doi:10.1016/S0022-2836(03)00506-0 J. Mol. Biol. (2003) 329, 645–654

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Response Delays and the Structure of Transcription Networks

Nitzan Rosenfeld and Uri Alon*

Department of Molecular Cell Biology and Department of Physics of Complex Systems Weizmann Institute of Science Rehovot 76100, Israel

Sensory transcription networks generally control rapid and reversible gene expression responses to external stimuli. Developmental transcription networks carry out slow and irreversible temporal programs of gene expression during development. It is important to understand the design principles that underlie the structure of sensory and developmental transcription networks. Cascades, which are chains of regulatory reactions, are a basic structural element of transcription networks. When comparing databases of sensory and developmental transcription networks, a striking difference is found in the distribution of cascade lengths. Here, we suggest that delay times in the responses of the network present a design constraint that influences the network architecture. We experimentally studied the response times in simple cascades constructed of well-characterized repressors in *Escherichia coli*. Accurate kinetics at high temporal resolution was measured using green fluorescent protein (GFP) reporters. We find that transcription cascades can show long delays of about one cell-cycle time per cascade step. Mathematical analysis suggests that such a delay is characteristic of cascades that are designed to minimize the response times for both turning-on and turning-off gene expression. The need to achieve rapid reversible responses in sensory transcription networks may help explain the finding that long cascades are very rare in databases of *E. coli* and *Saccharomyces cerevisiae* sensory transcription networks. In contrast, long cascades are common in developmental transcription networks from sea urchin and from *Drosophila melanogaster*. Response delay constraints are likely to be less important for developmental networks, since they control irreversible processes on the timescale of cell-cycles. This study highlights a fundamental difference between the architecture of sensory and developmental transcription networks.

Keywords: design principles; systems-biology; network motifs; computational biology

*Corresponding author

Introduction

Gene expression is governed by networks of transcription interactions between transcription factors and the genes they regulate. These networks are complex, and unifying design principles are needed to help make sense of their structure.[1–14] One way to approach this is to attempt to understand the structure of transcription networks based on databases of experimentally verified interactions. For *Escherichia coli*, this was pioneered by the work of Collado-Vides and associates,[6] based on an extensive transcription database. Recently, an approach for identifying the basic building blocks of networks, termed network motifs, has been presented,[15,16] and applied to transcription networks. Network motifs were defined as patterns that occur more often in the real network than in randomized networks.[17–20] Three types of recurring network motif circuits were found to describe most of the *E. coli* and *Saccharomyces cerevisiae* transcription networks.[15,16,21] Each of these network motifs was suggested to have a specific function in information processing.

One important feature of the network architecture is the distribution of transcription cascade lengths. Transcription cascades are defined by a set of transcription factors that regulate each other

Abbreviations used: IPTG, isopropyl-β-D-thiogalactopyranoside; GFP, green fluorescent protein.

E-mail address of the corresponding author: urialon@weizmann.ac.il

sequentially.[21,22] The first-step transcription factor activates or represses the second-step transcription factor, which in turn activates or represses the third-step transcription factor, and so on. In a study of the architecture of the transcription network of *E. coli*, it was observed that it has a strikingly shallow architecture, with most genes regulated by overlapping cascades of length one to two.[15] Here, we find that two different databases of transcription interactions in yeast (YPD,[23] and a genome-wide location analysis experiment[21]) also have a shallow architecture where long cascades are much less common than short cascades. What is the reason for a design that favors short cascades?

To address this, we experimentally study the response times of simple transcription cascades. To study the properties of transcription cascades in a setting that minimizes unknown coupling to the rest of the cell, we constructed a two-step cascade made of well-characterized bacterial repressors.[22] A similar approach has been used for studying oscillators,[24] toggle switches,[25] negative autoregulation circuits[26,27] and logic-gates[28] in bacteria. We find that cascades can show long delays. We suggest that response delays in transcription cascades constitute one of the design constraints that underlie the observed transcription network architecture. We show that developmental transcription networks, in which response delays may not be an important constraint, have a different architecture, with many long cascades.

Results

Long cascades are rare in databases of *E. coli* and *S. cerevisiae* transcription interactions

We analyzed a database of experimentally verified direct transcription interactions in *E. coli*,[15] which includes 578 interactions and 423 operons. The database consists of operons regulated by at least one transcription factor. For each operon, we traced the maximal number of transcription steps back to a transcription factor that is not regulated by any other transcription factor. We find that the most common cascades in *E. coli* are one-step cascades. Two-step cascades are less common, and longer cascades are rare (Figure 1(a)).

A similar cascade distribution was found in a database of transcription interactions for the yeast *S. cerevisiae*, based on transcription interactions from the YPD literature-based database,[23] which includes 1052 interactions and 678 genes. Again, one-step cascades are the most common, and the occurrence of longer cascades decreases sharply with length (Figure 1(b)). We analyzed the network of interactions found by a genome-wide location analysis experiment in yeast,[21] which includes 3969 interactions and 2341 genes. This dataset does not have the methodological biases that may occur in databases of interactions collected from the literature, such as the possibility of unsaturated genetics. In this database, there are more long cascades than in the literature database, which may reflect multi-step processes in yeast.[29] On the whole, short cascades are much more frequent than long ones (Figure 1(c)).[21]

Long cascades are common in transcription networks for sea urchin and Drosophila early development

Developmental networks in metazoans are well known to contain elaborate transcription cascades.[4] In order to compare such networks to the sensory networks of micro-organisms on an equal footing, we enumerated the cascades in the rather limited developmental databases that are currently available. We analyzed the transcription network governing endomesoderm specification during sea urchin development.[30] This network contains 44 genes and 82 interactions. We find that long cascades are about as common as short cascades (Figure 1(d)). We analyzed the network of the established direct transcription interactions involved in early development in *Drosophila melanogaster*, based on the direct interactions listed in the GeNet database. This network is comprised of 45 direct interactions and 25 genes, including gap genes (*hunchback*, *kruppel*, *knirps*, *giant*, *torso*, *tailless*, and *caudal*), pair-rule genes (*hairy*, *odd-paired*, *odd-skipped*, *fushi tarazu*, *even-skipped*, *paired*, *runt*, and *sloppy-paired*), homeotic genes (*Ultrabithorax*, *empty-spiracles*, *Abdominal-A*, and *Antennapedia*), and segment polarity genes (*decapentaplegic*, *engrailed*, *wingless*, and *gooseberry-neuro*), as well as *bicoid* and *achete*. In this network, long cascades are at least as common as short ones (Figure 1(e)).

