

Theory on the Dynamics of Feedforward Loops in the Transcription Factor Networks

Rajamanickam Murugan*

Department of Biotechnology, Indian Institute of Technology Madras, Chennai, India

Abstract

Feedforward loops (FFLs) consist of three genes which code for three different transcription factors A, B and C where B regulates C and A regulates both B and C. We develop a detailed model to describe the dynamical behavior of various types of coherent and incoherent FFLs in the transcription factor networks. We consider the deterministic and stochastic dynamics of both promoter-states and synthesis and degradation of mRNAs of various genes associated with FFL motifs. Detailed analysis shows that the response times of FFLs strongly dependent on the ratios ($w_h = \gamma_{pc}/\gamma_{ph}$ where $h = a, b, c$ corresponding to genes A, B and C) between the lifetimes of mRNAs ($1/\gamma_{mh}$) of genes A, B and C and the protein of C ($1/\gamma_{pc}$). Under strong binding conditions we can categorize all the possible types of FFLs into groups I, II and III based on the dependence of the response times of FFLs on w_h . Group I that includes C1 and I1 type FFLs seem to be less sensitive to the changes in w_h . The coherent C1 type seems to be more robust against changes in other system parameters. We argue that this could be one of the reasons for the abundant nature of C1 type coherent FFLs.

Citation: Murugan R (2012) Theory on the Dynamics of Feedforward Loops in the Transcription Factor Networks. PLoS ONE 7(7): e41027. doi:10.1371/journal.pone.0041027

Editor: Christof Markus Aegerter, University of Zurich, Switzerland

Received: January 31, 2012; **Accepted:** June 16, 2012; **Published:** July 20, 2012

Copyright: © 2012 Rajamanickam Murugan. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The author has no support or funding to report.

Competing Interests: The author has declared that no competing interests exist.

* E-mail: rmurugan@gmail.com

Introduction

Transcription factors (TFs) regulate the quantitative levels of many proteins inside a cell [1–4]. TF networks consist of several fundamental building blocks such as auto regulatory loops, flip-flops, feedback loops, single input modules, cascades, feed-forward loops (FFL) and dense overlapping regulons [5–7]. Positive auto regulatory loops play critical roles in the maintenance of cellular memory [3] and reprogramming whereas a negative auto regulatory loop seems to speed up the response time against an external stimulus [8–9]. The response time of a gene/motif is the amount of time that is required to achieve half of the steady-state concentration of resultant protein product which is often referred to as rise-time [3–4].

Feedforward loops consist of three different genes namely A, B and C which code for three different TFs. Here the protein of gene A regulates the transcription of both B and C whereas both the proteins of genes A and B regulate the transcription of C (Fig. 1). As shown in Figure 1, totally there are three such regulatory connections in a FFL network motif and eight such regulatory combinations viz. (PPP, PNP, NNN, NPP, PNN, PPP, NNP, and NPN). The first one in a combination “FGH” denotes the type of regulation of the transcription of gene B by the protein of gene A, the second one stands for the type of regulation of C by B and the third one denotes the type of regulation of C by TF protein A. Here the type of regulation can be either positive or negative where P denotes positive and N denotes negative. Here PPP is classified as a coherent type FFL whereas PNP is classified as incoherent type. A FFL motif is said to be a coherent type if the direct effect of the general transcription factor (A) on the effector operons (C) has the same sign (negative or positive) as its net

indirect effect through the specific transcription factor (B). Here PPP, NPN, PNN and NNP (termed as C1, C2, C3 and C4) are coherent types and PNP, NNN, PPN and NPP (termed as I1, I2, I3 and I4) are incoherent types [10–11]. Here A is the general transcription factor that directly regulates the effector operon of gene C and also indirectly regulates gene C through B. For example, in PNN coherent type, gene A positively regulates B which in turn negatively regulates C and therefore the net effect of regulation of C by TF gene A indirectly through B is negative. Since A directly regulates C via negative mode, PNN is called as a coherent type.

FFLs perform several important cellular tasks in various biological systems [11–15]. It has been shown that PPP type FFL can act as a sign sensitive delay [14]. For example when A and B regulate C via an AND type logic, then PPP type FFL shows a delay in the expression of gene C following induction of A by an external signal, but no delay following deactivation of A [14]. It seems that coherent types constitute ~85% of the naturally occurring FFL motifs. Magnan and Alon [10] have comprehensively studied the dynamical and kinetic behavior of various types of FFLs under the condition that the promoters of both genes A and B were triggered by external stimuli/signals. Upon analyzing literature-based databases of experimentally verified direct transcription interactions for *E. coli* [6] and *S. cerevisiae* [7], they have discovered that PPP (C1) and PNP (I1) type FFLs are more abundant in nature than others [10]. Although all the FFLs are biologically feasible [10], it is still not clear why these two types were preferably selected by nature against other types in these organisms.

Most of the earlier theoretical and experimental studies on FFLs assumed a quasi-equilibrium condition for the binding-unbinding

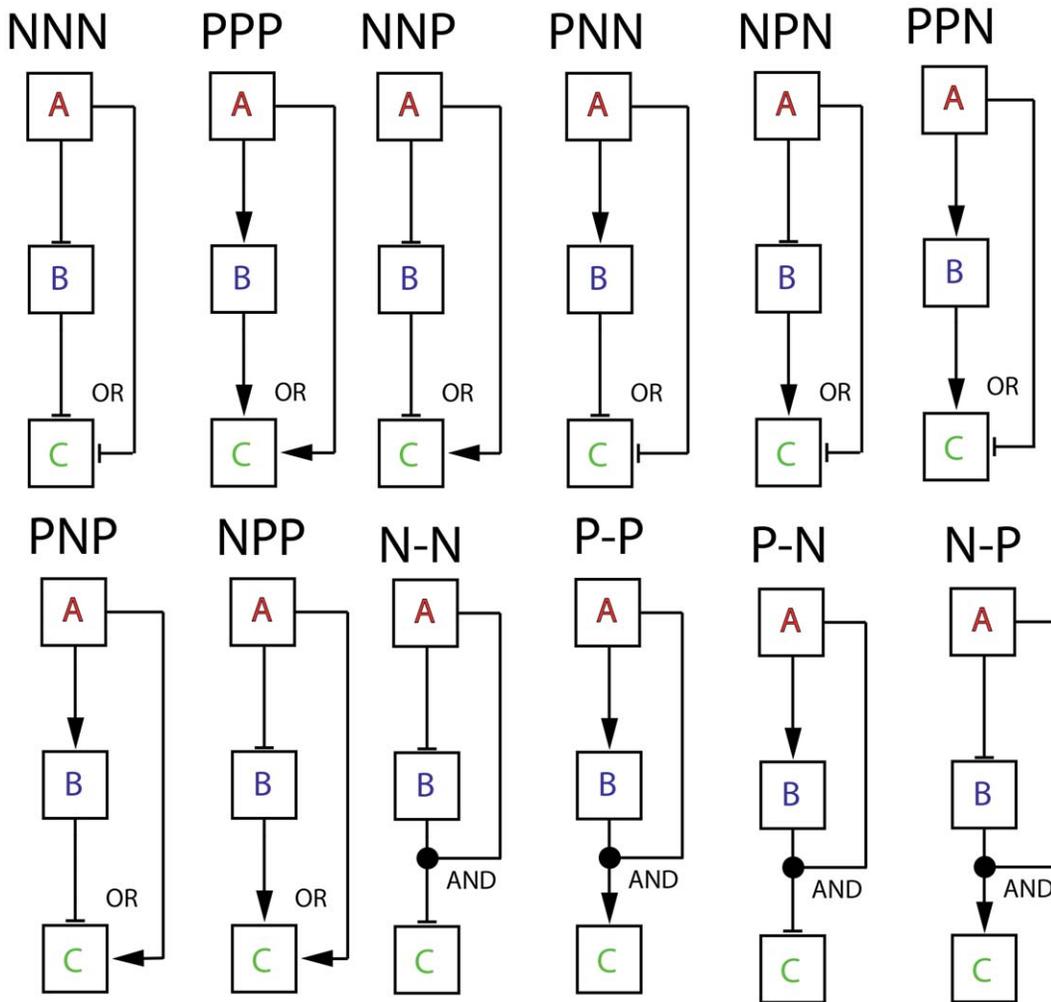


Figure 1. Various types of feedforward loops (FFLs) considered. FFLs consist of three genes which code for three different transcription factors A, B and C where B regulates C and A regulates both B and C. There are three regulatory connections in FFLs. Since each of these regulatory connections can be either positive or negative, totally there are eight different FFLs. We use three letter codes as “FGH” where ‘F’ denotes the type of regulation of B by A and ‘G’ denotes the type of regulation of C by B and ‘H’ denotes the type of regulation of C by A. Here PPP, NPN, PNN and NNP (termed as C1, C2, C3 and C4) are coherent type and PNP, NNN, PPN and NPP (termed as I1, I2, I3 and I4) are incoherent type. When dimer of A and B regulate C through AND-logic, there are four possible FFLs as P-P, P-N, N-N and N-P. doi:10.1371/journal.pone.0041027.g001

dynamics of regulatory TF proteins at various promoters [10] and a steady-state condition for the dynamics of synthesis and degradation of mRNAs. These assumptions are valid [9] only when the timescales associated with the synthesis and degradation of TF proteins are much slower (several orders of magnitude) than the timescales associated with the binding-unbinding of regulatory TFs at the respective promoters and synthesis and degradation of mRNAs. Recently the role of mRNA stability in tuning the kinetics of gene induction has been studied in detail by Elkon et.al [16]. It seems that the rapidity of induction negatively correlates with the stability of mRNAs. Further, the dynamics of mRNAs can be approximated to be in a steady state only when the ratio $w = \gamma_p / \gamma_m$ (= lifetime of mRNA/lifetime of protein) is closer to zero which is not true for most of the protein coding genes [9]. Here γ_p and γ_m are the decay rate constants associated with the respective protein and mRNA. In prokaryotic systems $w \sim 0.1$ and in eukaryotic systems such as yeast w seems to vary [17–19] approximately from 0.1 to 1 with a median of ~ 0.3 . The response times associated with various TFs in a given network are strongly dependent on w

of the respective genes and the rate of protein production will be in turn dependent on the response time [9]. TF regulatory networks whose response times are close to or lesser than the generation time of the cell and also less sensitive to the variation in the values of w corresponding to the component genes are more desirable. In this paper using a combination of theoretical and simulation tools we will (a) formulate a detailed model of various types of FFLs that includes the binding-unbinding dynamics of regulatory TFs at various promoters and synthesis and degradation of mRNAs, (b) investigate the effect of variation in w and other system parameters on the response times and overall dynamics of different type of FFLs using the detailed model and (c) explain why some of the FFLs are more abundant in nature than others.

Results

Theoretical Formulation

Feedforward loops consist of three genes coding for transcription factors (TF) A, B and C (Fig. 1). The corresponding

concentrations of mRNAs are (m_a , m_b and m_c) and the concentrations of the protein products are p_a , p_b and p_c all are measured in mol/lit (M). Corresponding steady-state concentrations are denoted as (m_{as} , m_{bs} , m_{cs} , p_{as} , p_{bs} and p_{cs}). In the absence of any regulation or when the promoters are turned-on completely, these steady state values will be ($m_{hs} = k_{mh}/\gamma_{mh}$, $p_{hs} = k_{mk}k_{pk}/\gamma_{mk}\gamma_{pk}$ where subscripts $h = a, b, c$ denote the TF genes A, B and C respectively). Here the transcription rate of TF gene ‘ h ’ is k_{mh} ($M s^{-1}$) and the respective translation rate is k_{pk} (s^{-1}). The decay rate constant associated with the mRNA of gene ‘ h ’ is γ_{mh} (s^{-1}) and the decay rate constant corresponding to the protein product is γ_{pk} (s^{-1}). Here subscripts $h, g = a, b, c$ respectively denote gene A, B and C. The gene associated with the transcription factor A is controlled/triggered by external signal which may be an arbitrary time dependent pulse function (we denote this as $\chi(t)$) or an exponentially decaying one. There is a *cis*-acting element associated with the promoter of gene B where TF protein of A can bind and hence up/down regulate the expression of B via distal action that is mediated by either tracking or looping modes [20]. There are *cis*-acting elements associated with promoter of C where the protein products of both genes A and B or the dimer of A–B can bind and hence can up/down regulate C respectively in a “OR” or “AND” logic mode. In an A-OR-B mode the presence of either protein A or B is enough to up/down regulate the promoter of gene C. The protein products of TF gene A and B can also up/down regulate C in a “AND” type logic when the dimer of protein products A–B binds with the promoter of gene C. In this case the presence of both A and B is essential for up/down regulation of gene C. There are eight numbers of regulatory combinations with A-OR-B type logic and four different combinations are possible with A-AND-B type logic (Fig. 1). The fraction occupancy of promoter of TF gene ‘ h ’ by the respective regulatory TF protein ‘ g ’ is denoted as $X_{hg} \in (0, 1)$ which is the ratio x_{hg}/d_{hz} where x_{hg} is the cellular concentration of the promoter of gene ‘ h ’ that is bound with the protein of TF gene ‘ g ’ and d_{hz} is the total concentration of the promoter of gene ‘ h ’ inside the cellular volume. There are at least two types of inter-molecular interactions [22] involved in the binding of TF proteins at the corresponding *cis*-regulatory sequences namely (a) a weak non-specific electrostatic interactions between the negatively charged backbone of DNA and the positively charged side chains of the aminoacids which are present at the DNA binding domains of TF proteins and (b) the specific hydrogen bonding interactions at the site-specifically bound DNA protein interface. The strength of electrostatic interactions will be modulated by the presence of water molecules at the DNA protein interface. In such well hydrated conditions at the DNA-protein interface, the net electrostatic interactions can be either attractive or repulsive owing to the presence of multiple electrical double layers around each charged group [22]. The free energy barrier associated with the fluctuating dynamics of TF proteins within this electrostatic field is comparable with that of the thermal free energy that in turn helps the TF proteins to freely slide along the DNA within this electrostatic capturing domain without physical dissociation [23]. Dissociation or unbinding ($X_{hg} = 0$) of TF protein from the *cis*-acting site happens when all the specific hydrogen bonds are broken and the TF protein completely escapes from this electrostatic force field or capturing domain. When the TF protein is well within the electrostatic field then depending on the net inter-molecular interactions at the DNA-protein interface we find that $X_{hg} \in (0, 1)$. This is reasonable since there may be a situation where the site-specific hydrogen bonding network present at the interface of *cis*-regulatory DNA sequence and TF is broken due to thermal induced fluctuations but the TF protein is still present

