A Hill type equation reveals the regulation principle of target protein expression led by p53 pulsing

Xiaomin Shi

Department of Mathematics and International Center for Quantum and Molecular Structures, Shanghai University, Shanghai 200444 China

email: shixm1@vip.126.com

Abstract The central dogma indicates the basic direction of gene expression pathway. For activated gene expression, the relationship between the various links that are from the binding of transcription factor (TF) DNA to synthesis of proteins is still unclear. There is a consensus that at steady state the protein levels are largely determined by mRNA level. How to find this steady state? What factors determine it? Taking TF p53 as an example, on the basis of the previously discovered Hill type equation that can characterize the average steady state of mRNA expression under p53 pulsing, and demonstrating the regulatory ability of duration and frequency beyond saturation. I proved that the average steady state of target protein expression driven by p53 pulsing can also be governed by this equation. Therefore, at steady state the average mRNA and protein expression under TF pulsing are exactly the same. This consensus has been successfully demonstrated. For gene *BAX* of p53 target, the observed fold changes of mRNA and protein are 1.72, 1.28, while the predictions for both mRNA and protein from the obtained equation are 1.35. The regulatory principle of gene expression revealed by this equation is that TF dynamics and binding affinity control the expression levels; and the half-life of mRNA and protein regulate the expression stability. By introducing two quantitative indicators, we can understand the degree of accumulation and stability of protein expression. Furthermore, this generalized equation broadens the application of classical biochemical theory.

Keyword Hill equation, p53 target protein expression, stability of gene expression, protein accumulation, transcription factor p53 oscillations

Introduction

According to the central dogma of molecular biology, gene expression pathway is from DNA to RNA to protein (1-2) , which is a process of transcribing gene encoded information into mRNA and translating it into functional proteins. In gene expression, mRNAs and proteins provide the useful readouts connecting genes and phenotypes (3). Generally speaking, there are two different gene expression pathways. One is basal pathway that regulates homeostatic expression; the other is activated or repressed pathways that dynamically regulate expression in response to environmental stimuli (4-8).

For basal gene expression, the protein levels at steady state depends on the transcription rate, the mRNA half-life, the translation rate constant, and the protein half-life (9,10). Moreover, the situations with high transcription rates and low translation rate constants should be ruled out (10). Under environmental stimuli, for gene expression activated by transcription factor (TF), the researches on transcripteomics and proteomics indicate that the protein level at the steady state largely contributed by the mRNA level (3,11,12). And there is little relationship with the synthesis and protein half-life.

Here, the steady state can be defined as a relatively long-term process experienced by cells (11). Therefore, both cell proliferation and differentiation can be considered as the steady state (11). Cell apoptosis and senescence are also under the steady states (13,14). The steady state achieved after stimulation is different from that before the perturbation (15).

For the steady state before stimulation in dendritic cells, the mRNA levels explain 68% the change in protein expression, and for the approximate steady state after stimulation with Lipopolysaccharide, mRNA levels explain 90% of the changes in protein expression (12,16). In addition, in the development of *C elegans*, the fold changes of mRNA and protein expression are almost exactly the same. The transcripts fold change is 2.02, and the proteins is 2.05 (17). Therefore, we can speculate that the fold changes of mRNA and protein expression at steady state are equal. Could we theoretically find this special steady state of protein expression? What factors determine the protein expression levels at steady state?

The gene expression under environmental stimuli is driven by TF which is a mapping of the corresponding stimulus. TF dynamics encodes the specific stimuli (18). The binding component of the TF and DNA constitute the decoder of TF dynamics. This initiates the corresponding target gene expression and fulfills the relevant function. Some TFs exhibit oscillatory dynamical behaviors, in which the duration, frequency, and amplitude can all serve as the encoding components of gene expression, thereby leads to different cell fate (19,20).

The tumor suppressor p53 is a TF of being the most extensively studied. In response to DNA damage, p53 concentrations exhibit a series of pulses with fixed duration, frequency, and amplitude, while the number of p53 pulses increases with the degree of DNA damage caused by γ irradiation (21). Changing p53 dynamics from pulsed to sustained behavior leads to alter cell fate from DNA repair to senescence (22). For target gene expression driven by p53 pulsing, there is an interesting phenomenon that the levels of mRNA and protein expression are very close. For example, for the gene *MDM2*, it was observed 24 hours after simulation that the fold change of mRNA is 2.2 (23) or 2.0 (24), while the protein expression is 1.8 that can be regarded as the average steady state over the cell population (23). Therefore, without loss of generality, I will take target gene expression under p53 pulsing as an example

to seek the principle to govern target protein expression at steady state.