Experiment on a two-step cascade shows a delay of about one cell-cycle per step

To study the response kinetics of a cascade, we used a synthetic repressor cascade in *E. coli* (Figure 2). In this cascade, the LacI repressor, made chromosomally by the cell, represses a gene encoding the Tet repressor. The Tet repressor, in turn, represses the *tet* promoter that controls a reporter gene (*gfp*) encoding the green fluorescent protein (GFP) (Figure 2(a) and (c)).[31] The promoter activity of the first step in the cascade is monitored in a separate reporter strain carrying the *lac* promoter controlling *gfp* and a blank *tet* promoter plasmid (Figure 2(b)). The system is induced by adding external saturating isopropyl-β-D-thiogalactopyranoside (IPTG) to inactivate the LacI repressor. Fluorescence and cell density during exponential growth were measured automatically from both reporter strains in a multi-well fluorimeter at about a five minute resolution.[32,33] After induction, the fluorescence from the first step reporter rises, and fluorescence from the second step is repressed after a delay (Figure 3).

We repeated this experiment at two different

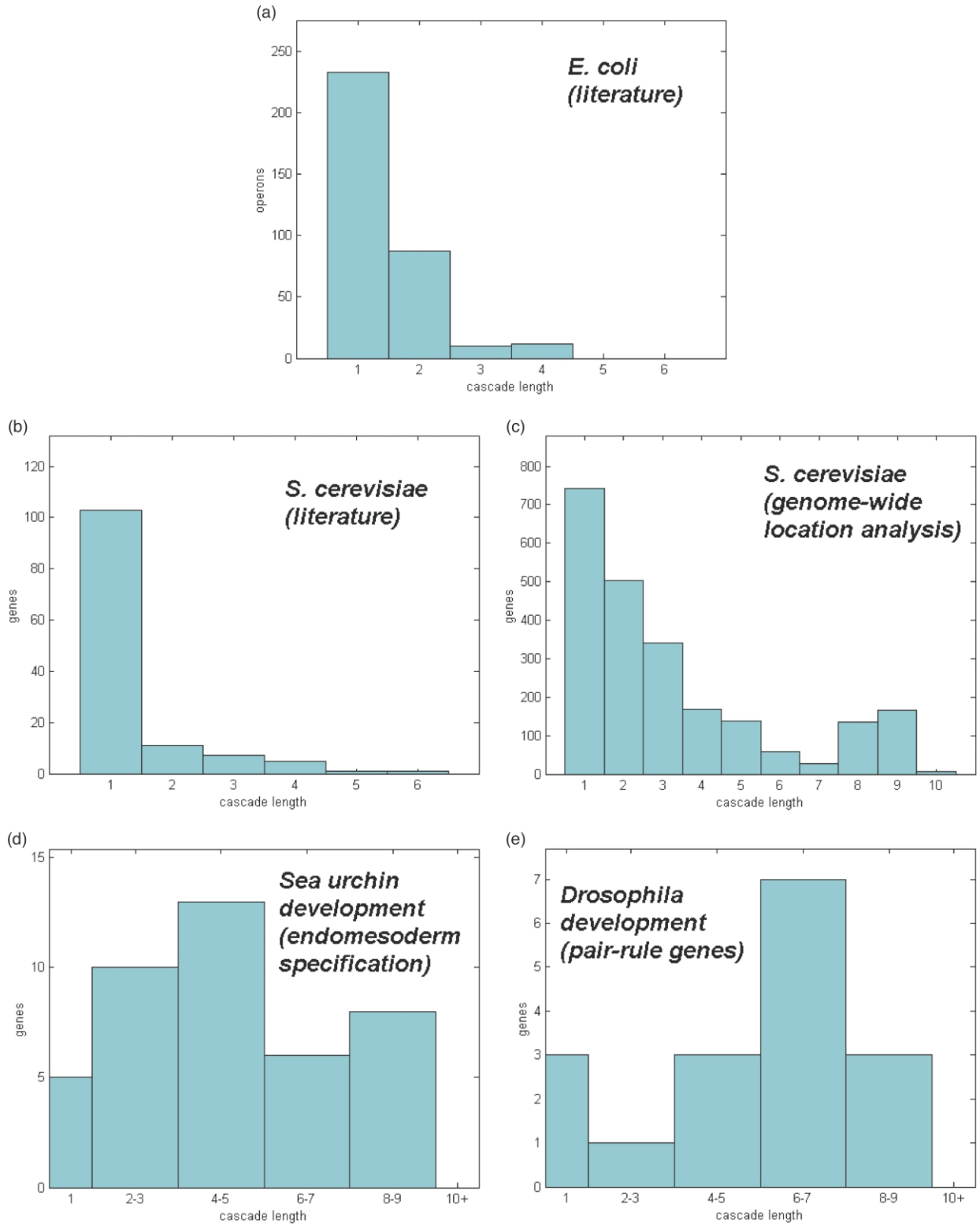

Figure 1. Distribution of cascade lengths in databases of transcription interactions for (a) *E. coli* literature,[15] (b) *S. cerevisiae* literature,[23] (c) *S. cerevisiae* genome-wide location analysis,[21] (d) endomesoderm development in sea urchin,[30] (e) *Drosophila* early development literature (GeNet database). The cascade length for each gene is defined as the maximal number of transcription steps that link it to a transcription factor that is not regulated by any other transcription factor.

