbon source foraging strategy by Escherichia coli. J Biol Chem **280:**15921-7.

35. **Loewen, P. C., and R. Hengge-Aronis.** 1994. The role of the sigma factor sigma S (KatF) in bacterial global regulation. Annu Rev Microbiol **48:**53-80.

36. **Luscombe, N. M., M. M. Babu, H. Yu, M. Snyder, S. A. Teichmann, and M. Gerstein.** 2004. Genomic analysis of regulatory network dynamics reveals large topological changes. Nature **431:**308-12.

37. **Magnusson, L. U., A. Farewell, and T. Nystrom.** 2005. ppGpp: a global regulator in Escherichia coli. Trends Microbiol **13:**236-42.

38. **Mangan, S., and U. Alon.** 2003. Structure and function of the feed-forward loop network motif. Proc Natl Acad Sci U S A **100:**11980-5.

39. **Merrikh, H., A. E. Ferrazzoli, A. Bougdour, A. Olivier-Mason, and S. T. Lovett.** 2009. A DNA damage response in Escherichia coli involving the alternative sigma factor, RpoS. Proc Natl Acad Sci U S A **106:**611-6.

40. **Neidhardt, F. C., P. L. Bloch, and D. F. Smith.** 1974. Culture medium for enterobacteria. J Bacteriol **119:**736-47.

41. **Newman, E. B., R. D'Ari, and R. T. Lin.** 1992. The leucine-Lrp regulon in E. coli: a global response in search of a raison d'etre. Cell **68:**617-9.

42. **Newman, E. B., and R. Lin.** 1995. Leucine-responsive regulatory protein: a global regulator of gene expression in E. coli. Annu Rev Microbiol **49:**747-75.

43. **Paul, B. J., M. B. Berkmen, and R. L. Gourse.** 2005. DksA potentiates direct activation of amino acid promoters by ppGpp. Proc Natl Acad Sci U S A **102:**7823-8.

44. **Paul, B. J., W. Ross, T. Gaal, and R. L. Gourse.** 2004. rRNA transcription in *Escherichia coli*. Annu Rev Genet **38:**749-70.

45. **Potrykus, K., and M. Cashel.** 2008. (p)ppGpp: still magical? Annu Rev Microbiol **62:**35-51.

46. **Rhodius, V. A., W. C. Suh, G. Nonaka, J. West, and C. A. Gross.** 2006. Conserved and variable functions of the sigmaE stress response in related genomes. PLoS Biol **4:**e2.

47. **Rutherford, S. T., C. L. Villers, J. H. Lee, W. Ross, and R. L. Gourse.** 2009. Allosteric control of Escherichia coli rRNA promoter complexes by DksA. Genes Dev **23:**236-48.

48. **Ryals, J., R. Little, and H. Bremer.** 1982. Control of rRNA and tRNA syntheses in Escherichia coli by guanosine tetraphosphate. J Bacteriol **151:**1261-8.

49. **Sarubbi, E., K. E. Rudd, and M. Cashel.** 1988. Basal ppGpp level adjustment shown by new *spoT* mutants affect steady state growth rates and *rrnA* ribosomal promoter regulation in *Escherichia coli*. Mol Gen Genet **213:**214-22.

50. **Sokawa, Y., J. Sokawa, and Y. Kaziro.** 1975. Regulation of stable RNA synthesis and ppGpp levels in growing cells of Escherichia coli. Cell **5:**69-74.

51. **Tani, T. H., A. Khodursky, R. M. Blumenthal, P. O. Brown, and R. G. Matthews.** 2002. Adaptation to famine: a family of stationary-phase genes revealed by microarray analysis. Proc Natl Acad Sci U S A **99:**13471-6.

52. **Traxler, M. F., D. E. Chang, and T. Conway.** 2006. Guanosine 3',5'-bispyrophosphate coordinates global gene expression during glucose-lactose diauxie in Escherichia coli. Proc Natl Acad Sci U S A **103:**2374-9.

53. **Traxler, M. F., S. M. Summers, H. T. Nguyen, V. M. Zacharia, G. A. Hightower, J. T. Smith, and T. Conway.** 2008. The global, ppGpp-mediated stringent response to amino acid starvation in Escherichia coli. Mol Microbiol **68:**1128-48.

54. **VanBogelen, R. A., P. M. Kelley, and F. C. Neidhardt.** 1987. Differential induction of heat shock, SOS, and oxidation stress regulons and accumulation of nucleotides in *Escherichia coli*. J Bacteriol **169:**26-32.

55. **Vinella, D., C. Albrecht, M. Cashel, and R. D'Ari.** 2005. Iron limitation induces SpoT-dependent accumulation of ppGpp in Escherichia coli. Mol Microbiol **56:**958-70.

56. **Vrentas, C. E., T. Gaal, M. B. Berkmen, S. T. Rutherford, S. P. Haugen, D. G. Vassylyev, W. Ross, and R. L. Gourse.** 2008. Still looking for the magic spot: the crystallographically defined binding site for ppGpp on RNA

polymerase is unlikely to be responsible for rRNA transcription regulation. J Mol Biol **377:**551-64.

57. **Wanner, B. L., R. Kodaira, and F. C. Neidhart.** 1977. Physiological regulation of a decontrolled lac operon. J Bacteriol **130:**212-22.

58. **Weber, H., T. Polen, J. Heuveling, V. F. Wendisch, and R. Hengge.** 2005. Genome-wide analysis of the general stress response network in Escherichia coli: sigmaS-dependent genes, promoters, and sigma factor selectivity. J Bacteriol **187:**1591-603.

59. **Wendrich, T. M., G. Blaha, D. N. Wilson, M. A. Marahiel, and K. H. Nierhaus.** 2002. Dissection of the mechanism for the stringent factor RelA. Mol Cell **10:**779-88.

60. **Wren, J. D., and T. Conway.** 2006. Meta-analysis of published transcriptional and translational fold changes reveals a preference for low-fold inductions. Omics **10:**15-27.

61. **Xiao, H., M. Kalman, K. Ikehara, S. Zemel, G. Glaser, and M. Cashel.** 1991. Residual guanosine 3',5'-bispyrophosphate synthetic activity of *relA* null mutants can be eliminated by *spoT* null mutations. J Biol Chem **266:**5980-90.

62. **Yang, W., L. Ni, and R. L. Somerville.** 1993. A stationary-phase protein of Escherichia coli that affects the mode of association between the trp repressor protein and operator-bearing DNA. Proc Natl Acad Sci U S A **90:**5796-800.

63. **Zaslaver, A., A. Bren, M. Ronen, S. Itzkovitz, I. Kikoin, S. Shavit, W. Liebermeister, M. G. Surette, and U. Alon.** 2006. A comprehensive library of fluorescent transcriptional reporters for Escherichia coli. Nat Methods **3:**623-8.

64. **Zaslaver, A., A. E. Mayo, R. Rosenberg, P. Bashkin, H. Sberro, M. Tsalyuk, M. G. Surette, and U. Alon.** 2004. Just-in-time transcription program in metabolic pathways. Nat Genet **36:**486-91.

65. **Zhi, J., E. Mathew, and M. Freundlich.** 1999. Lrp binds to two regions in the dadAX promoter region of Escherichia coli to repress and activate transcription directly. Mol Microbiol **32:**29-40.

Chapter 5: Synthesis and Conclusions

Introduction

The experiments detailed in this thesis have explored the role of the signaling molecule ppGpp in regulating global gene expression during carbon and amino acid starvation. Regulation by ppGpp was also considered in light of the activity of other global regulators such as RpoS, Crp, and Lrp. This final chapter seeks to consider lessons learned from this work in its totality and place them in the context of the larger body of knowledge regarding ppGpp. Thus, the main points from each chapter are reviewed and overarching trends discussed. The experimental results are further considered with regards to proposed mechanisms of action of ppGpp. Finally, remaining questions for continued inquiry are outlined.

