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

Correspondence: shixm1@vip.126.com

Abstract
The central dogma indicates the basic direction of gene expression pathways. For activated gene expression, the quantitative relationship between various links from the binding of transcription factor (TF) to DNA to protein synthesis remains unclear and debated. There is consensus that at a steady state, protein levels are largely determined by the mRNA level. How can we find this steady state? Taking p53 as an example, based on the previously discovered Hill-type equation that characterizes mRNA expression under p53 pulsing, I proved that the same equation can be used to describe the average steady state of target protein expression. Therefore, at the steady state, the average fold changes of mRNA and protein expression under TF pulsing are the same. This consensus has been successfully demonstrated. For the gene *BAX* of the p53 target, the observed fold changes in mRNA and protein were 1.72, 1.28, respectively; the mRNA and protein expressions calculated using the Hill-type equation were 1.35. Therefore, using this equation, we are not only able to fine-tune gene expression, but also to predict the proteome from the transcriptome. Furthermore, by introducing two quantitative indicators, we can understand the degree of accumulation and stability of protein expression.

Keywords: Hill equation, p53 target protein expression, steady state, stability of gene expression, protein accumulation, transcription factor pulsing

Introduction

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

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

Here, steady state can be defined as a relatively long-term process experienced by cells (11). Therefore, both cell proliferation and differentiation can be considered steady states (11). Apoptosis and senescence also occur in a steady state (13, 14). The steady state achieved after stimulation differs from that before perturbation (15).

For the steady state before stimulation in dendritic cells, mRNA levels explain 68% of the change in protein expression, and for the approximate steady state after stimulation with lipopolysaccharide, mRNA levels explain 90% of the change in protein expression (12, 16). In addition, during the development of *C elegans*, the fold changes in mRNA and protein expression were almost identical. The transcript fold change was 2.02, and the proteins was 2.05 (17). Therefore, we can speculate that the fold-changes in mRNA and protein expression in the steady state are equal. Could we theoretically determine this special steady state of protein expression? Which factors determine protein expression levels in steady state?

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

The tumor suppressor p53 is the most extensively studied TF. In response to DNA damage, p53 concentrations exhibit a series of pulses of fixed duration, frequency, and amplitude, whereas 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 an altered 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 10 h or 24 h after stimulation that the fold change in mRNA was 2.2 (10 h)(23) or 2.0 (24 h)(24), while the protein expression was 1.8 (24 h), which 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 governing target protein expression at a

steady state.

Following the genetic information flow specified by the central dogma, we obtained a modified Hill equation to characterize the average p53 DNA-binding probability (25). I also found a Hill-type equation that could predict the fold-change in 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 is the period, A is the amplitude, β is the maximal fold change in mRNA expression, and K_A is the dissociation constant.

Here, I will complete the last step in describing the central dogma, from mRNA to protein. In response to pulsed p53 input dynamics, p21 mRNA dynamics show pulsing expression; however, p21 protein dynamics exhibit rising expression (27). The half–life of mRNA and protein determines the stability of expression dynamics (23, 24, 28). However, the relationship between steady-state mRNA and protein expression levels remains unclear. Therefore, I tried to prove that the average steady-state fold changes in mRNA and protein expression under p53 pulsing are equal, *i.e.*, the average fold change of protein at steady state is equal to $\bar{m}$.

Methods

Mathematical model of p53 target protein expression dynamics and its analytical solution

To achieve this goal, we must develop a useful and accurate mathematical model for gene expression under pulsed TF dynamics (16, 29). Some models do not include basal transcription term,(30) but some early models provided necessary terms for environmental stimuli to activate gene expression.(31) 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 square wave 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 denotes the rate constants of translation and degradation, respectively. The initial protein concentration was determined using $protein(0) = p_0$.

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

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

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

where $m(t)$ and $P(t)$ represent the fold changes in mRNA and protein expression, respectively, and $\beta = \beta' / \beta_0'$ denotes the maximal fold change in transcription. It seems that I have never seen gene expression levels appear in a dimensionless form in the differential equation. This is a crucial step. Because from equation 5, we already know that the steady state levels of mRNA and target protein expression are the same. Because Equation 4 has been solved, the solution from Reference (26) is rewritten as

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

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

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

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.

