doi:10.1016/j.jmb.2003.09.049 J. Mol. Biol. (2003) 334, 197–204

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

The Coherent Feedforward Loop Serves as a Sign-sensitive Delay Element in Transcription Networks

S. Mangan, A. Zaslaver and U. Alon*

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

*Corresponding author

Recent analysis of the structure of transcription regulation networks revealed several "network motifs": regulatory circuit patterns that occur much more frequently than in randomized networks. It is important to understand whether these network motifs have specific functions. One of the most significant network motifs is the coherent feedforward loop, in which transcription factor X regulates transcription factor Y, and both jointly regulate gene Z. On the basis of mathematical modeling and simulations, it was suggested that the coherent feedforward loop could serve as a sign-sensitive delay element: a circuit that responds rapidly to step-like stimuli in one direction (e.g. ON to OFF), and at a delay to steps in the opposite direction (OFF to ON). Is this function actually carried out by feedforward loops in living cells? Here, we address this experimentally, using a system with feedforward loop connectivity, the L-arabinose utilization system of *Escherichia coli*. We measured responses to step-like cAMP stimuli at high temporal resolution and accuracy by means of green fluorescent protein reporters. We show that the arabinose system displays sign-sensitive delay kinetics. This type of kinetics is important for making decisions based on noisy inputs by filtering out fluctuations in input stimuli, yet allowing rapid response. This information-processing function may be performed by the feedforward loop regulation modules that are found in diverse systems from bacteria to humans.

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

Introduction

Cells compute responses to external stimuli using networks of regulatory interactions. A major goal of biology is to understand the dynamics of these complex networks.[1–3] Great simplification would occur if the network could be studied in terms of recurring circuit elements,[4,5] each with a defined signal-processing function.

Recently, an approach for discovering significantly recurring patterns in networks was introduced.[2,3] This was based on detecting network motifs: connectivity patterns that occur much more frequently than in randomized "control" networks built of the same components as the real network. The transcription network of *Escherichia coli* was found to contain several highly significant motifs.[2] It was then found that *Saccharomyces cerevisiae* shares the same network motifs.[3,6] One of the most significant motifs in both *E. coli* and yeast is the feedforward loop (FFL).[2] The FFL appears in hundreds of diverse, non-homologous gene systems.[2,3,6]

The FFL is composed of a transcription factor X, which regulates a second transcription factor Y, such that X and Y jointly regulate gene Z (Figure 1(a)). The transcription factors X and Y usually have inducers, Sx and Sy, respectively, which are small molecules, protein partners or covalent modifications that activate or inhibit their transcriptional activity (Figure 1(a)). Each of the three transcription interactions in the FFL can be either positive (activation) or negative (repression). There are therefore eight possible structural configurations of connection signs. Four of these configurations are termed "coherent": the sign of the direct regulation path (from X to Z) is the

Abbreviations used: GFP, green fluorescent protein; FFL, feedforward loop.

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

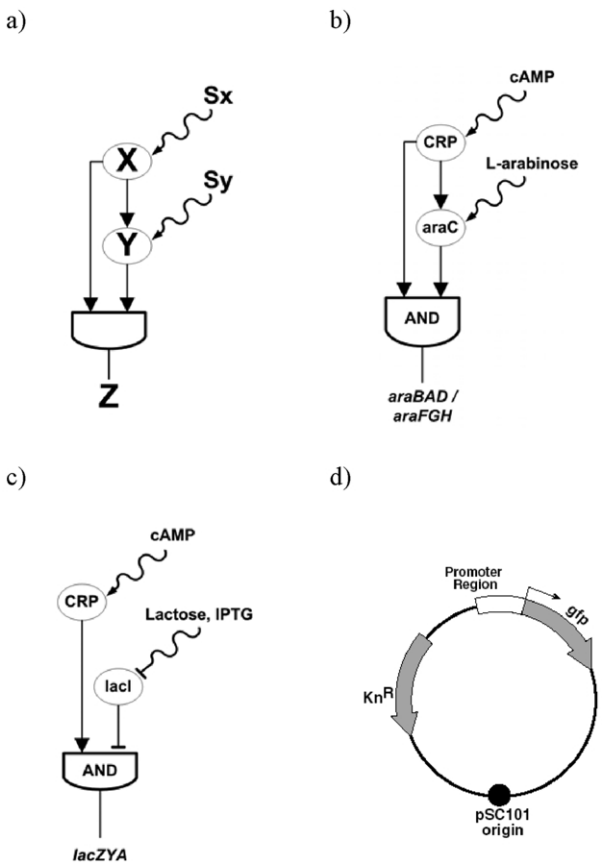

Figure 1. (a) The coherent type-1 feedforward loop and the inducers Sx and Sy; (b) the ara system; (c) the lac system; (d) the low-copy reporter plasmid used in the measurements.

same as the overall sign of the indirect regulation path (from X through Y to Z).[2] The other four structures are termed "incoherent": the signs of the direct and indirect regulation paths are opposite. The FFL with three positive interactions, termed type-1 coherent FFL, is by far the most common configuration in *E. coli*.[7] To understand the regulation of Z, one needs also to specify the *cis*-regulatory logic[8–10] that combines the two inputs X and Y, such as AND-gate logic, in which both X and Y are needed, and OR-gate logic in which either X or Y is sufficient to activate Z.