Chapter 2 Summary

The experiments presented in Chapter 2 examined the role of ppGpp in coordinating gene expression in a classic experimental system: glucose-lactose diauxie (8). In this setup, cellular growth leads to the exhaustion of glucose, a transitional period of growth arrest (diauxie), and a resumption of growth on lactose. We experimentally defined the RpoS and Crp regulons that were induced during diauxie. This included genes of the general stress response and genes involved in carbon scavenging, respectively. We further found that a Δ*relA* strain had a compromised ability to induce these responses, concomitant with an attenuated ability to down-regulate genes involved in ribosome synthesis.

Taken together, the data suggest that ppGpp is required for mounting a rapid reorganization of global gene expression in times of sudden nutrient exhaustion. A model that places ppGpp at the top of the global regulation network and integrates this view with other known aspects of catabolite (de)repression is presented. This work demonstrated a new fundamental connection, namely that efficient carbon scavenging, which is managed by cyclic AMP/Crp, also requires ppGpp. Given that Crp controls one of the largest regulons within *E. coli*, this finding illustrates a hierarchy operating at the highest level of global regulation.

Chapter 3 Summary

Experiments presented in Chapter 3 focus on the extent of gene regulation controlled by ppGpp in response to amino acid starvation. Using isoleucine starvation as a model system, we observed that the WT down-regulated gene expression involved in a range of processes. These changes included the down-regulation of genes related to macromolecular synthesis of all kinds (DNA, RNA, protein, peptidoglycan, fatty acids, etc.), and the induction of the general stress response. The WT also comprehensively restructured its metabolic gene expression in a way consistent with production of isoleucine. This observation was further verified using a high-throughput assay designed to examine the metabolic proteome. In contrast, a $ppGpp^0$ strain was crippled in its ability to effect these changes. $ppGpp^0$ cells were larger than WT cells at all times, a phenotype that correlated with aberrantly high expression of genes involved in macromolecular synthesis of all kinds. Despite producing an identical

amount of protein to the WT under isoleucine starvation, the ppGpp^0 strain made 50% more biomass than the WT. Only about of half of this extra biomass was attributable to RNA, suggesting that ppGpp plays a central role in limiting the production of diverse macromolecular structures to a level compatible with the nutritional environment.

Chapter 4 Summary

Chapter 4 is devoted to understanding the regulatory relationships which govern gene induction during the stringent response. We used the isoleucine starvation system developed in Chapter 3 to define the Lrp and RpoS regulons in addition to the ppGpp regulon. The overwhelming majority of both the Lrp and RpoS regulons required ppGpp for their induction. However, despite mutual dependence on ppGpp, activation of the Lrp regulon is temporally separated from activation of the RpoS regulon. This observation led us to consider a model in which the metabolic genes controlled by Lrp only require a low level of ppGpp for their induction while the RpoS regulon might require a higher threshold ppGpp level for their activation. In keeping with this model, we found that accelerating the rate of ppGpp accumulation reduced the difference in response times of the Lrp and RpoS regulons. We also found that under harshly introduced amino acid starvation conditions Lrp activity can influence the ppGpp level, implying a feedback loop with the potential to influence expression of the RpoS regulon. From a fundamental perspective, the role of the Lrp system is to reestablish a new metabolic homeostasis, while the role of the RpoS regulon is to prepare the

cell for long-term survival under stressful conditions. The discrete tuning of these systems (*vis a vis* ppGpp level) supplies a rationale for how cells balance metabolic and stress survival processes across a range of environments, from feast to famine.

Role of ppGpp in global gene expression

Given the results described in all sections of this work, what can be said about the role of ppGpp in regulating global gene expression? Several overarching trends are evident: 1) The number of genes and processes influenced by ppGpp is extremely large, 2) ppGpp is required for cells to 'interpret' the overall quality of the surrounding environment, 3) The stringent response is a variable response. Implications of these trends are considered below.

In virtually all experimental conditions examined here, the absence of ppGpp (or alteration of ppGpp metabolism by *relA* mutation) led to aberrant expression of very large numbers of genes (i.e. >30% of the genome), (Ch 2: Fig 2, Ch 3: Fig 2). We demonstrated that these genes fell into a variety of regulons including those governed by Crp, RpoS, and Lrp. Comprehensive overlap between ppGpp and these regulons across different conditions suggests that ppGpp is at the apex of global regulation in *E. coli*. This overlap also implies that ppGpp rarely acts alone to mediate gene expression, especially induction. In this sense, if cells use regulators to properly partition transcriptional space, then ppGpp might be better understood as a blunt tool rather than a scalpel. This

observation has important implications for future study of genes influenced by ppGpp. Indeed, as part of a largely reductionist strategy in studying genetic regulation, the effect of ppGpp on individual promoters has omitted the possible effect of ppGpp function combined with more 'typical' DNA binding regulators, such Crp or Lrp (discussed in detail below). Finally, the large number of genes affected by ppGpp implicates its involvement in a much wider range of individual cellular processes than previously thought. As such, cells lacking ppGpp are profoundly altered at a systems level.

Bacteria adjust their growth rate, cell size, and ribosome load according to the nutritional richness (or poorness) of their environment. Microbiologists have been aware of these relationships for decades (12). Among the first recognized quantifiable phenotypes of 'relaxed' mutants was their tendency to continue synthesis of stable RNAs in the face of amino starvation (12). It was later observed that $ppGpp^0$ cells were larger than normal cells (19). Observations reported in Chapter 3 suggest that the role of ppGpp extends well beyond regulating these fundamental aspects of microbial growth. During amino acid starvation, ppGpp regulates total biomass production, including macromolecular synthesis outside of stable RNA transcription. The WT strain successfully moderated expression of genes involved in cell division, and DNA, fatty acid, and membrane/cell wall biosynthesis. Without ppGpp, cells consistently failed to execute these changes. Thus, $ppGpp^0$ cells are seemingly stuck in 'growth mode,' with virtually no ability to moderate basic vegetative functions even in environments incapable of supporting growth. Mutation of a bacterial genetic

regulator usually results in an inability to respond to a certain environmental condition or variable (6). In contrast, it seems that ppGpp^0 cells do not even 'realize' that they *should* be responding. All of these observations point to a central conclusion: ppGpp is both the primary signal and the primary effector used by *E. coli* cells to adjust their physiology to the surrounding environment.

Examining the role of ppGpp in regulating gene expression under different types of starvation (carbon starvation in Chapter 2, and amino acid biosynthesis in Chapters 3-4) suggests that the stringent response is variable in its genetic content. Many types of starvation provoke ppGpp accumulation: amino acid, carbon, fatty acid, iron, nitrogen, phosphate, etc. Based on the experiments described here, the stringent response can be thought of as having a core set of genes that always respond when growth arrests, and a condition-specific component triggered by accessory cues. For example, carbon starvation and isoleucine starvation both led to down-regulation of the translation apparatus genes and the induction of the general stress response. These two groups of genes make up the 'core response.' In contrast, carbon starvation initiated Crp-dependent catabolite derepression and carbon foraging, whereas isoleucine starvation triggered Lrp-dependent induction of amino acid biosynthetic pathways. Regulation via Crp and Lrp are examples of condition-specific responses. In both cases, the condition-specific responses also required ppGpp for their proper development.