Solving Equation (5) is challenging. Reference (32) provides a method for solving such equation. Assuming $\alpha \neq \mu$, we can obtain analytical solutions for target protein expression in response to p53 pulsing as follows:

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

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

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

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. Here, a detailed derivation of A_i and B_i is presented. From Equation. (7), we consider the following:

$$P_{i-1}(T) = P_i(0), \quad P_i(\Delta+) = P_i(\Delta-), \tag{8}$$

thus there are

$$A_i - B_i = \frac{\alpha m_d}{\mu - \alpha} e^{\mu\Delta}, \quad B_{i-1} e^{-\mu T} - B_i = \frac{\alpha m_d}{\mu - \alpha}(e^{\mu\Delta} - 1). \tag{9}$$

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

$$A_1 = \frac{\alpha m_d}{\mu - \alpha}, \quad B_1 = -\frac{\alpha m_a}{\mu - \alpha}(e^{\mu\Delta} - 1). \tag{10}$$

Then from iterative equation 41, we can obtain:

$$B_i = B_1 e^{-(i-1)\mu T} - \frac{\alpha m_d}{\mu - \alpha} \frac{(e^{\mu\Delta} - 1)(1 - e^{-(i-1)\mu T})}{1 - e^{-\mu T}}. \tag{11}$$

Substituting B_i into Equation 9 yields A_i. A detail solving process for Equation 5 is in Supplementary Material.

Results

Target protein expression at steady state exhibits oscillations

The steady state is the relatively long-term behavior of the cells. Neither Reference (32) nor (25) conducted this steady-state analysis. Reference (26) began this type of analysis. By inspecting Equation 7, for sufficient p53 input pulses, let $i \to \infty$, we can obtain the steady state of the target protein expression dynamics:

$$\begin{aligned} P_{ss}(\xi) &= 1 + m_d + \frac{\alpha m_d}{\mu - \alpha} \frac{1 - e^{-\mu(T-\Delta)}}{1 - e^{-\mu T}} e^{-\mu\xi} - \frac{\mu m_d}{\mu - \alpha} \frac{1 - e^{-\alpha(T-\Delta)}}{1 - e^{-\alpha T}} e^{-\alpha\xi}, \xi \in [0, \Delta), \\ P_{ss}(\xi) &= 1 - \frac{\alpha m_d}{\mu - \alpha} \frac{e^{\mu\Delta} - 1}{1 - e^{-\mu T}} e^{-\mu\xi} + \frac{\mu m_d}{\mu - \alpha} \frac{e^{\alpha\Delta} - 1}{1 - e^{-\alpha T}} e^{-\alpha\xi}, \qquad \xi \in [\Delta, T]. \end{aligned} \tag{12}$$

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

$$P_{ss,\max} = P_{ss}(\Delta) = 1 + m_d + \frac{\alpha m_d}{\mu - \alpha} \frac{e^{-\mu\Delta} - e^{-\mu T}}{1 - e^{-\mu T}} - \frac{\mu m_d}{\mu - \alpha} \frac{e^{-\alpha\Delta} - e^{-\alpha T}}{1 - e^{-\alpha T}} \tag{13}$$

And the minimal fold change or valley is

$$P_{ss,\min} = P_{ss}(0) = 1 + m_d + \frac{\alpha m_d}{\mu - \alpha} \frac{1 - e^{-\mu(T-\Delta)}}{1 - e^{-\mu T}} - \frac{\mu m_d}{\mu - \alpha} \frac{1 - e^{-\alpha(T-\Delta)}}{1 - e^{-\alpha T}} \tag{14}$$

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

A Hill type equation can characterize the constant steady state

Let us now examine the characteristics of the steady state within the limits of α and μ. From Equation 9 and 10, when $\alpha T \ll 1, \mu T \ll 1$ was applied,

$$P_{ss,\max} = P_{ss,\min} = 1 + \gamma m_d = 1 + \frac{\Delta}{T} \frac{\beta A^n}{K_A^n + A^n} = \bar{m} \tag{15}$$

Equation 15 is the same as Equation 1, which is exactly the Hill-type equation found previously (26).