In a previous study,[2] based on numerical simulations, we suggested that the coherent FFL with AND logic is a processing element that functions as a persistence detector. Only a persistent stimulus of Sx can activate both X and Y, and lead to expression of Z. On the other hand, even a temporary removal of the Sx stimulus leads to a rapid turn-off of Z expression. An equivalent and more intuitive description is that the FFL is a sign-sensitive delay element[7]: it responds rapidly to step-like stimuli of Sx in one direction (ON to OFF), and at a delay to steps in the opposite direction (OFF to ON). By sign-sensitive delay, we mean that the response time to step-like stimuli is not symmetric and depends on the sign of the step.

Here, we present experimental results that support this premise. We select a representative gene system with an FFL connectivity, we show that this system is a coherent AND-gate FFL (in the sense that Y is regulated significantly by X, and that both X and Y are needed for Z expression). Then, we show that the system displays sign-sensitive delay kinetics. We discuss the biological function of sign-sensitive delay as a filter that can protect the target gene from fluctuations in the input stimuli.

Results

The experimental system

To experimentally study the FFL, we selected one of the best-characterized systems in *E. coli*, the L-arabinose (ara) utilization system (Figure 1(b)).[11–15] The ara system includes the catabolism operon *araBAD*, and transporters such as *araFGH*. Both *araBAD* and *araFGH* are regulated transcriptionally by two transcription factors, AraC and CRP. AraC acts as a transcriptional activator when it binds the sugar L-arabinose, and as a repressor in its absence. CRP acts as an activator when it binds the inducer cyclic AMP (cAMP). cAMP is a molecule that is produced within the cell upon glucose starvation (e.g. during growth on glycerol as sole carbon source), and whose production is suppressed during growth on glucose. In addition, CRP binds the *araC* promoter and enhances the transcription of AraC.[14,15] Therefore, in the presence of L-arabinose, the ara system has the connectivity of a type-1 coherent FFL, in which the inducer Sx is cAMP (Figure 1(b)).

As a control, we required a non-FFL system with the same input Sx. We chose the lactose (lac) utilization system, which has a simple AND-gate structure (Figure 1(c)). In the lac system,[16,17] CRP and LacI jointly regulate the *lacZYA* operon, but with no transcription regulation of LacI by CRP.

We generated low-copy reporter plasmids in which the *araC*, *araBAD*, *araFGH* and *lacZYA* promoters control the green fluorescent protein (*gfp*) gene.[10,18,19] The *gfp* variant used becomes fluorescent within a few minutes of transcription,[18] allowing rapid responses to be measured. Promoter activity was measured by means of an automated multiwell fluorimeter from cultures of strains bearing the reporter plasmids.[18,19] Both green fluorescent protein (GFP) fluorescence and optical absorbance ($A_{600\,\mathrm{nm}}$) were measured at a very high temporal resolution (about once per minute). These measurements were used to determine the dynamics of GFP concentration produced by the promoters. We applied steps of glucose and exogenous cAMP, to determine the temporal response of the ara FFL and the reference lac system to cAMP steps.

The ara system is a type-1 coherent FFL with AND logic

The ara system has FFL connectivity according to transcription databases.[2,20] In order to function as a type-1 coherent FFL, *araC* expression needs to be activated significantly by CRP in the presence of cAMP. We measured *araC* promoter activity with and without cAMP, during growth in minimal medium supplemented with saturating concentration of L-arabinose (Table 1(A) and (B)). Levels of cAMP were controlled in two ways: (a) cells growing on glucose (low endogenous cAMP) were compared to cells growing on glycerol (high endogenous cAMP production); (b) cells growing on glucose were compared to cells growing on glucose with saturating exogenous cAMP. In all cases, the cells were compared at the same A value. Both assays showed similar results (Table 1). The *araC* promoter exhibited a measurable basal level of expression in the presence of L-arabinose and in the absence of cAMP. The promoter activity increased at least four- to five-fold above this basal level in the presence of cAMP (Table 1(C)).

In addition, we measured the promoter activity of the *lacZYA*, *araBAD* and *araFGH* promoters under all combinations of zero or saturating inducers cAMP and arabinose/IPTG (Table 1(A) and (B)). We find that the expression in the presence of both inducers is at least an order of magnitude greater then the expression measured when either or both inducers are missing (Table 1(A) and (B)). This confirms previous studies[14,15] that suggest that these promoters behave as an AND-gate with respect to their inputs Sx = cAMP and Sy = L-arabinose (or Sy = IPTG in the case of lac[10]). These results demonstrate that the ara system is a type-1-AND-gate FFL, although not a perfect one, since *araC* has a significant basal level of expression.

Table 1. GFP/A of reporter strains during exponential growth ($A = 0.03$), under all combinations of zero or saturating inducers arabinose/IPTG and cAMP

cAMP	−	+Int	+Ext	−	+Int	+Ext
Arabinose /IPTG	−	−	−	+	+	+
A. GFP/A at 30 °C (× 10³, GFP/A units)						
araC	5	12	20	8	34	42
araBAD	<1	<1	<1	<1	110	60
araFGH	<1	<1	<1	<1	160	120
lacZYA	<1	25	48	42	250	360
B. GFP/A at 37 °C (× 10³, GFP/A units)						
araC	10	18	90	20	90	190
araBAD	<1	<1	<1	<1	200	580
araFGH	<1	<1	<1	<1	150	460
lacZYA	<1	40	60	55	170	460

C. araC fold induction by cAMP In the presence of arabinose

	30 °C		37 °C	
cAMP	Int	Ext	Int	Ext
	4.2	5.2	4.5	9.5

cAMP Ext is 20 mM cAMP added to cells growing on glucose minimal medium. cAMP Int is growth on glycerol as carbon source, resulting in endogenous production of cAMP. A, Growth at 30 °C; B, growth at 37 °C; C, fold induction of the *ara*C promoter by cAMP, in the presence of L-arabinose. The standard error in all GFP/A measurements, based on two to six repeated experiments was, at most, the greater of ± 10% or 2×10^3. Values smaller than 1×10^3 are marked as <1.

