Feedback Regulation of the Heat Shock Response in *E. coli*

H. Kurata *, H. El-Samad †, T.-M. Yi ‡, M. Khammash ‡, J. Doyle ‡

* Department of Biochemical Science and Engineering, Kyushu Institute of Technology
  Iizuka, 820-8502, Japan
† Electrical and Computer Engineering, Iowa State University
  Ames, Iowa 50011
‡ Control and Dynamical Systems, California Institute of Technology
  Pasadena, CA 91125

Abstract

Survival of organisms to extreme conditions has necessitated the evolution of stress response networks that detect and respond to environmental changes. Among the extreme conditions that cells must face is the exposure to higher than normal temperatures. In this paper, we propose a detailed biochemical model that captures the dynamical nature of the heat-shock response in *Escherichia coli*. Using this model, we show that both feedback and feedforward control are utilized to achieve robustness, performance, and efficiency of the response to the heat stress. We discuss the evolutionary advantages that feedback confers to the system, as compared to other strategies that could have been implemented to get the same performance.

1 Introduction

Organisms are subject to a plethora of environmental and metabolic stress conditions including physiochemical factors such as heat shock, metabolically harmful substances such as ethanol, and viral infection, to mention but a few. These sources of stress damage a wide variety of cellular structures and cause the proteins to denature or unfold. Denatured proteins aggregate into insoluble, nonfunctional inclusions, which disrupt normal cellular functions. Regulatory systems have evolved to detect such damage and initiate a response that increases the resistance of cells to the damage and aid in its repair. The most important of these protective systems is the heat shock response [1]. A wide variety of organisms from bacteria to man respond to environmental perturbations such as an increase in temperature by rapidly inducing the synthesis of a small set of highly conserved heat shock proteins (*hsp*). Many of these proteins serve as molecular chaperones that assist in the refolding of denatured proteins; other *hsp* are proteases that degrade and remove the denatured proteins. Disrupting this response renders the cell sensitive to the heat disturbance. At 30°C, *hsp* represent less than 5% of the total protein content in *E. coli*. After a shift to 42°C, the synthesis of the heat shock proteins are induced up to 15-fold [2]. The cell must maintain a fine balance between the protective effect of the *hsp* and the metabolic burden of overexpressing these proteins. What type of control strategy have the bacteria implemented to ensure both optimal growth at physiological temperatures and survival at elevated temperatures? To answer this question and others, we have devised a mathematical model of the heat shock response in *E. coli*. Although there is an abundance of experimental data on the heat shock system, there has been very few attempts to integrate this information into a quantitative understanding. One such attempt, that forms the basis for this paper, is the model proposed by Kurata et al. [4]. This model captures the major features of the heat shock response. The model proposed in this paper involves some modifications to [4]. This includes the addition of a detailed model for the protein folding dynamics and few other structural modifications to accommodate new biological information uncovered by some of the recent experimental results. The modified model is able to reproduce the dose-response of the production of *hsp* as a function of temperature, while taking into account all the known biology of the heat shock response. In addition, the model can explain the altered phenotype of mutants in which genes involved in the pathway have been deleted. We have identified both a feedforward and feedback component to the control mechanism that together ensure a fast, efficient and robust response.

2 Biology of the Heat Shock Response in *E. coli*

In *E. coli*, much of the regulation of the concentrations of proteins occurs at the level of transcription. The en-
zyme RNA polymerase (RNAP) bound to a regulatory sigma factor recognizes the promoter and then transcribes specific genes into messenger RNA (mRNA). The mRNA is translated by the ribosome into protein. At physiological temperatures (30°C to 37°C), RNAP is bound to the major sigma factor $\sigma^{70}$. The RNAP : $\sigma^{70}$ complex transcribes the genes necessary for growth at normal temperatures. When E. coli are exposed to high temperatures, the special heat shock sigma factor, $\sigma^{32}$ encoded by the $rpoH$ gene, is rapidly induced. $\sigma^{32}$ binds to RNA polymerase and directs the transcription of a small set (approximately 20) of heat shock genes [2]. The heat shock genes encode for molecular chaperones (GroEL, DnaK, DnaJ, GroES, GrpE, etc.) that are involved in refolding denatured proteins. Another class of heat shock proteins are proteases (Lon, FtsH, etc.) that function to degrade unfolded proteins. In an $rpoH$ null mutant that does not make $\sigma^{32}$, the heat shock proteins are not induced and the cells are viable only at temperatures below 20°C [9]. There are two mechanisms by which $\sigma^{32}$ levels are increased when the temperature is raised. First, the translation rate of the $rpoH$ mRNA increases immediately, resulting in a fast 10-fold increase in the concentration of $\sigma^{32}$ [5]. Second, $\sigma^{32}$ is recognized and sequestered by the hsp chaperone DnaK. The concentration of the $\sigma^{32} : $DnaK$ complex depends on the amount of DnaK that is bound to unfolded proteins. Raising the temperature produces an increase in the cellular levels of unfolded proteins which titrate DnaK away from $\sigma^{32}$, resulting in more $\sigma^{32}$ that is capable of binding to RNA polymerase and initiating the transcription of the heat shock genes [6]. The accumulation of high levels of heat shock proteins leads to the down regulation of the response. During this phase, the abundant chaperones efficiently refold most of the denatured proteins thereby decreasing the pool of unfolded protein, freeing up DnaK to sequester $\sigma^{32}$ from RNA polymerase. In addition, heat shock proteases such as FtsH degrade $\sigma^{32}$. The result is a decrease in the concentration of $\sigma^{32}$ to a new steady state concentration that is dictated by the balance between the temperature-dependent translation of the $rpoH$ mRNA and the level of $\sigma^{32}$ activity modulated by the hsp chaperones and proteases acting in a negative feedback fashion. These relationships are depicted in Figure 1.