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

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

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

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

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

$$P_{ss,\max} = 1 + m_d , \quad P_{ss,\min} = 1 . \tag{16}$$

Equation 16 is the same as the classical Hill equation, which governs the steady-state mRNA fold-change under sustained p53 dynamics (26). Therefore, in the limit of a very long or short half-life for mRNA and protein, 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 . Only 3 genes had complete data. α and μ determine the trajectory of target protein expression dynamics (23). The four scenarios discussed above correspond to the four sets of protein expression dynamics defined in reference (23).

As shown in Table 1, the mRNA decay rate constants of *PUMA*, *MDM2*, and *p21* are from (28), and the mRNA half-life of *BAX* is 38.77 h (33), thus, the mRNA decay rate is $\alpha = \ln 2/38.77 \approx 0.018\ \mathrm{h}^{-1}$ (23). The protein degradation rate constants of *PUMA*, *MDM2*, *p21*, and *BAX* are obtained from (23). The dissociation constants of *PUMA*, *MDM2*, *p21*, *BAX* ,and *GADD45A* are from (34). For maximal mRNA expressions of *p21*(*CDKN1A*), *GADD45A*, *MDM2*, and *BAX* , log (maximal mRNA fc is 3.9008(3h), 3.1949(3h), 2.7861(3h), and 1.2373(9 h)(24), thus the maximal fold changes are 14.937, 9.157, 6.898, and 2.358, respectively.

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
PUMA	Apoptosis	0.73 h^{-1}	0.056 h^{-1}	260 nM	Unknown
MDM2	Feedback inhibition	0.27 h^{-1}	0.7916 h^{-1}	12.3 nM	$2^{2.7861} = 6.898$
p21	Cell cycle arrest	0.265 h^{-1}	0.2546 h^{-1}	4.9 nM	$2^{3.9008} = 14.937$
BAX	Apoptosis	0.018 h^{-1}	0.0262 h^{-1}	73 nM	$2^{1.2373} = 2.358$
GADD45A	DNA repair	Unknown	Unknown	7.7 nM	$2^{3.1949} = 9.157$

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

The average fold change of the target protein over the i-th period can be calculated by (supplementary material) :

$$\bar{P}_i = \frac{1}{T}\int_0^T P_i(\xi_i)d\xi_i = 1 + \frac{\Delta}{T}m_d + \frac{\alpha m_d}{(\mu-\alpha)\mu T}(e^{\mu\Delta}-1)e^{-i\mu T} - \frac{\mu m_d}{(\mu-\alpha)\alpha T}(e^{\alpha\Delta}-1)e^{-i\alpha T}. \tag{17}$$

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

$$\bar{P} = 1 + \frac{\Delta}{T}m_d = 1 + \frac{\Delta}{T}\frac{\beta A^n}{K_A^n + A^n} \tag{18}$$

As we expect

$$\bar{P} = \bar{m} \tag{19}$$

Therefore, we proved that at steady state, the average fold changes in mRNA and protein were the same. This result also holds for the average levels of different

proteins and their corresponding mRNA within a single cell, which is consistent with the observed results (17). Similarly, for the average of the same proteins over the cell population, the average expression levels of mRNA and protein were also the same.

Next, we considered the target protein BAX as an example to examine the predictability of the Hill-type equation for protein expression. The observed fold-change of the BAX protein 24 h after stimulation was 1.28 (23), and the BAX mRNA was $2^{0.7802}=1.72$ (24). For any cell, assuming that Δ/T remains unchanged, the average BAX expression over the cell population can be calculated using Hill-type equation (26):

$$\bar{P}_{BAX}=\bar{m}_{BAX}=1+2^{1.23}\cdot 0.37\cdot\frac{60^{1.8}}{73^{1.8}+60^{1.8}}=1.35 \tag{20}$$

