Complexity in Regulation Generates Robustness in Bacterial Molecular Networks

Hiroyuki Kurata [1,2] and Kazunari Taira[1,3]

[1]Department of Chemistry and Biotechnology
The University of Tokyo, Tokyo 113-8656, JAPAN
Tel.:81-3-5841-7340, Fax: 81-3-5841-7279, E-mail: hkurata@bio.t.u-tokyo.ac.jp / taira@chembio.t.u-tokyo.ac.jp
[2] Research Advisor, Systems Biology Group, ERATO Kitano Symbiotic Systems Project, JST,
M-31 Suite 6A 6-31-15 Jingumae Shibuya-ku Tokyo, 150-0001
[3] National Institute for Advanced Interdisciplinary Research, AIST, MITI Tsukuba Science City 305-8562
*Corresponding to Hiroyuki Kurata

There is a question of how one can understand a biological system composed of heterogeneous and interactive subsystems. If the interaction among subsystems is sensitive, one cannot analyze smaller subsystems separately. If their interaction shows robustness, one can extract a subsystem out of the whole system, analyzing them one by one. In the heat shock response, σ^{32} (encoded the *rpoH* gene) plays a major role in controlling expression of heat shock protein genes that encode chaperones and proteases. The level of active σ^{32} is regulated by complex mechanisms: chaperone-mediated regulation of σ^{32} activity and stability, thermoregulated-induction of the *rpoH* mRNA, and protease-mediated degradation of σ^{32}. The numerical framework model shows that complexity in σ^{32} regulation generates robustness in the *E. coli* heat shock response, thereby increasing the robustness of the interconnected factors among subsystems. Complexity seems to impede isolating a smaller subsystem out of a whole biological system. Actually, complexity generates robustness among subsystems, thereby making it possible to extract a smaller subsystem out of the whole system and analyze it separately.

1. Introduction

Genome sequencing projects and further systematic functional analyses of complete gene sets will enable a computer to synthesize a biological system for a wide range of model organisms at a molecular interaction. Molecular interaction networks of biological systems, such as cell cycle [1], bacterial chemotaxis [2], λ-phage gene regulatory circuit [3], and circadian clock, have been extensively elucidated and analyzed using a mathematical model. Computer simulation has been carried out to test current hypotheses of the mechanisms, and predict the performance under a virtually limitless variety of conditions. The problem in simulating a biological system is that it is quite difficult to determine all of the biochemical parameters. Since the parameters vary with intracellular or extracellular conditions, it is hard to measure them by means of experiments and most of them will remain unknown.

On the other hand, control theory has been applied to analyze a biological system without insisting on the exact values of parameters in order to extract a design principle underlying molecular processes. The key words characterizing a biological system are complexity, robustness, and feedback control. A well-characterized gene regulatory circuit of phage λ has been extensively studied to elucidate the robustness generated by feedback controls[3]. In λ phage, negative and positive feedback controls for expression of λ-repressor (cI) and of repressor and operator (cro) play a key role in selecting either bacteriolysis or lysogeny. The mathematical simulation of the λ phage gene regulatory circuit predicted that these double feedback controls held the concentrations of cI and cro proteins within a certain region against the genetic changes in components or environmental perturbation, making the circuit system robust. This robust property was also experimentally demonstrated[4]. In bacterial chemotaxis, a great variation in the biochemical parameters varied the adaptation time, but the property of adaptation was held without disrupting its feature despite the variation in the parameters [5]. The robustness of adaptation behavior was shown to be a consequence of the network's connectivity and that the chemotaxis system did not require the fine-tuning of biochemical parameters.

There have been conceptual difficulties for analyzing a biological system with mathematical models. Among the most important has been the difficulty in isolating a smaller subsystem that could be analyzed separately. Complexity seems to impede isolating a smaller subsystem out of the whole system. Some subsystems have been extracted and simulated using mathematical models, but few researches have referred to the crucial question as to whether it is possible to isolate such a subsystem from the whole system composed of heterogeneous and interactive networks and to analyze it separately. To address this issue, we first established a numerical framework model of the *E. coli* heat shock response, and analyzed it with control theory in terms of robustness. We clarified how complexity in regulation generated the robustness of the factors interconnected between the heat shock response and other subsystems. A complex network seems to impede isolating a smaller subsystem out of a whole biological system. However, in the heat shock response complexity in the regulation generates robustness, thereby making it possible to extract such a subsystem out of the whole system.