Experiments described here also addressed stresses encountered on different timescales (i.e. quick–onset starvation of glucose exhaustion in Chapter

2 and rapid downshifts by filter collection in Chapter 4 vs. the drawn out starvation associated with isoleucine exhaustion in Chapters 3 and 4). During isoleucine exhaustion (not filtration and resuspension), ppGpp accumulated slowly over the course of ~100 min. This long timescale suggests that isoleucine starvation might approximate a traversal of the feast to famine gradient. The discrete timing and order of induction of the Lrp and RpoS regulons implies that the stringent response is variable in keeping with the severity of the nutritional situation. In the case of isoleucine starvation, the Lrp-mediated, condition-specific component of the response was activated before the culture showed a decrease in growth rate. Interestingly, the same is true for carbon starvation, wherein Crp-mediated carbon scavenging was initiated in the minutes before the onset of diauxie, and before the induction of the general stress response. Together these results imply that low levels of ppGpp are usually sufficient to allow the initiation of condition-specific metabolic responses. Should ppGpp continue to accumulate, the general stress response is initiated. Finally, it has been known for many years, that basal ppGpp level is correlated with the growth rate (11, 14, 15, 17). The proposed rheostatic model of the stringent response bears consideration as most experiments designed to interrogate the role of ppGpp utilize extremes of feast or famine. As such, there may be a general misconception that all genes that require ppGpp for their induction actually require a similar amount of ppGpp, or that all genes that require ppGpp for their induction will respond similarly when ppGpp accumulates. This view fails to

appreciate the nuance accommodated by the regulatory structure of the stringent response.

ppGpp and RNAP

Most positively acting transcriptional regulators bind DNA and function by 'recruiting' RNAP to the promoter of the genes to be induced (5). ppGpp is not a normal regulator in this sense, since it binds directly to (or inside) RNAP itself (7, 13). With this in mind, it is perhaps not surprising that $ppGpp^0$ strains exhibit such radically impacted transcription profiles. Based on these observations (and some interesting, yet rarely cited, work from Hans Bremer and coworkers in the early 1980s)(14), it might be suggested that ppGpp is required for normal promoter prioritization at a global level. One way to think about RNAP activity is to consider that this enzyme has two broad functions: 1) to produce stable RNAs, which are in fact structural cellular components required for life, and 2) mRNAs for all the other gene products in the cell. Stable RNAs must be produced at much higher levels than any other genes in the cell, except when growth is arrested. Even then however, the cell must retain a robust translational capacity. Accordingly, RNAP is optimized for binding and transcribing stable RNA genes, not mRNAs (of course, one could also argue that stable RNA gene promoters are optimized for transcription initiation). It might be suggested that ppGpp's principal role is to shift the primary activity of RNAP away from stable RNAs to the rest of the genome.

While a shift away from stable RNA synthesis implies that the promoter structure favored by ppGpp-bound RNAP is different than RNAP alone, the defining features of a promoter sequence induced by ppGpp has yet to be adequately described (7, 13). Perhaps the only common feature of promoters induced by ppGpp is the presence of an A/T rich 'discriminator region' located between the transcriptional start site and the -10 region promoter region (4, 9, 18). However, the commonality of discriminator regions is not broad enough to be genuinely predictive positive regulation by ppGpp (data not shown). The lack of a compelling structural model for induction by ppGpp based on RNAP-promoter interaction, and the broad overlap between the ppGpp, Lrp, and Crp regulons may suggest that a new model should be considered. As it seems that regulation in a wide range of regulons is disrupted in the absence of ppGpp, perhaps ppGpp makes RNAP more sensitive to regulation by DNA binding proteins. In this model, the interaction between RNAP and the regulatory proteins would be altered by ppGpp. While the work in this thesis provides no structural or mechanistic information, it at least suggests that investigation of an alternative model should be undertaken.

Going forward

A goal of systems biologists and bacterial physiologists alike is the creation of a model capable of describing a bacterial cell as a whole. A large component of such a model is the network of regulation that determines global transcriptional output (2, 3, 16). Probably due to the sheer number of genes

influenced by ppGpp and the lack of a an adequate structural model, ppGpp has not yet been incorporated into any large scale analysis of the *E. coli* transcription network. Hopefully the work described here is a step toward integrating this important regulator into these analyses. Many such models describe promoters as Boolean logic gates (1, 10). The publicly available array data from these studies can readily supply lists of genes that require ppGpp in AND-type functions. Furthermore, the data presented in Chapter 4 suggest ppGpp level thresholds that could be readily built into large scale physiological models.

Beyond the lofty (and somewhat esoteric) goal of integrating ppGpp into whole cell network simulations, I hope that this work advances critical understanding of the role of ppGpp in global regulation. As all free living bacteria have *relA* or *spoT* homologues, microbiologists in general can hopefully use the information provided here to facilitate breakthroughs in other medically or industrially important organisms.

References

1. **Alon, U.** 2007. Network motifs: theory and experimental approaches. Nat Rev Genet **8:**450-61.

2. **Balleza, E., L. N. Lopez-Bojorquez, A. Martinez-Antonio, O. Resendis-Antonio, I. Lozada-Chavez, Y. I. Balderas-Martinez, S. Encarnacion, and J. Collado-Vides.** 2009. Regulation by transcription factors in bacteria: beyond description. FEMS Microbiol Rev **33:**133-51.

3. **Barrett, C. L., C. D. Herring, J. L. Reed, and B. O. Palsson.** 2005. The global transcriptional regulatory network for metabolism in Escherichia coli exhibits few dominant functional states. Proc Natl Acad Sci U S A **102:**19103-8.

4. **Bougdour, A., and S. Gottesman.** 2007. ppGpp regulation of RpoS degradation via anti-adaptor protein IraP. Proc Natl Acad Sci U S A **104:**12896-901.

5. **Browning, D. F., and S. J. Busby.** 2004. The regulation of bacterial transcription initiation. Nat Rev Microbiol **2:**57-65.

6. **Dorman, C. J.** 2009. Nucleoid-associated proteins and bacterial physiology. Adv Appl Microbiol **67:**47-64.

7. **Haugen, S. P., W. Ross, and R. L. Gourse.** 2008. Advances in bacterial promoter recognition and its control by factors that do not bind DNA. Nat Rev Microbiol **6:**507-19.

8. **Jacob, F., and J. Monod.** 1961. Genetic regulatory mechanisms in the synthesis of proteins. J. Mol Biol **3:**318-56.

9. **Josaitis, C. A., T. Gaal, and R. L. Gourse.** 1995. Stringent control and growth-rate-dependent control have nonidentical promoter sequence requirements. Proc Natl Acad Sci U S A **92:**1117-21.

10. **Kaplan, S., A. Bren, A. Zaslaver, E. Dekel, and U. Alon.** 2008. Diverse two-dimensional input functions control bacterial sugar genes. Mol Cell **29:**786-92.

11. **Lazzarini, R. A., M. Cashel, and J. Gallant.** 1971. On the regulation of guanosine tetraphosphate levels in stringent and relaxed strains of Escherichia coli. J Biol Chem **246:**4381-5.

12. **Maaloe, O., and N. Kjeldgaard.** 1966. Control of Macromolecular Synthesis, vol. W. A. Benjamin, inc., New York.

13. **Potrykus, K., and M. Cashel.** 2008. (p)ppGpp: still magical? Annu Rev Microbiol **62:**35-51.

14. **Ryals, J., R. Little, and H. Bremer.** 1982. Control of rRNA and tRNA syntheses in Escherichia coli by guanosine tetraphosphate. J Bacteriol **151:**1261-8.

15. **Sarubbi, E., K. E. Rudd, and M. Cashel.** 1988. Basal ppGpp level adjustment shown by new *spoT* mutants affect steady state growth rates and *rrnA* ribosomal promoter regulation in *Escherichia coli*. Mol Gen Genet **213:**214-22.

16. **Sauer, U., M. Heinemann, and N. Zamboni.** 2007. Genetics. Getting closer to the whole picture. Science **316:**550-1.

17. **Sokawa, Y., J. Sokawa, and Y. Kaziro.** 1975. Regulation of stable RNA synthesis and ppGpp levels in growing cells of Escherichia coli. Cell **5:**69-74.

18. **Travers, A. A.** 1980. Promoter sequence for stringent control of bacterial ribonucleic acid synthesis. J Bacteriol **141:**973-6.

19. **Xiao, H., M. Kalman, K. Ikehara, S. Zemel, G. Glaser, and M. Cashel.** 1991. Residual guanosine 3',5'-bispyrophosphate synthetic activity of *relA* null mutants can be eliminated by *spoT* null mutations. J Biol Chem **266:**5980-90.