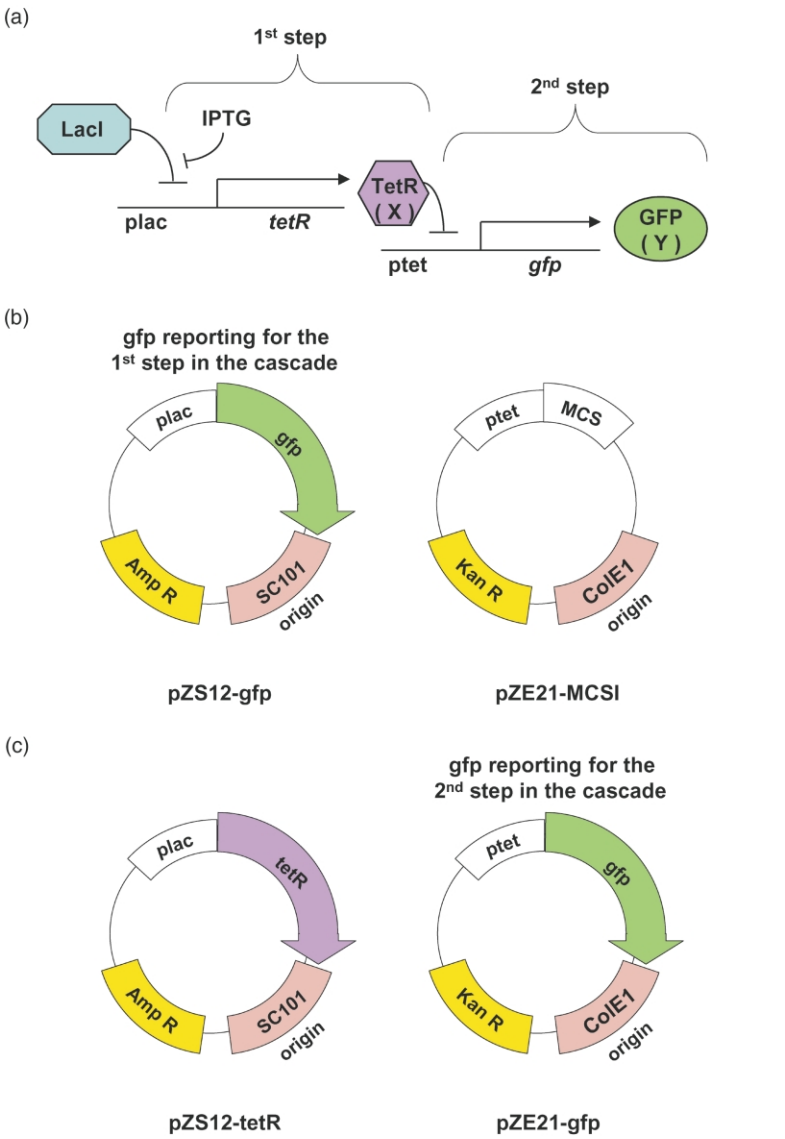

Figure 2. (a) Plasmid system used to implement the two-step transcription cascade. A chromosomally encoded LacI repressor represses the tetR gene. The TetR repressor represses the second step promoter, which controls a *gfp* reporter gene. The inducer IPTG binds and inactivates lacI, causing the expression of tetR and the repression of the reporter gene. (b) The first-step reporter strain bears a *gfp* gene under control of the *lac* promoter (step 1). As a control, this strain bears a compatible plasmid with the *tet* promoter upstream of a multicloning site (MCS). (c) The second-step reporter strain carries *tetR* under the *lac* promoter, and the second plasmid with *tet* promoter controlling *gfp*. The two strains allow the monitoring of the promoter activity of both cascade steps in parallel.

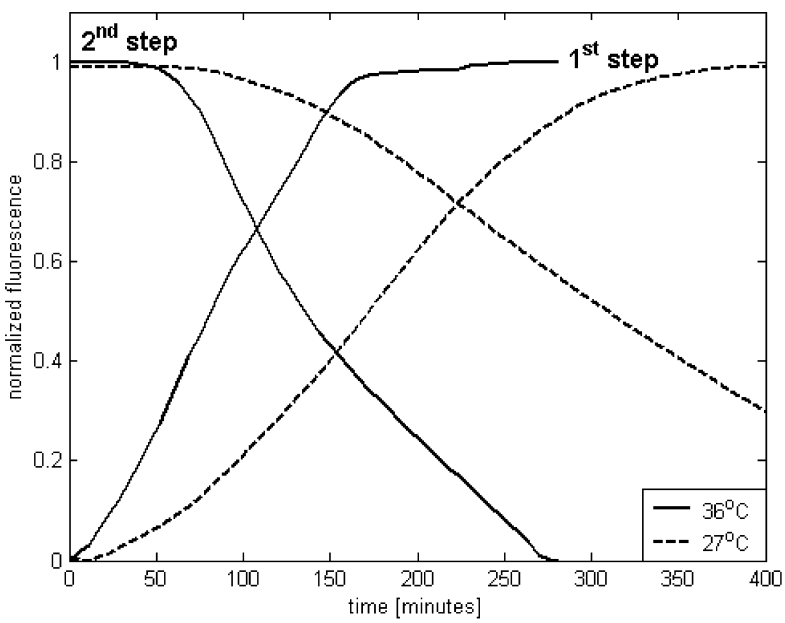

Figure 3. Normalized GFP fluorescence from the two cascade steps as a function of time. At time = 0 the inducer IPTG was added. The first step shows an increase in fluorescence, and the second step shows a delayed decrease in fluorescence. Continuous lines, 36 °C; broken lines, 27 °C.