The values for β and K_A are listed in Table 1. Hill coefficient $n=1.8$ (34). $A=60$ nM (25). The duty cycle of p53 pulsing γ can be written as

$$\gamma=\frac{\text{average pulsed p53 concentration over a period}}{\text{average sustained p53 concentration over a period}}$$

$$=\frac{\frac{1}{T}\int_0^T p53(t)dt}{\frac{1}{T}\int_0^T Adt}=\frac{\int_0^T p53(t)dt}{AT}=\frac{\mathrm{AUC}\{p53(t)|_0^T\}}{\mathrm{AUC}\{A|_0^T\}},$$

where $\mathrm{AUC}\{p53(t)|_0^T\}$ represents area under the curve $p53(t)$. For general periodic function, $p53(t)$, during $[0,T]$ can be approximated by a square wave function $S(t)$, as long as

$$\mathrm{AUC}\{p53(t)|_0^T\}=\mathrm{AUC}\{S(t)|_0^T\}.$$

Thus

$$\gamma=\frac{\int_0^T S(t)dt}{AT}=\frac{A\Delta}{AT}=\frac{\Delta}{T}.$$

Here, γ is taken as 0.37 to minimize the error between the predicted p21 mRNA fold change and the observed fold change(26).

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

We have previously obtained the relaxation time for mRNA dynamics (26). Similarly, the time of protein dynamic trajectory relaxation to the average steady state can be calculated by (25, 26, 32)

$$\tau_{pulsed}=\frac{T\sum_{i=1}^{\infty}\left|\bar{P}_i-\bar{P}\right|}{\bar{P}-1}=\frac{T}{\Delta|\mu-\alpha|}\left|\frac{\alpha}{\mu}\frac{e^{\mu\Delta}-1}{e^{\mu T}-1}-\frac{\mu}{\alpha}\frac{e^{\alpha\Delta}-1}{e^{\alpha T}-1}\right|. \tag{21}$$

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

$$\tau_{pulsed}\approx\frac{1}{\alpha}+\frac{1}{\mu}-\frac{T}{2}(1-\gamma). \tag{22}$$

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

$$\frac{\tau_{pulsed}}{T}\approx\frac{1}{\alpha T}+\frac{1}{\mu T}-\frac{1}{2}(1-\gamma) \tag{23}$$

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

$$\tau_{sus} \approx \frac{1}{\alpha} + \frac{1}{\mu} \tag{24}$$

Therefore, compared to the sustained input, pulsed p53 dynamics cause the target protein dynamics to reach the maximum more quickly. In other words, if the half-lives of mRNA and protein and $T(1-\gamma)/2$ have the same order of magnitude, oscillatory p53 input enhances the sensitivity of protein expression. On the other hand, as shown in Table 1, the gene *BAX* with a longer half-life of mRNA and protein requires multiple p53 pulse inputs to reach a steady state, which not only provides sufficient time for DNA repair but also accumulates sufficient expression levels required for triggering apoptosis.

The index of target protein accumulation under multiple p53 pulses

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

$$\lambda = \frac{P_{ss,\max}}{P_{1,\max}}, \tag{25}$$

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

$$P_{1,\max} = P_1(\Delta) = 1 + m_d + \frac{\alpha m_d}{\mu - \alpha} e^{-\mu\Delta} - \frac{\mu m_d}{\mu - \alpha} e^{-\alpha\Delta}. \tag{26}$$

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

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

Thus

$$\lambda = 1 + \gamma m_d. \tag{27}$$

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

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

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

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

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

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

Thus $\lambda = 1$ (28)

Therefore, proteins with longer half-lives have a higher accumulation; however, proteins with shorter half-lives have a lower accumulation.

The first three cases discussed above broadly correspond to the rising expression dynamics observed in reference (23), and the last corresponds to 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 multiple p53 pulses

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

$$m_{ss}(\xi)=1+m_d\left(1-\frac{e^{-\alpha\xi}(1-e^{-\alpha(T-\Delta)})}{1-e^{-\alpha T}}\right), \qquad \xi\in[0,\Delta),$$
$$m_{ss}(\xi)=1+m_d\frac{(e^{\alpha\Delta}-1)}{1-e^{-\alpha T}}e^{-\alpha\xi}, \qquad \xi\in[\Delta,T]. \tag{29}$$

The peak is

$$m_{ss,\max}=m_{ss}(\Delta)=1+m_d\frac{1-e^{-\alpha\Delta}}{1-e^{-\alpha T}}. \tag{30}$$

And, the valley is

$$m_{ss,\min}=m_{ss}(T)=1+m_d\frac{e^{\alpha\Delta}-1}{e^{\alpha T}-1}. \tag{31}$$