Following the genetic information flow specified by the central dogma, we obtained a modified Hill equation characterizing the average p53 DNA binding probability (25). I also found a Hill type equation that can predict the average fold change of target mRNA expression under p53 pulsing (26)

$$\bar{m} = 1 + \frac{\Delta}{T}\frac{\beta A^n}{K_A^n + A^n}, \tag{1}$$

where Δ denotes the duration, T the period, A the amplitude, β the maximal fold change of mRNA expression, and K_A the dissociation constant.

Here, I will complete the last step of describing the central dogma, which is from mRNA to protein. In response to pulsed p53 input dynamics, p21 mRNA dynamics shows the pulsing expression, however, p21 protein dynamics exhibits the rising expression (27). The half–life of mRNA and protein determines the stability of expression dynamics (23,24,28). But the relationship between the steady state levels of mRNA and protein expression is not clear. Thereby I try to prove that the average steady state fold change of mRNA and protein expression under p53 pulsing are equal, *i.e.*, the average fold change of protein at steady state equals $\bar{m}$. In order to achieve this goal, we must develop a useful and correct mathematical model concerning the gene expression under the pulsed TF dynamics (16,29). The ordinary differential equation for the mRNA dynamics is (26)

$$\frac{dmRNA(t)}{dt} = \beta_0' + \beta' \frac{S^n(t)}{K_A^n + S^n(t)} - \alpha mRNA(t), \tag{2}$$

where

$$S(t) = \begin{cases} A, & (i-1)T \le t < (i-1)T + \Delta \\ 0, & (i-1)T + \Delta \le t \le iT \end{cases} \quad i = 1,2,.... ,$$

$mRNA(t)$ denotes the target mRNA concentration, $S(t)$ is the piecewise constant function of TF p53 pulses, β_0' is the basal transcription rate, β' is the maximal transcription rate activated by p53, α is the mRNA decay rate constant, and n is the Hill coefficient. The initial condition is $mRNA(0) = m_0$, where m_0 denotes the basal mRNA concentration.

The equation for the protein expression dynamics is (9,10,23)

$$\frac{dprotein(t)}{dt} = \sigma mRNA(t) - \mu\, protein(t) \tag{3}$$

where $protein(t)$ represents the target protein concentration, σ and μ denote the rate constant of translation and degradation, respectively. The initial protein concentration is $protein(0) = p_0$.

For activated expression dynamics, considering $m_0 = \beta_0'/\alpha$, $p_0 = \sigma m_0/\mu$, let $m(t) = mRNA(t)/m_0$, $P(t) = protein(t)/p_0$, Equation 2 and Equation 3 become

$$\frac{dm(t)}{dt} = \alpha\left(1 + \beta\frac{[S(t)]^n}{K_A^n + [S(t)]^n} - m(t)\right), \quad m(0) = 1 , \tag{4}$$

$$\frac{dP(t)}{dt} = \mu(m(t) - P(t)), \qquad P(0) = 1 . \tag{5}$$

where $m(t)$ and $P(t)$ represent the fold change of mRNA and protein expression, respectively, and $\beta = \beta'/\beta_0'$ denotes the maximal fold change of transcription. Since

Equation 4 has been solved, rewriting the solution from Reference 26 as

$$m_i(\xi_i) = 1 + m_d\left(1 - \frac{e^{-\alpha\xi_i}}{1-e^{-\alpha T}}\left((1-e^{-\alpha(T-\Delta)}) + e^{-i\alpha T}(e^{\alpha\Delta}-1)\right)\right), \quad \xi_i \in [0,\Delta),$$

$$m_i(\xi_i) = 1 + m_d\frac{(e^{\alpha\Delta}-1)(1-e^{-i\alpha T})}{1-e^{-\alpha T}}e^{-\alpha\xi_i}, \quad \xi_i \in [\Delta, T]. \tag{6}$$

$$\xi_i = t-(i-1)T$$

where

$$m_d = \frac{\beta A^n}{K_A^n + A^n} \quad .$$

$m_i(\xi_i)$ represents the mRNA fold change during the i-th TF pulse.

To solve Equation 5 is challenging. Assuming $\alpha \neq \mu$, we can obtain the analytical solutions about target protein expression in response to p53 pulsing:

$$P_i(\xi_i) = 1 + m_d + A_i e^{-\mu\xi_i} - \frac{\mu m_d C(i)}{\mu-\alpha}e^{-\alpha\xi_i}, \xi_i \in [0,\Delta),$$

$$P_i(\xi_i) = 1 + B_i e^{-\mu\xi_i} + \frac{\mu m_d D(i)}{\mu-\alpha}e^{-\alpha\xi_i}, \quad \xi_i \in [\Delta, T], \tag{7}$$

$$\xi_i = t-(i-1)T$$

where

$$A_i = \frac{\alpha m_d}{\mu-\alpha}(e^{\mu\Delta} - \frac{(e^{\mu\Delta}-1)(1-e^{-i\mu T})}{1-e^{-\mu T}}),$$

$$B_i = -\frac{\alpha m_d}{\mu-\alpha}\frac{(e^{\mu\Delta}-1)(1-e^{-i\mu T})}{1-e^{-\mu T}},$$

$$C(i) = \frac{1-e^{-\alpha(T-\Delta)} + e^{-i\alpha T}(e^{\alpha\Delta}-1)}{1-e^{-\alpha T}},$$

$$D(i) = \frac{(e^{\alpha\Delta}-1)(1-e^{-i\alpha T})}{1-e^{-\alpha T}}.$$

$P_i(\xi_i)$ represents the protein fold change under the i-th p53 pulse. The detailed derivation process for A_i and B_i is described in Methods.

