Understanding biology by reverse engineering the control

Claire J. Tomlin, and Jeffrey D. Axelrod

PNAS 2005;102;4219-4220; originally published online Mar 14, 2005;
doi:10.1073/pnas.0500276102

Understanding biology by reverse engineering the control

Claire J. Tomlin*† and Jeffrey D. Axelrod‡

**Department of Aeronautics and Astronautics, Stanford University, Stanford, CA 94305-4035; and ‡Department of Pathology, Stanford University School of Medicine, Stanford, CA 94305-5324*

Faced with the problem of controlling a physical system, an engineer will identify a model for the system, and then use this model to design a process for automatically actuating some of the system's input signals in order that the behavior of the system follows a desired profile. The most common control processes use feedback. Consider the example of designing an autopilot function that controls an aircraft to fly at constant altitude (see Fig. 1). A mathematical model of the aircraft dynamics (the plant) is designed, which describes how the aircraft's altitude, pitch, and roll angles (the outputs) change when the elevators, ailerons, and throttle (the inputs) are manipulated. By using sensors to measure the outputs, the engineer uses the model to calculate how these measurements should be employed to automatically adjust the inputs, so that if the aircraft deviates off course, it is guided, quickly and smoothly, back to the desired altitude. This function, which maps the measurements to the input adjustment through the feedback path, is known as the controller. The signal that carries the desired profile is fed into the system in the feedforward path. The entire system composed of the functional parts plant, feedback, feedforward, and controller is referred to as the closed loop system. The hallmark of a good feedback control design is a resulting closed loop system that is stable and robust to modeling errors and parameter variation in the plant and achieves a desired output value quickly without unduly large actuation signals at the plant input. Some insightful recent papers advocate a similar modular decomposition of biological systems according to the well defined functional parts used in engineering (1–5) and, specifically, engineering control theory (6–10). Yi *et al.* (6) reported that adaptation in bacterial chemotaxis (11) appears to use integral feedback control, which is a basic strategy used all of the time to reduce a system's error in tracking a desired signal. In a recent issue of PNAS, El-Samad *et al.* (12) showed that the mechanism used in *Escherichia coli* to combat heat shock is just what a well trained control engineer would design, given the signals and the functions available.

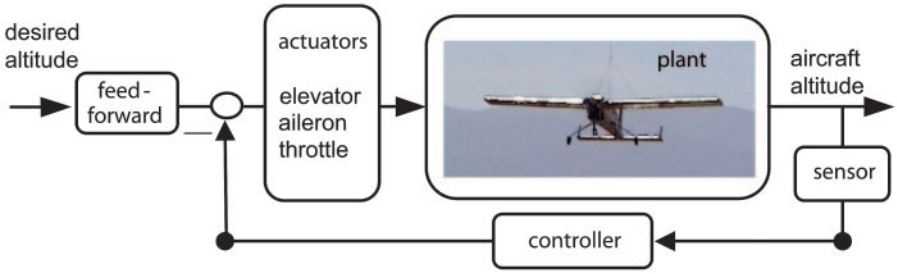

Fig. 1. Design of an autopilot function that controls aircraft altitude.

Living cells defend themselves from a vast array of environmental insults. One such environmental stress is exposure to temperatures significantly above the range in which an organism normally lives. Heat unfolds proteins by introducing thermal energy that is sufficient to overcome the noncovalent molecular interactions that maintain their tertiary structures (13). Evidently, this threat has been ubiquitous throughout the evolution of most life forms. Organisms respond with a highly conserved response that involves the induced expression of heat shock proteins. These proteins include molecular chaperones that ordinarily help to fold newly synthesized proteins and in this context help to refold denatured proteins. They also include proteases and, in eukaryotes, a proteolytic multiprotein complex called the proteasome, which serve to degrade denatured proteins that are otherwise harmful or even lethal to the cell. Sufficient production of chaperones and proteases can rescue the cell from death by repairing or ridding the cell of damaged proteins.

The challenge to the cell is that the task is gargantuan. A large portion of the entire cellular complement of proteins is at risk in heat shock, and therefore the level of chaperone and protease expression required is correspondingly large; in the bacterium *E. coli*, up to one quarter of the cellular protein after heat shock consists of induced heat shock proteins. Although cells might, in principle, maintain high levels of heat shock proteins at all times as insurance against heat shock or other stresses, the metabolic cost would confer a substantial disadvantage compared to cells that evolve a more efficient system. Instead, it has been found that heat shock protein expression is induced on a massive scale. Furthermore, it is easy to suppose that more rapid responses, and responses whose magnitude is well tuned to the level of denatured protein, would also be favored over slower or less accurate ones.

The biology of the heat shock response in *E. coli* has been well studied (14, 15). The RNA polymerase cofactor σ^{32} confers specificity for the heat shock genes; after heat shock stress, a rapid and profound increase in σ^{32} activity produces an appropriately scaled burst of protein expression. The critical regulation of σ^{32} activity depends on a feedforward mechanism that senses temperature and controls σ^{32} transcription and on several levels of feedback regulation that sense the levels of denatured cellular protein and sequester or degrade σ^{32}. Nonetheless, it has remained a considerable challenge to understand the contributions each of these mechanisms makes to the dynamic behavior of the intact system. What accounts for the rapidity of the response? What is most responsible for scaling the response to the amount of denatured cellular protein? How does the system achieve reliability?