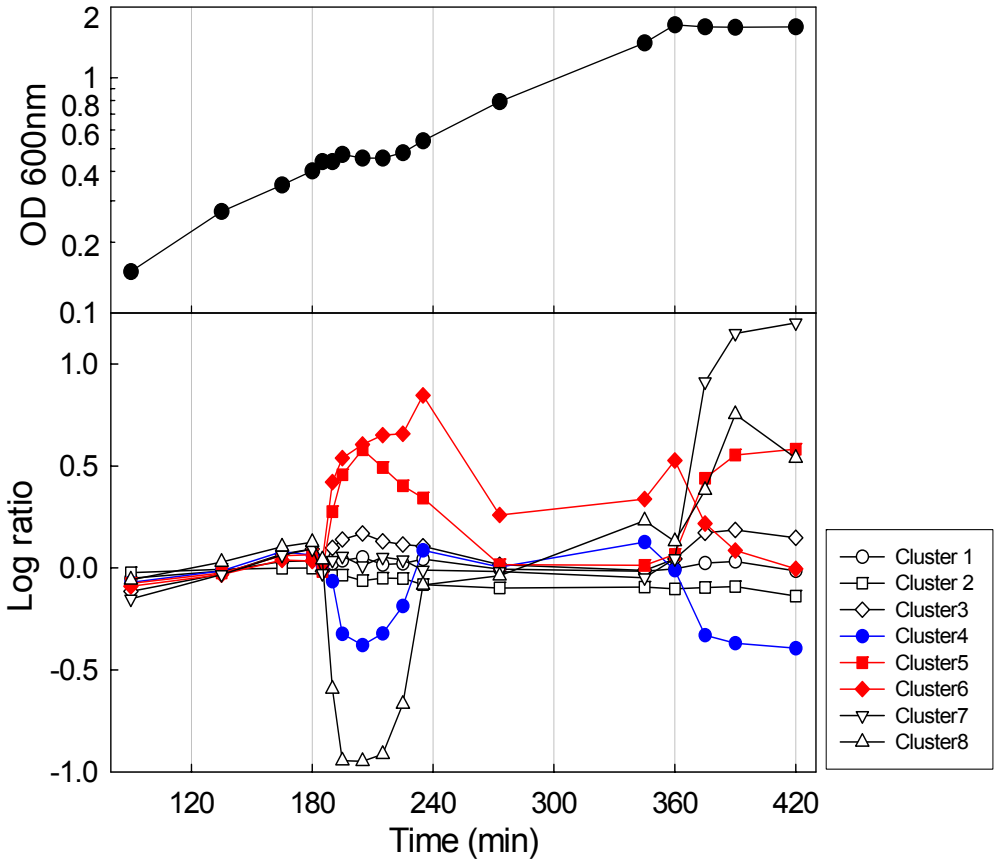
Note on appendices and high-thoughput data generated for this thesis:

The datasets generated for this thesis are publicly available, or are in the process of being made publicly available at the time of this writing. Array data from all chapters is available in an interactive manner at *E. coli* GenExpDB (http://genexpdb.ou.edu/), as part of the EcoliHub project and hosted by the Bioinformatics Core Facility at the University of Oklahoma Advanced Center of Genome Technology. Datasets associated with Chapter 3 can be downloaded from GEO (Gene Expression Omnibus: http://www.ncbi.nlm.nih.gov/geo/), which is hosted by NCBI.

Appendices 1 and 2 contain supplementary figures for Chapters 2 and 3, respectively.

Appendix 1: Supplementary data to accompany Chapter 2.

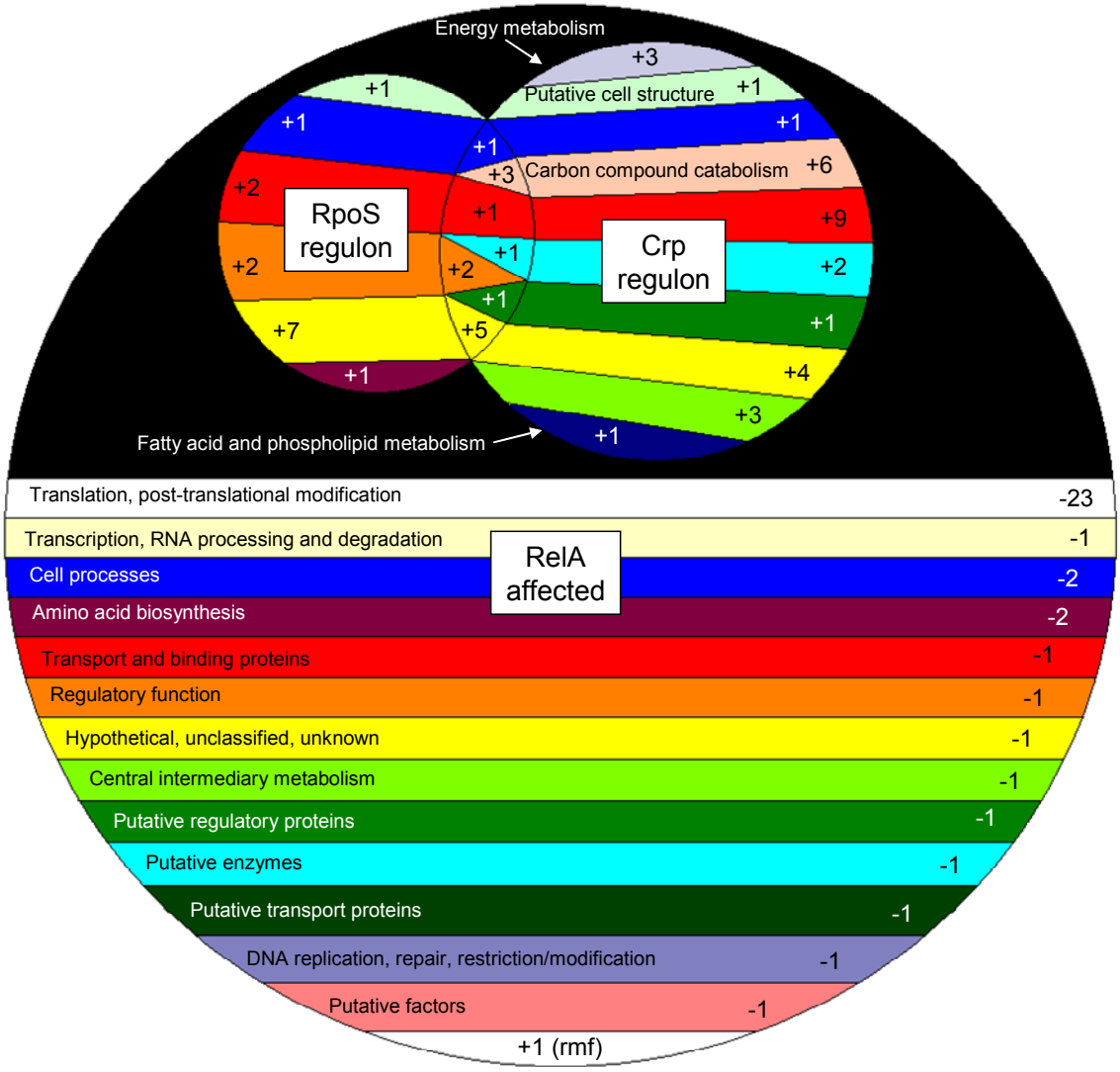
Supplementary Figure 1



Legend to Supplementary Figure 1. K-means cluster analysis (K=8) of diauxie gene expression profile data (from (8)). Top, growth curve of *E. coli* MG1655 on mixture of glucose (0.05%) and lactose (0.15%). Bottom, the average $\log_{10}$ expression ratio for each cluster is plotted for the entire diauxie experiment,