there within the electrostatic force field (partially bound condition). This partially bound promoter state is common in eukaryotic systems where a DNA loop connects the *cis*-acting module and promoter and holds the transcription initiation components so that they are nearby each other in three dimensional space within the electrostatic capturing domain. Further the binding-unbinding dynamics of TF protein ‘ g ’ at the promoter of TF gene ‘ h ’ will be observed as a continuous process in the timescale of the synthesis and decay of the corresponding mRNA and protein of TF gene ‘ h ’. This means that at the timescales of synthesis and decay of mRNAs and proteins, the thermally driven local fluctuations in the occupancy of the promoters by the TF proteins well within the electrostatic capturing domain will be averaged out. Under such conditions the averaged promoter state occupancy will be equal to the thermodynamic probability of finding the promoter to be occupied by the regulatory TF protein. Upon considering all these facts, one can conclude that it is appropriate to use a continuous type probability variable (such as X_{hg}) to denote the promoter state occupancy to account for the promoter of gene ‘ h ’ that is partially bound with the TF protein ‘ g ’ rather than a discrete Boolean type variable as described earlier [9,17–18]. The bimolecular collision rates associated with the binding of TF protein ‘ g ’ with the promoter of gene ‘ h ’ is denoted as k_{fgh} ($M^{-1}s^{-1}$) and the corresponding off-rates of these bimolecular site-specific DNA-protein complexes are denoted as k_{rgh} (s^{-1}). $K_{gh} = k_{rgh}/k_{fgh}$ (M) is the overall dissociation constant corresponding to the site-specific binding of TFs at their respective cognate sites. We have summarized the parameters of our detailed model in Tables 1 and 2. Since there are several system parameters, exploration of the entire parametric hyperspace will be a complicated one. To simplify the calculations further we introduce the following scaling scheme to project the dynamical variables onto a dimensionless space.

$$\tau = \gamma_{pc} t; M_k = m_k/m_{ks}; m_{ks} = k_{mk}/\gamma_{mk}; P_k = p_k/p_{ks}; p_{ks} = k_{mk}k_{pk}/\gamma_{mk}\gamma_{pk}; k = a, b, c$$

Here we measure the real time t in terms of numbers of lifetimes ($1/\gamma_{pc}$) of the protein product of gene C. When protein C is stable over several generations of the cell, then its rise-time that is required to achieve half of the steady state value will be equal to the generation time of the cell upon considering the dilution owing to the doubling of cell volume along the process of cell division [8–9]. In such conditions one can also transform the dimensionless τ in terms of numbers of generation times of the cell by dividing as $\tau/\ln 2$. With these definitions we can write the deterministic differential equations associated with temporal evolution of (M_h , P_h and X_{hg} where $h = b, c$ and $g = a, b, c$) of FFLs with A-OR/AND-B type regulatory logic imposed on the transcription of gene C as follows.

TF A:

$$\begin{aligned} w_a dM_a/d\tau &= \chi(\tau) - M_a \\ \rho_a dP_a/d\tau &= M_a - P_a - \sigma_{ab}(P_a(1 - X_{ba}) - \mu_{ab}X_{ba}) \\ &\quad - \sigma_{ac}(P_a(1 - X_{ca} - X_{cb}) - \mu_{ac}X_{ca}) \end{aligned} \quad (1)$$

The first one in Eqs 1 describes the dynamics of mRNA associated with the TF gene A whereas the second one describes the dynamics of protein synthesis, decay and binding-unbinding of protein A with the promoters of TF genes B and C. The

Table 1. Definition of parameters used to describe the dynamics of various FFLs.

Parameter	Gene A	Gene B	Gene C	Units and remarks
m_k	m_a	m_b	m_c	M, conc. of mRNAs
p_k	p_a	p_b	p_c	M, conc. of proteins
k_{mk}	k_{ma}	k_{mb}	k_{mc}	Ms^{-1} , transcription rate
k_{pk}	k_{pa}	k_{pb}	k_{pc}	s^{-1} , translation rate
γ_{mk}	γ_{ma}	γ_{mb}	γ_{mc}	s^{-1} , decay rate constant for mRNAs
γ_{pk}	γ_{pa}	γ_{pb}	γ_{pc}	s^{-1} , decay rate constants for proteins
m_{ks}	k_{ma}/γ_{ma}	k_{mb}/γ_{mb}	k_{mc}/γ_{mc}	M, steady state values of mRNAs in unregulated case
p_{ks}	$k_{pa} k_{ma}/\gamma_{pa}\gamma_{ma}$	$k_{pb} k_{mb}/\gamma_{pb}\gamma_{mb}$	$k_{pc} k_{mc}/\gamma_{pc}\gamma_{mc}$	M, steady state values of proteins in unregulated case
$M_k = m_k/m_{ks}$	m_a/m_{as}	m_b/m_{bs}	m_c/m_{cs}	dimensionless
$P_k = p_k/p_{ks}$	p_a/p_{as}	p_b/p_{bs}	p_c/p_{cs}	dimensionless
$W_k = \gamma_{pc}/\gamma_{mk}$	γ_{pc}/γ_{ma}	γ_{pc}/γ_{mb}	γ_{pc}/γ_{mc}	dimensionless
$\rho_k = \gamma_{pc}/\gamma_{pk}$	γ_{pc}/γ_{pa}	γ_{pc}/γ_{pb}	1	dimensionless
d_{zk}	d_{za}	d_{zb}	d_{zc}	M, total conc. of promoter
λ_{mh}	$1/m_{as}W_a$	$1/m_{bs}W_c$	$1/m_{cs}W_c$	M^{-1} , noise parameter
λ_{ph}	$1/p_{as} \rho_a$	$1/p_{bs} \rho_b$	$1/p_{cs} \rho_c$	M^{-1} , noise parameter

Note: This table describes the variables and parameters used in the numerical simulations of different FFLs. Here $k = a, b, c$ represent genes A, B and C respectively. The values d_{zk} represent the concentration of promoters of the gene k inside the cell. The values m_{ks} and p_{ks} are the steady-state numbers of mRNA and protein molecules associated with gene k in the absence of any type regulation.
doi:10.1371/journal.pone.0041027.t001

dimensionless perturbation parameters in Eqs 1 are defined as follows.

$$w_a = \gamma_{pc}/\gamma_{ma}; \sigma_{ab} = k_{fab}d_{cz}/\gamma_{pa}; \sigma_{ac} = k_{fac}d_{cz}/\gamma_{pa};$$

$$\mu_{ab} = K_{ab}/p_{as}; \mu_{ac} = K_{ac}/p_{as}; \rho_a = \gamma_{pc}/\gamma_{pa}$$

Here $\chi(\tau) \in (0,1)$ is a time dependent external signal that can turn on/off the expression of A. For a constitutive expression of TF gene A we can set $\chi(\tau) = H(\tau)$ and depending on the type of the signals and their decay properties this function can be modified. To investigate the properties of the response-times of various types

of FFLs we can set $\chi(\tau)$ as rectangular pulses/dips with predefined widths. Here $H(\tau)$ is the Heaviside step function such that $H(\tau) = 0$ when $\tau < 0$ and $H(\tau) = 1$ when $\tau > 0$. Rectangular pulses of signals at a given time point can be constructed with a combination of step functions. To introduce a rectangular pulse at the scaled time $\tau = \tau_p$ for a width of θ , we need to set $\chi(\tau) = H(\tau - \tau_p) - H(\tau - \tau_p - \theta)$. To generate a series of n numbers of rectangular pulses of signals at time points τ_i with widths θ_i where $i = 1, 2, \dots, n$ we need to set $\chi(\tau) = \sum_{i=1}^n (H(\tau - \tau_i) - H(\tau - \tau_i - \theta_i))$. Further we assume that the basal expression levels of all TF genes A, B and C are zero.

Table 2. Parameters used to describe the interactions between various components of FFLs.

Parameters	AB	BC	AC	Units and remarks
$\sigma_{hk} = k_{fhk} d_{zk}/\gamma_{ph}$	$k_{fab}d_{zb}/\gamma_{pa}$	$k_{fbc}d_{zc}/\gamma_{pb}$	$k_{fac}d_{zc}/\gamma_{pa}$	dimensionless
$K_{hk} = k_{fhk}/k_{rhk}$	k_{fab}/k_{rab}	k_{fbc}/k_{rbc}	k_{fac}/k_{rac}	M, dissociation constant connected with binding of protein 'h' with promoter 'k'
$v_{hk} = \gamma_{pc}/k_{fhk} p_{hs}$	$\gamma_{pc}/k_{fab} p_{as}$	$\gamma_{pc}/k_{fbc} p_{bs}$	$\gamma_{pc}/k_{fac} p_{as}$	dimensionless
$\mu_{hk} = K_{hk}/p_{hs}$	K_{ab}/p_{as}	K_{bc}/p_{bs}	K_{ac}/p_{as}	binding of protein 'h' with promoter of 'k'
x_{hk}	x_{ba}	x_{cb}	x_{ca}	M, conc. of promoter 'h' bound with 'k' protein
$X_{hk} = x_{hk}/d_{zh}$	x_{ba}/d_{zb}	x_{cb}/d_{zc}	x_{ca}/d_{zc}	occupancy of promoter 'h' by protein 'k'
k_{fhk}	k_{fab}	k_{fbc}	k_{fac}	$M^{-1}s^{-1}$, binding rate of 'h' with promoter 'k'
k_{rhk}	k_{rab}	k_{rbc}	k_{rac}	s^{-1} , dissociation rate of protein 'h' from promoter 'k'

Note: This table describes various parameters associated with the different types of regulatory interactions in FFLs. The value K_{hk} is the dissociation constant connected with binding of protein of gene 'h' with promoter of gene 'k'. The column AB denotes the regulation of the promoter of gene B by the protein product of A and so on.
doi:10.1371/journal.pone.0041027.t002

TF B:

$$\begin{aligned} v_{ba}dX_{ba}/d\tau &= P_a(1 - X_{ba}) - \mu_{ab}X_{ba} \\ w_b dM_b/d\tau &= \Omega_{\pm}(X_{ba}) - M_b \\ \rho_b dP_b/d\tau &= M_b - P_b - \sigma_{bc}(P_b(1 - X_{ca} - X_{cb}) - \mu_{bc}X_{cb}) \end{aligned} \quad (2)$$

The first one in Eqs 2 describes the binding-unbinding dynamics of TF protein A at the promoter of TF gene B. Second and third equations describe respectively the dynamics of synthesis and degradation of mRNA and protein products associated with TF gene B. The dimensionless perturbation parameters in Eqs 2 are defined as follows.

$$\begin{aligned} v_{ba} &= \gamma_{pc}/k_{fab}p_{as}; w_b = \gamma_{pc}/\gamma_{mb}; \sigma_{bc} = k_{fbc}d_{cz}/\gamma_{pb}; \\ \mu_{bc} &= K_{bc}/p_{bs}; \rho_b = \gamma_{pc}/\gamma_{pb} \end{aligned}$$

The function $\Omega_{\pm} \in (0, 1)$ will vary depending on the type of regulation. For a positive regulation of the promoter of TF gene B by TF protein of A, we find $\Omega_{+} = X_{ba}$ and for a negative type regulation we find $\Omega_{-} = 1 - X_{ba}$.