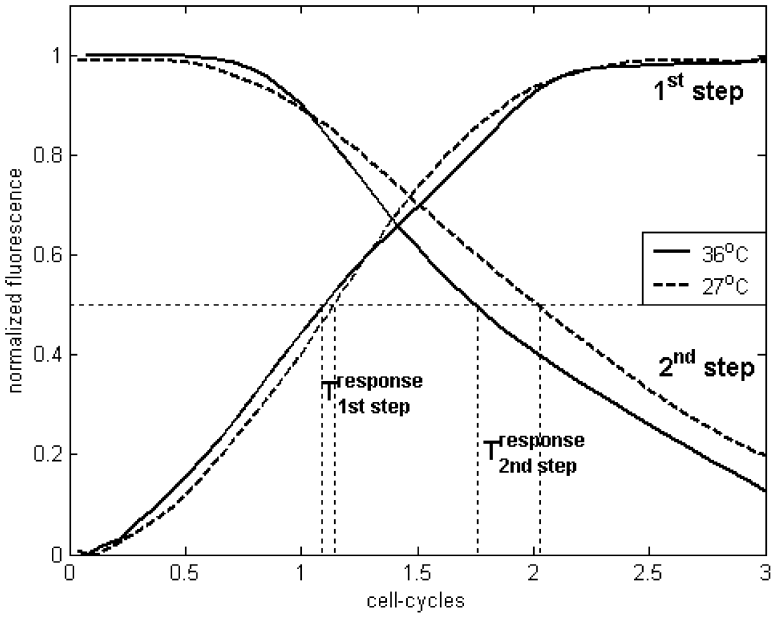

Figure 4. GFP fluorescence from the two cascade steps as a function of cell-cycles. Data for Figure 3 were plotted against time in units of cell-cycles. The response-time is the time at which the response reaches half of its maximal change. This is the time at which the curves intersect the horizontal line at relative fluorescence = 0.5. The response-time of the first step in the cascade is about 1.1 cell-cycles, and the response-time of the second step in the cascade is about 1.9 cell-cycles. Continuous lines, 36 °C; broken lines, 27 °C.

temperatures, 27 °C where the cell-cycle is 155 minutes, and 36 °C where the cell-cycle is 80 minutes. The kinetics is slower at 27 °C than at 36 °C (Figure 3). When plotting the response *versus* cell-cycles, determined by the number of optical absorbance doublings, we find that the kinetics overlap (Figure 4). The response-time of a system is defined as the time to reach half of the change between the pre-induced steady-state and the post-induced steady-state.[2,27] The response-time of the first step in the cascade is about 1.1 cell-cycles, and the response-time of the second step in the cascade is about 1.9 cell-cycles.

Mathematical analysis suggests that transcription cascades give rise to response delays of about one cell-cycle per step

We extend a previous mathematical treatment of transcription cascades[27,34,35] (see Materials and Methods), to ask what is the optimal response-time of a cascade given that it needs to respond rapidly to reversible changes. How should the cascade be designed, in order to minimize the sum of the response times in the ON and OFF directions?

For simplicity, we consider long-lived proteins with threshold-like transcription activation.[36] Assume that protein X begins to be produced at a constant rate upon induction, and is diluted out by cells dividing with a cell-cycle time of τ. When X is induced, its concentration in the cells rises, such that the distance to the steady-state level decreases exponentially with time (Figure 5(a), continuous line). The response-time of this first step in the cascade (time to reach half of its maximal level) is one cell-cycle (equation (4), in Materials and Methods). When the production of X is turned off, X is diluted. Its concentration decays exponentially to zero, and the response-time is one cell-cycle as well (Figure 5(b), continuous line, and see equation (7)).

Gene Y is activated when the concentration of X exceeds a threshold value. If the threshold is low, Y is turned on quickly (Figure 5(a)). However, it takes a long time until Y is turned off when the concentration of X decreases (Figure 5(b)). The opposite is true if the threshold is high; it takes a short time for Y to respond when the concentration of X decreases (Figure 5(f)) but a long time to respond when the concentration of X rises (Figure 5(e)). In Materials and Methods, we show that the optimal case, for which the sum of the response-times in both directions is minimal, is when the threshold is set at one-half of the steady-state concentration of X (Figure 5(c) and (d)). In this case, the response-time of Y in either direction (the time for Y to reach half of its maximal change) is two cell-cycles. Continuing in this fashion, it is easy to show that the optimal response-time for the nth step in a cascade is n cell-cycles.

These conclusions hold approximately also if the transcription regulation is a graded function rather than a threshold-like function. In the case of rapidly degradable proteins, the cell-cycle time should be replaced by the degradation half-life[27,37] (see Materials and Methods)†, and each step in the cascade contributes a delay of the order of the degradation half-life of the protein produced in that step.

Discussion

Synthetic networks built of well-characterized components are useful for analyzing the general

† See also math primer, in downloadable data, at http://www.weizmann.ac.il/mcb/UriAlon/

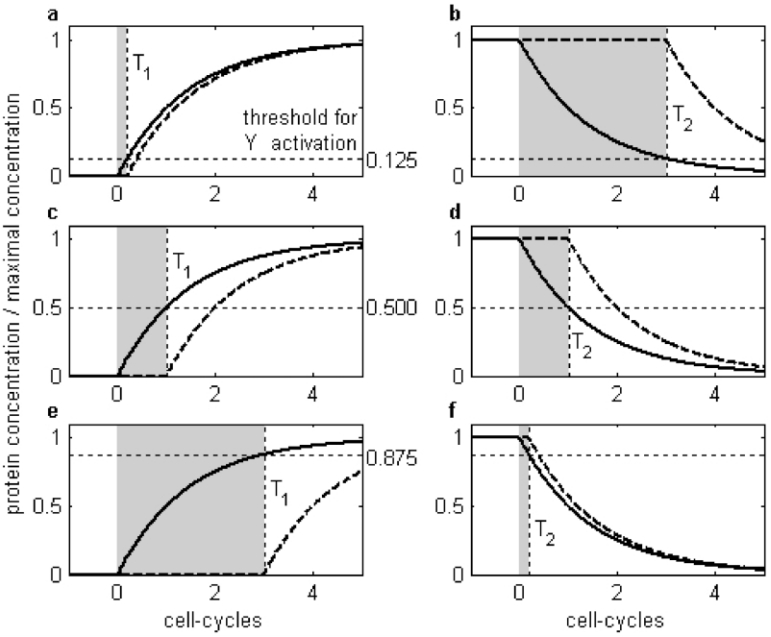

Figure 5. Mathematical model of the delay times between two steps in a cascade of genes producing long-lived proteins. The horizontal axes are time in cell-cycles, the vertical axes are relative protein concentration. Continuous lines, relative concentration of the first step in the cascade, the upstream protein X; heavy broken lines, relative concentration of the second step in the cascade, the downstream protein Y. Protein X production is turned on (a, c and e) or off (b, d and f) at time = 0. Y is expressed only when the concentration of X is above a threshold value (horizontal broken line). The time of onset of Y transcription after turn-on of X production, T_1, increases as the threshold rises (a, c and e). On the other hand, the time of shutoff of Y production after turn-off of X, T_2, decreases with increasing threshold (b, d and f). The optimal case in which $T_1 + T_2$ is smallest occurs at intermediate values of the threshold (c and d).