2. The heat shock response system

2.1 Molecular mechanism

E. coli grows well on nutrient rich medium at 37 ℃. In fast-growing *E. coli*, the major sigma factor, σ^{70}, binds RNA polymerase (RNAP) core enzyme, selecting the genes expressed in the course of growth. Heat shock (42 ℃) causes the amount of sigma factor σ^{32} expressed to increase and it replaces σ^{70} in binding to the RNAP core enzyme, resulting in the expression of the σ^{32}-regulated heat shock protein (HSP including the chaperones–DnaK/J, GrpE, GroEL, GroES and the proteases–Lon, FtsH) genes.

Figure 1 shows the molecular mechanism of the heat shock response[6]. The heat shock response aims at refolding of heat-denatured proteins through the action of chaperones and degradation of such proteins through the action of proteases, preventing the denatured proteins from forming aggregates[7]. The heat shock response depends primarily on the regulation of σ^{32} activity, stability and synthesis[8]. σ^{32} activity is varied by sequestering chaperone-bound σ^{32}, which limits its binding to the RNAP core enzyme[9]. The stability of σ^{32} is regulated through degradation of chaperone-bound σ^{32} [10] or through degradation of σ^{32} by FtsH protease[7]. The synthesis of σ^{32} is regulated at the translational level through degradation of the chaperone-bound nascent polypeptide on *rpoH* mRNA[12, 13] or through the translation of *rpoH* mRNA directly induced by heat[14, 15]. The *rpoH* mRNA forms a fairly stable secondary structure which prevents the initiation of translation at low temperature. Higher temperature disrupts such secondary structure, inducing translation.

The time course of the heat shock response is as follows. Heat shock results in an increase in the cellular levels of unfolded proteins that preferentially bind to members of the chaperone team[9]. The increase in unfolded protein-bound chaperones stabilizes σ^{32}, resulting in a shift from the chaperone-bound form of σ^{32} to the RNAP core enzyme-bound form, thus leading to HSP induction. In addition, heat shock directly induces translation of *rpoH* mRNA. Overexpressed chaperones reduce the level of active σ^{32} by sequestering or degrading the chaperone bound σ^{32}, and the FtsH protease degrades σ^{32}. The level of σ^{32} decreases, reaching a new steady state at high temperature.

2.2 Modeling

Mathematical equations that describe the heat shock response are provided according to the method that will described elsewhere. DnaK is chosen as a representative of the chaperone team, because it binds to denatured proteins or σ^{32} prior to the other chaperones [16, 17]. Lon is selected as a representative of the proteases serving to degrade denatured proteins. FtsH is the major protease involved in degrading σ^{32}. Since DnaK interacts with the same region C of σ^{32} protein and its nascent polypeptide on *rpoH* mRNA[12, 13], the binding association constant for the interaction between σ^{32} and DnaK and the rate of degradation of DnaK-bound σ^{32} are assumed to be the same as the association constant for the interaction between the σ^{32} nascent polypeptide and DnaK and the rate of degradation of DnaK-bound σ^{32} nascent polypeptide, respectively. This assumption contributes to simplifying the calculations.