Viewing the heat shock response as a control engineer would, El-Samad *et al.* (12) constructed and validated a full mathematical model (31 equations, 27 parameters) encoding the influence of each signal on the other and describing the dynamics of each signal over time. El-Samad *et al.* then decomposed this model into functional modules of plant (the refolding of denatured proteins), actuated plant input (the numbers of molecular chaperones), sensed plant output (the amount of denatured

See companion article on page 2736 in issue 8 of volume 102.

†To whom correspondence should be addressed. E-mail: tomlin@stanford.edu.

protein), controller (the level of σ^{32} activity), feedforward (temperature dependent translational efficiency of σ^{32} synthesis), and two feedback signals (both sequestration and degradation of σ^{32} from measurement of the amount of denatured protein). In so doing, they constructed a reduced order mathematical model (six equations, 11 parameters), in which each equation describes the dynamics of the module it defines. The advantage of this smaller model is that it may be mathematically analyzed and therefore may be used to understand the kinds of control laws that the biological system employs to regulate its behavior.

El-Samad *et al.* (12) used this analysis to motivate a series of simulation experiments on their larger mathematical model. The experiments involved disconnecting one and both feedback paths and disrupting the feedforward path. The comparison between the closed-loop system and the open-loop case (with feedforward intact and both feedback paths disconnected) is particularly interesting. As illustrated by El-Samad *et al.*, an operational control design for heat shock could simply involve a temperature sensor and an appropriate translational response in σ^{32} synthesis. However, such a system requires meticulous tuning of the σ^{32} synthesis rate, is very sensitive to parameter variations, and requires a large actuation signal at the plant input (large number of chaperones) at high temperatures. With the sequestration feedback path connected, the model indicates a more reliable system demonstrating robustness to parameter variation and a more efficient use of chaperones as actuators. With both feedback paths connected, the time it takes to fold the denatured proteins significantly decreases, supporting the hypothesis that directly controlling σ^{32} degradation in response to the number of denatured proteins has a large effect on the speed of response. El-Samad *et al.* also detailed the favorable properties of noise attenuation when feedback is used and described how the use of a temperature sensor in the feedforward path that is independent of the measurement of denatured proteins appears to have a strong effect on the resulting number of denatured proteins.

An operational control design for heat shock could simply involve a temperature sensor and an appropriate translational response.

The analysis in El-Samad *et al.* (12) is important not just because it captures the behavior of the system, but because it decomposes the mechanism into intuitively comprehensible parts. If the heat shock mechanism can be described and understood in terms of engineering control principles, it will surely be informative to apply these principles to a broad array of cellular regulatory mechanisms and thereby reveal the control architecture under which they operate.

In the postgenomic era, a flood of data concerning gene expression profiles, protein interaction networks, and inferred structure and function is becoming available. However, the physiologically relevant functions of the majority of proteins encoded in most genomes are either poorly understood or not understood at all. One can imagine that, by combining these data with measurements of response profiles, it may be possible to deduce the presence of modular control features, such as feedforward or feedback paths, and the kind of control function that the system uses. It may even be possible to examine the response characteristics of a given system, for example, a rapid and sustained output, as seen here, or an oscillation, and to draw inferences about the conditions under which a mechanism is built to function. This, in turn, could help in deducing what other signals are participating in the system behavior. The ultimate goal of this kind of analysis would be to predict, from the response characteristics, the overall function of the biological network. Learning to do this on a well understood system like heat shock response helps in designing the analysis techniques, which could then be applied to lesser known systems. Finally, characterizing biological mechanisms in this way may be the most effective means to better predict problems or performance when a foreign substance is introduced or, indeed, to facilitate reengineering the system to achieve desired behaviors. Certainly, these kinds of analyses promise to raise the bar for understanding biological processes.

1. Hartwell, L. H., Hopfield, J. J., Leibler, S. & Murray, A. W. (1999) *Nature* **402,** C47–C52.
2. Atkinson, M. R., Savageau, M. A., Myers, J. T. & Ninfa, A. J. (2003) *Cell* **113,** 597–607.
3. Hasty, J., McMillen, D., Isaacs, F. & Collins, J. J. (2001) *Nat. Rev.* **2,** 268–279.
4. McAdams, H. H. & Shapiro, L. (1995) *Science* **269,** 650–656.
5. Asthagiri, A. R. & Lauffenburger, D. A. (2000) *Annu. Rev. Biomed. Eng.* **2,** 31–53.
6. Yi, T.-M., Huang, Y., Simon, M. I. & Doyle, J. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 4649–4653.
7. Csete, M. & Doyle, J. C. (2000) *Science* **295,** 1664–1668.
8. Lauffenburger, D. A. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 5031–5033.
9. Ferrell, J. E., Jr. (2002) *Curr. Opin. Cell Biol.* **14,** 140–148.
10. Mangan, S. & Alon, U. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 11980–11985.
11. Alon, U., Surette, M. G., Barkai, N. & Leibler, S. (1999) *Nature* **397,** 168–171.
12. El-Samad, H., Kurata, H., Doyle, J. C., Gross, C. A. & Khammash, M. (2005) *Proc. Natl. Acad. Sci. USA* **102,** 2736–2741.
13. Lindquist, L. & Craig, E. A. (1988) *Annu. Rev. Genet.* **22,** 631–677.
14. Yura, T. & Nakahigashi, K. (1999) *Curr. Opin. Microbiol.* **2,** 153–158.
15. Lund, P. A. (2001) *Adv. Microb. Physiol.* **44,** 93–140.