3 The Model

Although a wide range of chaperones is available in the cell, we chose DnaK as a representative of the chaperone team. We considered FtsH as the major protease that degrades $\sigma^{32}$, although we assumed that its action is mediated through interaction with the $\sigma^{32} : $DnaK$ complex [11]. All other proteases that act through chaperones to degrade $\sigma^{32}$ are lumped into a separate term. It is believed that there are some proteases (like HsIVU for example) that degrade $\sigma^{32}$ independently of chaperones and in a temperature dependent fashion [9]. We considered HsIVU as the representative of these proteases. We assumed that all mRNA are degraded at the same rate, and that all proteins are degraded at the same rate. Transcription and translation of proteins were assumed to follow simple first order dynamics (law of mass action). Binding rates (association and dissociation) between proteins or between proteins and specific DNA promoters were assumed to be fast compared with the rate of synthesis and degradation of mRNAs and proteins. Therefore, we assumed that they reach their steady-state very fast compared to other reactions in the system. This is known as the Two-Phase Partitioning Method [3]. We also use mass-balance equations to relate the total quantity of a species in the system to its free concentration and the concentration of the different compounds where it appears. The resulting model is a set of Differential Algebraic Equations (DAEs), which are of the form:

$$X(t) = F(t; X; Y)$$
$$0 = G(t; X; Y)$$
$$X(t = t_0) = X_0$$
$$Y(t = t_0) = Y_0$$

where $X$ is a 11-dimensional vector whose elements are the differential variables and $Y$ is a 20-dimensional vector whose elements are algebraic variables. This form is known as a semi-explicit DAE. These equations are detailed in the Appendix. Numerical simulations were carried out using the DASSL DAE. For illustration purposes, a simplified block diagram of the model can be seen in Figure 2. In this diagram, the
folding dynamics were taken as a disturbance to the free DnaK pool. The parameters used in the model were determined using various sources. The binding and degradation constants were mostly taken from the literature of the heat shock response and from personal communications. The synthesis rates for different proteins, $\sigma$ factors, and chaperones were tuned to produce biologically plausible numbers of these quantities in the cell.

Figure 2: Control diagram of the heat shock response in *E. coli*

![Control diagram of the heat shock response in *E. coli*](image)

4 Simulation Results and Discussions

The level, activity and stability of $\sigma^{32}$ are the key regulatory elements of the heat shock response. Upon exposure to heat, the time profile of the level of $\sigma^{32}$ shows a fast but transient increase due to increased translation and stabilization of the otherwise very unstable $\sigma^{32}$. The transient stabilization of $\sigma^{32}$ is a result of the large number of unfolded proteins that titrate $\sigma^{32}$ away from DnaK. Therefore, $\sigma^{32}$ becomes inaccessible for degradation and more free to bind to RNAP. The increased level of $\sigma^{32}$ leads to the synthesis of large numbers of molecular chaperones and proteases, that in turn act as a negative feedback on the level of $\sigma^{32}$. Chaperones refold proteins efficiently and rapidly. After a sharp transient (Figure 3), folded protein levels rise again, leaving the chaperones free to bind to $\sigma^{32}$, thereby decreasing the production of unneeded heat shock proteins. $\sigma^{32}$ is also degraded by FtsH through interaction with DnaK and by other proteases. This negative feedback leads to the decrease in the level of $\sigma^{32}$ until it reaches a steady state 3-4 fold larger than that at low temperature. This could be seen in Figure 3 where the simulation results are shown for the the cellular level of $\sigma^{32}$ (temperature increases at t=400 min). The result of this simulation agrees qualitatively with experimental results [2, 6]. Figure 3 also shows the levels of the DnaK chaperones after heat shock. Upon temperature increase, the number of chaperones overshoots slightly and then reaches a steady state at the high temperature. This behavior has also been established experimentally in the heat shock literature [1, 2, 6].