TF C:

$$\begin{aligned} v_{ca}dX_{ca}/d\tau &= P_a(1 - X_{ca} - X_{cb}) - \mu_{ac}X_{ca}; \\ v_{cb}dX_{cb}/d\tau &= P_b(1 - X_{ca} - X_{cb}) - \mu_{bc}X_{cb} \\ w_c dM_c/d\tau &= \zeta_{a\pm b\pm}(X_{ca}, X_{cb}) - M_c \\ dP_c/d\tau &= M_c - P_c \end{aligned} \quad (3)$$

The first set of equations in Eqs 3 describe respectively the binding-unbinding dynamics of TF proteins A and B at the corresponding *cis*-regulatory elements associated with the promoter of gene C. Second and third equations describe the dynamics of synthesis and degradation of mRNA and protein products of C. The dimensionless perturbation parameters in Eqs 3 are defined as follows.

$$w_c = \gamma_{pc}/\gamma_{mc}; v_{ca} = \gamma_{pc}/k_{fac}p_{as}; v_{cb} = \gamma_{pc}/k_{fbc}p_{bs}$$

The function $\zeta_{a\pm b\pm}(X_{ca}, X_{cb}) = \zeta_{a\pm b\pm}$ varies depending on the type of regulation. Here the total fraction of promoter of C occupied by the proteins of either A or B is $X_c = X_{ca} + X_{cb}$. There are four different possibilities.

$$\begin{aligned} \zeta_{a+b+}(X_{ca}, X_{cb}) &= \zeta_{a+b+} = X_c; \zeta_{a-b-} = 1 - X_c; \\ \zeta_{a+b-} &= (1 + \bar{X}_c)/2; \zeta_{a-b+} = (1 - \bar{X}_c)/2 \end{aligned} \quad (4)$$

Here the subscript ‘‘a+b+’’ indicates the case where both the TF proteins A and B positively regulate C and other combinations are defined in the similar way. In Eqs 4 we have defined $\bar{X}_c = X_{ca} - X_{cb}$ and depending on the type of regulation imposed on the promoter of gene C we find the following limiting conditions of the promoter-state occupancy of TF gene C.

$$\begin{aligned} \lim_{P_a \text{ or } P_b \rightarrow \infty} \zeta_{a+b+} &= 1; \lim_{P_a \text{ or } P_b \rightarrow \infty} \zeta_{a-b-} = 0; \\ \lim_{P_a \text{ and } P_b \rightarrow \infty} (\zeta_{a+b-}, \zeta_{a-b+}) &= 1/2 \end{aligned} \quad (5)$$

Upon combining Eqs 1–4 we find that there are eight possible numbers of regulatory combinations as (PPP, NNN, PNP, NPN, NNP, PPN, NPP and PNN) in case of A-OR-B type FFL. Here PPP, NPN, PNN and NNP are coherent type FFLs (corresponding standard notations are namely C1, C2, C3, and C4) whereas PNP, NNN, PPN and NPP (corresponding standard notations are namely I1, I2, I3 and I4) are incoherent types FFLs. There will be an additional step corresponding to the dimerization of TF proteins A and B in case of FFLs with A-AND-B logic type regulation. The dynamics of protein-protein dimerization and dissociation can be described by the following differential equation.

$$\left. \begin{aligned} \varepsilon_{ab}dY_{ab}/d\tau &= P_a P_b - L_{ab} Y_{ab} - \sigma_y (Y_{ab}(1 - X_{cy}) - \mu_{yc} X_{cy}); \\ Y_{ab} &= y_{ab}/p_{as} \\ \varepsilon_{ab} &= \gamma_{pc}/\lambda_{fab}p_{bs}; L_{ab} = \lambda_{fab}/\lambda_{rab}p_{bs}; \sigma_y = k_{fyc}d_{zc}/\lambda_{fab}p_{bs}; \\ \mu_{yc} &= K_{yc}/p_{as} \end{aligned} \right\} \quad (6)$$

Here y_{ab} is the cellular concentration (M) of the dimer of the protein products of TF genes A and B, λ_{fab} and λ_{rab} are the corresponding forward bimolecular (M–1s–1) and reverse unimolecular (s–1) rate constants associated respectively with the dimerization and dissociation reactions. To be consistent with Eqs 6 the equations associated with the dynamics of synthesis and degradation of Pa and Pb will be modified as follows.

$$\begin{aligned} \rho_a dP_a/d\tau &= M_a - P_a - \sigma_{ab}(P_a(1 - X_{ba}) - \mu_{ab}X_{ba}) \\ &\quad - \sigma_{\lambda a}(P_a P_b - L_{ab} Y_{ab}) \\ \rho_b dP_b/d\tau &= M_b - P_b - \sigma_{\lambda b}(P_a P_b - L_{ab} Y_{ab}); \\ \sigma_{\lambda a} &= \lambda_{ab}p_{bs}/\gamma_{pa}; \sigma_{\lambda b} = \lambda_{ab}p_{as}/\gamma_{pb} \end{aligned} \quad (7)$$

Further when the dimer A-B binds with the promoter of gene C and hence up/down regulate then the related rate equations corresponding to the expression of gene C will be modified as follows.

$$\begin{aligned} v_{cy}dX_{cy}/d\tau &= Y_{ab}(1 - X_{cy}) - \mu_{cy}X_{cy}; v_{cy} = \gamma_{pc}/k_{fyc}p_{as} \\ w_c dM_c/d\tau &= \Psi_{y\pm}(X_{cy}) - M_c \end{aligned} \quad (8)$$

The function $\Psi_{y\pm} \in (0, 1)$ varies depending on the type of regulation of the promoter C by the A–B dimer. For a positive regulation we find $\Psi_{y+} = X_{cy}$ and for negative regulation $\Psi_{y-} = 1 - X_{cy}$. Upon combining Eqs 2 and 8 we find four different types of FFLs with A–AND–B type logic on the promoter of gene C viz. (P–P, P–N, N–P and N–N). Here one should note that P–P is similar to PPP type FFL however with A–AND–B gated logic at the promoter of TF gene C, and P–N corresponds to PNN type, N–P corresponds to NPP, and N–N corresponds to NNN. In a combination ‘‘K–H’’, ‘K’ is the type of regulation of promoter B by protein A and ‘H’ is the type of regulation of promoter C by A–B dimer.

Steady-state Analysis

When the dynamics of the variables X_{ik} and M_k are much faster than the rate of synthesis and degradation of the corresponding proteins P_k then we have the following limiting conditions for A-OR-B regulation.

$$\begin{aligned} \rho_a dP_a/d\tau &= \chi(\tau) - P_a; \rho_b dP_b/d\tau = \Omega_{\pm}(X_{bas}) - P_b; \\ dP_c/d\tau &= \zeta_{a\pm b\pm}(X_{cas}, X_{cbs}) - P_c \end{aligned} \quad (9)$$

Eqs 9 can be obtained by setting $w_k=0$ for all $k=(a, b, c)$ and $v_b=v_{ca}=v_{cb}=0$ in Eqs 1–3. The steady state values of the protein products P_{hs} where $h=a, b, c$ will vary depending on the type of regulation and protein-protein interactions between A and B. Here we have defined the steady state values of X_{hk} in the coupled Eqs 9 as follows.

$$\begin{aligned} X_{bas} &= P_a/(\mu_{ab} + P_a); X_{cas} = \mu_{bc}P_a/(\mu_{ac}P_b + \mu_{bc}P_a + \mu_{bc}\mu_{ac}); \\ X_{cbs} &= \mu_{ac}P_b/(\mu_{ca}P_b + \mu_{bc}P_a + \mu_{bc}\mu_{ac}) \end{aligned}$$

Similarly one can derive the limiting condition in the presence of A-AND-B type regulation of promoter C by the dimer of the protein products of genes A and B. In such conditions the differential equation associated with P_c in Eqs 9 will be modified as follows.

$$dP_c/d\tau = \Psi_{y\pm}(X_{cys}) - P_c; X_{cys} = P_aP_b/(P_aP_b + L_{ab}\mu_{yc}) \quad (10)$$

Here one should note that the dynamical variables (P_a, P_b and P_c) also represent the efficiency of TF genes A, B and C in raising their protein levels toward their unregulated steady state values ($phs = kmhkp/h\gamma mh\gamma ph$ where $h=a, b, c$) even in the presence of various types of positive/negative type regulation on their promoters. The maximum achievable steady state values of P_{hs} will be $P_{hs} = 1$. In case of PPP type A-OR-B FFL, P_a and P_b influence the rate of change in P_c in an additive way. Whereas in case of P-P type A-AND-B FFL P_a and P_b influence the rate of change of P_c in a multiplicative way. As a result of this multiplicative effect, P-P type FFL can effectively filter out the short/transient pulses of signals which are introduced at the promoter of TF gene A. When TF proteins A and B decay much faster than TF protein C, then we find that as ρ_a and ρ_b tend toward zero. Under such conditions, the expression for X_{cys} in Eqn 10 can be written as follows.

$$X_{cys} = X_{cys}(\tau) = \chi(\tau)\chi(\tau)/(\chi(\tau)\chi(\tau) + (\mu_{ab} + \chi(\tau))L_{ab}\mu_{yc}) \quad (11)$$

Eqn 11 suggests that when the binding parameters ($\mu_{ab}, L_{ab}, \mu_{yc}$) are much lesser than one and the promoter of gene A is turned on for sufficiently longer time periods then the TF gene C will be turned on to its maximum expression level since $X_{cys} \sim 1$ under such conditions. Similar to Eq 11 when ρ_a and ρ_b tend toward zero, then we can write the expression for the occupancy level of the promoter C by TF proteins of genes A or B in case of PPP type A-OR-B FFL as follows.

$$\begin{aligned} X_{cs} &= (\mu_{bc}\chi(\tau)(\mu_{ab} + \chi(\tau)) + \mu_{ac}\chi(\tau))/ \\ &(\mu_{ac}\chi(\tau) + (\mu_{ab} + \chi(\tau))(\mu_{bc}\chi(\tau) + \mu_{bc}\mu_{ac})) \end{aligned} \quad (12)$$

This equation suggests that when the binding parameters (μ_{bc}, μ_{ac}) are much lesser than one and the promoter of gene A is turned on for sufficiently longer time periods, then the TF gene C will be turned on to its maximum expression level since $X_{cs} \sim 1$ under such conditions. Using Eqs 11 and 12 one can derive the steady-state values (P_{as}, P_{bs}, P_{cs}) of scaled protein concentrations P_a, P_b , and P_c associated with Eqs 9 as follows.

$$\begin{aligned} P_{as} &= \chi(\tau); P_{bs} = \Omega_{\pm}(X_{bas}); \\ P_{cs} &= \zeta_{a\pm b\pm}(X_{cas}, X_{cbs}) \text{ or } \Psi_{y\pm}(X_{cys}) \end{aligned} \quad (13)$$

To evaluate P_{bs} one needs to substitute $P_{as} = \chi(\tau)$ in the appropriate expression of $\Omega_{\pm}(X_{bas})$ that in turn depends on the type of regulation of promoter of the TF gene B by the protein of gene A and subsequently these P_{as} and P_{bs} need to be substituted in the appropriate expression of P_{cs} that depends on the types of regulation of TF gene C by proteins A and B. While deriving Eqs 1–3 related to A-OR-B type FFLs we have assumed that monomeric units of proteins A or B interact with the promoters of B and C. Similar to the regulation of C by the dimer of TF proteins A–B in case of A-AND-B type FFLs, one can generalize Eqs 1–3 corresponding to A-OR-B type FFLs to include the regulation of the promoter of the TF genes B and C by the multimeric form of TF proteins A and B respectively [10]. So for we have set $w_k=0$ for all $k=a, b$ and c . One can show that the response-time associated with the synthesis of the terminal TF protein C upon induction of the promoter of TF gene A by an external signal is strongly dependent on w_c as follows. Consider a quasi-equilibrium situations for the promoter state occupancies and steady state situation for the synthesis and degradation of TF genes A and B so that $(w_a, w_b) = 0$ and $(v_{ca}, v_{cb}, v_{ba}) = 0$. When $(\rho_a, \rho_b) = 0$, then from Eqs 9 we can obtain the integral solution for the temporal evolution of the variable P_c for a given arbitrary $w_c \neq 0$ and the initial conditions $P_c = 0$ at $\tau = 0$ as follows.

$$P_c = (e^{-\tau/w_c}) \int_0^{\tau} \left(e^{s-w_c} \int_0^s \zeta_{a\pm b\pm}(X_{cas}, X_{cbs}) e^{s'/w_c} ds' \right) ds \quad (14)$$

When the TF gene C is turned on toward its maximum expression level then the input function $\zeta_{a\pm b\pm}(X_{cas}, X_{cbs}) \sim 1$ and we can write the integral solution for temporal evolution of P_c under such conditions as follows.

$$P_c = P_c(\tau) = 1 - (e^{-\tau} - w_c e^{-\tau/w_c}) / (1 - w_c) \quad (15)$$

From Eq 15 we find that $P_c = 1 - e^{-\tau}$ for w_c is zero. Since the inequality $e^{-\tau} < e^{-\tau/w_c}$ will be true for all the values of $w_c > 0$ and P_c will tend toward zero as w_c tends toward infinity, we can conclude that the response time associated with the expression of TF gene C upon induction of TF gene A will monotonically increase as w_c increases. The maximum amount of TF protein C that is synthesized for a given rectangular pulse at the promoter of TF gene A with a width θ will be $P_c(\theta)$ and one can measure the filtering efficiency of the FFL under consideration for an arbitrary pulse width $\theta > 0$ from the ratio $P_c(\theta)/P_{cs}$. The cutoff pulse width θ_c can be defined by the inequality $P_c(\theta_c)/P_{cs} < \delta$ where δ is proportional to the experimentally detectable limit of the TF protein C under consideration. When $w_c = 0$, then we find that $\theta_c \leq \ln(1/(1-\delta))$ and when $w_c = 1$ we find that $\theta_c \leq (-\text{LambertW}((-1+\delta)/e) - 1)$ where $\text{LambertW}(x)$ function is the solution of equation $yey = x$ for y . Here we set $\delta \sim 10^{-2}$ for simulation purposes. Upon substituting this value we find $\theta_c \leq 10^{-2}$ for $w_c = 0$ and $\theta_c \leq 0.15$ for $w_c = 1$. For other values of w_c , one needs to numerically solve the inequality $P_c(\theta_c)/P_{cs} < \delta$ for θ_c . For example, when $w_c = 0.1$ then we find that $\theta_c \leq 0.05$. These results suggest that the critical cut-off pulse width θ_c increases along with the parameter w_c .