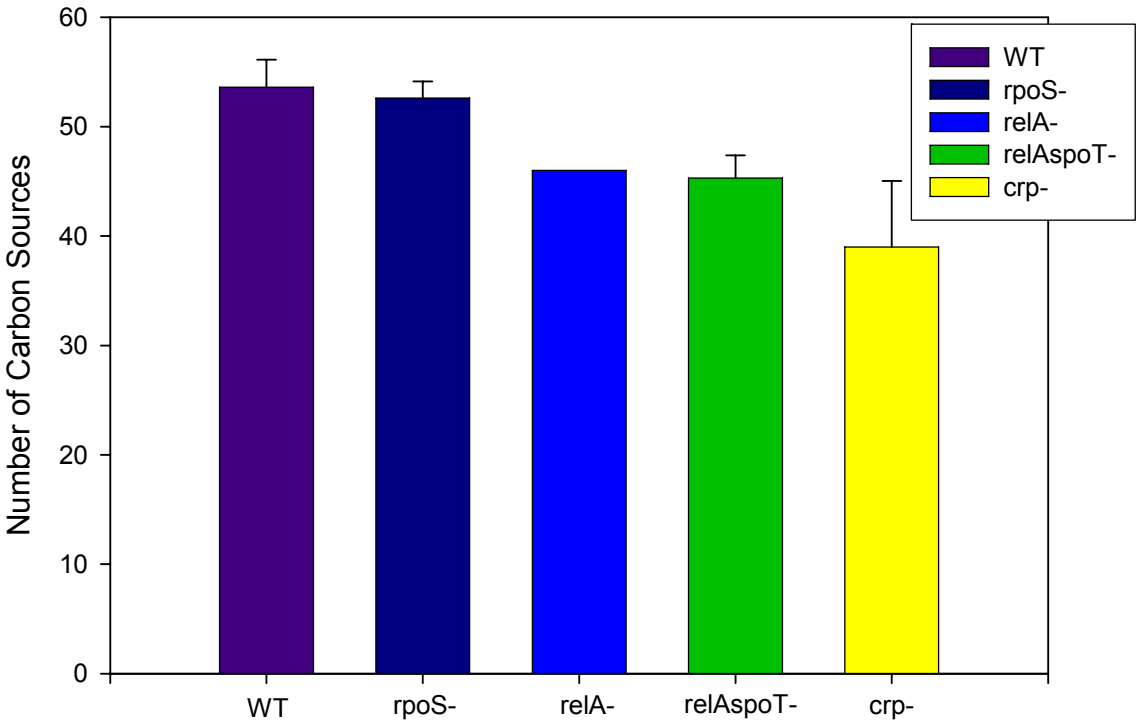
showing both the diauxic lag period when glucose is exhausted and the stationary phase when lactose is exhausted. Clusters 1-3 contain genes whose expression did not change significantly during diauxie. Cluster 4 contains stringent response genes. Cluster 5 contains stress survival genes. Cluster 6 contains *lac* operon and carbon scavenging genes. Clusters 7 and 8 contain nitrogen starvation genes.

Supplementary Figure 2.



Legend to Supplementary Figure 2. Nested Venn diagram illustrating overlapping regulatory control circuits and their constituent functional groups. Each functional group has a unique color; black does not signify a functional group. Numbers indicate the number of genes regulated within each group and the + or – designation indicates up- or down-regulation, respectively. The RelA circle encompasses both the RpoS and Crp regulons. *rmf* was highly up-regulated during the stringent response, but was not dependent upon RpoS or Crp.

Supplementary Figure 3.



Legend to Supplementary Figure 3. *relA* and *relA spoT* mutants exhibit impaired ability to adapt to a wide variety of carbon sources. GN2 microplates (Biolog) containing 95 different carbon sources were used to screen metabolic activity of

the mutant strains. Assays were conducted according to the manufacturer's instructions.

Appendix 2: Supplementary data to accompany Chapter 3.

This supplemental note is an extended consideration of the data presented in Supplemental Fig 1. The same data is summarized in Table 1 in Chapter 3.

Cell division, DNA replication, and DNA repair

Given the morphological anomalies associated with the lack of ppGpp, it is perhaps unsurprising that 15 cell division genes showed differential expression in the ppGpp^0 strain (Fig 7A). Of these, 9 were expressed >2-fold higher in the ppGpp^0 strain, while 6 were expressed >2-fold lower. Three genes involved in cell division were induced >2-fold in the WT in response to isoleucine starvation. These genes, *fic, dacC,* and *bolA*, are members of the RpoS-dependent general stress response (15, 22, 26). Among them, the *bolA* and *dacC* gene products are known to play a role in determining the coccoid morphology associated with stationary phase cells (15, 22). None of these genes were induced in the ppGpp^0 strain, likely because of failed induction/accumulation of RpoS and/or impacted ability of σ^S to compete for core RNAP (7, 14).

Most of the remaining cell division genes showed no change or modest down regulation in response to isoleucine starvation in the WT. Two genes were down regulated >2-fold in the WT (*yibP* and *tig*) in response to isoleucine starvation, but both of these genes were expressed >2-fold higher in the ppGpp^0 strain. Also notable among the cell division genes expressed higher in the ppGpp^0 strain, is *cgtA* (12.4-fold higher), which encodes an essential GTPase known to interact with SpoT in *Vibrio cholerae* (19). In that system, the activity of

CgtA apparently keeps the hydrolytic activity of SpoT high in order to maintain low amounts of ppGpp in nutrient rich environments. Similar results for *E. coli* have been recently reported (12). Furthermore, depletion of CgtA in *E. coli* was shown to result in defective chromosome partitioning and filamentous morphology (6). Though the implications of heightened expression of *cgtA* in the $ppGpp^0$ strain are not entirely clear, we note that since the lack of ppGpp results in filamentation even in the presence of presumably high levels of CgtA, it is possible that CgtA and ppGpp function in the same regulatory cascade to modulate normal cell division during times of nutrient limitation.

Sixteen genes involved in DNA replication showed a different pattern of expression in the $ppGpp^0$ strain, by comparison to the WT, with all but one gene (*cspD*) expressed more highly in the mutant (Fig. 7B). Many of these genes encode products directly involved in DNA replication including helicases (*recG, hrpB*), primase (*dnaG*, *priB*, *priA*), topoisomerases (*topB*, *parC*, *topA*, and *parE*), and DNA polymerase III, the primary replicative DNA polymerase (*dnaN*, *holD*, *holB*). *dnaA*, whose product controls replication initiation, was also expressed >2-fold higher in the $ppGpp^0$ strain. Taken together, the expression levels of these genes show a downward trend in the WT, consistent with decreased chromosome replication. No such trend is apparent in the $ppGpp^0$ strain with most of these genes being modestly or significantly induced. Recently, ppGpp has been shown to regulate DNA elongation through direct interaction with DNA primase in *B. subtilis* (28). The higher levels of expression observed in our experiments for genes involved in DNA replication, combined with possible direct

effects of ppGpp on DNA primase, are consistent with continued DNA replication in the ppGpp^0 strain despite the growth limitations imposed by amino acid starvation.

The WT strain did not exhibit increased expression of genes associated with the SOS response to DNA damage, however, this trend was quite apparent in the ppGpp^0 strain (Fig. 7C). Twenty-three DNA repair genes showed altered expression in the ppGpp^0 strain. The ppGpp^0 strain induced 16 DNA repair genes >2-fold in response to isoleucine starvation, including 13 members of the LexA regulon (5), i.e., *recN, uvrD, dinD*, *dinF, recA, ruvA, lexA*, *dinG, ydjQ, umuD*, *umuC, yebG,* and *sbmC*. Only 2 genes were more highly expressed in the WT, including *sbmC*, which encodes a DNA gyrase inhibitor known to protect cells against DNA damage (1). Interestingly, stationary phase induction of *sbmC* has been shown to be RpoS-dependent (18). Whereas the WT likely terminates DNA replication via a combination of DNA gyrase inhibition and lowered transcription of DNA replication genes, it appears that chromosome replication continues abnormally in the ppGpp^0 strain, ultimately resulting in DNA damage.