Results

Target protein expression at steady state exhibits oscillations

The steady state is a relatively long-term behavior in the cells. By inspecting Equation 7, for sufficient p53 input pulses, let $i \to \infty$, we can obtain the steady state of target protein expression dynamics:

$$P_{ss}(\xi) = 1 + m_d + \frac{\alpha m_d}{\mu-\alpha}\frac{1-e^{-\mu(T-\Delta)}}{1-e^{-\mu T}}e^{-\mu\xi} - \frac{\mu m_d}{\mu-\alpha}\frac{1-e^{-\alpha(T-\Delta)}}{1-e^{-\alpha T}}e^{-\alpha\xi}, \xi \in [0,\Delta),$$

$$P_{ss}(\xi) = 1 - \frac{\alpha m_d}{\mu-\alpha}\frac{e^{\mu\Delta}-1}{1-e^{-\mu T}}e^{-\mu\xi} + \frac{\mu m_d}{\mu-\alpha}\frac{e^{\alpha\Delta}-1}{1-e^{-\alpha T}}e^{-\alpha\xi}, \quad \xi \in [\Delta, T]. \tag{8}$$

$P_{ss}(\xi)$ reaches the maximum and minimum at Δ, and 0 or T, respectively. Therefore, the steady state of target protein expression dynamics is repetitive and invariant oscillation. The maximal fold change or peak of the oscillations is

$$P_{ss,\max} = P_{ss}(\Delta) = 1 + m_d + \frac{\alpha m_d}{\mu - \alpha}\frac{e^{-\mu\Delta} - e^{-\mu T}}{1 - e^{-\mu T}} - \frac{\mu m_d}{\mu - \alpha}\frac{e^{-\alpha\Delta} - e^{-\alpha T}}{1 - e^{-\alpha T}} \quad (9)$$

And the minimal fold change or valley is

$$P_{ss,\min} = P_{ss}(0) = 1 + m_d + \frac{\alpha m_d}{\mu - \alpha}\frac{1 - e^{-\mu(T-\Delta)}}{1 - e^{-\mu T}} - \frac{\mu m_d}{\mu - \alpha}\frac{1 - e^{-\alpha(T-\Delta)}}{1 - e^{-\alpha T}} \quad (10)$$

For the given target protein, β and K_A are relatively fixed, the amplitude depends on the p53 dynamic parameters and the half-life of mRNA and protein. Apparently, the smaller the amplitude, the more stable the steady state becomes.

A Hill type equation can characterize the constant steady state

Let us inspect the characteristics of the steady state in the limit of α and μ. From Equation 9 and 10, when $\alpha T \ll 1, \mu T \ll 1$, there is

$$P_{ss,\max} = P_{ss,\min} = 1 + \gamma m_d = 1 + \frac{\Delta}{T}\frac{\beta A^n}{K_A^n + A^n} = \bar{m} \quad (11)$$

Equation 11 is the same as Equation 1, which is exactly the Hill type equation we found previously (26).

Similarly, when $\alpha T \ll 1, \mu\Delta \ll 1$, *i.e.* $\mu T \ll 1$, there are

$$P_{ss,\max} = P_{ss,\min} = 1 + \gamma m_d .$$

And when $\mu T \ll 1$, $\alpha\Delta \ll 1$, *i.e.* $\alpha T \ll 1$, there are

$$P_{ss,\max} = P_{ss,\min} = 1 + \gamma m_d .$$

When $\alpha\Delta \ll 1, \mu\Delta \ll 1$, *i.e.* $\alpha T \ll 1, \mu T \ll 1$, there are

$$P_{ss,\max} = P_{ss,\min} = 1 + m_d \quad (12)$$

Equation 12 is the same as the classical Hill equation that governs the steady state mRNA fold change under sustained p53 dynamics (26). Therefore, in the limit of very longer or shorter half-life of mRNA and protein, the oscillations contract into a constant line that is very stable.

For several target genes of p53, Table 1 lists the observations for α, μ, β and K_A. There are only 3 genes with complete data. α and μ determine trajectory of target protein expression dynamics (23). The four scenarios discussed above correspond to the four sets of protein expression dynamics defined in Literature 23.