Stochastic Analysis

The set of Chemical Langevin equations (CLE) associated with the expression of TF genes A, B and C within the various types of A-OR-B type FFLs can be written as follows [26–28].

TF A:

$$\begin{aligned} w_a dM_a/d\tau &= \chi(\tau) - M_a + \sqrt{\chi(\tau)}\Gamma_\tau^{ma,o} - \sqrt{M_a}\Gamma_\tau^{ma,t} \\ \rho_a dP_a/d\tau &= M_a - P_a - \sigma_{ab}(P_a(1 - X_{ba}) - \mu_{ab}X_{ba}) \\ &\quad - \sigma_{ac}(P_a(1 - X_{ca} - X_{cb}) - \mu_{ac}X_{ca}) + Z_\tau^{abc} \end{aligned} \quad (16)$$

TF B:

$$\begin{aligned} v_{ba} dX_{ba}/d\tau &= P_a(1 - X_{ba}) - \mu_{ab}X_{ba} + Z_\tau^{xba} \\ w_b dM_b/d\tau &= \Omega_\pm(X_{ba}) - M_b + \sqrt{\Omega_\pm(X_{ba})}\Gamma_\tau^{mb,o} \\ &\quad - \sqrt{M_b}\Gamma_\tau^{mb,t} \\ \rho_b dP_b/d\tau &= M_b - P_b + \sqrt{M_b}\Gamma_\tau^{pb,o} - \sqrt{P_b}\Gamma_\tau^{pb,t} \\ &\quad - \sigma_{bc}Z_\tau^{xcb} \end{aligned} \quad (17)$$

TF C:

$$\begin{aligned} v_{ca} dX_{ca}/d\tau &= P_a(1 - X_{ca} - X_{cb}) - \mu_{ac}X_{ca} + Z_\tau^{xca} \\ v_{cb} dX_{cb}/d\tau &= P_b(1 - X_{ca} - X_{cb}) - \mu_{bc}X_{cb} + Z_\tau^{xcb} \\ w_c dM_c/d\tau &= \zeta_{a\pm b\pm}(X_{ca}, X_{cb}) - M_c \\ &\quad + \sqrt{\zeta_{a\pm b\pm}(X_{ca}, X_{cb})}\Gamma_\tau^{mc,o} - \sqrt{M_c}\Gamma_\tau^{mc,t} \\ \rho_c dP_c/d\tau &= M_c - P_c + \sqrt{M_c}\Gamma_\tau^{pc,o} - \sqrt{P_c}\Gamma_\tau^{pc,t} \end{aligned} \quad (18)$$

Eqs 16–18 suggest the presence of non-zero temporal correlations among the set of concentration variables $\langle X_{ba}P_a \rangle$, $\langle X_{ca}P_a \rangle$ and $\langle X_{cb}P_b \rangle$. The characteristic correlation times associated with these pairs of variables will be much lesser than that of the timescales associated with the synthesis and decay of the respective protein products. This follows from the fact that the timescale associated with the promoter state fluctuations is well separated from the timescale associated with the synthesis and decay of the TF proteins. Here one should note that $\rho_c = 1$ by definition. In this equation, various types of Z parameters are defined as follows.

$$\begin{aligned} Z_\tau^{xba} &= \sqrt{P_a(1 - X_{ba})}\Gamma_\tau^{xba,o} - \sqrt{\mu_{ab}X_{ba}}\Gamma_\tau^{xba,t} \\ Z_\tau^{xca} &= \sqrt{P_a(1 - X_{ca} - X_{cb})}\Gamma_\tau^{xca,o} - \sqrt{\mu_{ac}X_{ca}}\Gamma_\tau^{xca,t} \\ Z_\tau^{abc} &= \sqrt{M_a}\Gamma_\tau^{pa,o} - \sqrt{P_a}\Gamma_\tau^{pa,t} - \sigma_{ab}Z_\tau^{xba} - \sigma_{ac}Z_\tau^{xca} \\ Z_\tau^{xcb} &= \sqrt{P_b(1 - X_{ca} - X_{cb})}\Gamma_\tau^{xcb,o} - \sqrt{\mu_{bc}X_{cb}}\Gamma_\tau^{xcb,t} \end{aligned} \quad (19)$$

In Eqs 16–19, the term Γ is the dimensionless delta-correlated Gaussian white noise with the following mean and variance properties.

$$\begin{aligned} \Gamma_\tau^{kh,i} &= \sqrt{\lambda_{kh}}\zeta_t^{kh,i}; \lambda_{mh} = (m_{hs}w_h)^{-1}; \lambda_{ph} = (p_{hs}\rho_h)^{-1}; \\ \langle \Gamma_\tau^{kh,i}\Gamma_\tau^{kh,i} \rangle &= \lambda_{kh}\delta(\tau - \tau') \\ \langle \Gamma_\tau^{kh,i} \rangle &= 0; k, u = m, p; h, v = a, b, c; i = o, t; \\ \langle \Gamma_\tau^{uv,i}\Gamma_\tau^{kh,j} \rangle &= 0; u \neq k; v \neq h; i \neq j \\ \langle \Gamma_\tau^{xkh,i} \rangle &= \sqrt{\lambda_{xkh}}\langle \zeta_t^{xkh,i} \rangle = 0; k = b, c; h = a, b; \\ \langle \Gamma_\tau^{xuv,i}\Gamma_\tau^{xkh,j} \rangle &= 0; u \neq k; v \neq h; i \neq j \\ \langle \Gamma_\tau^{xkh,i}\Gamma_\tau^{xkh,i} \rangle &= \lambda_{xkh}\delta(\tau - \tau'); \lambda_{xkh} = (k_{fkh}d_{zck}p_{hs})^{-1} \end{aligned} \quad (20)$$

In case of A-AND-B type FFLs there is an additional source of fluctuations arises from the dimerization reaction between the TF proteins A and B. The modified equations in Eqs 16–18 in the presence of A-AND-B type logic can be written as follows.

$$\begin{aligned} \varepsilon_{ab} dY_{ab}/d\tau &= P_aP_b - L_{ab}Y_{ab} - \sigma_y(Y_{ab}(1 - X_{cy}) - \mu_{yc}X_{cy}) \\ &\quad + \sigma_yZ_\tau^{yc} + Z_\tau^{yab} \\ \rho_a dP_a/d\tau &= M_a - P_a - \sigma_{\lambda a}(P_aP_b - L_{ab}Y_{ab}) + \sqrt{M_a}\Gamma_\tau^{pa,o} \\ &\quad - \sqrt{P_a}\Gamma_\tau^{pa,t} - \sigma_{\lambda a}Z_\tau^{yab} \\ \rho_b dP_b/d\tau &= M_b - P_b - \sigma_{\lambda b}(P_aP_b - L_{ab}Y_{ab}) + \sqrt{M_b}\Gamma_\tau^{pb,o} \\ &\quad - \sqrt{P_b}\Gamma_\tau^{pb,t} - \sigma_{\lambda b}Z_\tau^{yab} \\ v_{cy} dX_{cy}/d\tau &= Y_{ab}(1 - X_{cy}) - \mu_{yc}X_{cy} + Z_\tau^{yc} \\ w_c dM_c/d\tau &= \Psi_{y\pm}(X_{cy}) - M_c + \sqrt{\Psi_{y\pm}(X_{cy})}\Gamma_\tau^{pc,o} \\ &\quad - \sqrt{M_c}\Gamma_\tau^{pc,t} \end{aligned} \quad (21)$$

In this equation, various types of Z parameters and the Gaussian noise terms Γ associated with the dimerization dynamics are defined as follows.

$$\begin{aligned} Z_\tau^{yc} &= \sqrt{Y_{ab}(1 - X_{cy})}\Gamma_\tau^{xcy,o} - \sqrt{\mu_{yc}X_{cy}}\Gamma_\tau^{xcy,t} \\ Z_\tau^{yab} &= \sqrt{P_aP_b}\Gamma_\tau^{ab,o} - \sqrt{L_{ab}Y_{ab}}\Gamma_\tau^{ab,t} \\ \Gamma_\tau^{ab,i} &= \sqrt{\lambda_{ab}}\zeta_\tau^{ab,i}; \lambda_{ab} = (\lambda_{fab}p_{bs})^{-1}; \langle \Gamma_\tau^{ab,i} \rangle = 0; \\ \langle \Gamma_\tau^{ab,i}\Gamma_\tau^{ab,i} \rangle &= \lambda_{ab}\delta(\tau - \tau'); i = o, t \\ \Gamma_\tau^{xcy,i} &= \sqrt{\lambda_{xye}}\zeta_\tau^{xcy,i}; \lambda_{xye} = (k_{fyc}d_{zcp_{as}})^{-1}; \\ \langle \Gamma_\tau^{xcy,i} \rangle &= 0; \langle \Gamma_\tau^{xcy,i}\Gamma_\tau^{xcy,i} \rangle = \lambda_{xye}\delta(\tau - \tau') \end{aligned} \quad (22)$$

Discussion

Response times of various type of FFLs are strongly dependent on the sets of parameters μ , \bar{v} , \bar{w} and weakly on $\bar{\sigma}$. The sets of

parameters $\bar{\nu}$ and $\bar{\sigma}$ characterize the temporal coupling between the mRNA/protein dynamics and the binding-unbinding dynamics of protein molecules with the promoter. The ordinary perturbation parameter $\bar{\sigma}$ characterizes the strength of temporal coupling between the protein level dynamics with the promoter state occupancies. The variable \bar{w} represents the ratios of the lifetimes of various mRNAs to lifetime of protein of TF gene C and reflects the coupling between the mRNA and protein degradation dynamics. The variable $\bar{\mu}$, the set of dissociation constants is inversely proportional to the TFs/promoters binding affinity and characterizes the strength of various regulatory connections. The parameter set $\bar{\sigma}$ does not affect the rise-time of the TF gene C significantly since the change in the number of TF protein molecules due to binding-unbinding at various promoters is negligible. In contrast, $\bar{\nu}$ significantly affects the response times since the effect of varying $\bar{\nu}$ is indirectly amplified through the corresponding mRNA dynamics. Furthermore, an increase in $\bar{\nu}$ would decrease the rate at which promoter state occupancies shift towards saturation (in positive type regulation) or free form (in negative type regulation) as the regulatory TFs level builds up.

The parameter set \bar{w} describes how best the rate of degradation of mRNAs of the TF genes A, B and C of FFLs are coupled to the rate of degradation of TF protein C. Lower values of $\bar{w} = \gamma_{pc} / \gamma_{mk}$ ($k = a, b, c$) occur only when the decay rate constants of various mRNAs (γ_{mk}) are much higher than the decay rate (γ_{pc}) of TF protein C. One should note that both the transcription and translation of various TF genes of prokaryotes are taking place in the cytoplasm whereas in case of eukaryotes the transcription is taking place inside the nucleus and the synthesized mRNA transcripts need to be spliced and then transported to cytoplasm through nuclear pores for translation. These differences in the cellular architecture warrants higher lifetimes ($1/\gamma_{mk}$) for eukaryotic mRNAs than the prokaryotic ones which results in the general observation that the values of \bar{w} (= lifetime of mRNA/lifetime of protein) associated with various genes in prokaryotes are lower than eukaryotes genes. It seems that a spectrum of various values of \bar{w} occurs in the protein coding genes of both prokaryotes and eukaryotes. In yeast, the values of \bar{w} seems to vary from 0.1 to 1 with a median of ~ 0.3 [17–19]. This means that we cannot ignore the dynamics of mRNAs while describing any type of TF network.