Nucleotide biosynthesis and degradation

Down regulation of nucleotide biosynthesis has been associated with the stringent response (2, 3, 25), and as expected, we observed this trend in isoleucine-starved WT cultures: 21 genes down regulated >2-fold. Nucleotide biosynthetic genes showed some of the greatest levels of down regulation in the entire data set, including *pyrI, pyrB,* and *carA* (down regulated 79-, 93-, and 78-

fold, respectively). Twenty-seven nucleotide biosynthetic genes showed abnormal expression in the strain lacking ppGpp (Fig. 7D); only 3 genes (*nrdA*, *nrdB*, and *thyA*) were expressed at a lower level, while the remaining 24 were expressed at levels >2-fold higher in the $ppGpp^0$ strain compared to the WT. While 9 of these genes were down regulated at least 2-fold in the $ppGpp^0$ strain, the greatest level of down regulation was only 23-fold (observed for *pyrI*) compared to 79-fold in the WT. Overall, the $ppGpp^0$ strain exhibited a greatly diminished ability to stringently down regulate genes involved in nucleotide biosynthesis. The strong down regulation of nucleotide biosynthetic genes in the WT reflects the greatly decreased need for nucleotides as cellular processes that consume nucleotides are curtailed (namely ribosome and chromosome synthesis). Given the continued production of rRNA and DNA in the $ppGpp^0$ strain, it might be expected that nucleotide pools would be diminished. In keeping with this notion, expression of *purR*, whose product autorepresses its own transcription (21), as well as a large number of genes involved in purine biosynthesis (16), was induced 4.5-fold, compared to the modest repression observed in the WT. The derepression of *purR* suggests that the hypoxanthine levels are low in the $ppGpp^0$ strain, implying that the supply of nucleotides had dwindled. Moreover, decreased nucleotide availability in the $ppGpp^0$ strain might stall DNA replication, thus prompting the DNA damage response noted above.

The WT showed both induction and down regulation of genes involved in nucleotide salvage/degradation (Fig. 7E), including 12 that were abnormally

expressed in the ppGpp^0 strain. Several of these genes were induced in the WT in response to isoleucine starvation, including *amn*, *udp*, *deoA*, and *add*. These four genes, and two more, *dgt* and *deoC*, were expressed >2-fold higher in the WT strain. All six of these genes are involved in degradation, i.e., catabolism of nucleotides to intermediates of central metabolism. Induction of these nucleotide degradation genes, down regulation of nucleotide biosynthesis genes (discussed above), and the observation that nucleotide utilization was enhanced as a result of isoleucine starvation in our Biolog assays (Fig. 5), lead us to conclude that the WT shifted from funneling precursors into nucleotide production during rapid growth, to catabolism of excess nucleotides upon growth arrest. Furthermore, the induction of the non-oxidative branch of the pentose phosphate pathway observed in the WT might further enhance the flux of nucleotide degradation products into central metabolism. These observations are consistent with stationary phase-induced nucleotide degradation and byproduct excretion described elsewhere (13, 20). The genes *codA*, *upp*, *cmk*, *gpt*, *apt*, and *gsk* were expressed >2-fold higher in the ppGpp^0 strain. Interestingly, all 6 of these genes are involved in the salvage of endogenous nucleotide precursors for the production of new nucleotides, as opposed to catabolism of nucleotides via the non-oxidative branch of the pentose phosphate pathway. This trend suggests that the ppGpp^0 strain actively attempted to salvage nucleotides in keeping with continued synthesis of nucleic acid. Thus, the expression data, combined with the loss of ability to utilize nucleotides as C-sources in the Biolog assays,

suggest that the switch from nucleotide production for nucleic acid synthesis to degradation of nucleotides is a ppGpp-dependent process.

Fatty acid biosynthesis, β-oxidation, cell wall/LPS biosynthesis, and glycogen synthesis

Phospholipids and their constituent fatty acids (FA) form a significant portion of the dry weight of *E. coli* (~9%) (17), and accordingly, cells invest a corresponding amount of carbon and energy in phospholipid biosynthesis (see (4) for a review). Seventeen genes involved in phospholipid biosynthesis showed abnormal expression in the ppGpp^0 strain (Fig. 7F). Considered in context, the *accC* and *accD* genes are involved in the initial production of malonyl-CoA from acetyl-CoA and CO_2, while *fabD*, *fabH*, *fabF*, and *acpT* gene products are involved in the cyclical elongation of FA. The product of the *plsC* gene catalyzes the second step in phosphatidic acid biosynthesis from glycerol-3-P. Phosphatidic acid is then activated to CDP-diacylglycerol by the *cdsA* gene product. The *plsX* gene product is also thought to play a role in phosphatidic acid synthesis. Of the 17 phospholipid metabolism genes under consideration here, 12 showed either modest down regulation or no change in the WT (*accC*, *mdoB*, *cdh*, *acpT*, *fabH*, *plsX*, *fabF*, *fabA*, *accD*, *fabD*, *plsC*, and *cdsA*). In the ppGpp^0 strain, all 12 of these genes were expressed at levels >2-fold higher than in the WT, including 5 (*plsX*, *fabF*, *fabA*, *plsC*, and *cdsA*) that were induced >2-fold in response to isoleucine starvation. Of the remaining 5 genes, 4 were significantly induced in the WT in response to isoleucine starvation (*pgsA, ybhO*,

uspA and *cfa*), but were expressed at lower levels in the ppGpp^0 strain. Induction of the *cfa* and *ybhO* genes in the WT reflects the conversion of unsaturated FAs in membrane lipids into cyclopropane derivatives and the accumulation of cardiolipin, both of which are associated with entry into stationary phase (9, 23, 27).

Six genes involved in β-oxidation of FA were aberrantly expressed in the ppGpp^0 strain (Fig. 7G), 3 of which (*yfcX*, *aidB*, and *fadE*) were induced in the WT, but not in the ppGpp^0 strain. The only gene expressed higher in the ppGpp^0 strain than in the WT was *fadD*, which encodes acyl-CoA synthetase. These data fit the known shift from large scale phospholipid synthesis to FA breakdown (4), concomitant with the assumption of smaller cell sizes during transition to stationary phase. Gene expression in the ppGpp^0 strain points in the opposite direction, with possibly enhanced synthesis of FA (described above) and lower-than-normal levels of FA degradation.

Though stringent control of peptidoglycan synthesis at the level of enzyme inhibition has been observed (11), the genes involved in synthesis of peptidoglycan were not comprehensively down regulated in the WT. However, 13 genes involved in peptidoglycan biosynthesis were abnormally regulated in the ppGpp^0 strain, with 9 of them expressed >2-fold higher than the WT (Fig. 7H). These included two genes for penicillin binding proteins (*mrdA* and *dacD*), and two N-acetylmuramoyl-L-alanine amidases (*amiC* and *amiA*) involved in post-division cell separation. *mltA* and *mpl,* whose products are involved with murein turnover, were also expressed higher in the ppGpp^0 strain. *mreC* which encodes

a subunit of the MreBCD transmembrane complex was down regulated in the WT, but not in the ppGpp^0 strain. Genes expressed lower in the ppGpp^0 strain include *murE* and *murF*, which catalyze the final cytoplasmic steps in the formation of peptidoglycan precursors. Taken together, these results suggest the ppGpp^0 strain may have an abnormally high rate of peptidoglycan turnover. We hypothesize that this response of the ppGpp^0 culture may be linked to the exhaustion of glutamate, which serves as a precursor for peptidoglycan synthesis.

Twenty-four genes involved in the synthesis of lipopolysaccharide and other components of the outer membrane (OM) were expressed abnormally in the ppGpp^0 strain and all but one of these (*gutQ*) were expressed >2-fold higher in the mutant when compared to the WT. Most of these genes are involved in O-antigen synthesis, lipid A synthesis, or synthesis of other LPS precursors or transporters. Moreover, 8 of these genes (*rfaS*, *lpxC*, *rfaH*, *htrB*, *rfaK*, *ddg*, *lpxK*, and *msbB*) were induced >2-fold in the ppGpp^0 strain in response to isoleucine starvation. Fourteen surface antigen and OM protein encoding genes were down regulated >2-fold in the WT in response to isoleucine starvation, while only one of these (*ugd*) was significantly down regulated in the ppGpp^0 strain. These results indicate that the down regulation of genes involved in LPS synthesis and other OM structures in response to isoleucine starvation is directly or indirectly dependent on ppGpp. When considered with the expression patterns observed for FA metabolism genes, the down regulation of OM component genes, including those for LPS synthesis, suggests that inner and outer membrane

assembly is largely halted in the WT in response to isoleucine starvation. Moreover, these processes likely continue unabated in the ppGpp^0 strain, adding to the abnormally high biomass produced by the strain lacking ppGpp.