Table 1 . The observed values of α, μ, β and K_A corresponding to genes

Gene	Function	mRNA decay rate constant	Protein degradation rate constant	Dissociation constant	Maximal mRNA fold change	Refs
PUMA	Apoptosis	0.73 h^{-1}	0.056 h^{-1}	260 nM	Unknown	23,28,33
MDM2	Feedback inhibition	0.27 h^{-1}	0.7916 h^{-1}	12.3 nM	$2^{2.78} = 6.87$	23,24,28,33
p21	Cell cycle arrest	0.265 h^{-1}	0.2546 h^{-1}	4.9 nM	$2^{3.90} = 14.93$	23,24,28,33
BAX	Apoptosis	0.0188 h^{-1}	0.0262 h^{-1}	73 nM	$2^{1.23} = 2.35$	23,24,30,33
GADD45A	DNA repair	Unknown	Unknown	7.7 nM	$2^{3.19} = 9.13$	33,24

The average fold change of mRNA and protein expression are the same at steady state

The average fold change of target protein over the i-th period can be calculated by

$$\bar{P}_i = \frac{1}{T}\int_0^T P_i(\xi_i)d\xi_i = 1 + \frac{\Delta}{T}m_d + \frac{\alpha m_d}{(\mu-\alpha)\mu T}(e^{\mu\Delta}-1)e^{-i\mu T} - \frac{\mu m_d}{(\mu-\alpha)\alpha T}(e^{\alpha\Delta}-1)e^{-i\alpha T}. \tag{13}$$

Let $i \to \infty$, there is the average steady state

$$\bar{P} = 1 + \frac{\Delta}{T}m_d = 1 + \frac{\Delta}{T}\frac{\beta A^n}{K_A^n + A^n} \tag{14}$$

As we expect

$$\bar{P} = \bar{m} \tag{15}$$

Thereby we proved that at steady state the average fold change of mRNA and protein are the same. This result also holds true for the average levels of different proteins and corresponding mRNA within a single cell, which consistent with the observed results (17). Similarly, for the average of the same protein over the cell population, the average expression levels of mRNA and protein also are the same.

Next, let us take target protein BAX as an example to examine the predictability of the obtained Hill type equation for protein expression. The observed fold change of the BAX protein 24 hour after stimulation is 1.28 (Reference 23, Fig4C), and the BAX mRNA is $2^{0.78} = 1.72$ (24). For any cell, assuming that Δ/T remains unchanged, the average BAX expression over the cell population can be calculated by the Hill type equation (26)

$$\bar{P}_{BAX} = \bar{m}_{BAX} = 1 + 2^{1.23} \cdot 0.37 \cdot \frac{60^{1.8}}{73^{1.8} + 60^{1.8}} = 1.35 \tag{16}$$

The values of β and K_A are given by Table 1.

Longer mRNA or protein half-life determine the relaxation time of target protein expression dynamics

We obtained previously the relaxation time of mRNA dynamics (26). Similarly, the time of protein dynamic trajectory relaxation to the average steady state can be calculated by (25,26,31)

$$\tau_{pulsed} = \frac{T\sum_{i=1}^{\infty}\left|\bar{P}_i - \bar{P}\right|}{\bar{P} - 1} = \frac{T}{\Delta\left|\mu-\alpha\right|}\left|\frac{\alpha}{\mu}\frac{e^{\mu\Delta}-1}{e^{\mu T}-1} - \frac{\mu}{\alpha}\frac{e^{\alpha\Delta}-1}{e^{\alpha T}-1}\right|. \tag{17}$$

When $\alpha T \ll 1$, $\mu T \ll 1$, τ_{pulsed} can be expanded in the Taylor series:

$$\tau_{pulsed} \approx \frac{1}{\alpha} + \frac{1}{\mu} - \frac{T}{2}(1-\gamma). \tag{18}$$

Thus, the number of p53 pulses required to reach the average steady state is

$$\frac{\tau_{pulsed}}{T} \approx \frac{1}{\alpha T} + \frac{1}{\mu T} - \frac{1}{2}(1-\gamma) \tag{19}$$

For sustained p53 input dynamics, $\Delta = T$, thereby

$$\tau_{sus} \approx \frac{1}{\alpha} + \frac{1}{\mu} \tag{20}$$

Therefore, compared to the sustained input, the pulsed p53 dynamics makes the target protein dynamics reach the maximum more quickly. That is to say, if the half-life of mRNA and protein and $T(1-\gamma)/2$ have the same order of magnitude, the oscillatory p53 input enhances the sensitivity of protein expression. On the other hands, as shown in Table 1, the gene *BAX* with longer half-life of mRNA and protein must require multiply p53 pulses input to reach it steady state, which not only provides sufficient

time for DNA repair, but also accumulate enough expression levels required for triggering apoptosis.