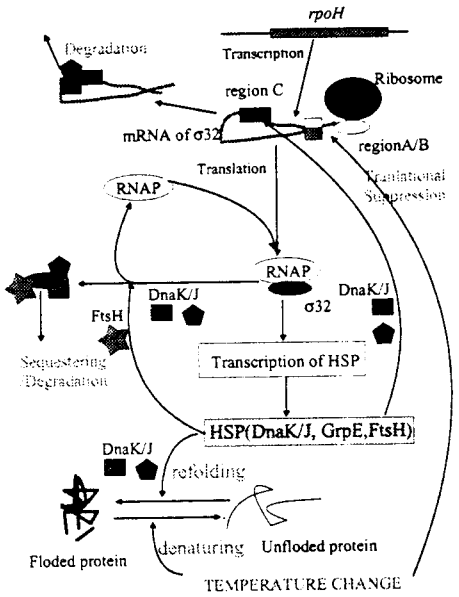

Figure 1 Schematic diagram of the heat shock response

2.3 System analysis

We define the complexity, performance (yield and efficiency), and robustness to analyze the heat shock response system. Complexity in the regulation of σ^{32} is generated from three processes, (1) feedback control: chaperone (DnaK)-mediated sequestering of σ^{32} which limits its binding to the RNAP core enzyme, and chaperone (DnaK)-mediated degradation of the σ^{32} and its nascent polypeptide on the *rpoH* mRNA, (2) feedforward control: heat-induced translation of the *rpoH* mRNA, (3) autogenous control (named as such because of the degradation of σ^{32} seemed to involve a self regulatory control loop): degradation of σ^{32} by FtsH protease whose expression is dependent on σ^{32}. The degree of complexity is adjusted by changing the combination of these three processes. The parameters regarding the regulation are shown in Table 1.

Performance of the *E. coli* heat shock response system is characterized by yield and efficiency. Since the aim of the heat shock response is considered to be to reduce the amounts of free unfolded proteins, the yield is defined by:

$$\text{Yield} = 1 - \frac{\text{free } P_{un}}{\text{Total } P} \quad \text{.....(1)}$$

An excess amount of chaperone refolds proteins

sufficiently, but it is a burden to the system. Efficiency is defined by:

$$\text{Efficiency} = 1 - \frac{\text{free DnaK}}{\text{Total DnaK}} \quad \ldots\ldots(2),$$

where free DnaK is neither involved in a refolding process nor in binding to σ^{32}. Efficiency means how efficiently the heat shock response is regulated by a minimum amount of chaperone (DnaK).

3. Methods

Calculations involving differential equations and those involving simultaneous nonlinear equations were performed by the Runge-Kutta method and by the Newton-Raphson method, respectively. Computer programs in C language and Message-Passing Interface were executed on sixty-four CPUs of a super parallel computer SR2201 (HITACHI, TOKYO). The typical CPU time for finding a numerical solution of the model system was of the order of minutes. Heat shock occurred at 50 min. The mathematical simulation sampled the yield and efficiency at 50 min and 150 min at low and high temperatures, respectively, when the system reached the steady state.

Table 1 A list of parameters employed

Feedback control		
Parameter	Definition	Unit
Kfb[1]	binding association constant between σ^{32} and DnaK	M^{-1}
rfb[1]	degradation rate of DnaK-bound σ^{32}	min^{-1}
Kfb[2]	binding association constant between mRNA(σ^{32}) and DnaK	M^{-1}
rfb[2]	degradation rate of DnaK-bound mRNA(σ^{32})	min^{-1}
Feedforward control		
Parameter	Definition	Unit
η	translation efficiency of the *rpoH* mRNA	-
Autogenous control		
Parameter	Definition	Unit
Kac[1]	binding association constant between σ^{32} and FtsH	M^{-1}
rac[1]	degradation rate of FtsH-bound σ^{32}	min^{-1}
Disturbance to the heat shock response		
Parameter	Definition	Unit
Kdisturb[1]	binding association constant between $\sigma^{disturb}$ and RNAP	M^{-1}
Kdisturb[2]	binding association constant between DnaK and DnaKdisturb	M^{-1}

4. Results

4.1 Simulation of the heat shock response

To simulate the heat shock response, we employed some kinetic data and determined the values of the parameters regarding the heat shock response. As shown in Figure 2, the numerical model reproduced the main features of the *E. coli* heat shock response: (1) when the temperature is shifted up, the rate of chaperone synthesis increases greatly within several minutes; (2) heat shock results in an increase in the level of σ^{32} level within the initial few minutes, and then it decreases to a new steady-state level.