If excess carbon is available during the transition to stationary phase, glycogen can be accumulated (as is the case under isoleucine starvation). Glycogen can account for up to 3% of *E. coli* dry cell weight, and serves as a carbon source during periods of carbon starvation. Accumulation of glycogen is known to be a ppGpp-dependent process (10, 24), and this trend is clearly observable in the array data, with seven genes involved in glycogen metabolism being expressed >2-fold lower in the ppGpp^0 strain (Fig. 7J). Moreover, *glgS*, known to be a member of the RpoS-dependent general stress response (8), was not induced in the ppGpp^0 strain as it was in the WT. Taken together, these results suggest that glycogen probably does not contribute to the higher biomass produced by the ppGpp^0 strain.

References

1. **Baquero, M. R., M. Bouzon, J. Varea, and F. Moreno.** 1995. sbmC, a stationary-phase induced SOS Escherichia coli gene, whose product protects cells from the DNA replication inhibitor microcin B17. Mol Microbiol **18:**301-11.

2. **Cashel, M., D. R. Gentry, V. J. Hernandez, and D. Vinella.** 1996. The stringent response, p. 1458-1496. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.

3. **Chang, D. E., D. J. Smalley, and T. Conway.** 2002. Gene expression profiling of *Escherichia coli* growth transitions: an expanded stringent response model. Mol Microbiol **45:**289-306.

4. **DiRusso, C. C., and T. Nystrom.** 1998. The fats of Escherichia coli during infancy and old age: regulation by global regulators, alarmones and lipid intermediates. Mol Microbiol **27:**1-8.

5. **Fernandez De Henestrosa, A. R., T. Ogi, S. Aoyagi, D. Chafin, J. J. Hayes, H. Ohmori, and R. Woodgate.** 2000. Identification of additional genes belonging to the LexA regulon in Escherichia coli. Mol Microbiol **35:**1560-72.

6. **Foti, J. J., N. S. Persky, D. J. Ferullo, and S. T. Lovett.** 2007. Chromosome segregation control by Escherichia coli ObgE GTPase. Mol Microbiol **65:**569-81.

7. **Gentry, D. R., V. J. Hernandez, L. H. Nguyen, D. B. Jensen, and M. Cashel.** 1993. Synthesis of the stationary-phase sigma factor sigma s is positively regulated by ppGpp. J Bacteriol **175:**7982-9.

8. **Hengge-Aronis, R., and D. Fischer.** 1992. Identification and molecular analysis of glgS, a novel growth-phase-regulated and rpoS-dependent gene involved in glycogen synthesis in Escherichia coli. Mol Microbiol **6:**1877-86.

9. **Hiraoka, S., H. Matsuzaki, and I. Shibuya.** 1993. Active increase in cardiolipin synthesis in the stationary growth phase and its physiological significance in Escherichia coli. FEBS Lett **336:**221-4.

10. **Ishiguro, E. E., and W. D. Ramey.** 1978. Involvement of the relA gene product and feedback inhibition in the regulation of DUP-N-acetylmuramyl-peptide synthesis in Escherichia coli. J Bacteriol **135:**766-74.

11. **Ishiguro, E. E., and W. D. Ramey.** 1976. Stringent control of peptidoglycan biosynthesis in Escherichia coli K-12. J Bacteriol **127:**1119-26.

12. **Jiang, M., S. M. Sullivan, P. K. Wout, and J. R. Maddock.** 2007. G-protein control of the ribosome-associated stress response protein SpoT. J Bacteriol **189:**6140-7.

13. **Kim, Y., C. M. Lew, and J. D. Gralla.** 2006. Escherichia coli pfs transcription: regulation and proposed roles in autoinducer-2 synthesis and purine excretion. J Bacteriol **188:**7457-63.

14. **Kvint, K., A. Farewell, and T. Nystrom.** 2000. RpoS-dependent promoters require guanosine tetraphosphate for induction even in the presence of high levels of sigma(s). J Biol Chem **275:**14795-8.

15. **Lange, R., and R. Hengge-Aronis.** 1991. Growth phase-regulated expression of bolA and morphology of stationary-phase Escherichia coli

cells are controlled by the novel sigma factor sigma S. J Bacteriol **173:**4474-81.

16. **Meng, L. M., M. Kilstrup, and P. Nygaard.** 1990. Autoregulation of PurR repressor synthesis and involvement of purR in the regulation of purB, purC, purL, purMN and guaBA expression in Escherichia coli. Eur J Biochem **187:**373-9.

17. **Neidhardt, F. C., J. L. Ingraham, and M. Schaechter.** 1990. Physiology of the Bacterial Cell, 1st ed, vol. Sinauer Associates, Inc., Sunderland, MA.

18. **Oh, T. J., I. L. Jung, and I. G. Kim.** 2001. The Escherichia coli SOS gene sbmC is regulated by H-NS and RpoS during the SOS induction and stationary growth phase. Biochem Biophys Res Commun **288:**1052-8.

19. **Raskin, D. M., N. Judson, and J. J. Mekalanos.** 2007. Regulation of the stringent response is the essential function of the conserved bacterial G protein CgtA in Vibrio cholerae. Proc Natl Acad Sci U S A **104:**4636-41.

20. **Rinas, U., K. Hellmuth, R. Kang, A. Seeger, and H. Schlieker.** 1995. Entry of Escherichia coli into stationary phase is indicated by endogenous and exogenous accumulation of nucleobases. Appl Environ Microbiol **61:**4147-51.

21. **Rolfes, R. J., and H. Zalkin.** 1990. Autoregulation of Escherichia coli purR requires two control sites downstream of the promoter. J Bacteriol **172:**5758-66.

22. **Santos, J. M., M. Lobo, A. P. Matos, M. A. De Pedro, and C. M. Arraiano.** 2002. The gene bolA regulates dacA (PBP5), dacC (PBP6) and ampC (AmpC), promoting normal morphology in Escherichia coli. Mol Microbiol **45:**1729-40.

23. **Taguchi, M., K. Izui, and H. Katsuki.** 1980. Augmentation of cyclopropane fatty acid synthesis under stringent control in Escherichia coli. J Biochem (Tokyo) **88:**1879-82.

24. **Taguchi, M., K. Izui, and H. Katsuki.** 1980. Augmentation of glycogen synthesis under stringent control in Escherichia coli. J Biochem (Tokyo) **88:**379-87.

25. **Turnbough, C. L., Jr.** 1983. Regulation of Escherichia coli aspartate transcarbamylase synthesis by guanosine tetraphosphate and pyrimidine ribonucleoside triphosphates. J Bacteriol **153:**998-1007.

26. **Utsumi, R., S. Kusafuka, T. Nakayama, K. Tanaka, Y. Takayanagi, H. Takahashi, M. Noda, and M. Kawamukai.** 1993. Stationary phase-

specific expression of the fic gene in Escherichia coli K-12 is controlled by the rpoS gene product (sigma 38). FEMS Microbiol Lett **113:**273-8.

27. **Wang, A. Y., and J. E. Cronan, Jr.** 1994. The growth phase-dependent synthesis of cyclopropane fatty acids in Escherichia coli is the result of an RpoS(KatF)-dependent promoter plus enzyme instability. Mol Microbiol **11:**1009-17.

28. **Wang, J. D., G. M. Sanders, and A. D. Grossman.** 2007. Nutritional control of elongation of DNA replication by (p)ppGpp. Cell **128:**865-75.