The Index of target protein accumulation under multiply p53 pulses

In order to understand the degree of target protein expression accumulation in response to multiply p53 input pulses, the index of protein accumulation can be defined as

$$\lambda = \frac{P_{ss,\max}}{P_{1,\max}}, \tag{21}$$

which is the ratio of the maximal protein fold change at steady state to the maximal protein fold change during 1-th pulse.

$P_{ss,\max}$ is given by Equation 9, and

$$P_{1,\max} = P_1(\Delta) = 1 + m_d + \frac{\alpha m_d}{\mu - \alpha} e^{-\mu\Delta} - \frac{\mu m_d}{\mu - \alpha} e^{-\alpha\Delta}. \tag{22}$$

When $\alpha T \ll 1$, $\mu T \ll 1$, there are

$$P_{ss,\max} = 1 + \gamma m_d, \quad P_{1,\max} = 1.$$

Thus

$$\lambda = 1 + \gamma m_d. \tag{23}$$

Similarly, when $\alpha T \ll 1$, $\mu\Delta \ll 1$, *i.e.* $\mu T \ll 1$, there is

$$\lambda = 1 + \gamma m_d.$$

When $\alpha\Delta \ll 1$, *i.e.* $\alpha T \ll 1$, and $\mu T \ll 1$, there is

$$\lambda = 1 + \gamma m_d.$$

When $\alpha\Delta \ll 1$, i.e. $\alpha T \ll 1$and $\mu\Delta \ll 1$,*i.e.* $\mu T \ll 1$, there are

$$P_{ss,\max} = 1 + m_d, \quad P_{1,\max} = 1 + m_d.$$

Thus $\lambda = 1$ (24)

Therefore, the proteins with longer half-life have higher accumulation; however, the proteins with shorter half-life have lower accumulation.

The first three cases discussed above correspond broadly to the rising expression dynamics observed in Literature 23, and the last one is the oscillatory expression dynamics. In other words, the rising expressions have higher accumulation, and the pulsing expressions have lower accumulation.

Peak to valley ratio can measure the stability of mRNA and protein expression dynamics under multiply p53 pulses

From Equation 6, it can be seen that mRNA expression dynamics at steady state also exhibits the oscillatory behaviors. Let $i \to \infty$, the steady state for mRNA dynamics can be written as

$$\begin{aligned} m_{ss}(\xi) &= 1 + m_d \left(1 - \frac{e^{-\alpha\xi}(1 - e^{-\alpha(T-\Delta)})}{1 - e^{-\alpha T}}\right), \qquad \xi \in [0, \Delta), \\ m_{ss}(\xi) &= 1 + m_d \frac{(e^{\alpha\Delta} - 1)}{1 - e^{-\alpha T}} e^{-\alpha\xi}, \qquad \xi \in [\Delta, T]. \end{aligned} \tag{25}$$

The peak is

$$m_{ss,\max} = m_{ss}(\Delta) = 1 + m_d \frac{1 - e^{-\alpha\Delta}}{1 - e^{-\alpha T}}. \tag{26}$$

And, the valley is

$$m_{ss,\min} = m_{ss}(T) = 1 + m_d \frac{e^{\alpha\Delta} - 1}{e^{\alpha T} - 1}. \quad (27)$$

For the given gene, the peak to valley ratio for target mRNA expression can be defined as

$$\rho_m = \frac{m_{ss,\max}}{m_{ss,\min}}. \quad (28)$$

The closer ρ_m is to 1, the more stable mRNA expression dynamics. In particular, when $\alpha T \ll 1$, there are

$$m_{ss,\max} = m_{ss,\min} = 1 + \gamma m_d. \quad (29)$$

Thus

$$\rho_m = 1. \quad (30)$$

Therefore, the longer mRNA half-life is, the more stable mRNA expression dynamics is. Similarly, when $\alpha\Delta \ll 1$, *i.e.* $\alpha T \ll 1$, there are

$$m_{ss,\max} = 1 + m_d, \quad m_{ss,\min} = 1. \quad (31)$$

Thereby

$$\rho_m = 1 + m_d \quad (32)$$

Therefore, target mRNA dynamics with shorter half-life is unstable.

Similarly, we can define the peak to valley ratio that characterizes the stability of target protein expression dynamics as

$$\rho_p = \frac{P_{ss,\max}}{P_{ss,\min}}. \quad (33)$$

As can be seen from the previous discussion, for the 4 extreme cases of α and μ, $\rho_p = 1$. However, for the limit of very short mRNA half-life, $\rho_m = 1 + m_d > 1$. Therefore, the translation of mRNA enhances the stability. By calculating the peak to valley ratio of each target gene, we can compare the stability of expression of different genes.