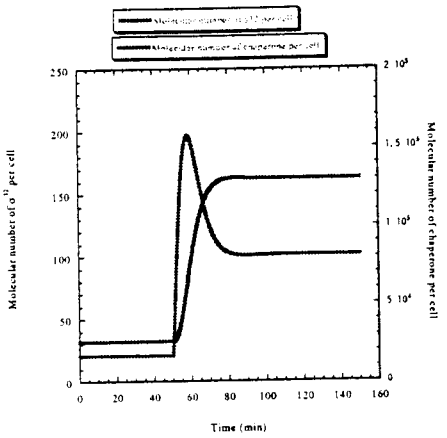

Figure 2 Time course for σ^{32} and DnaK

4.2 Two-dimensional parameter space of feedback control

To determine the robustness in the heat shock response, the two-dimensional feedback control (σ^{32}/DnaK) parameter space, consisting of the binding association constant between DnaK and σ^{32} and the degradation rate of DnaK-bound σ^{32}, was searched to provide the required performance (*e.g.*, efficiency > 0.8 and yield > 0.995), because feedback control plays a major role in the heat shock response, and is commonly employed by all bacteria, whereas feedforward control is optional, used only by γ-proteobacteria including *E. coli* [6] and autogenous control cannot function in the heat shock response by itself. Robustness is characterized by the wideness of the parameter space satisfying the required performance. The target values of yield and efficiency were set as 0.995 and 0.8 to present a well-visualized space of the two-dimensional parameters. We tested various values of yield and efficiency to simulate the heat shock response. For example, when yield and efficiency were 0.999 and 0.8, respectively, the parameter space that satisfied the requirement was so narrow that it was hard to clearly show that complexity increased the parameter space at the condition that the value of each parameter logarithmically varied by 2-fold although these values never changed our conclusion.

Figure 3 indicates that complexity in σ^{32} regulation

enlarges the feedback control parameter space. For simplicity, the effects of the Lon protease on the heat shock response are neglected, and this will be verified later. By comparing (B) with (A), the addition of feedforward control increases the feedback control parameter space, *i.e.*, increases the robustness. Feedforward control is effective in increasing the efficiency at low temperature by decreasing the rate of σ^{32} synthesis without affecting the performance at high temperature. In (C), the addition of autogenous control further enlarges the feedback control parameter space. It has been hard to understand the self-regulatory-loop-like function where FtsH protease whose expression is dependent on σ^{32} directly degrades σ^{32}, but system analysis indicated that autogenous control has the capability to increase the robustness of the feedback control parameters. These simulations led to the hypothesis that complexity in σ^{32} regulation generates the robustness of the interaction between σ^{32} and DnaK.

To elucidate the robustness of feedforward and autogenous controls, we investigated the region of their parameters that made the system performance enhanced. A wide range of translation efficiency enlarges the feedback control parameter space. Since heat shock enhances translation efficiency several-fold compared with the level at low temperatures in *E. coli* [14], feedforward control must be effective in increasing the robustness *in vivo*. Autogenous control was effective in enhancing the robustness, when either the binding association constant Kac[1] or the degradation rate rac[1] was relatively weak.

4.3 Disturbance to the interconnected factors among subsystems

We also explored the issue of robustness for the factors interconnected among subsystems. The factor DnaK is not only responsible for regulating the heat shock response but also commonly employed in other subsystems such as DNA replication, cell division[18], ribosome synthesis[19], and osmotic pressure changes[20]. On the other hand, at least seven sigma factors including σ^{32} are responsible for selective gene expression, competing for binding to the RNAP core enzyme, thus the binding of σ^{32} to the RNAP core enzymes is strongly influenced by the other sigma factors. DnaK and RNAP core enzyme can be regarded as two of the interconnecting factors among subsystems. The robustness in the heat shock response was investigated under conditions where DnaK and the RNAP core enzyme were exposed to disturbance (Figure 4). Regardless of the strength of the disturbance to DnaK or to the RNAP core enzyme, the addition of feedforward and autogenous controls increased the robustness in the heat shock response, showing that complexity in σ^{32} regulation generates the robustness for the factors crosstalking to other subsystems.