The ara system acts as a sign-sensitive delay with respect to cAMP steps

We studied the temporal responses of the ara system to cAMP steps. We generated an ON step of cAMP by adding saturating cAMP to cells growing exponentially on glucose minimal medium. A cAMP OFF step was generated by adding saturating glucose to cells growing exponentially in glycerol minimal medium. In order to affect the Y transcription factor in the ara and lac systems appropriately, saturating concentrations of L-arabinose and IPTG were included in the medium. We calculated the promoter activity (rate of GFP production per cell) from the GFP and absorbance measurements,[19] and measured the time it took the promoter activity to reach 50% of its maximal level. This corresponds to the onset time of transcription following the step stimulus.

We find that the ON response of both *araBAD* and *araFGH* was significantly slower than the ON response of *lacZYA*. The OFF response of *araBAD* and *araFGH* was more rapid, and identical with that of *lacZYA*. In contrast, the *araC* promoters responded just as rapidly as *lacZYA* to both ON and OFF steps. Representative ON and OFF kinetics are shown in Figures 2 and 3. Repeated measurements showed a significant and reproducible delayed response to cAMP ON steps for *araBAD* relative to *lacZYA* of about 13 minutes at 30 °C, and six minutes at 37 °C (Table 2). This delay corresponds to about 0.2 cell-division times

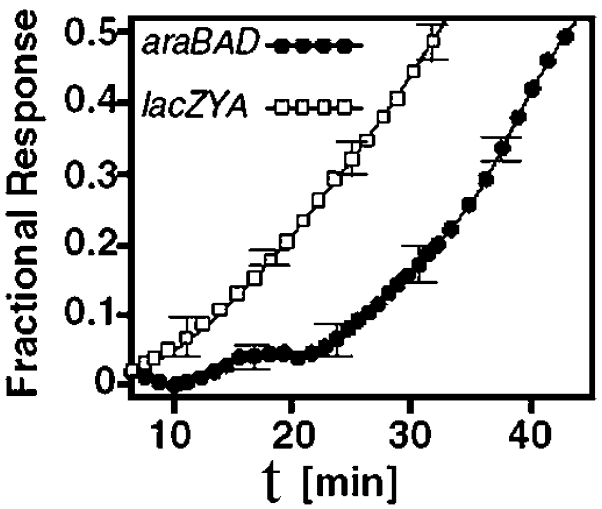

Figure 2. GFP/A, normalized by its maximal level, of *araBAD* and *lacZYA* reporters growing on glucose minimal medium, following a cAMP ON step (10 mM cAMP was added at time $t = 0$). Error bars represent standard error of triplicate wells. The kinetics of *araFGH* (not shown) is similar to the *araBAD* kinetics.

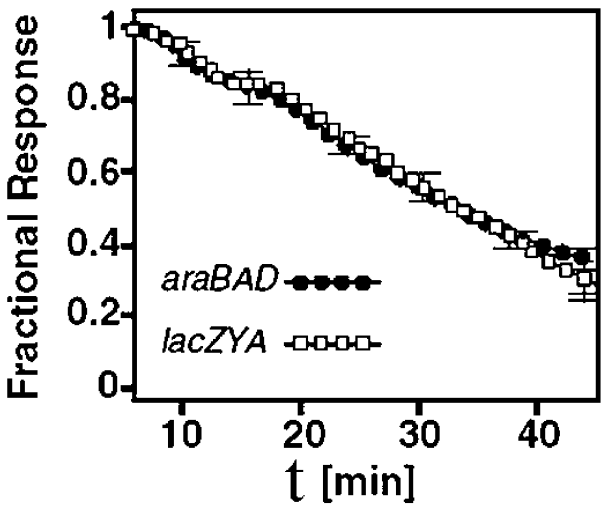

Figure 3. GFP/*A*, normalized by its maximal level, of *araBAD* and *lacZYA* reporters growing on glycerol minimal medium, following a cAMP OFF step (0.4% glucose was added at time $t = 0$). Error bars represent standard error of triplicate wells. The kinetics of *araFGH* (not shown) is similar to the *araBAD* kinetics.

at both temperatures. Similarly, *araFGH* showed a delay of about 20 minutes at 30 °C, and 13 minutes at 37 °C relative to *lacZYA*. This delay corresponds to about 0.3 and 0.4 cell-division times at 30 °C and 37 °C. In contrast, the response to OFF steps was indistinguishable for all promoters, consisting of simple exponential decay with equal timing. The asymmetric behavior of *araBAD* and *araFGH*, with delayed responses to cAMP ON steps but not to OFF steps (relative to the *lacZYA* control), is the hallmark of sign-sensitive delay.