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

$$\rho_m=\frac{m_{ss,\max}}{m_{ss,\min}}. \tag{32}$$

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

$$m_{ss,\max}=m_{ss,\min}=1+\gamma m_d. \tag{33}$$

Thus

$$\rho_m=1. \tag{34}$$

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

$$m_{ss,\max}=1+m_d, \quad m_{ss,\min}=1. \tag{35}$$

Thereby

$$\rho_m=1+m_d \tag{36}$$

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

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

$$\rho_p=\frac{P_{ss,\max}}{P_{ss,\min}}. \tag{37}$$

As can be seen from the previous discussion, for the three extreme cases of α and μ, and $\rho_p=1$. However, for the short half-lives of mRNA and target protein, from Equation 16, there is $\rho_p=1+m_d>1$, therefore, target protein is unstable in this case. When mRNA has short half-life, $\rho_m=1+m_d$, therefore, translation of mRNA does not improve its stability. By calculating the peak-to-valley ratio of each target gene, we can compare the stability of expression of different genes.

The index of mRNA accumulation can be defined as

$$\lambda_m=\frac{m_{ss,\max}}{m_{1,\max}} \tag{38}$$

From Equation 6, we obtain

$$m_{1,\max}=1+m_d(1-e^{-\alpha\Delta}). \tag{39}$$

For $\alpha T\ll 1$, there is $\lambda_m=1+\gamma m_d$, and for $\alpha T\gg 1$, there is $\lambda_m=1$. Therefore, stable mRNA has a high degree of accumulation.

As shown in Table 2, for the four combinations of half life, both long mRNA and

short protein half life, as well as short mRNA and long protein half life can produce stable proteins. Only short mRNA and short protein half life produces unstable proteins. Stable proteins always have high accumulation.

Table 2. The four combinations of mRNA and protein half lives determine protein stability

mRNA half life	Protein Half life	mRNA stability, ρ_m	Protein stability, ρ_p	mRNA accumulation, λ_m	Protein Accumulation, λ_p
Long	Long	Stable, 1	Stable, 1	High, $1+\gamma m_d$	High, $1+\gamma m_d$
Long	Short	Stable, 1	Stable, 1	High, $1+\gamma m_d$	High, $1+\gamma m_d$
Short	Long	Unstable, $1+m_d$	Stable, 1	Low, 1	High, $1+\gamma m_d$
Short	Short	Unstable, $1+m_d$	Unstable, $1+m_d$	Low, 1	Low, 1

The regulatory principle of protein expression dynamics under p53 pulsing

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

$$\bar{P} = 1 + \frac{\Delta}{T}\frac{\beta A^n}{K_A^n + A^n} \approx 1 + \beta\frac{\Delta}{T} \quad . \tag{40}$$

Therefore, target protein expression with high p53 DNA-binding affinity is insensitive to amplitude. Only through duration and frequency can fine-tuning of p21 protein expression be achieved, demonstrating the regulatory ability of duration and frequency beyond saturation (Fig 1). In addition, it is important to note that $\Delta/T < 1$, namely, the frequency must be less than that of the $1/\Delta$. Thus, a situation in which too high a frequency causes the protein expression to decrease is avoided. Furthermore, the results from the experiment showed that gene expression increased proportionally with TF frequency (35), which agrees with the prediction from the equation.

The maximum mRNA expression of β was determined using binding affinity. The higher the affinity, the greater the maximal mRNA level (24, 34). 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 (34), and the corresponding β are 14.937, 9.157, 6.898, 2.358, respectively (24). Therefore, the maximal gene expression level is mainly determined by the TF DNA-binding affinity, whereas the stability of gene expression is determined by the half-life of mRNA and protein. Therefore, regulation of binding affinity has become a key issue in gene expression control. Gurdon *et al.* can greatly improve the binding affinity by extending the residence time of the TF target (36), allowing the duration and frequency of TF dynamics to finely tune gene expression.