In order for a biochemical model to be valid, it should be able to reproduce both the normal and the altered phenotype behaviors. As mentioned in the previous discussion, our model is able to reproduce the levels of the various quantities involved in the heat shock response in *E. coli*. The model is also able to reproduce the behavior of different types of mutant cells exposed to heat. For example, it is able to reproduce the levels of $\sigma^{32}$ in the *FtsH* null mutant. It has been established that in *FtsH*-deficient mutant cells, there is accumulation of $\sigma^{32}$ at normal temperatures ($\sim$ 20-fold the normal level) as well as at high temperatures ($\sim$ 30-fold the normal level) [11]. This is due to the stabilization of $\sigma^{32}$ in the absence of *FtsH*. The $\sigma^{32}$ level in *FtsH*-deficient mutant is shown in Figure 4. It is worth noting here that the increased numbers of $\sigma^{32}$ cause the production of a larger number of chaperones as compared to wild type ($\sim$ 11000) (Figure 4), which in turn leads to good folding of proteins (Figure 4). However, this comes at the expense of a huge metabolic burden on the cell.

Based on these simulation results, an immediate conclusion can be drawn. Although the folding of proteins is the main objective of this whole machinery,
it cannot be done at the expense of metabolic cost. In the absence of *FtsH*, protein folding is enhanced but the metabolic cost that this entails on the cell, through the upregulation of chaperones, is extensive. Therefore, both the feedforward and the feedback components are necessary. While feedforward control is providing a fast response and improving the system performance through efficient folding of proteins, feedback is stabilizing the system and increasing its efficiency, through degradation and sequestration of $\sigma^{32}$ respectively. There seems to be an ongoing optimization where neither performance, nor efficiency, are sacrificed. This is very important to cell viability. We were able to show that the feedback structure confers to the system both remarkable stability and robustness against parametric uncertainty. Robustness is of extreme importance in cellular biochemical networks where uncertainty is an inherent property. The robustness issue will be detailed elsewhere.

One may wonder about the evolutionary advantage of the above described architecture, as compared to other possible arrangements for dealing with higher than normal temperatures. For example, what is the possible evolutionary advantage of the feedback regulated heat-shock response over cells with thicker walls and supracellular structures that would shield the protein content from heat? What are the possible advantages over proteins that are naturally more robust to increased temperatures? Such proteins would conceivably function without denaturing at wide temperature ranges. Although these defense mechanisms against heat are appealing and are in fact implemented to some extent in thermophilic organisms, they do have drawbacks that make them poor alternatives to the existing architecture. Thicker walls, for example, introduce a whole set of unjustified complexities in organisms growing at normal temperatures, and do not offer protection without interfering with the influx of nutrients and outflow of catabolites and without compromising on protein catalytic activity. Robust proteins that do not denature at higher temperature are also less desirable, due to their lower functional performance at normal growth temperatures. It has been established experimentally that protein residues that contribute to catalysis or ligand binding are not optimal for thermal stability. As an illustration, thermostable proteins from thermophilic bacteria are often less active at low temperatures than their mesophilic counterparts [13]. Therefore, the best strategy from a design perspective seems to be to achieve high performance, not through robust components, but through barely stable and imprecise components, and then to achieve robustness through added feedback control loops. It is interesting that this is also a strategy that is used extensively in the design of manmade engineering systems.

Finally, we would like to point out that although our analysis is carried out in a deterministic setting, the small number of the $\sigma^{32}$ molecules might suggest some stochastic effects at play. This issue arises in small systems where the number of molecules cannot be approximated as continuously varying quantities that vary deterministically with time. Instead, one should adopt a probabilistic approach. In order to address this issue, we used the Stochastic Simulation Algorithm developed by Gillespie [14] to simulate our model in a complete stochastic setting. Through this approach, we could show that stochastic fluctuations do not alter the general behavior described by the model. These fluctuations are essentially filtered by the transcription and translation machinery. These results will be presented elsewhere.