The parameter set $\bar{\mu}$ describes the strengths of various binding events associated with the FFLs under consideration. Higher values of $\bar{\mu}$ represent strong binding condition and lower values represent weak binding condition. Most of the interactions of TF proteins with the *cis*-acting DNA elements of associated promoters seems to be much stronger and $\bar{\mu}$ will be generally in the order of $\sim 10^{-3}$. When the steady state values of protein numbers in the absence of regulation under *in vivo* conditions is $p_{ks} \sim 10^3$ then the value of $\bar{\mu} \sim 10^{-3}$ indicates that a single TF protein is enough to occupy the associated promoter for 50% of the observation times.

The parameter set $\bar{\nu}$ represents the strength of coupling between the promoter state occupancies and the rates of synthesis and degradation of TF proteins. The effect of $\bar{\nu}$ on the rate of synthesis and degradation of TFs will be generally mediated through the rate of changes in the respective mRNA levels. The speed at which a regulated promoter reaches its quasi-equilibrium state is inversely proportional to the value of $\bar{\nu}$. Higher values of $\bar{\nu}$ will slow down the rate at which promoter state occupancy reached its quasi-equilibrium state that in turn can increase the response times of positively regulated promoters and decrease the response times of negatively regulated promoter. Decrease in $\bar{\nu}$ can also lead to overshooting of protein production in negatively auto regulated loops [9]. Small changes in $\bar{\nu}$ can significantly affect the dynamics of the associated TF protein since these changes are subsequently

amplified by the dynamics of mRNAs. In most of the prokaryotic cases $\bar{\nu}$ will be in order of $\sim 10^{-4}$ and it strongly dependent on the volume of the cell or nucleus in case of eukaryotic systems. Response times are not much influenced [9] by the parameter set $\bar{\sigma}$ even in a wide range (10^{-3} – 10^3). Here $\bar{\sigma}$ represents the direct coupling between the promoter state occupancies and the rate of synthesis and degradation of TF proteins. Variation in $\bar{\sigma}$ does not influence the response times much since the number of protein molecules associated with the binding-unbinding events are much less compared to the steady state values.

The dependency of the response times on the parameters \bar{w} seems to be strongly influenced by the set of binding parameters $\bar{\mu}$. Under weak binding conditions such as $\bar{\mu} \geq 1$, we find that each of FFLs under consideration show different type of variations as \bar{w} changes. Figure 2A suggests that the response times of various FFLs under weak binding conditions seems to be in the descending order of P-P, N-P, PPP, NPP, NNP, PNP, PPN, NPN, N-N, P-N, NNN and PNN. We further observe that almost all the FFLs show similar type of variation in response time with respect to changes in \bar{w} in the dynamic range $\bar{w} \in (0.1, 1)$. From Figures 2B, 2C and 2D we find that this scenario significantly changes as the binding strength increases. Under strong binding conditions such as $\bar{\mu} \ll 1$, the entire set of FFLs can be approximately categorized into three subgroups (Figure 2D) based on the behavior of overall response times with respect to changes in \bar{w} viz. Group I: $\{\{PPN, NNP\}, \{PPP, NPP, N-P\}, \{P-P, PNP\}\}$, Group II: $\{NPN, \{N-N, NNN\}, PNN\}$ and Group III: P-N. Here the order of response times of various subgroups is Group I > Group II > Group III. One can write the segregation pattern in the standard terminology of FFLs as $\{\{I3, C4\}, \{C1, I4, N-P\}\}, \{P-P, I1\}\}, \{C2, \{N-N, I2\}, C3\}, P-N$. Though Group II and Group III type FFLs possess lesser response times than Group I their response times increase almost linearly upon an increase \bar{w} over the entire range of investigation on a log-log scale.

Within Group I we find three different subsets of FFLs having similar type of response times and their response times are in the order as $\{I3, C4\} > \{C1, I4, N-P\} > \{P-P, I1\}$. As we have pointed out in the introduction section, a FFL will be an efficient one when (1) the associated response times with respect to an input signal at the promoter of TF gene A is reasonably low or close to the generation time of the cell and (2) also it is less sensitive to the changes in \bar{w} over the physiological dynamic range (0.1, 1). Those FFLs which satisfy these two criteria will be the efficient ones and naturally selected. Upon applying these two criteria on the subsets of Group I type FFLs, we find that C1 with both A-AND-B and A-OR-B gated logics and I1 are the preferred FFLs on overall range of \bar{w} since response times of I3 and C4 type FFLs are higher than the generation time of the cell. We find from Figure 2D that the increase in the response times is $< 200\%$ for Group I type FFLs upon increasing \bar{w} from 0.1 to 1 whereas it is $> 400\%$ for Group II and III type FFLs. These results agree well with the earlier observations on the abundance of various FFLs in *E. coli* [10,13] and yeast [10,14]. Though Group II and III FFLs possess less response times, these FFLs are not selected by nature since the same speeding-up functionality can be achieved through a much simpler negative auto regulatory loops associated with the TF gene C [8–9].

Figures 2E and 2F suggest that the members of Groups I–III are not robust against variation of the parameter set $\bar{\nu}$ under strong binding conditions. Results show that the P-P type FFL will be moved from third subset of Group I to the second one upon an order of increase in $\bar{\nu}$. Further we find that the FFLs such as NPP, P-P, PPP and N-P behave in a similar way with respect to the changes in \bar{w} at higher values of $\bar{\nu}$. The segregation pattern of

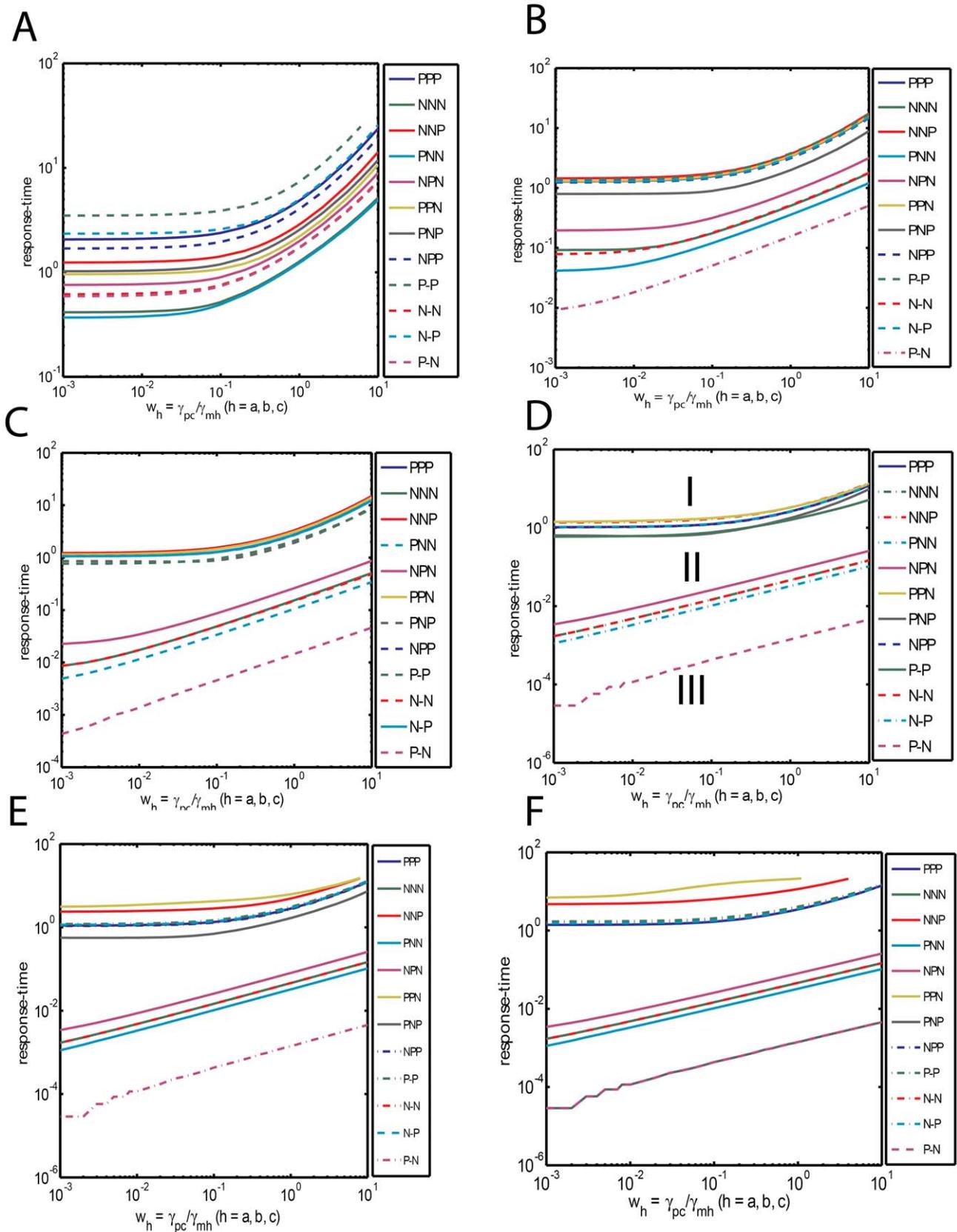


Figure 2. Dependency of response times on various sets of system parameters viz. $\vec{w} = (w_a, w_b, w_c)$, $\vec{\mu} = (\mu_{ab}, \mu_{ac}, \mu_{bc}, \mu_{ca}, \mu_{cb}, \mu_{ba})$ and $\vec{v} = (v_{ba}, v_{ca}, v_{cb}, v_{ab}, v_{bc}, v_{ca})$. Response times are expressed in terms of number of generation times. A. Dependency of response times of various types of FFLs under weak binding conditions. Here the simulation setting are $\bar{\mu} = 1$, $\bar{v} = 0.0003$, $\bar{\sigma} = 4$, $\Delta\tau = 5 \times 10^{-6}$, the total simulation time was set to $T = 25$

generation times and \bar{w} was iterated in the interval (0.001, 10) with $\Delta\bar{w}=0.001$. Under weak binding condition we observe that each of the considered FFLs behaves in a different way from others. B. Dependency of response times of FFLs \bar{w} . Settings are same as A with $\bar{\mu}=0.1$. C. Dependency of response times of FFLs \bar{w} . Settings are same as A with $\bar{\mu}=0.01$. D. Dependency of response times of FFLs \bar{w} . Settings are same as A however with $\bar{\mu}=0.001$. Under this strong binding conditions all the considered FFLs segregate into three Groups namely I, II and III. It seems that PPP (C1) and PNP (I1) type FFLs show less variation with respect to changes in \bar{w} and also their response times are comparable with that of the unregulated C. E. Influence of increase in \bar{v} on the dependency of response times of FFLs on \bar{w} . Here $\bar{v}=0.003$. This is the physiological value of \bar{v} for a typical yeast cell nucleus whose volume is ~ 10 times higher than a bacterial cell. In this case, P-P type FFL shifts from the third subgroup of Group I to the second subgroup. F. Influence of changes in \bar{v} on the dependency of response times of FFLs on \bar{w} . Here $\bar{v}=0.03$. This is the physiological value for a typical human cell nucleus whose volume is ~ 100 times larger than a bacterial cell. In this case, PNP (I1) type incoherent FFL shifts to Group III. doi:10.1371/journal.pone.0041027.g002

various FFLs based on the behavior of the response times over changes in \bar{w} under such conditions seems to be {C4, I3, {C1, P-P, N-P, I4}, I1}, {C2, {N-N, I2}, C3}, P-N. Since the value \bar{v} is inversely proportional to the speed at which the promoter state reaches the steady state, an increase in \bar{v} would decrease the transcription and translational rates. This will in turn increase the response times of positively regulated promoters and decrease the response times of negatively regulated promoters. Upon considering all these results together one can conclude that the FFLs of types PPP, P-P and PNP are robust against changes in both \bar{v} and \bar{w} . This also could be one of the reasons [14] associated with the observation [10] that these particular coherent C1 type FFLs with both OR and AND-logics are more abundant in nature than the other types. These results are summarized in Table 3.