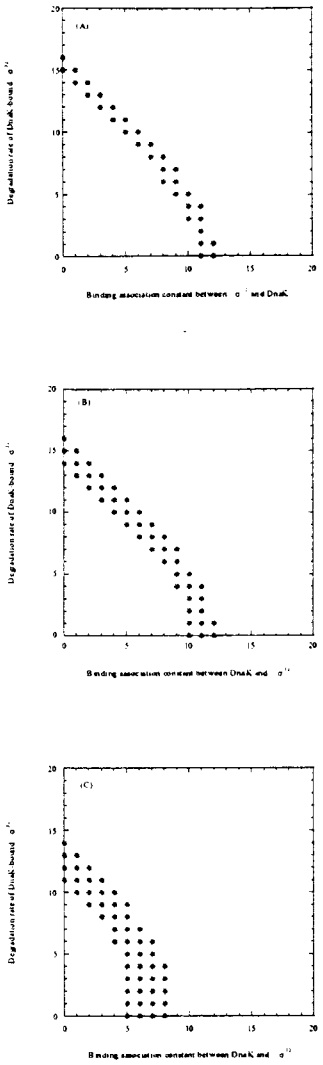

Figure 3 Two-dimensional space of the feedback control parameters. The whole space as follows: the x-axis (binding association constant between σ^{32} and DnaK) is $10^4 \times 2^x M^{-1} (x = 0,1,2,3,.....19)$, the y-axis (degradation rate of DnaK-bound σ^{32}) is $0.01 \times 2^y \min^{-1} (y = 0,1,2,3,.....19)$, is searched to provide the performance (efficiency > 0.8 and yield > 0.995), while the feedforward control parameter (translation efficiency) and the autogenous control parameters are varied. The efficiency and yield were calculated at low and high temperature. (A) Regulation by feedback control. (B) Regulation by feedback and feedforward controls. The translation efficiency is employed to provide the most enlarged space for the feedback control parameters. (C) Regulation by feedback, feedforward, and autogenous controls. The feedforward and autogenous control parameters are used to provide enlarged space for the feedback control parameter.

4.4 Verification of the hypothesis

To verify the hypothesis that complexity in σ^{32} regulation generates robustness, we must consider the following effects: (1) Lon protease degradation of denatured proteins, (2) variations in the ratio of folded proteins to denatured proteins, (3) changes in the parameter values in response to changes in temperature. The addition of feedforward and autogenous controls increases the number of feedback control parameter combinations regardless of the strength of the Lon protease effects (data not shown). The integration of the Lon protease effects into the model does not affect my hypothesis. The question of how much protein is denatured as a result of a change in temperature remains unclear, because of the difficulty in experimental measurement. When the ratio of folded protein to unfolded protein was relatively large, the addition of feedforward and autogenous controls increased the number of the feedback control parameter combinations, supporting the hypothesis. When the parameter values were either increased or decreased at high temperature, the feedback control parameter space was searched to provide the required performance. The addition of feedforward control always enlarged the feedback control parameter space (data not shown), whereas the effects of autogenous control depended on temperature. When the rate of σ^{32} degradation by FtsH decreased at high temperature, autogenous control enlarged the feedback control parameter space (data not shown).

5. Discussion

Complexity in σ^{32} regulation shows the capability to generate robustness in the heat shock response, thereby increasing the robustness for the factors interconnected among subsystems. Feedforward control can generate robustness in any case discussed. Note, however, that autogenous control depends on the ratio of folded proteins to denatured proteins, the σ^{32} /FtsH (autogenous control) parameters, and the heat-induced changes in the parameters. To demonstrate the role of autogenous control *in vivo*, biological experiments are now underway.