The regulatory principle of protein expression dynamics under p53 pulsing

Let us examine the Hill type equation, for very high binding affinity, *i.e.* $K_A \ll A$, there is

$$\overline{P} = 1 + \frac{\Delta}{T} \frac{\beta A^n}{K_A^n + A^n} \approx 1 + \beta \frac{\Delta}{T} \quad . \quad (34)$$

Therefore, target protein expression with high p53 DNA binding affinity is insensitive to the amplitude. Only through the duration and frequency can fine tuning of protein expression be achieved. In addition, it is important to note that $\Delta/T < 1$, namely, the frequency must be less than $1/\Delta$. Thus a situation where too high frequency causes protein expression to decrease is avoided. Furthermore, the results from experiment show that gene expression increases proportionally with the TF frequency (32), which agrees with prediction from the equation.

The maximal mRNA expression β is determined by binding affinity. The higher the affinity is, the greater the maximal mRNA level is (24,33). For example, as shown in Table 1, for genes *p21*, *GADD45*, *MDM2*, *BAX*, K_A are 4.9 nM, 7.7 nM, 12.3 nM, 73 nM (33), and the corresponding β are 14.93, 9.13, 6.87, 2.35, respectively (24). Therefore, the maximal gene expression level is mainly determined by TF DNA binding affinity, while the stability of gene expression is determined by the half-life of

mRNA and protein. Therefore, how to regulate binding affinity has become a key issue in the gene expression control. John B. Gurdon *et al.* can greatly improve the binding affinity by extending the residence time of TF target (34), allowing the duration and frequency of TF dynamics to finely tune gene expression.

Discussion

TF is just a node of a complex network, which is like a bridge connecting the cellular signal transduction networks and the transcription networks. Taking the famous TF p53 as an example, I proved that the average steady state of mRNA and protein expression fold change driven by p53 pulsing are the same. The obtained Hill type equation may be also expected by previous methods (3,11,13,16). By this equation, not only do we understand the regulation principle of gene expression clearly, but also we can put this equation into practice. In addition, we gave two quantitative indicators to determine the degree of accumulation and stability of protein expression under multiply TF input pulses.

For the activated gene expression in response to stimulation, this obtained equation does not depend on the rate constant of translation and degradation. However, for the basal gene expression, Equation 2 and 3 become

$$\frac{dmRNA(t)}{dt} = \beta_0' - \alpha mRNA(t) \tag{35}$$

$$\frac{dprotein(t)}{dt} = \sigma mRNA(t) - \mu protein(t) \tag{36}$$

It must be noted that $mRNA(t)$ and $protein(t)$ cannot be dimensionless here to eliminate the parameters σ and μ. So the steady state equations for basal gene expression are

$$\beta_0' - \alpha mRNA(t) = 0, \tag{37}$$

$$\sigma mRNA(t) - \mu protein(t) = 0. \tag{38}$$

i.e.

$$mRNA_{ss} = \frac{\beta_0'}{\alpha}, \quad protein_{ss} = \frac{\sigma}{\mu} mRNA_{ss}, \tag{39}$$

where $mRNA_{ss}$ and $protein_{ss}$ represent the concentration of basal mRNA and protein at steady state, respectively. All symbols and definitions are shown in Table 2. Therefore, for the basal gene expression system, the absolute levels of protein expression at steady state are determined by the rate constant of translation and degradation, and mRNA level, respectively, which agree with the previous results (9,10,11,12,16).

Gene expression dynamics is random and complex. The Hill type equation reflects the deeply simplicity principle about gene expression dynamics. The history of mechanics, physics and chemistry tells that the essence of nature is simplicity. Complex phenomena evolved from the principle of simplicity. According to the regulation principle revealed by this equation, we can control gene expression from the upstream of genetic information flow described by central dogma. Only four parameters need to be adjusted, namely, the duration, frequency, and amplitude of TF, and binding affinity, we can control the steady state levels of mRNA and protein expression at the same time.

Owing to the fact that the levels of mRNA and protein expression at steady state are equal, over the years, a large amount of transcriptome data has been accumulated, therefore, by this equation we may predict the proteome simply by transcriptome.

The classical Hill equation applies when the TF dynamics is a constant. When $\Delta = T$, the equation is reduced to the classical Hill equation. Therefore, this equation broadens the application of classic biochemical theory. Through the derivation process, this generalized equation, which reveals the regulation mechanism of gene expression, may be applied to any gene expression pathway driven by TF dynamics.

Hill wrote: "My object was rather to see whether an equation of this type can satisfy all the observations, than to base any direct physical meaning on n and K (35-36)." More than one hundred years later, we have not forgotten his caveat (36).