When the binding parameters $(\mu_{abs}, \mu_{ycc}, L_{ab})$ are much lesser than one, then we find from Eq 11 that the TF gene C will be turned on toward its maximum expression level. This means that strong binding conditions are required for the following types of molecular interactions to achieve the maximum production of TF protein C viz. (1) protein-protein interactions between A and B, (2) binding of the dimer A-B at the promoter of C and (3) binding of protein A at the promoter of B. Conditions (1–3) also suggest that for an efficient filtering activity of the P-P type A-AND-B FFL against transient input signals at the promoter of the TF gene A, the inequality conditions $(\mu_{abs}, \mu_{ycc}, L_{ab}) \gg 1$ are necessary which in turn will decrease the maximum achievable steady state value of protein C as shown in Figure 3. In other words there exists a critical value of $\bar{\mu} = \bar{\mu}_c$ for a given signal with a pulse width θ above which the response of gene C for a pulse of

signal at the promoter of gene A will be practically zero. In Figure 3, this critical value occurs at $\bar{\mu}_c = 1$ for a pulse width of $\theta = 0.3$. On the other hand, for a given value of $\bar{\mu}$ there exist a cutoff pulse width $\theta = \theta_c$ below which the response of TF gene C is practically zero which is demonstrated in Figure 4. From Figure 4 we find that there also exists a transition region in which the TF gene C responds to the width of the input pulse that is given at the promoter of A in a graded manner. From Eqs 10 we find that

$$P_c = e^{-\tau} \int_0^{\tau} X_{cys}(s) e^{\theta s} ds \text{ and when } \chi(\tau) = H(\tau) - H(\tau - \theta), \text{ then}$$

we can conclude that the TF gene C will respond to the variation of width θ of the activation signal in a graded manner in the strong binding limit as $(\mu_{abs}, \mu_{ycc}, L_{ab})$ tend toward zero. The maximum achievable P_c will increase proportional to the width of the activation signal. Whereas in the weak binding limit as the binding parameters $(\mu_{abs}, \mu_{ycc}, L_{ab})$ tend toward infinity, the response of TF gene C with respect to the pulse width θ seems to be approximately a sharp type. Under such conditions there exists a critical pulse width (θ_c) with a sharp transition region above which the maximum achievable P_c will be $P_c \sim 1$ and below which the maximum achievable P_c will be $P_c \sim 0$.

It seems that the parameters $(\theta_{ab}, \theta_{ycc}, L_{ab})$ associated with various binding events need to be fine-tuned to achieve both high efficiency in the filtering activity as well as maximum possible steady-state values of protein P_c upon inducing the promoter of gene A by a persistent signal. Results from stochastic simulations at various values of \bar{v} and w_c are shown in Figure 5 and summarized in Table 4. These results suggest that the coefficient of variation in the response times associated with various FFLs under strong

Table 3. Summary of segregation patterns at weak and strong binding conditions.

Parameters	Segregation pattern of response times in $w \in (0.001, 10)$	Abundances of various types of FFLs as in Ref [10]	Source
Condition I	P-P > N-P > C1 > I4 > C4 > I1 > I3 > C2 > N-N > P-N > I2 > C3		
Condition II	{I3, C4} > {C1, I4, N-P} > {P-P, I1}, {C2} > {N-N, I2} > C3, P-N	C1 > I1 > C3 > C2 > {C4, I3, I4}	Prokaryotes
Condition III	{C4 > I3} > {C1, P-P, N-P, I4} > I1, {C2} > {N-N, I2} > C3, P-N	C1 > I1 > C2 > I2 > I3	Eukaryotes (yeast)
Condition IV	{C4 > I3} > {C1, P-P, N-P, I4}, {C2} > {N-N, I2} > C3, {P-N, I1}	I1 will be much lower than C1 type.	Higher eukaryotes

Note: This table summarizes the behavior of various FFLs under strong and weak binding conditions as well as fast and slow promoter-state dynamics. Under weak binding conditions ($\bar{\mu} \geq 1$) each FFL behaves differently from each other with respect to changes in \bar{w} . Here the settings for **Condition I:** weak binding and fast promoter state dynamics ($\bar{v}=0.0003, \bar{\sigma}=4, \bar{\mu}=1$). **Condition II:** strong binding and fast promoter state dynamics ($\bar{v}=0.0003, \bar{\sigma}=4, \bar{\mu}=0.001$). **Condition III:** strong binding and slow promoter state dynamics ($\bar{v}=0.003, \bar{\sigma}=4, \bar{\mu}=0.001$). **Condition IV:** strong binding and slow promoter state dynamics ($\bar{v}=0.03, \bar{\sigma}=4, \bar{\mu}=0.001$). Under strong binding conditions $\bar{\mu} \leq 0.001$, the entire set of FFLs segregates approximately into three subgroups I, II and III. Here P-P (C1 type FFL with AND type logic on TF gene C) and I1 behaves similarly and therefore the advantages of I1 type FFL whose response time is lower than the generation time will be shared by P-P type which will be added up to the C1 type FFL with OR type gated logic to TF gene C. However this pattern seems to be weakly dependent on \bar{v} . When \bar{v} increases as in case of eukaryotic cell, then P-P behaves similar to C1 type and I1 type FFL will have the entire advantage of having lower response times than other subgroups of Group I. As a result, I1 type FFL will be more abundant in eukaryotes than prokaryotes. All these results are not dependent on changes in $\bar{\sigma} \in (0.1, 10^3)$, increasing \bar{v} beyond 0.03 or decreasing below 0.0003. One should note that the physiological value of $\bar{\sigma}$ in prokaryotes will be $\bar{\sigma} \sim 4$. The overall response time of first subgroup (I3 and C4) of Group I is higher than the generation time of the cell. The response times of the second subgroup (C1, I4 and N-P) are closer to the generation time whereas the third subgroup possess lesser response times than the others. The response times of the FFLs in Group-I are more robust against changes in $\bar{w} \in (0.1, 1)$ over the physiological values than Group II and III. Comparison with the relative abundances of naturally occurring FFLs, one can conclude that those FFLs are naturally selected when their response times are (a) robust against changes in $\bar{w} \in (0.1, 1)$ and (b) closer to or lesser than the generation time.

doi:10.1371/journal.pone.0041027.t003

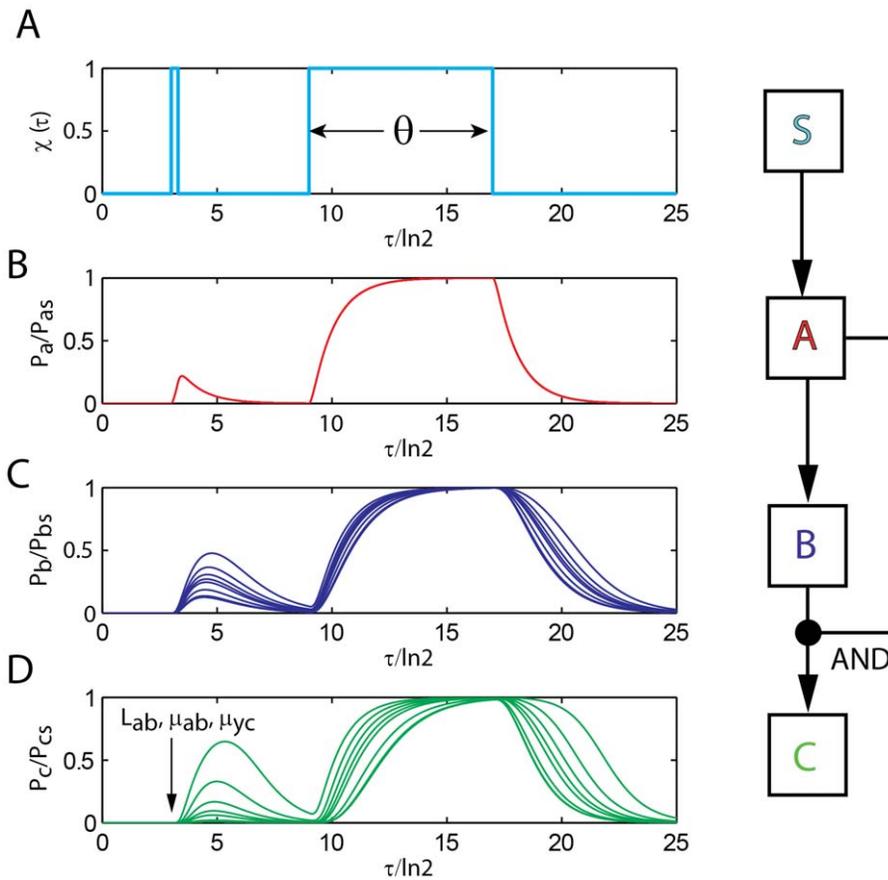


Figure 3. Dependency of filtering efficiency of P-P type (C1 type with AND-logic) FFL on the set of binding parameters $\bar{\mu}$. Here the general settings are: $\bar{w}=0.12$, $\bar{v}=0.0003$, $\sigma=4$, $T=25$ generation times, and $\bar{\mu}$ was varied as (0.1, 0.2, 0.3, 0.4, 0.5, 1, 3 and 5). A. Input signal at the promoter of TF gene A. This has one short rectangular pulse with a width of $\theta=0.3$ and a large one with $\theta=8$ all are measured in terms of number of generation times. B. TF gene A responds similarly for all values of $\bar{\mu}$ to both the signals irrespective of the binding strengths. The response seems to be proportional to the pulse width θ without any delay. C. As binding strength increase, the response of B also increases proportionately. D. There exists a cutoff value of $\bar{\mu}$ above which the expression level of TF gene C is practically zero. With the current settings, this cutoff seems to occur at $\bar{\mu}=1$. Arrow shows the increasing direction of $\bar{\mu}$. doi:10.1371/journal.pone.0041027.g003

binding conditions ($\bar{\mu} < 0.001$) are robust against changes in w_c as well as \bar{v} . When $\bar{v} < 0.003$ (as in case of prokaryotes and yeast), then based on the overall coefficient variation in the response times all the FFLs can be categorized into at least three different groups viz. $\{I1, P-N\} > \{C3, N-N, C2, I2\} > \{C1, C4, I3, P-P, N-P, I4\}$. When $\bar{v} > 0.003$ (as in case of higher eukaryotes such as human and plants), then the segregation pattern associated with the overall coefficient of variation in the response times seems to be as $\{I1, P-N\} > \{C3, N-N, C2, I2\} > \{C4, I3\} > \{C1, P-P, N-P, I4\}$. The coefficient of variation in the response time of the first group seems to be $> 100\%$ and in the second group it is 1% and in the third/fourth group of FFLs it is 0.1%. The coherent C1 FFL with both OR/AND type logic shows the least amount of variation in the response times among all the FFLs under consideration.

Earlier results ([10], supplementary materials) based on the analysis of literature-based databases of experimentally verified direct transcription interactions for *E. coli* [6] suggested a distribution pattern of various FFLs as 83% coherent (out of which 80% were C1 type, 11% C3, 6% C2 and 3% C4) and 17% incoherent (out of which 72% were I1, 14% I2 and 14% I4) types. Similar analysis on *S. cerevisiae* [6,10] showed a distribution of FFLs as 55% coherent (out of which 84% were C1 and 16% C2) and 45% incoherent (out of which 84% were I1, 12% I2 and 4% I3)

types. This data suggested an overall distribution of various FFLs in the transcription factor networks of bacteria as (67% C1, 12% I1, 10% C3, 5% C2, 1% C4, 1% I3 and 1% I4). The overall distribution of various FFLs in the TF networks of yeast is (47% C1, 38% I1, 9% C2, 5% I2 and 1% I3). Comparison of the order of preferences of various FFLs in bacteria (C1 > I1 > C3 > C2 > {C4, I3, I4}) and yeast (C1 > I1 > C2 > I2 > I3) with the Group I, II and III types of FFLs, one can conclude that those FFLs of Group I (C1, I1) whose response times are close to or less than the generation time of the cell apart from the robustness against changes in \bar{w} are more preferably selected. The next set of preferable FFLs (C2, C3 and I2) is from Group II. The set of FFLs (I3, I4 and C4) are the least preferable ones. Although I3 and C4 fall within Group I, their response times are much higher than the generation time. It seems that the relative abundance of I1 type FFL is more in eukaryotes than the prokaryotes. Here one should note that the value of the parameter set $\bar{v}(v_{hk} = \gamma_{pc}/k_{ph}\phi_{hs}) = 0.0003$ that is used in our simulations (Fig. 2A–D) is derived for a prokaryote organism whose cell volume is $\sim 10^{-18} \text{ m}^3$. With this setting we find from our simulations that the subset (P-P and I1) behaves in a similar way within Group I which is evident from the segregation pattern of various FFLs at strong binding conditions $\{\{I3, C4\}, \{C1, I4, N-P\}, \{P-P, I1\}\}, \{C2, \{N-N, I2\}, C3\}, P-N$.