5 Conclusion

It is hard to think of a cell or an organism that does not get exposed to stressors many times during its life. Learning how to better cope with these stressors is important since it gives the organism an evolutionary advantage. The investigation of the mechanisms that has evolved to deal with stress is of tremendous importance, not only because it would uncover the mechanisms of cell resistance, but also because it provides the basis for many practical applications in
biotechnology and medicine. Bearing this in mind, we attempted to model the behavior of *E. coli* under heat stress and identify, at least qualitatively, the major feedback loops that confer to the system its stability, performance and efficiency. The result of this analysis is a feedforward-feedback interconnection that is fully justified by control theory and verified by the biology of the heat shock response.

Acknowledgments

The authors would like to acknowledge support from the EPRI/DOD Complex Interactive Network Program, the Kitano ERATO Complex project, the New Energy and Industrial Technology Development Organization (NEDO), and ISU SPRIGS Grant.

6 Appendix

6.0.1 Notation: In the model equations, the following notations are followed:

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>[X]</td>
<td>Total concentration of species X</td>
</tr>
<tr>
<td>[X]_f</td>
<td>Free concentration of species X</td>
</tr>
<tr>
<td>X : Y</td>
<td>Complex formed by the binding of X and Y</td>
</tr>
<tr>
<td>mRNA(X)</td>
<td>Messenger RNA that encodes for protein X</td>
</tr>
</tbody>
</table>

Also these abbreviations are used:

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pg</td>
<td>Promoter for housekeeping genes</td>
</tr>
<tr>
<td>ph</td>
<td>hsp gene promoter</td>
</tr>
<tr>
<td>T</td>
<td>Temperature</td>
</tr>
</tbody>
</table>

6.0.2 Transcription and Translation Equations:

\[
\begin{align*}
\frac{d[mRNA(DnaK)]}{dt} &= K_{tr1} \cdot [\sigma^{32} : RNAP : ph] \\
&- \alpha_{mRNA} \cdot [mRNA(DnaK)] \\
\frac{d[DnaK]}{dt} &= K_{TL} \cdot [mRNA(DnaK)] \\
&- \alpha_{prot} \cdot [DnaK] \\
\frac{d[mRNA(FtsH)]}{dt} &= K_{tr2} \cdot [\sigma^{32} : RNAP : ph] \\
&- \alpha_{mRNA} \cdot [mRNA(FtsH)] \\
\frac{d[FtsH]}{dt} &= K_{TL} \cdot [mRNA(FtsH)] \\
&- \alpha_{prot} \cdot [FtsH] \\
\frac{d[mRNA(protease)]}{dt} &= K_{tr3} \cdot [\sigma^{32} : RNAP : ph] \\
&- \alpha_{mRNA} \cdot [mRNA(protease)] \\
\frac{d[protease]}{dt} &= K_{TL} \cdot [mRNA(protease)] \\
&- \alpha_{prot} \cdot [protease]
\end{align*}
\]

6.0.3 Algebraic Binding Equations:

\[
\begin{align*}
[\sigma^{70} : RNAP] &= K_{1} \cdot [\sigma^{70}] \cdot [RNAP] \\
[\sigma^{32} : RNAP] &= K_{2} \cdot [\sigma^{70}] \cdot [RNAP] \\
[RNAP : D] &= K_{3} \cdot [RNAP] \cdot [D] \\
[\sigma^{32} : DnaK : FtsH] &= K_{4} \cdot [\sigma^{32}] \cdot [DnaK] \cdot [FtsH] \\
[\sigma^{32} : DnaK] &= K_{5} \cdot [\sigma^{32}] \cdot [DnaK] \\
[\sigma^{32} : protease] &= K_{6} \cdot [\sigma^{32}] \cdot [DnaK] \cdot [protease] \\
[\sigma^{32} : HslVU] &= K_{7} \cdot [\sigma^{32}] \cdot [HslVU] \\
[Panfolded] &= K_{8} \cdot [Panfolded] \cdot [DnaK] \\
[\sigma^{32} : RNAP : ph] &= K_9 \cdot [\sigma^{32}] \cdot [RNAP] \\
& \cdot ([ph]_I - [\sigma^{32}] \cdot [RNAP : ph]) \\
[\sigma^{70} : RNAP : pg] &= K_{10} \cdot [\sigma^{70}] \cdot [RNAP] \\
& \cdot ([pg]_I - [\sigma^{70}] \cdot [RNAP : pg]) \\
[\sigma^{70} : RNAP : D] &= K_{11} \cdot [\sigma^{70}] \cdot [RNAP] \cdot [D] \\
[\sigma^{32} : RNAP : D] &= K_{12} \cdot [\sigma^{32}] \cdot [RNAP] \cdot [D]
\end{align*}
\]