properties of network architecture.[24–28,38] Here, we analyzed the kinetics of a two-step cascade. The response-time,[2] which is the time taken to reach half of the change upon induction, was found to be about one cell-cycle per cascade step. Our experiment further suggests that the cell-division time is the basic time unit in the cascade response, for long-lived proteins. A mathematical model suggests that in order to minimize the response times in both the ON and OFF directions, the cascade should be designed so that the delays are of about one cell-cycle per cascade step.

The slow responses of multi-step cascades may explain the observation that the known *E. coli* transcription network displays primarily short cascades (Figure 1(a)). The exceptional long cascades in the network typically correspond to processes that take a few cell-cycles to complete. For example, a three-step transcription cascade controls flagella biosynthesis.[32,39] The genes in this cascade are termed class 1, class 2 and class 3, according to their cascade step. In the flagella system, the class 1 transcription factor FlhDC activates the class 2 genes that make up the basal-body motor and hook. One of these, FliA, activates class 3 genes that make up the flagellum filament and the chemotaxis navigation system. To complete the *de -novo* synthesis of a flagellum requires two to three cell-cycles,[32,40] which is roughly one cell-cycle per cascade step. Thus, a long transcription cascade is well suited for controlling this system. In contrast, the *E. coli* SOS DNA repair system[33,41,42] and the arginine biosynthesis system[43] both require rapid responses and both show a one-step cascade architecture.[15] We note that even within a single-step transcription cascade, subtle timing differences in expression can be achieved by using a hierarchy of activation (or repression) thresholds for the various genes controlled by a single transcription factor.[32,33,44]

In contrast to transcription networks, protein-level signaling networks are well known for having long cascades such as map-kinase cascades.[45] Since protein signaling interactions are typically much faster than transcription interactions, response delays in long signaling protein cascades need not be a serious design constraint.

The transcription network of the eukaryote *S. cerevisiae* is found to show a cascade distribution similar to that of the prokaryote *E. coli*, one-step cascades being the most common. Though long cascades exist, for example in the cell-cycle system,[29] the majority of genes are regulated by short cascades. The arguments regarding slow cascade responses should apply to all cell types. One important point is that in eukaryotes, many regulatory proteins have a life-time shorter than the cell-cycle. In contrast, prokaryotic transcription factors are often relatively long-lived (with important exceptions as in the heat-shock system[46]). For rapidly degradable gene products, the cell-cycle in all of the above arguments should be replaced by the protein half-life (see footnote on previous page).[27,37] Thus, the incremental response time in each cascade step is of the order of the lifetime of the protein produced in that step. Increased protein degradation can speed responses, though at a cost of increased production. Other mechanisms for speeding responses include negative autoregulation.[2,27]

In the single-celled organisms considered above, transcription cascades mainly orchestrate responses to fluctuating external conditions (we term these sensory networks). The decisions made by sensory networks are mostly reversible. In contrast, multi-cellular organisms have extensive transcription networks that are responsible for the cell-fate decisions made during development.[30,47–51] Since developmental networks often control processes on the order of many cell-cycle times, response-time considerations may be less important than in sensory networks. Such delays may even prove useful during development by providing the cells with a measure of the number of cell-cycles that have elapsed.† Furthermore, since developmental networks often control processes of deciding cell-fates, which are seldom reversed, they need not optimize both turn-on and turn-off times and can achieve fast timing in one direction by having appropriate activation thresholds (Figure 5). Therefore, developmental transcription networks need not be biased against long cascades. We find that developmental networks from sea urchin and from *Drosophila* indeed have a cascade structure that is very different from that of sensory networks. Developmental networks display many long cascades. Long cascades are also found in developmental-like processes in single-celled organisms, as in the sigma-factor cascade regulating sporulation in the bacterium *Bacillus subtillis*.[52] The difference in the cascade distributions suggests a fundamental difference between the architecture of sensory and developmental transcription factor networks. It would be fascinating to discover additional constraints and the design principles they induce in biological networks.

Materials and Methods

Measuring cascade lengths in network databases

For each gene in the network, we traced back along the network by an exhaustive search of all the possible direct transcription cascades that affect this gene, and found the longest cascade that affects each gene. When a closed transcription loop was encountered, the cascade was terminated. The *E. coli* network[15] contains no closed loops, though in rare cases two transcription factors are encoded on the same operon, such as marA and marR, and regulate their own promoter, thus effectively regulating each other. The two yeast networks contain several loops (the YPD[23] network contains one two-gene and one three-gene loop, and the genome-wide location network[21] contains two two-gene loops and one three-gene loop). The much smaller developmental networks contain several closed loops (which may function as irreversible switches[25,30,36]): the sea urchin network[30] has two closed two-gene loops and the even smaller drosophila network contains four two-gene loops and one three-gene loop. The drosophila network is from the GeNet database (www.csa.ru/Inst/gorb_dep/inbios/genet/s8prrl.htm).