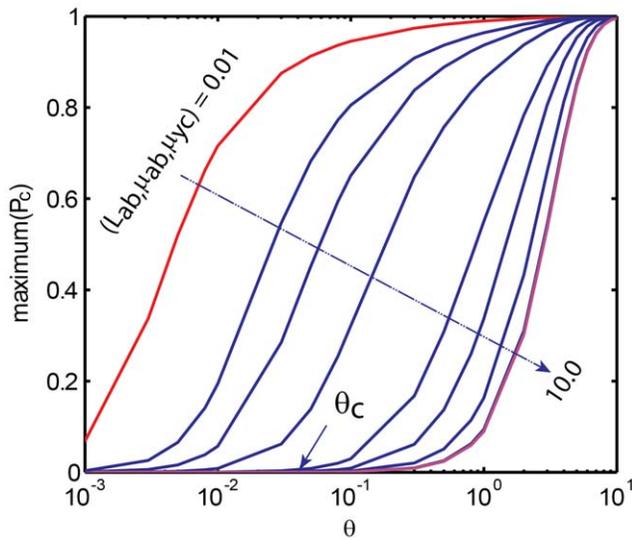


Figure 4. Dependency of the maximum achievable protein product of TF gene C on the width of the input pulse at the promoter of A in case of P-P A-AND-B type FFL (C1 with AND-logic). Here the settings are $\bar{w}=0.12$, $\bar{v}=0.0003$, $\bar{\sigma}=4$, $T=25$ generation times, and $\bar{\mu}$ was varied as (0.01, 0.03, 0.05, 0.1, 0.3, 0.5, 1, 5 and 10). Beyond $\bar{\mu}>5$, there is not much change in the variation of maximum of P_c with respect to the pulse width. The cutoff pulse width θ_c seems to be strongly dependent on the binding parameters $\bar{\mu}$. doi:10.1371/journal.pone.0041027.g004

We argue that some of the advantages of I1 such as lesser response time than the generation time apart from the robustness against changes in \bar{w} will be shared by P-P type FFL that is in turn added to the total abundances of C1. This follows from the fact that P-P is also a C1 type FFL with AND gated logic type regulation on the promoter of TF gene C. Stochastic simulations results suggested that the coefficient of variation in the response times is lesser in C1 type FFLs than the other FFLs which agrees well with the observation that these types are relatively more abundant in nature. On the other hand I1 type FFLs possess highest coefficient of variation in the response times. Overall results suggest that the coefficient of variation in the response times is relatively higher whenever the type of regulation imparted by TF genes A and B on the TF gene C is negative. This result is in line with the observations on the negatively self-regulated TF networks in which the coefficient of variation in the response times was shown [9] to be $>50\%$. These results suggest that I1 type FFL as well as negative self-regulated motifs can speed up the response time at the cost of increasing the fluctuations in their response times.

Since the volume of eukaryotic nucleus is much larger than a bacterial cell volume over several orders of magnitudes (volume of yeast nucleus is $\sim 10^1$ times larger than a bacterial cell volume [25] whereas human cell nucleus is $\sim 10^2$ times larger than a bacterial cell volume), the overall bimolecular collision rate associated with the binding of TFs at their cognate sites (k_{fb}) inside the nucleus will be much lower in case of eukaryotes than prokaryotes owing to the dilution effects. This means that the physiological values of \bar{v} will be higher in eukaryotes than prokaryotes. Upon an increase in \bar{v} as in case of eukaryotes such as yeast ($=0.003$), P-P type behaves in a similar way (Fig. 2E–F) as that of C1 whereas I1 type FFL recovers the full advantage over its lower response time within Group I compared to C1 which is evident from the altered segregation pattern {C4, I3, {C1, P-P, N-P, I4}, I1}, {C2, {N-N, I2}, C3}, P-N. This in turn will result in the higher relative abundances of I1

type FFL in eukaryotes than the prokaryotic TF networks. Further increase in \bar{v} as in case of nucleus of human cell ($=0.03$) shift the I1 type FFL from Group I to Group III. This result predicts that I1 type FFLs will be much lesser in abundance than the C1 type in the TF network of human and other higher animals and plants. One also should note that we are still not able to explain the relative scarcity of I4 and N-P type FFLs even though they fall well within Group I. As pointed out by Magnan and Alon in reference [10], the relative scarcity could be due to the reduced functionality of type 3 and type 4 FFLs with AND gated logic (here it is applicable to N-P) since they response at most one of the triggering signals given at the promoters of both the genes A and B.

We have considered the sets of similar parameters (\bar{w} , \bar{v} , $\bar{\mu}$, $\bar{\sigma}$, $\bar{\rho}$) associated with TF genes A, B and C as variable units rather than individual ones for all these calculations and simulations. The TF network of an organism consists of several FFL motifs and each of these FFLs is constituted with different subsets of a pool of TF genes. Each subset of TFs can be represented as points in the parametric space of our model. Upon considering the entire pool of TF genes, one should note that each of the members of these sets of parameters can take a spectrum of values with a probability distribution with definite mean and variance. There are two types of variability of parameters among the TF genes of FFLs viz. variation of the parameters of TF gene A/B/C across various FFLs found in the entire TF network and variation of the parameters of TF genes A, B, C within a given FFL. Here we have assumed that both of them are approximately the same. The main results of our analysis will not be affected much due to the second type of variation since we have used the mean values of the parameters associated with the entire set of TF genes of an organism to represent genes A, B and C of various types FFLs.

Methods

We use the following Euler type iterative numerical scheme to integrate the deterministic differential Eqs 1–3 associated with various A-OR/AND-B type FFLs in the dimensionless time and concentrations space.

TF A:

$$\begin{aligned} M_{a,n+1} &= M_{a,n} + w_a^{-1}(\chi(\tau_n) - M_{a,n})\Delta\tau \\ P_{a,n+1} &= P_{a,n} + \rho_a^{-1} \\ &\left(M_{a,n} - P_{a,n} - \sigma_{ab}(P_{a,n}(1 - X_{ba,n}) - \mu_{ab}X_{ba,n}) \right. \\ &\left. - \sigma_{ac}(P_{a,n}(1 - X_{ca,n} - X_{cb,n}) - \mu_{ac}X_{ca,n}) \right) \Delta\tau \end{aligned} \quad (23)$$

TF B:

$$\begin{aligned} X_{ba,n+1} &= X_{ba,n} + v_{ba}^{-1}(P_{a,n}(1 - X_{ba,n}) - \mu_{ab}X_{ba,n})\Delta\tau \\ M_{b,n+1} &= M_{b,n} + w_b^{-1}(\Omega_{\pm}(X_{ba,n}) - M_{b,n})\Delta\tau \\ P_{b,n+1} &= P_{b,n} + \rho_b^{-1}(M_{b,n} - P_{b,n} - \sigma_{bc}(P_{b,n}(1 - X_{ca,n} - X_{cb,n}) \\ &- \mu_{bc}X_{cb,n}))\Delta\tau \end{aligned} \quad (24)$$

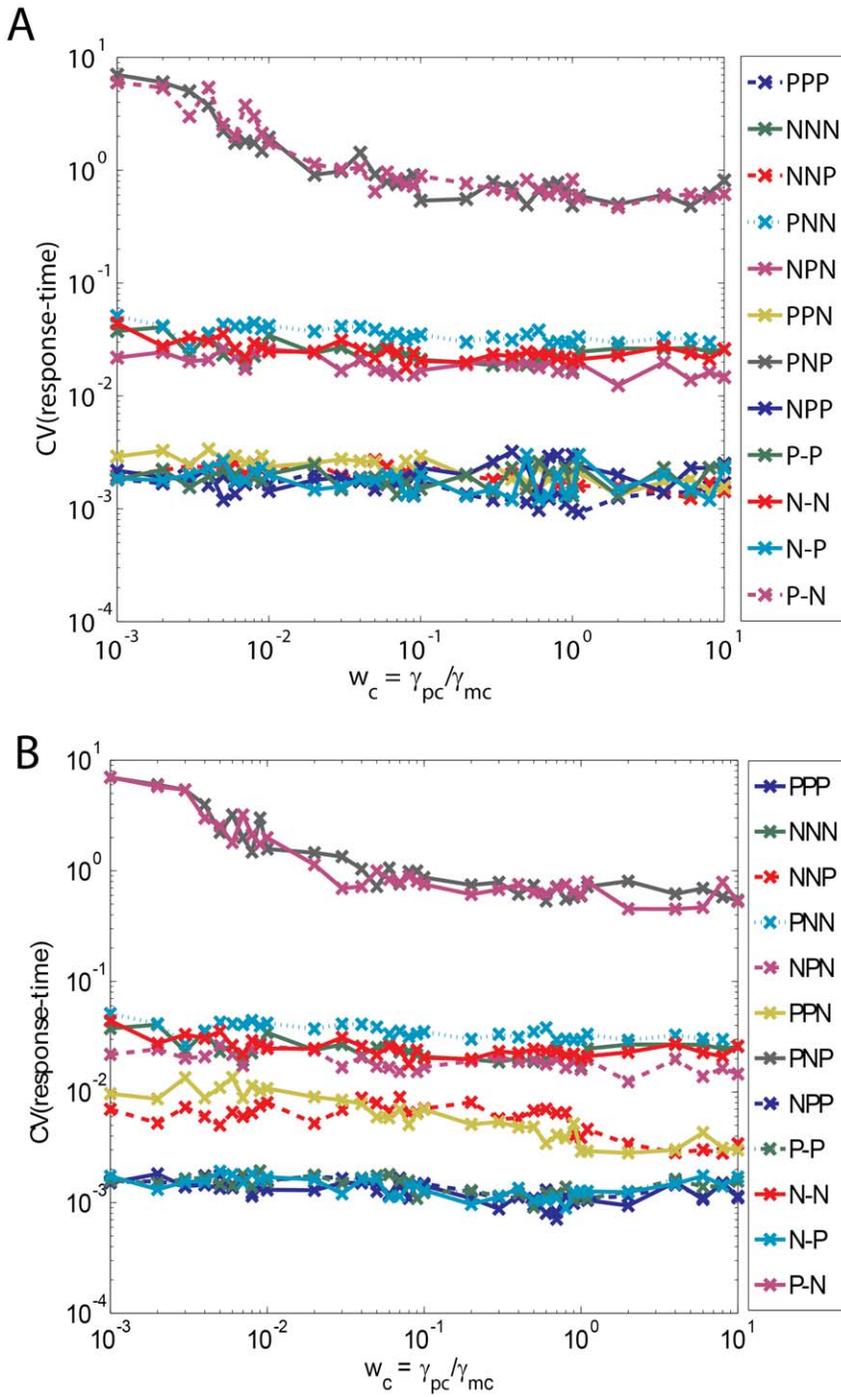


Figure 5. Coefficient of variation (CV) associated with the fluctuations in the response times of various types of FFLs under strong binding conditions ($\bar{\mu} = 0.001$). CV was calculated over 10^5 numbers of stochastic trajectories. The coherent C1 type FFL possesses lower CV of response times than other FFLs whereas I1 type possesses highest CV of response times. A. Here the settings are $\bar{v} \leq 0.003$ that is applicable to both prokaryotes and eukaryotes such as yeast, \bar{w} was iterated from 0.001 to 10, and $\bar{\sigma} = 4$, $T = 25$ generation times. B. Here the settings are $\bar{v} \geq 0.03$ that is applicable to higher eukaryotes such as human, \bar{w} was iterated from 0.001 to 10, and $\bar{\sigma} = 4$, $T = 25$ generation times. doi:10.1371/journal.pone.0041027.g005

Table 4. Segregation patterns of CV (response time) at weak and strong binding conditions.

Parameters	Segregation pattern of CV (response-time) in $\bar{w} \in (0.001, 10)$	Abundances of various types of FFLs as given in Ref [10]	Source
Condition II	{I1, P-N} > {C3, N-N, C2, I2} > {C1, C4, I3, P-P, N-P, I4}	C1 > I1 > C3 > C2 > {C4, I3, I4}	Prokaryotes
Condition III	{I1, P-N} > {C3, N-N, C2, I2} > {C1, P-P, C4, I3, I4, N-P}	C1 > I1 > C2 > I2 > I3	Eukaryotes (yeast)
Condition IV	{I1, P-N} > {C3, N-N, C2, I2} > {C4, I3} > {C1, P-P, N-P, I4}	I1 will be lower than C1 type.	Higher eukaryotes

Note: This table summarizes the results from the stochastic simulation. Parameter settings for **Condition II:** strong binding and fast promoter state dynamics ($\bar{v}=0.0003$, $\bar{\sigma}=4$, $\bar{\mu}=0.001$). **Condition III:** strong binding and slow promoter state dynamics ($\bar{v}=0.003$, $\bar{\sigma}=4$, $\bar{\mu}=0.001$). **Condition IV:** strong binding and slow promoter state dynamics ($\bar{v}=0.03$, $\bar{\sigma}=4$, $\bar{\mu}=0.001$). Here CV (defined as the ratio standard deviation/mean) represents the coefficient of variation in the response times (the time required to attain half of the steady state value of the transcription factor protein C in the FFLs).
doi:10.1371/journal.pone.0041027.t004

TF C:

$$\begin{aligned}
 X_{ca,n+1} &= X_{ca,n} + v_{ca}^{-1} (P_{an}(1 - X_{ca,n} - X_{cb,n}) - \mu_{ac} X_{ca,n}) \Delta\tau \\
 X_{cb,n+1} &= X_{cb,n} + v_{cb}^{-1} (P_{bn}(1 - X_{ca,n} - X_{cb,n}) - \mu_{bc} X_{cb,n}) \Delta\tau \\
 M_{c,n+1} &= M_{c,n} + w_c^{-1} (\zeta_{a \pm b \pm} (X_{ca,n}, X_{cb,n}) - M_{c,n}) \Delta\tau \\
 P_{c,n+1} &= P_{c,n} + (M_{c,n} - P_{c,n}) \Delta\tau
 \end{aligned} \quad (25)$$