6.0.4 Mass Balance Equations:

\[
\begin{align*}
[RNAP] &= [RNAP_I] + [\sigma^{70}] \cdot [RNAP] \\
&+ [\sigma^{32}] \cdot [RNAP] + [RNAP : D] \\
&+ [\sigma^{70} : RNAP : D] + [\sigma^{32} : RNAP : D] \\
&+ [\sigma^{70} : RNAP : pg] + [\sigma^{32} : RNAP : ph] \\
[\sigma^{70}] &= [\sigma^{70}_I] + [\sigma^{70}] \cdot [RNAP] \\
&+ [\sigma^{32}] \cdot [RNAP : D] + [\sigma^{70} : RNAP : pg] \\
[\sigma^{32}] &= [\sigma^{32}_I] + [\sigma^{32}] \cdot [RNAP : D] \\
&+ [\sigma^{70} : RNAP] + [\sigma^{32} : RNAP : D] \\
&+ [\sigma^{32} : DnaK : FtsH] + [\sigma^{32}] \cdot [DnaK]
\end{align*}
\]
6.0.5 Parameter values:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_1$</td>
<td>$10^9$ M$^{-1}$</td>
</tr>
<tr>
<td>$K_2$</td>
<td>$10^7$ M$^{-1}$</td>
</tr>
<tr>
<td>$K_3$</td>
<td>$10^7$ M$^{-1}$</td>
</tr>
<tr>
<td>$K_4$</td>
<td>$10^7$ M$^{-1}$</td>
</tr>
<tr>
<td>$K_5$</td>
<td>$5 \times 10^6$ M$^{-1}$</td>
</tr>
<tr>
<td>$K_6$</td>
<td>$10^6$ M$^{-1}$</td>
</tr>
<tr>
<td>$K_7$</td>
<td>$10^6$ M$^{-1}$</td>
</tr>
<tr>
<td>$K_8$</td>
<td>$10^6$ M$^{-1}$</td>
</tr>
<tr>
<td>$K_9$</td>
<td>$10^6$ M$^{-1}$</td>
</tr>
<tr>
<td>$K_{10}$</td>
<td>$10^6$ M$^{-1}$</td>
</tr>
<tr>
<td>$K_{11}$</td>
<td>$10^6$ M$^{-1}$</td>
</tr>
<tr>
<td>$K_{12}$</td>
<td>$10^6$ M$^{-1}$</td>
</tr>
<tr>
<td>$K_{tr1}$</td>
<td>$100$ min$^{-1}$</td>
</tr>
<tr>
<td>$K_{tr2}$</td>
<td>$20$ min$^{-1}$</td>
</tr>
<tr>
<td>$K_{tr3}$</td>
<td>$13$ min$^{-1}$</td>
</tr>
<tr>
<td>$K_{tr4}$</td>
<td>$30$ min$^{-1}$</td>
</tr>
<tr>
<td>$K_{tr5}$</td>
<td>$25$ min$^{-1}$</td>
</tr>
<tr>
<td>$K_{trf}$</td>
<td>$20$ min$^{-1}$</td>
</tr>
<tr>
<td>$v(T)$</td>
<td>0.05 at $37^\circ$C &amp; 0.35 x $5$ at $42^\circ$C</td>
</tr>
<tr>
<td>$K(T)$</td>
<td>75 at $37^\circ$C &amp; $150$ at $42^\circ$C min$^{-1}$</td>
</tr>
<tr>
<td>$K_{fold}$</td>
<td>$15 \times 10^3$ min$^{-1}$</td>
</tr>
<tr>
<td>$\alpha_{mRNA}$</td>
<td>0.5 min$^{-1}$</td>
</tr>
<tr>
<td>$\alpha_{prot}$</td>
<td>0.03 min$^{-1}$</td>
</tr>
<tr>
<td>$\alpha_{prot, T}$</td>
<td>0.5 at $T=37^\circ$ &amp; 1 at $T=42^\circ$ min$^{-1}$</td>
</tr>
<tr>
<td>$\alpha_{HsVU}(T)$</td>
<td>0.7 at $T=37^\circ$ &amp; 2 at $T=42^\circ$ min$^{-1}$</td>
</tr>
</tbody>
</table>

[RNA] = Total RNA core enzyme concentration = 5.1054 x $10^{-6}$ M

[D$_1$] = Nonspecific DNA binding site = 1.18 x $10^{-2}$ M

$\sigma_0$ = $1.778 \times 10^{-6}$ M

$pp$ = $1.016 \times 10^{-5}$ M

References