Bacterial strains and plasmids

E. coli strain Dh5α which expresses lacI was used in all experiments. Plasmid pairs pZS12-gfp + pZE21-MCS, pZS12-TetR + pZE21-gfp were used for the cascade in Figure 2(b) and (c).[24] All plasmids were based on the modular system described by Lutz & Bujard.[53] For the GFP variant used in this study, the biosynthetic delay between transcription onset and production of the first fluorescent GFP molecules is estimated at several minutes.[32,54] For pZS12-TetR and pZS12-gfp, the gene of interest (*tetR* or *gfp*[31]) was cloned into a vector containing the low-copy SC101 origin of replication, ampicillin-resistance gene, and $P_L$lacO1 promoter.[53] pZE21-MCSI contained a ColE1 origin, kanamycin-resistance gene, P_L-tetO1 promoter, and a multiple-cloning site, which was replaced by *gfp* to create pZE21-gfp. This plasmid configuration was chosen because it allows functional modulation of step 2 by IPTG. Other plasmid choices, such as placing tetR on a high-copy number plasmid, were non-functional, in the sense that they failed to cover much of the range of tet-promoter activity under various levels of IPTG (data not shown).[22]

Growth conditions and measurements

Cultures (2 ml) inoculated from single colonies were grown overnight in defined medium (M9 salts, 0.05% (w/v) Casamino acids, 0.5% (v/v) glycerol, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 1.5 μM thiamine and antibiotics: 50 μg/ml of kanamycin, 100 μg/ml of ampicillin) at 37 °C with shaking at 300 rpm. The cultures were diluted 1:100 (v/v) into fresh medium in a flat-bottom 96-well plate, at a final volume of 200 μl per well, and were grown at 36 °C or 27 °C with shaking. The bacteria were allowed to reach 1/10 stationary absorbance, and then diluted 1:10 (v/v) into fresh medium at the same temperature, with and without 1 mM IPTG to induce the lac promoter. Cultures were grown in a Wallac Victor2 multiwell fluorimeter set at 36 °C or 27 °C, and assayed with an automatically repeating protocol of shaking (2 mm double-orbital, normal speed, 20 seconds) absorbance (A) measurements (600 nm filter, 0.1 second, absorbance through approximately 0.5 cm of fluid), fluorescence readings (485 nm and 535 nm filters, 0.5 second, CW lamp energy 10,000 units), and a delay (100 seconds).[32,33] The time between repeated measurements was 5.75 minutes. Triplicates were averaged, and the fluorescence of the induced cultures was divided by the fluorescence of the uninduced culture. The relative fluorescence was normalized for each strain to a maximum level of one and a minimum of zero. The day-to-day reproducibility error of the normalized fluorescence is about 0.1.

Optimal response times for rapid turn-on and turn-off

Assume that protein X is produced at a rate of S_x, is degraded with a half-life of $\tau_{deg,x}$, and is diluted out by cells dividing with a cell-cycle time of τ. The basic equation describing X, the concentration of protein X,

† Drosophila early development occurs in the absence of cell divisions, and the time-scale may therefore be set by the degradation times of the regulatory proteins.

is†:[27,35]

$$dX/dt = S_x - \alpha_x x \quad (1)$$

in which the first term describes production and the second degradation plus dilution, with an effective dilution rate of $\alpha_x = \log(2)/\tau_x$, where the effective life-time is $\tau_x = (1/\tau + 1/\tau_{deg,x})^{-1}$. For rapidly degradable proteins with half-life $\tau_{deg,x}$ much smaller than τ, the effective life-time is approximately equal to the degradation half-life $\tau_x \approx \tau_{deg,x}$. For long-lived proteins with $\tau_{deg,x} \gg \tau$, the effective life-time is $\tau_x \approx \tau$. The biosynthetic delays between transcription onset and production of the first active product can usually be neglected because it is much shorter than typical protein life-times.

The solution, assuming that $X(t=0) = 0$, is (Figure 5(a), continuous line):

$$X(t) = X_{max}[1 - \exp(-\alpha_x t)] \quad (2)$$

where the steady-state level is $X_{max} = S_x/\alpha_x$. The response-time of this first step in the cascade, which is the time needed to reach $X_{max}/2$:

$$X(T_{ON,1}) = X_{max}[1 - \exp(-\alpha_x T_{ON,1})] = X_{max}/2 \quad (3)$$

is one effective life-time unit (one cell-cycle for long-lived proteins):

$$T_{ON,1} = \tau_x \text{ (first step)} \quad (4)$$

When the production of X is turned off, its concentration decays exponentially (Figure 5(b), continuous line):

$$X(t) = X_{max}\exp(-\alpha_x t) \quad (5)$$

and the response-time:

$$X(T_{OFF,1}) = X_{max}\exp(-\alpha_x T_{OFF,1}) = X_{max}/2 \quad (6)$$

is one effective life-time unit as well (one cell-cycle for long-lived proteins):

$$T_{OFF,1} = \tau_x \text{ (first step)} \quad (7)$$

Suppose now that X activates the transcription of a downstream gene Y, with a degradation half-life of $\tau_{deg,y}$, an effective life-time of:

$$\tau_y = (1/\tau + 1/\tau_{deg,y})^{-1}$$

and an effective dilution rate of $\alpha_y = \log(2)/\tau_y$. Assume, for simplicity, that X activates the transcription of Y through a step-like activation function,[34,36] so that Y is produced at a rate S_y when X is at a greater than a threshold concentration $X_{threshold}$, and at rate zero otherwise. More elaborate models, using Hill functions, produce the same qualitative results. The concentration of protein Y (Y) changes according to:

$$dY/dt = F(X) - \alpha_y Y, \text{ where } F(X < X_{threshold}) = 0 \text{ and } F(X > X_{threshold}) = S_y \quad (8)$$