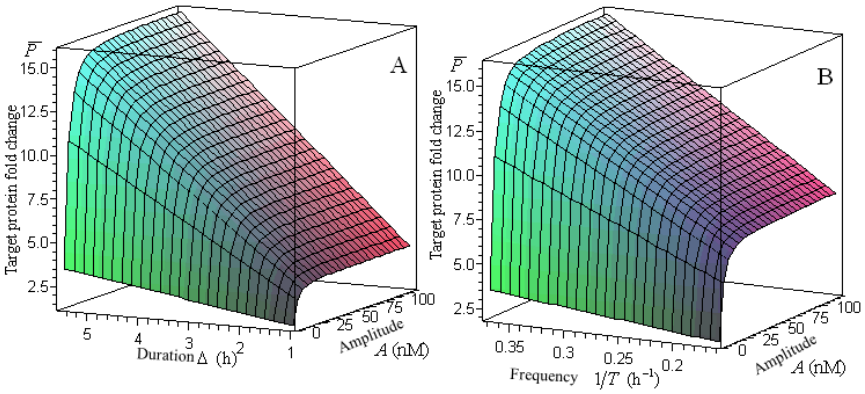

Figure 1. The roles of duration and frequency in p21 protein expression. Target protein expression nonlinearly reaches saturation with increased amplitude, however, it is linearly increased beyond saturation as duration and frequency increase, thus the cells achieved fine-tuning of gene expression using duration and frequency with high binding affinity.
(A). $T = 5.5$ h, $K_A = 4.9$ nM, $n = 1.8$. (B). $\Delta = 2.75$ h, $K_A = 4.9$ nM, $n = 1.8$.

The fold change in basal expression of mRNA and protein are equivalent at steady state

For basal gene expression, let $S(t) = 0$, Equation 3 and 4 that describing the fold changes of mRNA and protein expression become

$$\frac{dm(t)}{dt} = \alpha\left(1 - m(t)\right), \quad m(0) = 1, \tag{41}$$

$$\frac{dP(t)}{dt} = \mu(m(t) - P(t)), \; P(0) = 1. \tag{42}$$

Therefore, at steady state the fold change of basal mRNA and protein expression are

$$m_{basal} = P_{basal} = 1, \tag{43}$$

where m_{basal} and P_{basal} are the fold change of basal mRNA and protein expression, respectively. Let $\Delta = 0$ or $A = 0$ in Hill-type equation 18, we can also obtain Equation 41.

Similarly, for basal gene expression, let $S(t) = 0$, Equation 1 and 2 that describing the concentrations of mRNA and protein expression become

$$\frac{dmRNA(t)}{dt} = \beta_0' - \alpha mRNA(t), \tag{44}$$

$$\frac{dprotein(t)}{dt} = \sigma mRNA(t) - \mu protein(t). \tag{45}$$

The steady state is

$$mRNA_{ss} = \frac{\beta_0'}{\alpha}, \; protein_{ss} = \frac{\sigma}{\mu} mRNA_{ss} = \frac{\beta_0'\sigma}{\alpha\mu}, \tag{46}$$

where $mRNA_{ss}$ and $protein_{ss}$ represent the concentration of basal mRNA and protein at a steady state, respectively. All symbols and definitions used in this article are listed in Table 3. 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 levels, respectively. For any given gene, it must be point out that the above system only has one steady state.

The fold changes of basal mRNA and protein expression at steady state are

$$P_{basal} = \frac{protein_{ss}}{p_0} = \frac{mRNA_{ss}}{m_0}\frac{\sigma m_0}{\mu p_0} = m_{basal} = 1, \tag{47}$$

which is same as Equation 41.

Discussion

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

multiple TF input pulses.

For repressed gene expression, It was also proved that the fold changes of mRNA and target protein expression are the same at steady state, and a Hill-type equation with negative coefficient can describe the fold change of target gene expression (37)(Full text has not been submitted). Therefore, whether for basal and activated or repressed expression, we may reach such a consensus that the fold changes of mRNA and protein are the same at steady state and can be characterized by the Hill-type equation.

The experimental data was observed under that experimenters were not aware that the fold changes of mRNA and protein are the same at the steady state. After completing theoretical proof, I tried to search for the experimental data to support my findings. This just confirms what Einstein said that "It is the theory which decides what we can observe". In addition, the experimental data has errors and comes from different laboratories. The single-cell experiments in the future may measure data according to the requirements of the Hill-type equation to verify theoretical findings.