Simulations on the effect of *araC* basal level

We performed simulations to check the effect of the basal level of *araC* on the sign-sensitive delay

Table 2. Delay of onset of expression changes of the *araC*, *araBAD* and *araFGH* promoters, relative to the *lacZYA* promoter following cAMP steps

	araC	*araBAD*	*araFGH*
A. ON step delay at 30 °C (minutes)			
T–Tz	−1.0	12.6	19.6
SE	2	2	2
B. OFF step delay at 30 °C (minutes)			
T–Tz	−1.5	2.6	3.4
SE	6	2	2
C. ON step delay at 37 °C (minutes)			
T–Tz	−1.5	6.3	13.5
SE	3	1	2
D. OFF step delay at 37 °C (minutes)			
T-Tz	−2.5	−0.7	0.0
SE	3	2	3

T–Tz, time after *lacZYA*. SE, standard error of two to six repeats. A, Response to ON step at 30 °C; B, response to OFF step at 30 °C; C, response to ON step at 37 °C; D, response to OFF step at 37 °C.

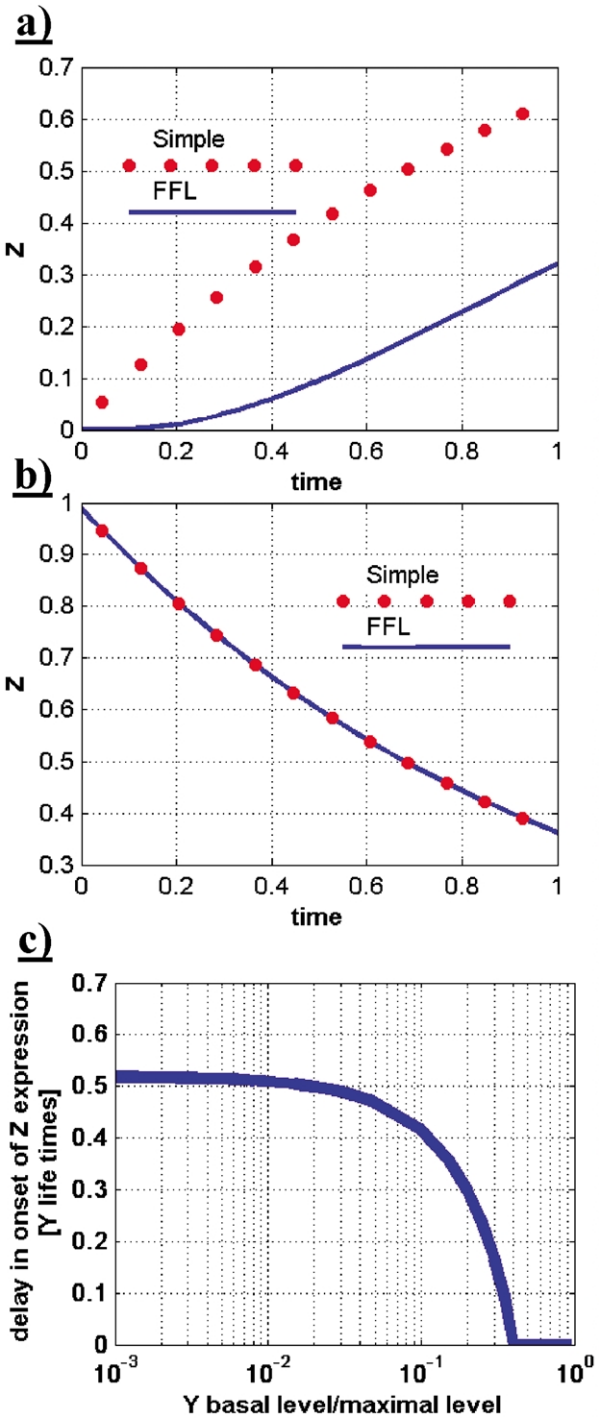

Figure 4. Simulation of type-1 coherent AND FFL. (a) Response to ON step of Sx, showing concentration of Z as a function of time. Inducer Sy is present throughout. Also shown are kinetics of a simple AND-gate system (no regulation of Y by X). (b) Response to OFF step of Sx. (c) Delay in onset of Z promoter activity as a function of Y basal level. Model parameters used are:

$$\alpha_1 = \alpha_2 = 1, \quad \beta_1 = \beta_2 = 1, \quad H = 2,$$

$$K_{xy} = K_{xz} = K_{yz} = 0.5 \quad B_y = 0$$

in (a) and (b), B_y is varied in (c).

behavior (Figure 4). We find that the delay after an ON step decreases as the basal level of Y increases. In other words, at large basal level of Y, the response becomes more symmetric and the sign-sensitive delay effect is reduced. We find that at the measured basal level of Y (about 0.2 of its maximal level), the sign-sensitive delay function is still quite strong (Figure 4c; see also Materials and Methods). In addition, for long-lived proteins, the model suggests that the delay of the FFL should depend on the cell-division time (the timescale that controls the response time of Y and Z[21,22]). This is in agreement with the presently observed increase of the delay at 30 °C compared to 37 °C, in which cell division time is shorter (Table 2).

Discussion

We chose the most common coherent FFL configuration,[7] the type-1 coherent FFL, for experimental study. We employed the ara feedforward loop system, and experimentally studied its dynamic response to cAMP input variations. We found that the ara system responded as a sign-sensitive delay element, with delayed responses to ON steps of cAMP and rapid responses to OFF steps.