To understand control in a biological system, we compared it with the control of an artificial system. In feedback control of an artificial thermostat process, the process generally detects a process temperature with a thermometer, calculating the gain, regulating the power of a coolant/heater to adjust it to a target temperature. Similar to that, the *E. coli* heat shock response measured the amount of denatured proteins with chaperones, controlling the σ^{32} level, producing the chaperone to refold unfolded proteins. Different from the thermostat process, chaperone is a bifunctional hardware: sensing and refolding (thermometer and coolant/heater). The advantages for the bifunction may be that the resource of biomolecules is efficiently saved. In artificial processes, feedback or feedforward control is generally used to enhance the stability or robustness. However, the autogenous control has not been hard to understand the effectiveness of its role in heat shock response system. We tried to understand what kinds of roles the autogenous control played. In many biological systems, the product with an increased concentration directly suppresses its synthesis, which is called autogenous or negative feedback control. From this viewpoint, autogenous control can be a type of feedback control. In the light of controlling the concentration of unfolded proteins, chaperone-mediated feedback control is similar to an artificial feedback control because chaperone controls its synthesis rate, monitoring the concentration of unfolded proteins. However, autogenous control never senses the amount of unfolded proteins, which differs from a generally employed artificial feedback control, thus it is a unique feature in a biological system.

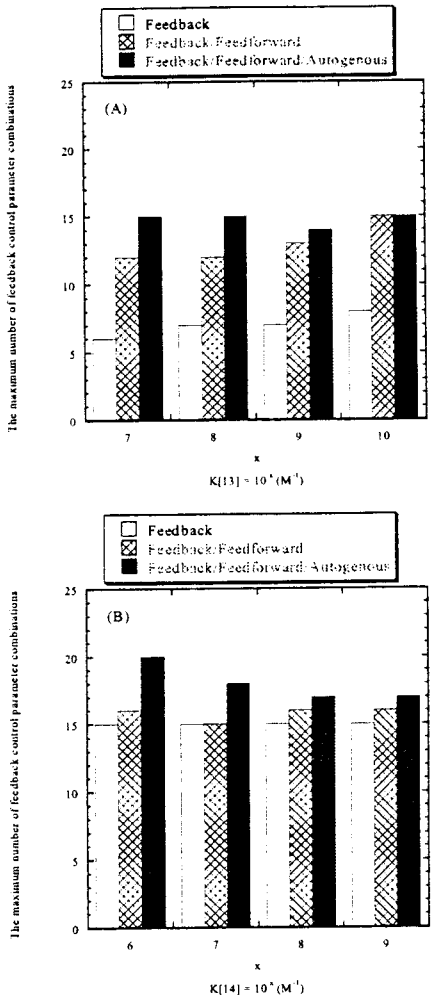

Figure 4 Disturbances to RNAP (A) or DnaK (B). Despite disturbance, complexity increases the robustness.

The problem in synthesizing/analyzing a biological system using numerical models is that it is quite difficult to determine all of the biochemical parameters. Since the parameters vary with intracellular or extracellular conditions, it is hard to measure them by means of experiments and they will remain unknown. In chemotaxis and bateriophage λ gene regulatory circuit, most of the parameters have been measured to simulate the exact behavior of such biological systems. In the heat shock response, some data was obtained from the previous work, but most of the biochemical parameters are assumed so that simulation reproduces the main features of the heat shock response. Is there any problem regarding the parameters ? In the heat shock response, instead of measuring the parameter values, we varied the values of the parameters that have not been measured (data not shown). As a result, the variation in parameters never affected our hypothesis.

An important approach to understanding a biological system would be to extract some of the design principles underlying the biological architecture rather than to obtain full knowledge of the molecular details. Synthesis of numerical framework model and analysis with control theory will be a powerful method. The important approach to understanding a biological system would be to extract some of the design principles underlying the biological architecture rather than to obtain full knowledge of the molecular details. Computational system analysis with modern control theory and subsequent verification through biological experiments are expected to be an important methodology in the post genome era.