In Reference (32) and (25), for slow pulsing or fast dissociation, an equation similar to Equation 1 is introduced, and cannot be directly derived from the trajectory of the system. This may be an average steady state in an unstable system and cannot be observed. In the limit of slow pulsing or fast dissociation, the peak-to-valley ratio tends infinity. However, for fast pulsing or slow dissociation, another steady state can be derived not only in the trajectory but also in the average equation. Therefore, it is observable. In this case the peak-to-valley ratio is 1, so the system is stable. This equation is called modified Hill equation which belongs to the ligand-receptor association and dissociation system upon pulsing signals. Correspondingly, Equation 1 is called Hill -type equation which comes from the synthesis and degradation of products upon oscillating signals (26). At present, there are only two types of Hill-type equation.

There are many models of gene expression. Some have been cited in the previous research from TF to DNA to mRNA. Gene expression exhibits stochastic characteristics.(38-40). The impact of the stochastic effects on gene expression has also been explained in the research from DNA to mRNA. The classic Hill equation has long been used in the modeling of biological systems (41-46). In mathematical form, the Michaelis-Menten equation is equivalent with the Hill equation if the Hill coefficient is unity((47). The purpose of this article is to find a principle of gene expression, and the obtained results need to withstand practical testing, and the parameters in the results are measurable. Therefore, we can only flexibly develop a minimal model (48) . Evidently, the basic purpose has been achieved. Hill initially only used the equation to fit experimental data. Now this Hill type equation reveals the principle of gene expression upon TF pulsing. If Hill could see this equation, he would definitely feel very pleased.

Enistein said: “There is no logical path to these laws; only intuition, resting on sympathetic understanding of experience, can reach them. In this methodological uncertainty, one might suppose that there were any number of possible systems of theoretical physics all equally well justified; and this opinion is no doubt correct, theoretically. But the development of physics has shown that at any given moment, out of all conceivable constructions, a single one has always proved itself decidedly superior to all the rest. Nobody who has really gone deeply into the matter will deny that in practice the world of phenomena uniquely determines the theoretical system, in spite of the fact that there is no logical bridge between phenomena and their theoretical principles; this is what Leibnitz. described so happily as a ‘pre-established

harmony(49, 50).'"

Although gene expression dynamics is random and complex, the regulation principle is always deterministic and simple. The Hill-type equation reflects the principle of deep simplicity of gene expression dynamics. The history of mechanics, physics, and chemistry indicates that the essence of nature is simple(51). Complex phenomena have evolved on the basis of the principle of simplicity. According to the regulation principle revealed by this equation, we can control gene expression from upstream of the genetic information flow described by the central dogma. Only four parameters need to be adjusted, namely, the duration, frequency, and amplitude of TF and binding affinity, and we can control the steady-state levels of mRNA and protein expression.

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

The classical Hill equation is applied when the TF dynamics are constant. When $\Delta = T$ is applied, 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 regulatory mechanism of gene expression, can be applied to any activated 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 (52, 53)." More than 100 years later, we have not forgotten his caveat (53).

Table 3. 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	-
$\bar{m}$	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_P	Index of protein accumulation	-
ρ_P	Peak to valley ratio for protein	-
γ	Duty cycle of p53 pulsing	-
λ_m	Index of mRNA accumulation	-
ρ_m	Peak to valley ratio for mRNA accumulation	-
m_{basal}	Basal mRNA fold change	-
P_{basal}	Basal protein fold change	-

Acknowledgements. I thank anonymous reviewers for several helpful comments.

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

Data Availability. Data used in this article come from [17, 23-26, 28, 30, 31] or are included in the text.

Author contributions. X.S.: conceptualization, formal analysis, investigation, methodology, writing—original draft, writing—review, and editing.

Conflict of interest declaration. I declare I have no competing interests.

Funding. This study received no funding.