Biological function of sign-sensitive delay

The presence of the FFL in the ara system raises the question of its biological function. One simple function performed by the ara FFL is computing AND-logic between the two input stimuli, to decide when the ara system should turn on. But for this purpose, a simple AND regulatory unit would be enough. The feedforward connectivity, and in particular the regulation of Y by X, presumably gives the bacteria some advantage, otherwise it would be lost by evolutionary forces. One explanation of the regulation of *araC* by CRP is that it allows modulating the level of AraC, thereby supplying more copies when the system is active. We believe that there could be an additional aspect that involves signal-processing: the bacterium may have an advantage in a rapidly varying environment if it responds asymmetrically when signals turn ON *versus* OFF.

What is the advantage of an asymmetric response to cAMP signals in the ara system? Our ability to address this is limited by lack of information on the natural ecology of *E. coli*. The following scenario emerges from the little that is known of this ecology. There are two main phases in the life of *E. coli*, one inside and one outside of its mammalian host.[23–26] The ecology inside the host is thought to be rich in sugars such as L-arabinose,[25] whereas glucose appears in rare pulses that are quickly depleted.[26] The ara system is therefore expressed during a large fraction of the time, since both of its inducers, arabinose and cAMP, are present. When glucose appears, resulting in a cAMP OFF step, it is utilized immediately by the cells. At this time, it is important to quickly shut down the expression of the ara genes to avoid needless production of about 1% of the cells protein. The ara proteins are diluted by cell growth. Glucose pulses are short, since glucose is taken up efficiently by both bacteria and host. Therefore, when glucose is depleted, resulting in a cAMP ON step, the cell is still likely to possess significant amounts of ara proteins. Hence, a delay in the turn-on of ara expression, in response to cAMP ON steps, can be tolerated, and is in fact advantageous due to savings in protein production. Thus the sign-sensitive delay function of the FFL may reflect the ecology in which it evolved.

In accord with these considerations, we note that coherent FFL appears in utilization systems of other sugars that are thought to be common in the host,[23–26] such as fucose and maltose.[2,7] It does not appear in sugars thought to be rare,[23–26] such as lactose. More generally, the FFL can protect the target genes from transient cAMP ON signals, allowing them to respond only to persistent stimuli.

Dependence of FFL function on biochemical parameters

The kinetic behavior of the FFL circuit is governed by the biochemical parameters of its three transcriptional regulation reactions. A simple mathematical model suggests that sign-sensitive delay is obtained for a wide range of parameters[7] (see Materials and Methods). The delay time in response to ON steps, t_{ON}, is determined by the time it takes for Y to reach levels sufficient to activate Z. We find that the sign-sensitive delay can be quite large even if Y has a significant basal level of expression (Figure 4(c)). The delay t_{ON} increases with increasing ratios of K_{zy}/Y_{max} where K_{zy} is the activation coefficient of Z by Y, and Y_{max} is the maximum level of Y. The precise value of t_{ON} can, in principle, be tuned by changing these parameters, for example by mutations that affect the binding of Y to the Z promoter.

The present study focused on a system with AND logic at the Z promoter. One can demonstrate readily that coherent FFLs with OR gates display sign-sensitive delay, but with reversed signs: a delay occurs in response to OFF steps, and rapid responses for ON steps of Sx.[7] Thus, OR-gate coherent feedforward-loops can protect the target genes from transient inactivation of the input signal.

Coherent feedforward loop in transcription networks of microorganisms and animals

Many of the coherent feedforward loops in the transcription network of *E. coli* appear in systems that respond to glucose (ara, mal), nitrogen (rpoN, glnA) and drugs (rob, mar). Interestingly, systems that respond to the same stimuli in *S. cerevisiae* also display coherent feedforward loops.[7] These *E. coli*

and yeast systems are generally non-homo logous.[27,28] This raises the possibility of convergent evolution to the same regulatory circuit, perhaps because it is needed for the same signal-processing functions in response to key fluctuating stimuli.[2,28]

Coherent FFLs also appears in developmental transcription networks. In the specification process of touch-cell fate in *Caenorhabditis elegans*, the activator unc-86 activates mec-3, and both act in a complex to activate genes such as mec-4.[29] An additional example is an FFL involving Ttx-3 and ceh-23, which is involved in interneuron cell fate determination.[30] There is evidence for several coherent FFLs in sea urchin development, such as one involving Otx and gata-E in endomesoderm development.[31] An example in human cells is E2F, which regulates p73 (and indirectly also p53), and both regulate apoptosis genes such as apaf-1.[32] An additional example involves interferon-γ induction in T-cells.[33] In the coherent FFLs found in eukaroytes, the activators X and Y often act as a protein complex to activate Z.

FFL in other biological networks

The FFL is found in the network of synaptic connections of *C. elegans* neurons.[3,34] It is found also in protein–protein interaction networks and signal transduction pathways. One example is the PIP2 (phospho-inositol 4,5-bisphosphate) signaling pathway:[35] upon signaling, the PIP2 molecule is broken into two parts, one part binds protein kinase C (PKC) directly, while the other part stimulates release of Ca^{2+} that binds PKC as well. It would be intriguing to test experimentally whether the coherent FFL carries out sign-sensitive delay function also in these networks.