To investigate the filtering efficiency of P-P type A-AND-B FFL we can use the signal function $\chi(\tau) = \sum_{i=1}^2 (H(\tau - \tau_i) - H(\tau - \tau_i - \theta_i))$ at the promoter of TF gene A with two well separated rectangular pulses with respectively lower and higher pulse widths. For the purpose of computing the response times we can reset the input function at the promoter of gene A as $\chi(\tau) = H(\tau - T)$ where T is the total simulation time. The numerical scheme for the modified Eqs 6–8 for Y_{ab} , P_a , P_b , X_{cy} , and M_c in case of A-AND-B type FFLs can be written as follows.

$$\begin{aligned}
 Y_{ab,n+1} &= Y_{ab,n} + \varepsilon_{ab}^{-1} (P_{a,n} P_{b,n} - L_{ab} Y_{ab,n}) \Delta\tau \\
 P_{a,n+1} &= P_{a,n} + \rho_a^{-1} \\
 &\quad \left(M_{a,n} - P_{a,n} - \sigma_{ab} (P_{a,n} (1 - X_{ba,n}) - \mu_{ab} X_{ba,n}) \right. \\
 &\quad \left. - \sigma_{\lambda a} (P_{a,n} P_{b,n} - L_{ab} Y_{ab,n}) \right) \Delta\tau \\
 P_{b,n+1} &= P_{b,n} + \rho_b^{-1} \\
 &\quad (M_{b,n} - P_{b,n} - \sigma_{\lambda b} (P_{a,n} P_{b,n} - L_{ab} Y_{ab,n})) \Delta\tau \\
 X_{cy,n+1} &= X_{cy,n} + v_{cy}^{-1} (Y_{ab,n} (1 - X_{cy,n}) - \mu_{yc} X_{cy,n}) \Delta\tau \\
 M_{c,n+1} &= M_{c,n} + w_c^{-1} (\Psi_{y \pm} (X_{cy,n}) - M_{c,n}) \Delta\tau
 \end{aligned} \quad (26)$$

The total scaled simulation time T will be divided into N equal intervals such that $\Delta\tau = T/N$. For simulation purpose we set $\Delta\tau = 5 \times 10^{-5}$ and the corresponding $\Delta t = 0.2s$ for a lifetime $1/\gamma_{pc} \sim 60$ mins. The initial conditions are $X_{hg,0} = 0$, $M_{h,0} = 0$ and $P_{h,0} = 0$ for $h = a, b, c$ and $g = b, c$. To compute the response-time of the TF gene C we set an absorbing boundary at $P_c/P_{cs} = 1/2$ where P_{cs} is the steady state value of scaled concentration P_c which

in turn depends on the type of FFL under consideration as given in Eq 13. We measure the concentrations in terms of number of molecules inside the cell. Considering a bacterial cell (volume $\sim 10^{-18} \text{ m}^3$) [21] we set $d_{hz} = 1$, $m_{hs} \sim 10^2$ molecules and $p_{hs} \sim 10^3$ molecules [22,23] where $h = a, b$ and c respectively denotes TF genes A, B and C. Concentration of a single TF molecule inside a bacterial cell will be ~ 2 nM. Concentration of a single TF molecule inside the nucleus of yeast cell [24,25] will be ~ 200 pM and inside the nucleus of human cell it is ~ 20 pM. We measure the timescales in terms of the lifetime of the TF protein C by dividing t by γ_{pc} . Since the dynamics of binding-unbinding of the transcription factors with the respective promoters is a typical diffusion-controlled site-specific DNA-protein interaction, under *in vivo* conditions of bacterial cell we find that $k_{fkh} \sim 10^{-3} \text{ molecules}^{-1} \text{ s}^{-1}$ [22–24]. Here we have assumed an *in vivo* three dimensional diffusion controlled collision rate $\sim 10^6 \text{ M}^{-1} \text{ s}^{-1}$. In case of nucleus of yeast cell we find $k_{fkh} \sim 10^{-4} \text{ molecules}^{-1} \text{ s}^{-1}$ and in case of nucleus of human cell we find $k_{fkh} \sim 10^{-5} \text{ molecules}^{-1} \text{ s}^{-1}$. Here the TF protein $k = (a, b)$ binds with the promoters of $h = (b, c)$. It will be a complicated task to explore the entire parametric space. To simplify the analysis further we can consider the sets of parameters viz. $\bar{v} = (v_{ba}, v_{ca}, v_{cb}, \varepsilon_{ab}, v_{cy})$, $\bar{w} = (w_a, w_b, w_c)$, $\bar{\rho} = (\rho_a, \rho_b)$, $\bar{\mu} = (\mu_{ab}, \mu_{ac}, \mu_{bc}, \mu_{yc}, L_{ab})$ and $\bar{\sigma} = (\sigma_{ab}, \sigma_{bc}, \sigma_{ac}, \sigma_{\lambda a}, \sigma_{\lambda b})$ as the parametric units of our numerical simulations. When all the TF proteins A, B and C decay with similar decay rate constants then we find that $\bar{\rho} = 1$. Using the steady state values of various proteins and mRNA numbers in the absence of any regulation one obtains $\bar{v} \sim 3 \times 10^{-4}$ (~ 0.003 for nucleus of yeast cell and ~ 0.03 for nucleus of human cell) and $\bar{\sigma} \sim 4$ (~ 0.4 for yeast cell nucleus and ~ 0.04 for nucleus of human cell) where we have used a protein decay rate $\gamma_{pc} \sim 3 \times 10^{-4} \text{ s}^{-1}$ (for a protein lifetime of ~ 60 min) to transform the real time variables to τ space variables. The dependency of various properties of TF gene C such as response-time, delay with respect to the induction signal at the promoter of gene A and maximum achievable P_c on the parameter sets $\bar{\mu}$, \bar{v} , $\bar{\sigma}$ and \bar{w} were explored. The values of \bar{w} seems to vary across the spectrum of genes from $\bar{w} \sim 10^{-1}$ in case of prokaryotes to $\bar{w} \sim 1$ in case of eukaryotes [17–19]. Together with all these values we iterated the parameter set \bar{w} inside the range (0.001, 10) with a step size of $\Delta\bar{w} = 0.001$. For further exploratory purposes we considered four different binding conditions from low to high viz. $\bar{\mu} = (0.001, 0.01, 0.1, 1)$ and three different values of $\bar{v} = (0.0003, 0.003, 0.03)$.

Similar to Eqs 21–24 we use Euler type numerical scheme to integrate the Chemical Langevin Eqs 16–20 where we replace Γ with Gaussian distributed random numbers with zero mean and unit variance. Assuming an overall *in vivo* steady state mRNA levels as $m_{hs} \sim 10^2$ and proteins levels as $p_{hs} \sim 10^3$ (where $h = a, b, c$,

measured in number of molecules) we find that $\lambda_{mh} \sim 10^{-2}/w_h$ molecule $^{-1}$ and $\lambda_{ph} \sim 10^{-3}/\rho_h$ molecule $^{-1}$. Further we also find that $\lambda_{ab} \sim 1$ s, $\lambda_{xkh} \sim 10^{-3}$ molecule $^{-1}$ s ($k = b, c$ and $h = a, b$) and $\lambda_{xye} \sim 10^{-3}$ molecule $^{-1}$ s. We use the reflecting boundaries (0, 1) for the scaled concentration variables (P_h, M_h, X_{ab} , and X_{nk}) where $n = b, c$ and $k = a, b$ and the absorbing boundary condition $P_c = 1/2$ to compute the mean first passage time (response-time) from the stochastic simulations. Averaging was done over 10^5 trajectories at each w_c value and the coefficient of variation of response time was computed as CV (response-time) = standard deviation of response-time/mean of response-time.

Conclusions

Feedforward loops (FFLs) consist of three genes which code for three different transcription factors A, B and C where B regulates C and A regulates both B and C. We have developed a detailed model to describe the dynamical behavior of various types of coherent and incoherent FFLs in the transcription factor networks. We considered the deterministic and stochastic dynamics of both promoter-states and mRNAs of various genes coding for the transcription factors associated with the FFL motifs. Detailed

analysis showed that the response times of FFLs strongly dependent on the ratios ($w_h = \gamma_{pc}/\gamma_{ph}$ where $h = a, b, c$) between the lifetimes of mRNAs of A, B and C ($1/\gamma_{mh}$) and the protein of C ($1/\gamma_{pc}$). When the binding of transcription factors A and B with the *cis*-acting elements of respective promoters is very strong, then we could categorize all the possible FFLs into Group I, II and III based on the dependency of the response times of FFLs on w_h . Though the response times of Group I FFLs were higher than II and III, they seem to be less sensitive to the changes in w_h within the naturally occurring dynamic range (0.1, 1). We have further shown that among the members of the Group I FFLs, the coherent C1 type was more robust against changes in other system parameters which could be one of the reasons why C1 type coherent FFLs are more abundant in nature than the others.

Author Contributions

Conceived and designed the experiments: RM. Performed the experiments: RM. Analyzed the data: RM. Contributed reagents/materials/analysis tools: RM. Wrote the paper: RM.

References

- Levin B (2011) Genes X. Boston: Jones and Bartlett Publishers. 546–572
- Ptashne M, Gann A (2001) Genes and Signals. New York: Cold Spring Harbor Laboratory Press. 60–72.
- Ptashne M, Gann A (1997) Transcriptional activation by recruitment. Nature 386: 569–577
- Alberts B, Bray D, Lewis J, Roberts K, Watson JD (1994) Molecular Biology of the Cell. New York: Garland. 331–400.
- Alon U (2006) An Introduction to Systems Biology. London: CRC Press. 41–71.
- Shen-Orr SS, Milo R, Mangan S, Alon U (2002) Network motifs in the transcriptional regulation network of Escherichia coli. Nat Genet 31: 64–68.
- Milo R, Itzkovitz S, Kashtan N, Levitt R, Shen-Orr S, et al. (2004) Superfamilies of evolved and designed networks. Science 303: 1538–1542.
- Rosenfeld N, Elowitz MB, Alon U (2002) Negative autoregulation speeds the response times of transcription networks. J Mol Biol 323: 785–793.
- Murugan R, Kreiman G (2011) On the minimization of fluctuations in the response times of auto regulatory gene networks. Biophys J 101: 1297–1306.
- Mangan S, Alon U (2003) Structure and function of the feed-forward loop network motif. Proc Natl Acad Sci USA 100: 11980–11985.
- Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, et al. (2002) Transcriptional regulatory networks in Saccharomyces cerevisiae. Science 298: 799–804.
- Mangan S, Itzkovitz S, Zaslaver A, Alon U (2006) The incoherent feed-forward loop accelerates the response-time of the gal system of Escherichia coli. J Mol Biol 356: 1073–81.
- Wall ME, Dunlop MJ, Hlavacek WS (2005) Multiple functions of a feed-forward-loop gene circuit. J Mol Biol 349: 501–14.
- Mangan S, Zaslaver A, Alon U (2003) The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. J Mol Biol 334: 197–204.
- Eichenberger P, Fujita MS, Jensen T, Conlon EM, Rudner DZ, et al. (2004) The program of gene transcription for a single differentiating cell type during sporulation in Bacillus subtilis. PLoS Biol 2: e328.
- Elkon R, Zlotorynski E, Zeller KI, Agami R (2010) Major role for mRNA stability in shaping the kinetics of gene induction. BMC Genomics 11: 259–267.
- Shahrezaei V, Swain PS (2008) Analytical distributions for stochastic gene expression. Proc Natl Acad Sci USA 105: 17256–17261.
- Shahrezaei V, Ollivier JF, Swain PS (2008) Colored extrinsic fluctuations and stochastic gene expression. Mol Syst Biol 4: 196–205.
- Newman JRS, Ghaemmaghami S, Breslow D, Noble M, DeRisi JL, et al. (2006) Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. Nature 441: 840–846.
- Murugan R (2010) Theory on the mechanism of distal-action of transcription factors: Looping of DNA versus tracking along DNA. J Phys A: Math & Theor 43: 415002–415015.
- Goodsell DS (1991) Inside a living cell. Biochem Sci 16: 203–206.
- Murugan R (2007) Generalized theory of site-specific DNA-protein interactions. Phys Rev E 76: 011901–011909.
- Murugan R (2010) Theory of site-specific DNA-protein interactions in the presence of conformational fluctuations of DNA binding domains. Biophys J 99: 353–359.
- Lomholt M, Broek VB, Kalisch SM, Wuite GJ, Metzler R (2009) Facilitated diffusion with DNA coiling. Proc Natl Acad Sci USA 106: 8204–8208.
- Neumann FR, Nurse P (2007) Nuclear size control in fission yeast. J Cell Biol 179: 593–600.
- Gardiner CW (2004) Handbook of Stochastic Methods. Springer, Berlin.
- Gillespie D (1977) Exact Stochastic Simulation of Coupled Chemical Reactions. J Phys Chem 81: 2340–2361.
- Gillespie D (2000) The chemical Langevin equation. J Chem Phys 113: 297–306.