References

1. Tyson, J. J., Novak, B., Odell, G. M., Chen, K. & Thron, C. D. Chemical kinetic theory: understanding cell-cycle regulation. *TIBS* 21, 89-95 (1996).
2. Bray, D., Bourret, R. B. & Simon, M. I. Computer simulation of the phosphorylation cascade controlling bacterial chemotaxis. *Mol. Biol. Cell* 4, 469-482 (1993).
3. Arkin, A., Ross, J. & McAdams, H. H. Stochastic kinetic analysis of developmental pathway bifurcation in phage l-infected Escherichia coli cells. *Genetics* 149, 1633-1648 (1998).
4. Little, J. W., Shepley, D. P. & Wert, D. W. Robustness of a gene regulatory circuit. *The EMBO J.* 18, 4299-4307 (1999).
5. Barkai, N. & Leibler, S. Robustness in simple biochemical networks. *Nature (London)* 387, 913-917 (1997).
6. Yura, T. Regulation and conservation of the heat-shock transcription factor σ^{32}. *Genes to Cells* **1**, 277-284 (1996).
7. Gross, C. in *Escherichia coli and SalmonellaCellular and Molecular Biology* (ed. Neidhardt, F. C.) 1382-1399 (ASM Press, Washington, DC., 1996).
8. Straus, D. B., Walter, W. A. & Gross, C. A. The heat shock response of *E. coli* is regulated by changes in the concentration of σ^{32}. *Nature* **329**, 348-351 (1987).
9. Gamer, J., Bujard, H. & Bukau, B. Physical interaction between heat shock proteins DnaK, DnaJ, and GrpE and the bacterial heat shock transcription factor σ^{32}. *Cell* **69**, 833-842 (1992).
10. Tilly, K., Spence, J. & Georgopoulos, C. Modulation of stability of the *Escherichia coli* heat shock regulatory factor σ^{32}. *J. Bacteriol.* **171**, 1585-1589 (1989).
11. Tomoyasu, T. *et al.* *Escherichia coli* FtsH is a membrane-bound, ATP-dependent zinc-metalloprotease with activity for the heat-shock transcription factor σ^{32}. *EMBO J.* **14**, 2551-2560 (1995).
12. Nagai, H., Yuzawa, H., Kanemori, M. & Yura, T. A distinct segment of the σ^{32} polypeptide is involved in DnaK-mediated negative control of the heat shock response of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **91**, 10280-10284 (1994).
13. McCarty, J. S. *et al.* Regulatory region C of the *E. coli* heat shock transcription factor, σ^{32}, constitutes a DnaK binding site and is conserved among eubacteria. *J. Mol. Biol.* **256**, 829-837 (1996).
14. Yuzawa, H., Nagai, H., Mori, H. & Yura, T. Heat induction of σ^{32} synthesis mediated by mRNA secondary structure: a primary step of the heat shock response in *Escherichia coli*. *Nucl. Acids Res.* **21**, 5449-5455 (1993).
15. Morita, M. T. *et al.* Translational induction of heat shock transcription factor σ^{32}: evidence for a buit-in RNA thermosensor. *Genes & Development* **13**, 655-665 (1999).
16. Schmidt, D., Baici, A., Gehring, H. & Christen, P. Kinetics of molecular chaperone action. *Science* **263**, 971-973 (1994).
17. Gamer, J. *et al.* A cycle of binding and release of the DnaK, DnaJ and GrpE chaperones regulates activity of the *Escherichia coli* heat shock transcripton factor σ^{32}. *EMBO J.* **15**, 607-617 (1996).
18. Tsuchido, T. R., Vanbogelen, R. A. & Neidhardt, F. C. Heat shock response in *Escherichia coli* influences cell division. *Proc. Natl. Acad. Sci. USA* **83**, 6959-6963 (1986).
19. Alix, J.-H. & Guerin, M.-F. Mutant DnaK chaperones cause ribosome assembly defects in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **90**, 9725-9729 (1993).
20. Jishage, M. & Ishihama, A. Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*: Intracellular levels of σ^{70} and σ^{38}. *J. Bacteriol.* **177**, 6832-6835 (1995).

Acknowledgements

I am very grateful for discussion Dr. H. Ohtake (Hiroshima University), and Dr. T. Yura (HSP Research Institute).