Supplemental Fig 1

Legend for all:



A



Cell Division

GENE	Locus	WT	relAspoT	Mutant-WT
xerC	b3811	-0.398	-1.453	-1.055
ybiP	b3813	-1.576	-0.082	1.494
fig	b0436	-1.339	0.514	1.853
gidA	b3741	-0.721	1.148	1.869
tolC	b3035	0.29	-1.174	-1.464
ftsX	b3462	0.073	1.262	1.189
sufI	b3017	-0.244	1.02	1.264
cedA	b1731	0.13	1.262	1.132
yjbQ	b4058	0.721	-0.696	-1.417
cptA	b3183	-0.972	2.634	3.606
nlpI	b3163	0.171	2.409	2.238
rnnJ	b3179	0.177	1.231	1.054
bolA	b0435	3.239	0.314	-2.925
fic	b3361	3.007	0.233	-2.774
dacC	b0639	2.704	0.13	-2.574

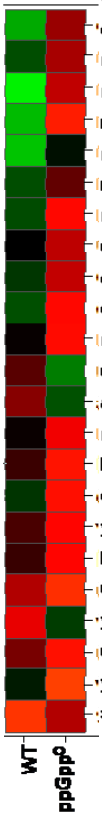
B



DNA Replication

GENE	Locus	WT	relAspoT	Mutant-WT
priB	b4201	-2.838	0.298	3.136
recG	b3652	-1.826	-0.289	1.537
topB	b1783	-1.445	0.462	1.907
hda	b2406	-1.166	0.348	1.501
holD	b4372	-1.369	0.559	1.928
parC	b3019	-0.912	1.053	1.965
dnaG	b3066	-1.117	1.072	2.189
topA	b1274	-0.39	1.289	1.679
priA	b3035	-0.848	0.589	1.437
parE	b3030	-0.688	0.417	1.105
dnaN	b3701	-0.305	0.993	1.298
holB	b1099	-0.275	1.61	1.885
hrpB	b0148	-0.415	0.757	1.172
dnaA	b3702	-0.167	1.069	1.236
cspD	b0880	3.875	0.431	-3.444
cspF	b1558	-0.438	3.979	4.417

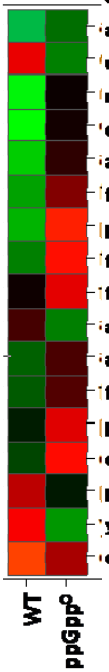
C



DNA Repair

GENE	Locus	WT	relAspoT	Mutant WT
dinI	b1081	-1	0.901	1.901
ruvC	b1863	-0.453	0.987	1.44
recF	b3700	-1.428	1.135	2.561
mutM	b3635	-1.103	2.08	3.183
recQ	b3822	-1.157	-0.069	1.088
mutL	b4170	-0.454	0.581	1.035
recN	b2616	-0.442	1.661	2.103
uvrD	b3813	0	1.113	1.113
dinD	b3645	-0.318	1.145	1.463
dinF	b4044	-0.485	1.65	2.115
recA	b2699	0.046	1.72	1.674
uvrC	b1913	0.522	-0.74	-1.262
alkA	b2088	0.795	-0.481	-1.276
ruvA	b1861	0.081	1.427	1.308
lexA	b4043	0.347	1.899	1.552
dinG	b0799	-0.307	1.803	2.11
ydjQ	b1741	0.43	1.587	1.157
ligA	b2411	0.334	1.598	1.262
umuD	b1183	1.045	2.462	1.417
ynel	b1525	1.37	-0.357	-1.727
umuC	b1184	0.708	1.837	1.129
yebG	b1848	-0.154	2.833	2.987
sbmC	b2009	2.574	1.049	-1.525

F Fatty Acid/Phospholipid Biosynthesis



GENE	Locus	WT	relAspoT	Mutant-WT
accC	b3256	-3.088	-0.648	2.44
uspA	b3495	1.379	-0.757	-2.136
mdoB	b4359	-1.722	0.068	1.79
cdh	b3818	-1.537	0.117	1.654
acpT	b3475	-1.206	0.274	1.48
fabH	b1091	-0.953	0.765	1.718
plsX	b1090	-1.06	2.155	3.215
fabF	b1095	-0.756	1.794	2.55
fabA	b0954	0.083	1.344	1.261
acpS	b2563	0.408	-0.751	-1.159
accD	b2316	-0.571	0.441	1.012
fabD	b1092	-0.535	0.484	1.019
plsC	b3018	-0.165	1.321	1.486
cdsA	b0175	-0.408	1.719	2.127
pgsA	b1912	1.109	-0.143	-1.252
ybhO	b0789	1.465	-0.894	-2.359
cfa	b1661	2.885	0.991	-1.874

G Fatty Acid β-oxidation



GENE	Locus	WT	relAspoT	Mutant-WT
fadB	b3846	0.343	-0.951	-1.294
yfcY	b2342	0.882	-0.456	-1.338
fadD	b1805	0.166	2.151	1.985
aldB	b4187	1.648	0.012	-1.636
fadE	b0221	1.735	0.467	-1.268
yfcX	b2341	1.386	-0.548	-1.934

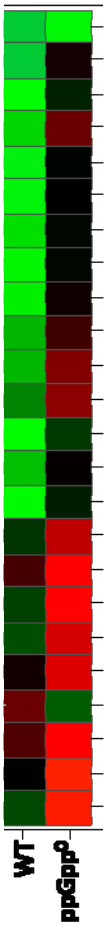
H Peptidoglycan



GENE	Locus	WT	relAspoT	Mutant-WT
murF	b0086	-0.753	-2.032	-1.279
murE	b0085	-0.221	-1.852	-1.631
mreC	b3250	-2.308	-0.029	2.279
mrdA	b0635	-1.904	0.703	2.607
amiC	b2817	-0.725	0.877	1.602
uppS	b0174	-0.874	1.036	1.91
yfiC	b2575	-0.91	0.724	1.634
mpl	b4233	-0.862	0.351	1.013
amiA	b2435	-0.377	0.824	1.201
dacD	b2010	0.008	1.026	1.018
mltA	b2813	-0.324	1.633	1.957
ampD	b0110	1.37	-1.077	-2.447
slt	b4382	0.74	-0.438	-1.178

I

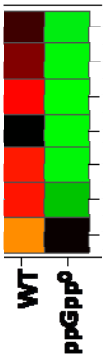
Surface Antigen/Outer Membrane



GENE	Locus	WT	relAspoT	Mutant-WT
ugd	b2028	-2.663	-1.639	1.044
htrL	b3616	-2.705	0.12	2.825
rfbB	b2041	-1.623	-0.201	1.322
lpxB	b0182	-1.259	0.63	1.889
rfbX	b2037	-1.805	-0.026	1.779
rfbC	b2036	-1.735	-0.006	1.729
wbbL	b2031	-1.319	-0.046	1.273
kdtA	b3633	-1.453	-0.041	1.412
wbbH	b2035	-1.414	0.094	1.508
wzzE	b3785	-1.075	0.376	1.451
rfe	b3784	-1.077	0.765	1.842
rfaZ	b3624	-0.783	0.828	1.611
glf	b2036	-1.519	-0.336	1.183
wbbK	b2032	-1.141	0.029	1.17
ybbF	b0524	-1.466	-0.176	1.312
rfaS	b3629	-0.306	1.125	1.431
lpxC	b0096	0.402	1.506	1.104
rfaH	b3642	-0.383	1.552	1.935
htrB	b1054	-0.467	1.258	1.725
rfaK	b3623	0.106	1.306	1.2
gutQ	b2708	0.828	-0.568	-1.196
ddg	b2378	0.463	1.515	1.052
malB	b1855	-0.014	2.211	2.225
lpxK	b0915	-0.416	2.061	2.479

J

Glycogen Metabolism



GENE	Locus	WT	relAspoT	Mutant-WT
glgX	b3431	0.367	-1.906	-2.273
glgB	b3432	0.768	-1.667	-2.435
glgP	b3428	1.588	-1.536	-3.124
malP	b3417	-0.009	-1.595	-1.586
glgC	b3430	2.02	-1.685	-3.705
glgA	b3429	1.928	-1.155	-3.083
glgS	b3049	4.416	0.057	-4.358