Summary

The present study demonstrated experimentally that the FFL network motif carries out a signal-processing function. It would be important to experimentally study other systems with FFL connectivity. The present approach could be used to study the signal-processing roles of other network motifs.

Materials and Methods

Plasmids and strains

Promoter regions were PCR amplified from MG1655 genomic DNA with the following start and end genomic coordinates:[36] *lacZYA* (365438–365669), *araC* (69973–70452), *araBAD* (70452–69973), *araFGH* (1984067–1984952). This included the entire region between open reading frames (ORFs) with an additional 50–150 bp into each of the flanking ORFs. The promoter regions were sub-cloned into *Xho* I and *Bam* HI sites upstream of a promoterless *gfpmut2*[37] gene in a low-copy pSC101-origin plasmid as described.[19] The plasmids were transformed into *E. coli* strain MG1655[36] (wild-type for *ara* and *lac*). The *araC* and *araBAD* operons are divergent operons that share the same intragenic regulatory region, represented by the same fragment in opposite orientations in our reporter plasmids. This promoter region includes the O1, I1 and I2 AraC sites but not the O2 AraC site, which participates in a DNA loop with the O1 site in the absence of arabinose.[15] The O2 binding site of *araC* was not included for two reasons: first, the O2 site has been well documented to bind AraC only in the absence of arabinose.[15,38] Since saturating arabinose is present in the kinetic measurements, the absence of O2 is not expected to affect the results of the induction kinetics by cAMP. The second reason is technical: the O2 site lies inside the coding region of the *araC* gene. Including it would mean expressing a sizable fragment of the AraC protein from our reporter plasmid, which would potentially affect the system. In contrast, the reporter plasmid for *araFGH* includes the entire wild-type control region, without missing any known binding site. Furthermore, *araFGH* does not share any site with the *araC* operon, and thus represents a more typical FFL configuration.

Media

The medium combinations used are denoted as follows: M0 is minimal medium (M9 supplemented with 10 mg/ml of thiamine, 1 mM $MgSO_4$, 0.1 mM $MgCl_2$, 20 mg/ml of thymine, 50 mg/ml of each of the 20 amino acids except tryptophan, 0.5% (w/v) glycerol, 25 mg/ml of kanamycin. Overnight medium MON is M0 supplemented with 0.4% (w/v) glucose. Medium M1 is M0 containing 10 mM arabinose and 1 mM IPTG. M1G is M1 containing 0.8% (w/v) glucose. M1GC is M1 containing 0.8% glucose and 10 mM cAMP. Supplement stocks: G, 6% (w/v) glucose; C, 150 mM cAMP; W, doubly distilled water. To prepare C and M1GC, 150 mM cAMP was dissolved in 0.1 M NaOH.

Culture and measurements

Single colonies grown on agar plates were inoculated in 10 ml of MON, and grown overnight at 37 °C with shaking at 250 rpm. Cultures were washed by centrifugation and fresh M0 medium was added. The cultures were diluted 1:50 (v/v) into 150 μl in a 96 well plate prepared with all combinations of reporter strains and media (M1, M1G, M1GC). After covering each well with 100 μl of mineral oil (Sigma M-3516) to prevent evaporation, the plate was placed in a Wallac Victor2 multiwell fluorimeter at 30 °C or 37 °C with a repeating protocol that included shaking (2 mm orbital, normal speed, 30 seconds), absorbance (A) measurements (600 nm, 0.1 second), and fluorescence (GFP) readings (excitation 485 nm, emission 535 nm, 0.2 seconds, CW lamp energy 15,000).[18,19,21] The protocol repeated the set of {Shake,GFP,OD,GFP}. The time between repeated GFP measurements was about one minute. When the cultures reached mid-exponential growth ($A_{600} = 0.05$), 10 μl of supplements (G, C and W) were added to create the stimuli step functions (resulting in a two to five minute pause in measurements). The plate was returned to the fluorimeter, and the same protocol was repeated for several additional hours. Each plate contained two or three replicates of each combination of strain-medium-supplement, where the strains were: *araC*, *araBAD*, *araFGH*, *lacZYA*, promoterless vector, and a reference

well with medium only, and the media: M1, M1G, M1GC. The supplements G, C and W were added as follows: M1 received supplement G (as OFF step), and W (for control); M1G received supplement C (as ON step), and W (for control); M1GC received only supplement W (for control). The outcome measurements of each experiment were: the initial kinetics of each strain in the initial media (M1, M1G and M1GC); The kinetics after a step stimulus (G or C); reference kinetics without stimulus (W). The exponential cell division time in the various media was between 0.5 hours and 0.7 hour at 37 °C, and 0.8–1 hours at 30 °C.

Data processing

Background absorbance from wells containing medium only was subtracted from all absorbance readings. Background GFP fluorescence, determined by the GFP reading from a strain bearing the promoterless vector at the same condition and absorbance, was subtracted from the GFP readings. GFP fluorescence per cell was GFP(t)/A(t). To normalize the step response from GFP fluorescence into dimensionless values between zero and one, we subtracted the GFP fluorescence under the appropriate lower base-line medium conditions, and divided by the upper-baseline condition results. The lower baseline condition for M1 + G is M1G, and the upper baseline is M1. The lower baseline condition for M1G + C is M1G and the upper baseline is M1GC. The promoter activity, rate of GFP production per cell, was calculated as (dGFP(t)/dt)/A(t).[19] Note that the lac system has a higher basal expression than the ara promoters. The delay of the ara promoters relative to lac is, however, not due to the time it takes them to cross the detection threshold. The time it takes the ara promoters to cross the detection threshold of the present measurement system (about 1000 GFP/A units) can be estimated to be of the order of one minute.