Table2. Variable and parameter definitions

Symbol	definition	Units
A	Amplitude of p53 pulses, or sustained constant p53 dynamics	nM
n	Hill coefficient	-
Δ	Duration of p53 pulses	h
T	Period of p53 pulses	h
K_A	Dissociation constant	nM
$P(t)$	Protein fold change	-
$mRNA(t)$	mRNA concentration	nM
$protein(t)$	Protein concentration	nM
$m_i(t)$	mRNA fold change during the i-th pulse	-
$m(t)$	mRNA fold change	-
$m_{ss}(\xi)$	mRNA fold change at steady state	-
$m_{ss,\max}$	Maximal mRNA fold change at steady state	-
$m_{ss,\min}$	Minimal mRNA fold change at steady state	-
m_d	Net mRNA fold change at steady state under sustained p53	-
$P_{ss}(\xi)$	Protein fold change at steady state	-
$P_{ss,\max}$	Maximal protein fold change at steady state	-
$P_{ss,\min}$	Minimal protein fold change at steady state	-
$\overline{P_i}$	Average protein fold change over the i-th pulse	-
$\overline{m}$	Average mRNA fold change at steady state	-
$\overline{P}$	Average protein fold change at steady state	-
m_0	Basal mRNA concentration	nM
p_0	Basal protein concentration	nM
$mRNA_{ss}$	Basal mRNA concentration at steady state	nM
$protein_{ss}$	Basal protein concentration at steady state	nM
β	Maximal mRNA fold change	-
β'	Maximal transcription rate	nMh^{-1}

β_0'	Basal transcription rate	nMh^{-1}
τ_{sus}	Relaxation time to steady state under sustained p53	h
τ_{pulsed}	Relaxation time to steady state under p53 pulsing	h
α	Rate constant of mRNA decay	h^{-1}
σ	Rate constant of protein translation	h^{-1}
μ	Rate constant of protein degradation	h^{-1}
$P_{1,\max}$	Maximal protein fold change during the 1-th pulse	-
λ	Index of protein accumulation	-
ρ	Peak to valley ratio	-

Methods

We use a minimal model for gene expression dynamics under p53 pulsing. In this way, we can find the analytical solution of this model. The solution is given in the Introduction. Here a detailed derivation of A_i and B_i is given. From Equation 7, considering

$$P_{i-1}(T) = P_i(0)\,,\quad P_i(\Delta+) = P_i(\Delta-)\,, \tag{40}$$

thus there are

$$A_i - B_i = \frac{\alpha m_d}{\mu-\alpha} e^{\mu\Delta}\,,\quad B_{i-1}e^{-\mu T} - B_i = \frac{\alpha m_d}{\mu-\alpha}(e^{\mu\Delta}-1)\,. \tag{41}$$

Since $P_1(0) = 1$, we can derive:

$$A_1 = \frac{\alpha m_d}{\mu-\alpha}\,,\quad B_1 = -\frac{\alpha m_a}{\mu-\alpha}(e^{\mu\Delta}-1)\,. \tag{42}$$

Then from the difference equation 41, we can obtain:

$$B_i = B_1 e^{-(i-1)\mu T} - \frac{\alpha m_d}{\mu-\alpha}\frac{(e^{\mu\Delta}-1)(1-e^{-(i-1)\mu T}}{1-e^{-\mu T}}\,. \tag{43}$$

Substituting B_i into Equation 41 yields A_i.

Author contributions

Xiaomin Shi conceived and performed the study and wrote the article.

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

References

1. Crick F. (1958) On protein synthesis. *Symp. Soc. Exp. Bio*,139-163.
2. Crick F. (1970) Central dogma of molecular biology. *Nature* 227, 561-563.
3. Buccitelli,C., Selbach, M. (2020) mRNAs, proteins and the emerging principles of gene expressions control. *Nat. Rev. Genet*. 21, 630-644.