The time when Y begins to be expressed is the time T_1 when X crosses the threshold (Figure 5(a), (c) and (e)). This occurs when:

$$X(T_1) = X_{max}[1 - \exp(-\alpha_x T_1)] = X_{threshold} \quad (9)$$

The solution for T_1 is:

$$T_1 = -(1/\alpha_x)\log(1 - X_{threshold}/X_{max}) \quad (10)$$

After time T_1, Y rises with a constant production rate S_y, and has simple exponential kinetics (as in equation (2)):

$$Y(t) = Y_{max}[1 - \exp(-\alpha_y(t - T_1))] \quad (11)$$

with $Y_{max} = S_y/\alpha_y$. The time to reach half-way to the maximal level, Y_{max}, is shifted by T_1 with respect to the response-time of X (equation (2)), and is equal to T_1 plus one effective protein life-time unit ($T_1 + \tau$ for long-lived protein Y):

$$T_{ON,2} = \tau_y + T_1 \text{ (second step)} \quad (12)$$

Conversely, Y is turned off when X crosses the threshold from above. Assume that at $t = 0$, protein X is at its steady-state level $X = X_{max}$, and then the production of X is turned off. The concentration of X decays as equation (4). It crosses the threshold $X_{threshold}$ at a time T_2 given by (Figure 5(b), (d) and (f)):

$$T_2 = -(1/\alpha_x)\log(X_{threshold}/X_{max}) \quad (13)$$

At time T_2, the production of Y ceases and it decays by dilution. Since the kinetics of Y is like that of X (equation (5)), only shifted by a time T_2, the corresponding OFF time is given by:

$$T_{OFF,2} = \tau_y + T_2 \text{ (second step)} \quad (14)$$

The sum of the ON and OFF response-times is:

$$T_{sum} = T_{ON,2} + T_{OFF,2} = 2\tau_y + T_1 + T_2 \quad (15)$$

The bigger $X_{threshold}$ is, the bigger is T_{ON} and the smaller is T_{OFF} (Figure 5). The activation threshold where T_{sum} is minimal can be found by differentiating T_{sum} by $X_{threshold}$, and seeking $d(T_{sum})/d(X_{threshold}) = 0$. The optimal threshold that gives minimal T_{sum} is:

$$X_{threshold}^{optimal} = X_{max}/2 \quad (16)$$

At this optimal solution the response-times in the ON and OFF directions are equal to each other, and both equal to the sum of the effective life-times of the two proteins:

$$T_{ON,2,optimal} = T_{OFF,2,optimal} = \tau_x + \tau_y \quad (17)$$

For long-lived proteins, $\tau_x = \tau_y = \tau$, and we obtain:

$$T_{ON,2,optimal} = T_{OFF,2,optimal} = \tau_x + \tau_y = 2\tau \quad (18)$$

Continuing in this fashion, it is easy to show that the response-time for the nth step in the cascade of long-lived proteins is n cell-cycles:

$$T_{ON,n,optimal} = T_{OFF,n,optimal} = n\tau \quad (19)$$

Acknowledgements

We thank M. Elowitz and S. Leibler for plasmids and discussions. We thank S. Shen-Orr and P. Bashkin for their contributions. We thank O. Hobart, M. Surette and all members of our laboratory for discussions. This work was supported by the Minerva fellowship, the Israel Science Foundation and the Human Frontiers Science Project. We dedicate this work to the memory of Yasha (Yaakov) Rosenfeld.