Mathematical model of FFL with basal expression of Y

To evaluate the effect of Y basal level on FFL function, we performed simulations using standard mass-action equations:[2,39,40]

$$dY/dt = B_y + \beta_1 f(X, K_{xy}) - \alpha_1 Y,$$

$$dZ/dt = \beta_2 f(X, K_{xz}) f(Y, K_{yz}) - \alpha_2 Z$$

The parameters α_1 and α_2 are the compound degradation and dilution rates of proteins Y and Z;[7,21,22] B_y is the basal production rate of Y; $f(u, K) = u^H/(K^H + u^H)$ is the promoter activation function with Hill coefficient H; K_{ij} are the activation thresholds of gene j by transcription factor i. The Sx stimulus was represented by setting X = 1 (ON) or X = 0 (OFF). We note that the present model is a simple sketch of the system and does not include effects such as negative auto-regulation of Y, DNA loops,[15,43] or delays due to mRNA production. This model does not include effects such as negative autoregulation of Y. Theoretical and experimental work[21] demonstrated that negative autoregulation acts to speed transcription responses. Therefore, one would expect that if the negative autoregulation of AraC had any effect, it would be to reduce the delay.

It is instructive to analytically solve the equations in the idealized case of sharp activation function:[7,41] $f(u, K) = 1$ if $u > K$ and $f(u, K) = 0$ otherwise. In this case, the delay in the onset of Z expression following an ON step, t_{ON}, is given by:[7]

$$t_{ON} = \log((Y_m - Y_o)/(Y_m - K_{zy}))/\alpha_1$$

where $Y_o = B_y/\alpha_1$ is the basal Y level, and $Y_m = (B_y + \beta_1)/\alpha_1$ is the maximal Y level. Note that typical *E. coli* sugar utilization systems express tens of thousands of copies of the proteins per cell during full activation. This comprises of the order of 1% of the several million protein units in the *E. coli* cell.[42]

Transcription network databases

We used literature-based databases of direct transcription interactions for *E. coli*[2] and *S. cerevisiae*†.[3] We enumerated the FFLs using methods described earlier.[2,3] In *E. coli*, five FFLs have a dual-regulation transcription factor, which behaves as an activator in the presence of an inducer and as a repressor in the absence of an inducer. AraC is such a dual transcription factor.[14,15] We considered these cases as activators in enumerating the FFL structures. In the present study, since the inducer L-arabinose is present in all of the kinetic measurements, AraC can be considered an activator.

Acknowledgements

We thank C. Bargmann, D. Ginsberg, O. Hobert, S. Leibler, M. Magnasco, S. Quake, R. Schleif, E. Sontag, and all members of our laboratory for discussions. This work was supported by the Israel Science Foundation, a Minerva Junior Research Group, and the Human Frontiers Science project.

References

1. Hartwell, L. H., Hopfield, J. J., Leibler, S. & Murray, A. W. (1999). From molecular to modular cell biology. *Nature*, **402**, 47–52.
2. 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.
3. 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.
4. Ouzounis, C. A. & Karp, P. D. (2000). Global properties of the metabolic map of *Escherichia coli*. *Genome Res.* **10**, 568–576.
5. Rao, C. V. & Arkin, A. P. (2001). Control motifs for intracellular regulatory networks. *Annu. Rev. Biomed. Eng.* **3**, 391–419.
6. 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.
7. Mangan, S. & Alon, W. (2003). Structure and function of the feed-forward loop network motif. *Proc. Natl Acad. Sci. USA*, **100**, 11980–11985.
8. Yuh, C. H., Bolouri, H. & Davidson, E. H. (1998).