4. Howeroft, T., K., *et al.* (2003) Distinct transcriptional pathways regulate basal and activated major hislocompatibility complex class I expression. *Mol. Cellular Biol.* 23, 3377-3391.
5. Richard, N.,D., Walder, J. A., Maurer, R. A. (1989) A protein kinase inhibitor gene reduces both basal and multi hormone-stimulated prolactim gene transcription. *J. Biol. Chem* 264, 431-436.
6. Willy, P.,J., Kobayashi, R., Kadonaga, J.T. (2000) A basal transcription factor that activates or represses transcription. *Science* 290, 982-984.
7. Ranish, J.,A., Hahn, S. (1996)Transcription: basal factors and activation. *Current Opinion Genet. Develop* . 6, 151-158.
8. Parvin, J., P., Timers, H.,T.,M., Sharp, P.,A. (1992) Promoter specificity of basal transcrition factors. *Cell* 68, 1135-1144.
9. Schwanhausser, B., Busse, D., Li, N., *et al.* (2011) Global quanlification of mammlian gene expression control. *Nature* 437, 337.
10. Hausser, J., Mayo, A., Keren, L., Alon, U. (2019) Central dogma rates and trade-off between precision and economy in gene expression. *Nat. Commun.* 10, 68.
11. Liu, Y., Beyer, A., Aebersold, R. (2016) On the dependency of cellular protein levels on mRNA abundance. *Cell* 165:535.
12. Jovanovic, M., Rooney, M., Mertins, P., *et al.* (2015) Dynamic profiling of the protein life cycle in resonse to pathogens. *Science* 347, 1239038.
13. Smgh, R., Letai, A., Sarosiek, K. (2019) Regulation of apoptosis in health and disease:the balancing act of BCL-2 family proteins. *Nat. Rev. Mol. Cell Bio*. 20, 175-193.
14. Micco, R. D., KHzhanovsky, V., Baker, D. (2021) Cellular senescence in ageing: from mechanism to therapeutic opporturnities. *Nat. Rev. Mol. Cell Bio.* 22, 75-95.
15. Vogel, C., Marcotte, E. M. (2013) Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Gnet.* 13, 227-232.
16. Li ,J.,J., Biggin, M.,D. (2015) Statistics requantitates the centrl dogma. *Science* 347, 1066.
17. Grun, D., Kirchner, M., *et al.* (2014) Conservation of mRNA and protein expression during development of C. elegans. *Cell Reports* 6, 565-577.
18. Purvis, J. E., Lahav, G. (2013) Encoding and decoding cellular information through signaling dynamics, *Cell* **152**, 945-56.
19. Levine, J. H., Lin, Y., Elowitz, M. B. (2013) Functional roles of pulsing in genetic circuits. *Science* **342**:1193-200.
20. Itaru, I., Akihiro,I., Yukiko, H., *et al.* (2013) Oscillatory control of factors determining multipotency and fate in mouse neural progenitors. *Science* 342, 1203-1208.
21. Lahav, G., Rosenfeld, N., Sigal, A., Geva-Zatorsky, N., Levine, A. J., Elowitz, M. B., Alon, U. (2004) Dynamics of the p53-Mdm2 feedback loop in individual cells, *Nat Genet.* **36**, 147-50.
22. Purvis, J. E., Karhohs, K. W., Mock, C., Batchelor, E., Loewer, A., Lahav, G. (2012) p53 dynamics control cell fate. *Science* **336**, 1440-4.
23. Hanson, R. L., Porter, J. R., Batcherlor, E. (2019) Protein stability of p53 targets determines their temporal expression dynamics in response to p53 pulsing. *J. Cell Biol*, 218, 1282-1297.
24. Hafner, A., Stewart-Ornstein, J., Purvis, J. E., Forrester, W. C., Bulyk, M. L., Lahav, G. (2017) p53 pulses lead to distinct patterns of gene expression albeit similar DNA-binding dynamics, *Nat Struct Mol Biol.* **24**, 840-847.
25. Shi, X., Reimers, J. R. (2018) Understanding non-linear effects from Hill-type dynamics with application to decoding of p53 signaling. *Sci Rep.* **8**, 2147.
26. Shi, X. (20210 Hill type equation can predict target gene expression driven by p53 pulsing. *FEBS Open Bio*. **11**, 1799-808.
27. Hatner, A., Reyos, J., *et al.* (2020) Quantifying the central dogma in the p53 pathway in live single cells. *Cell Syst.* 10, 495-505.
28. Porter, J. R., Fisher, B. E., Batchelor, E. (2016) p53 Pulses Diversify Target Gene Expression Dynamics in an mRNA Half-Life-Dependent Manner and Delineate Co-regulated Target Gene Subnetworks, *Cell Syst.* **2**, 272-282.
29. Banwarth Kuhn, M., Smdi, S. (2020) How and way to build a mathamaticle model: A case study using prion aggregation. *J. Biol. Chem*. 295, 5022-5035.

30. Friedel, C.,C., Dolken, L., Ruzsics, Z., *et al.* (2009) Conserved principles of mammalian transcriptional regulation revealed by RNA half-life. *Nucleic. Acids Res*. 37, e115.
31. Salazar, C., Politi, A. Z., Hofer, T. (2008) Decoding of calcium oscillations by phosphorylation cycles: analytic results. *Biophys J*. **94**, 1203-15.
32. Cai, L., Dalai, C.,K., Elowitz, M., B. (2008) Frequency–modulated nuclear localization bursts coordinate gene regulation. *Nature* 455, 485-490.
33. Weinberg, R. L., Veprintsev, D. B., Bycroft, M., Fersht, A. R. (2005) Comparative binding of p53 to its promoter and DNA recognition elements, *J Mol Biol*, **348**, 589-96.
34. Gurdon, J.,B., Javed, K., Vodnala, M., Garrett, N.(2020) Long-term association of a transcription factor with its chromatin binding site can stabilize gene expression and cell fate commitment. *Proc Natl Acad Sci U S A.* **117**, 15075-84.
35. Hill, A. V. (1910) The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves, *J Physiology (London)* **40**, iv-vii
36. Colquhoun, D. (2006) The quantitative analysis of drug-receptor interactions: a short history. *Trends Pharmacol Sci.* **27**, 149-57.