References

1. Savageau, M.A. (1971). Parameter sensitivity as a

criterion for evaluating and comparing the performance of biochemical systems. *Nature*, **229**, 542–544.
2. Savageau, M. A. (1974). Comparison of classical and autogenous systems of regulation in inducible operons. *Nature*, **252**, 546–549.
3. Savageau, M. A. (1977). Design of molecular control mechanisms and the demand for gene expression. *Proc. Natl Acad. Sci. USA*, **74**, 5647–5651.
4. Arnone, M. I. & Davidson, E. H. (1997). The hardwiring of development: organization and function of genomic regulatory systems. *Development*, **124**, 1851–1864.
5. Savageau, M. (2001). Design principles for elementary gene circuits: elements, methods, and examples. *Chaos*, **11**, 142–159.
6. Thieffry, D., Huerta, A. M., Perez-Rueda, E. & Collado-Vides, J. (1998). From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in *Escherichia coli*. *Bioessays*, **20**, 433–440.
7. Hartwell, L. H., Hopfield, J. J., Leibler, S. & Murray, A. W. (1999). From molecular to modular cell biology. *Nature*, **402**, C47–C52.
8. McAdams, H. H. & Arkin, A. (2000). Towards a circuit engineering discipline. *Curr. Biol.* **10**, R318–R320.
9. Barkai, N. & Leibler, S. (1997). Robustness in simple biochemical networks. *Nature*, **387**, 913–917.
10. Alon, U., Surette, M. G., Barkai, N. & Leibler, S. (1999). Robustness in bacterial chemotaxis. *Nature*, **397**, 168–171.
11. Yi, T. M., Huang, Y., Simon, M. I. & Doyle, J. (2000). Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proc. Natl Acad. Sci. USA*, **97**, 4649–4653.
12. Barkai, N. & Leibler, S. (2000). Circadian clocks limited by noise. *Nature*, **403**, 267–268.
13. Vilar, J. M., Kueh, H. Y., Barkai, N. & Leibler, S. (2002). Mechanisms of noise-resistance in genetic oscillators. *Proc. Natl Acad. Sci. USA*, **99**, 5988–5992.
14. Wyrick, J. J. & Young, R. A. (2002). Deciphering gene expression regulatory networks. *Curr. Opin. Genet. Dev.* **12**, 130–136.
15. Shen-Orr, S. S., Milo, R., Mangan, S. & Alon, U. (2002). Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nature Genet.* **31**, 64–68.
16. Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D. & Alon, U. (2002). Network motifs: simple building blocks of complex networks. *Science*, **298**, 824–827.
17. Wagner, A. & Fell, D. A. (2001). The small world inside large metabolic networks. *Proc. R. Soc. ser. B*, **268**, 1803–1810.
18. Guelzim, N., Bottani, S., Bourgine, P. & Kepes, F. (2002). Topological and causal structure of the yeast transcriptional regulatory network. *Nature Genet.* **31**, 60–63.
19. Newman, M. E., Strogatz, S. H. & Watts, D. J. (2001). Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. ser. E*, **64**, 026118.
20. Maslov, S. & Sneppen, K. (2002). Specificity and stability in topology of protein networks. *Science*, **296**, 910–913.
21. Lee, T. I., Rinaldi, N. J., Robert, F., Odom, D. T., Bar-Joseph, Z., Gerber, G. K. *et al.* (2002). Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science*, **298**, 799–804.
22. Yokobayashi, Y., Weiss, R. & Arnold, F. H. (2002). Directed evolution of a genetic circuit. *Proc. Natl Acad. Sci. USA*, **99**, 16587–16591.
23. Costanzo, M. C., Crawford, M. E., Hirschman, J. E., Kranz, J. E., Olsen, P., Robertson, L. S. *et al.* (2001). YPD, PombePD and WormPD: model organism volumes of the BioKnowledge library, an integrated resource for protein information. *Nucl. Acids Res.* **29**, 75–79.
24. Elowitz, M. B. & Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, **403**, 335–338.
25. Gardner, T. S., Cantor, C. R. & Collins, J. J. (2000). Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, **403**, 339–342.
26. Becskei, A. & Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature*, **405**, 590–593.
27. Rosenfeld, N., Elowitz, M. B. & Alon, U. (2002). Negative autoregulation speeds the response times of transcription networks. *J. Mol. Biol.* **323**, 785–793.
28. Guet, C. C., Elowitz, M. B., Hsing, W. & Leibler, S. (2002). Combinatorial synthesis of genetic networks. *Science*, **296**, 1466–1470.
29. Simon, I., Barnett, J., Hannett, N., Harbison, C. T., Rinaldi, N. J., Volkert, T. L. *et al.* (2001). Serial regulation of transcriptional regulators in the yeast cell cycle. *Cell*, **106**, 697–708.
30. Davidson, E. H., Rast, J. P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C. H. *et al.* (2002). A genomic regulatory network for development. *Science*, **295**, 1669–1678.
31. Cormack, B. P., Valdivia, R. H. & Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene*, **173**, 33–38.
32. Kalir, S., McClure, J., Pabbaraju, K., Southward, C., Ronen, M., Leibler, S. *et al.* (2001). Ordering genes in a flagella pathway by analysis of expression kinetics from living bacteria. *Science*, **292**, 2080–2083.
33. Ronen, M., Rosenberg, R., Shraiman, B. I. & Alon, U. (2002). Assigning numbers to the arrows: parameterizing a gene regulation network by using accurate expression kinetics. *Proc. Natl Acad. Sci. USA*, **99**, 10555–10560.
34. McAdams, H. H. & Arkin, A. (1998). Simulation of prokaryotic genetic circuits. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 199–224.
35. Monod, J., Pappenheimer, A. & Cohen-Bazire, G. (1953). La cinetique de la biosynthese de la b-Galactosidase chez *E. coli* consideree comme fonction de la croissance. *Biochim. Biophys. Acta*, 9.
36. Thomas, R. (1973). Boolean formalization of genetic control circuits. *J. Theor. Biol.* **42**, 563–585.
37. Schimke, R. (1969). On the roles of synthesis and degradation in regulation of enzyme levels in mammalian tissues. *Curr. Top. Cell. Regul.* **1**, 77–124.
38. Judd, E. M., Laub, M. T. & McAdams, H. H. (2000). Toggles and oscillators: new genetic circuit designs. *Bioessays*, **22**, 507–509.
39. Aldridge, P. & Hughes, K. T. (2002). Regulation of flagellar assembly. *Curr. Opin. Microbiol.* **5**, 160–165.
40. Aizawa, S. I. & Kubori, T. (1998). Bacterial flagellation and cell division. *Genes Cells*, **3**, 625–634.
41. Little, J. W. & Mount, D. W. (1982). The SOS regulatory system of *Escherichia coli*. *Cell*, **29**, 11–22.
42. Courcelle, J., Khodursky, A., Peter, B., Brown, P. O. & Hanawalt, P. C. (2001). Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient *Escherichia coli*. *Genetics*, **158**, 41–64.
43. Makarova, K. S., Mironov, A. A. & Gelfand, M. S.

(2001). Conservation of the binding site for the arginine repressor in all bacterial lineages. *Genome Biol.* **2**, 13.
44. Laub, M. T., McAdams, H. H., Feldblyum, T., Fraser, C. M. & Shapiro, L. (2000). Global analysis of the genetic network controlling a bacterial cell cycle. *Science*, **290**, 2144–2148.
45. Ferrell, J. E., Jr & Machleder, E. M. (1998). The biochemical basis of an all-or-none cell fate switch in Xenopus oocytes. *Science*, **280**, 895–898.
46. Bourke Arnvig, K., Pedersen, S. & Sneppen, K. (2000). Thermodynamics of heat-shock response. *Phys. Rev. Letters*, **84**, 3005–3008.
47. Lawrence, P. (1992). *The Making of a Fly,* Blackwell Science Inc, Oxford.
48. Maduro, M. F. & Rothman, J. H. (2002). Making worm guts: the gene regulatory network of the *Caenorhabditis elegans* endoderm. *Dev. Biol.* **246**, 68–85.
49. Kim, S. K., Lund, J., Kiraly, M., Duke, K., Jiang, M., Stuart, J. M. *et al.* (2001). A gene expression map for *Caenorhabditis elegans*. *Science*, **293**, 2087–2092.
50. Rothenberg, E. V. & Anderson, M. K. (2002). Elements of transcription factor network design for T-lineage specification. *Dev. Biol.* **246**, 29–44.
51. Tomancak, P., Beaton, A., Weiszmann, R., Kwan, E., Shu, S., Lewis, S. E. et al. (2002). Systematic determination of patterns of gene expression during Drosophila embryogenesis. Genome Biol, 3, research0088.1–research0088.14.
52. Stragier, P. & Losick, R. (1990). Cascades of sigma factors revisited. *Mol. Microbiol.* **4**, 1801–1806.
53. Lutz, R. & Bujard, H. (1997). Independent and tight regulation of transcriptional units in *Escherichia coli via* the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucl. Acids Res.* **25**, 1203–1210.
54. Cluzel, P., Surette, M. & Leibler, S. (2000). An ultrasensitive bacterial motor revealed by monitoring signaling proteins in single cells. *Science,* **287**, 1652–1655.

Edited by M. Yaniv

(Received 30 January 2003; received in revised form 1 April 2003; accepted 9 April 2003)