† The databases are available at http://www.weizmann.ac.il/mcb/UriAlon

Genomic *cis*-regulatory logic: experimental and computational analysis of a sea urchin gene. *Science*, **279**, 1896–1902.
9. Buchler, N. E., Gerland, U. & Hwa, T. (2003). On schemes of combinatorial transcription logic. *Proc. Natl Acad. Sci. USA*, **100**, 5136–5141.
10. Setty, Y., Mayo, A., Surette, M. & Alon, U. (2003). Detailed map of a *cis*-regulatory input function. *Proc. Natl Acad. Sci. USA*, **100**, 7702–7707.
11. Wilcox, G., Meuris, P., Bass, R. & Englesbe, E. (1974). Regulation of L-arabinose operon Bad *in vitro*. *J. Biol. Chem.* **249**, 2946–2952.
12. Casadaban, M. J. (1976). Regulation of regulatory gene for arabinose pathway, Arac. *J. Mol. Biol.* **104**, 557–566.
13. Johnson, C. M. & Schleif, R. F. (1995). *In vivo* induction kinetics of the arabinose promoters in *Escherichia coli*. *J. Bacteriol.* **177**, 3438–3442.
14. Schleif, R. (1996). Two positively regulated systems, ara and mal. In *Escherichia coli* and *Salmonella* (Neidhardt, F. C., ed.), vol. 1, pp. 1300–1309, ASM Press, Washington, DC.
15. Schleif, R. (2000). Regulation of the L-arabinose operon of *Escherichia coli*. *Trends Genet.* **16**, 559–565.
16. Beckwith, J. (1987). The lactose operon. In *Escherichia coli* and *Salmonella*: cellular and Molecular Biology (Neidhardt, F. C., ed.), vol. 2, pp. 1444–1452, ASM Press, Washington, DC.
17. Vilar, J. M. G., Guet, C. C. & Leibler, S. (2003). Modeling network dynamics: the lac operon, a case study. *J. Cell Biol.* **161**, 471–476.
18. 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.
19. 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.
20. Salgado, H., Santos-Zavaleta, A., Gama-Castro, S., Millan-Zarate, D., Diaz-Peredo, E., Sanchez-Solano, F. *et al.* (2001). RegulonDB (version 3.2): transcriptional regulation and operon organization in *Escherichia coli* K-12. *Nucl. Acids Res.* **29**, 72–74.
21. Rosenfeld, N., Elowitz, M. B. & Alon, U. (2002). Negative autoregulation speeds the response times of transcription networks. *J. Mol. Biol.* **323**, 785–793.
22. Rosenfeld, N. & Alon, U. (2003). Response delays and the structure of teranscription networks. *J. Mol. Biol.* **329**, 645–654.
23. Savageau, M. (1974). Genetic regulatory mechanisms and the ecological niche of *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **71**, 2453–2455.
24. Savageau, M. A. (1977). Design of molecular control mechanisms and the demand for gene expression. *Proc. Natl Acad. Sci. USA*, **74**, 5647–5651.
25. Savageau, M. (1989). Are there rules governing patterns of gene regulation? *Theoretical Biology-Epigenetics and Evolutionary Order* (Goodwin, B. C. & Saundres, P. T., eds), pp. 42–66, Edinburgh University Press.
26. Savageau, M. (2001). Design principles for elementary gene circuits: elements, methods and examples. *Chaos*, **11**, 142–159.
27. Babu, M. M. & Teichmann, S. A. (2003). Evolution of transcription factors and the gene regulatory network in *Escherichia coli*. *Nucl. Acids Res.* **31**, 1234–1244.
28. Conant, G. C. & Wagner, A. (2003). Convergent evolution of gene circuits. *Nat. Genetics*, **34**, 264–266.
29. Duggan, A., Ma, C. & Chalfie, M. (1998). Regulation of touch receptor differentiation by the *Caenorhabditis elegans* mec-3 and unc-86 genes. *Development*, **125**, 4107–4119.
30. Altun-Gultekin, Z., Andachi, Y., Tsalik, E. L., Pilgrim, D., Kohara, Y. & Hobert, O. (2001). A regulatory cascade of three homeobox genes, ceh-10, ttx-3 and ceh-23, controls cell fate specification of a defined interneuron class in *C. elegans*. *Development*, **128**, 1951–1969.
31. Davidson, E. H., Rast, J. P., Oliveri, P., Ransick, A., Calestani, C. & Yuh, C. H. (2002). A provisional regulatory gene network for specification of endomesoderm in the sea urchin embryo. *Dev. Biol.* **246**, 162–190.
32. Moroni, M. C., Hickman, E. S., Denchi, E. L., Caprara, G., Colli, E., Cecconi, F. *et al.* (2001). Apaf-1 is a transcriptional target for E2F and p53. *Nature Cell. Biol.* **3**, 552–558.
33. Mullen, A. C., Hutchins, A. S., High, F. A., Lee, H. W., Sykes, K. J., Chodosh, L. A. & Reiner, S. L. (2002). Hlx is induced by and genetically interacts with T-bet to promote heritable T(H)1 gene induction. *Nature Immunol.* **3**, 652–658.
34. White, J. G. (1985). Neuronal connectivity in *Caenorhabditis elegans*. *Trends Neurosci.* **8**, 277–283.
35. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., & Watson, J. D. *Molecular Biology of the Cell*. 4th edit, Gerland Science, New York, pp. 859–861.
36. Blattner, F. R., Plunkett, G., 3rd, Bloch, C. A., Perna, N. T., Burland, V., Riley, M. *et al.* (1997). The complete genome sequence of *Escherichia coli* K-12. *Science*, **277**, 1453–1474.
37. Cormack, B. P., Valdivia, R. H. & Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene*, **173**, 33–38.
38. Lobell, R. B. & Schleif, R. F. (1990). DNA looping and unlooping by AraC protein. *Science*, **250**, 528–532.
39. McAdams, H. H. & Arkin, A. (1998). Simulation of prokaryotic genetic circuits. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 199–224.
40. Tyson, J. J., Chen, K. C. & Novak, B. (2003). Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr. Opin. Cell. Biol.* **15**, 221–231.
41. Thomas, R. & D'Ari, R. (1990). *Biological Feedback*, CRC Press, Boca Raton, FL.
42. Bremer, H. & Dennis, P. P. (1987). Modulation of chemical composition and other parameters of the cell by growth rate. In *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology (Neidhardt, F. C., ed.), vol. 2, pp. 1553–1569, ASM Press, Washington, DC.
43. Vilar, J. M. & Leibler, S. (2003). DNA looping and physical constraints on transcription regulation. *J. Mol. Biol.* **331**, 981–989.

Edited by M. Yaniv

(Received 23 June 2003; received in revised form 15 September 2003; accepted 24 September 2003)