References

1. F. H. CRICK, On protein synthesis. *Symp Soc Exp Biol* **12**, 138-163 (1958).

2. F. Crick, Central dogma of molecular biology. *Nature* **227**, 561-563 (1970).
3. C. Buccitelli, M. Selbach, mRNAs, proteins and the emerging principles of gene expression control. *Nat Rev Genet* **21**, 630-644 (2020).
4. T. K. Howcroft, A. Raval, J. D. Weissman, A. Gegonne, D. S. Singer, Distinct transcriptional pathways regulate basal and activated major histocompatibility complex class I expression. *Mol Cell Biol* **23**, 3377-3391 (2003).
5. R. N. Day, J. A. Walder, R. A. Maurer, A protein kinase inhibitor gene reduces both basal and multihormone-stimulated prolactin gene transcription. *J Biol Chem* **264**, 431-436 (1989).
6. P. J. Willy, R. Kobayashi, J. T. Kadonaga, A basal transcription factor that activates or represses transcription. *Science* **290**, 982-985 (2000).
7. J. A. Ranish, S. Hahn, Transcription: basal factors and activation. *Curr Opin Genet Dev* **6**, 151-158 (1996).
8. J. D. Parvin, H. T. Timmers, P. A. Sharp, Promoter specificity of basal transcription factors. *Cell* **68**, 1135-1144 (1992).
9. B. Schwanhäusser *et al.*, Global quantification of mammalian gene expression control. *Nature* **473**, 337-342 (2011).
10. J. Hausser, A. Mayo, L. Keren, U. Alon, Central dogma rates and the trade-off between precision and economy in gene expression. *Nat Commun* **10**, 68 (2019).
11. Y. Liu, A. Beyer, R. Aebersold, On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell* **165**, 535-550 (2016).
12. M. Jovanovic *et al.*, Immunogenetics. Dynamic profiling of the protein life cycle in response to pathogens. *Science* **347**, 1259038 (2015).
13. R. Singh, A. Letai, K. Sarosiek, Regulation of apoptosis in health and disease: the balancing act of BCL-2 family proteins. *Nat Rev Mol Cell Biol* **20**, 175-193 (2019).
14. R. Di Micco, V. Krizhanovsky, D. Baker, F. d'Adda di Fagagna, Cellular senescence in ageing: from mechanisms to therapeutic opportunities. *Nat Rev Mol Cell Biol* **22**, 75-95 (2021).
15. C. Vogel, E. M. Marcotte, Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat Rev Genet* **13**, 227-232 (2012).
16. J. J. Li, M. D. Biggin, Gene expression. Statistics requantitates the central dogma. *Science* **347**, 1066-1067 (2015).
17. D. Grün *et al.*, Conservation of mRNA and protein expression during development of C. elegans. *Cell Rep* **6**, 565-577 (2014).
18. J. E. Purvis, G. Lahav, Encoding and decoding cellular information through signaling dynamics. *Cell* **152**, 945-956 (2013).
19. J. H. Levine, Y. Lin, M. B. Elowitz, Functional roles of pulsing in genetic circuits. *Science* **342**, 1193-1200 (2013).
20. I. Imayoshi *et al.*, Oscillatory control of factors determining multipotency and fate in mouse neural progenitors. *Science* **342**, 1203-1208 (2013).
21. G. Lahav *et al.*, Dynamics of the p53-Mdm2 feedback loop in individual cells. *Nat Genet* **36**, 147-150 (2004).
22. J. E. Purvis *et al.*, p53 dynamics control cell fate. *Science* **336**, 1440-1444 (2012).
23. R. L. Hanson, J. R. Porter, E. Batchelor, Protein stability of p53 targets determines their temporal expression dynamics in response to p53 pulsing. *J Cell Biol* **218**, 1282-1297 (2019).
24. A. Hafner *et al.*, p53 pulses lead to distinct patterns of gene expression albeit similar DNA-binding dynamics. *Nat Struct Mol Biol* **24**, 840-847 (2017).
25. X. Shi, J. R. Reimers, Understanding non-linear effects from Hill-type dynamics with application to decoding of p53 signaling. *Sci Rep* **8**, 2147 (2018).
26. X. Shi, A Hill type equation can predict target gene expression driven by p53 pulsing. *FEBS Open Bio* **11**, 1799-1808 (2021).
27. A. Hafner *et al.*, Quantifying the Central Dogma in the p53 Pathway in Live Single Cells. *Cell Syst* **10**, 495-505.e494 (2020).
28. J. R. Porter, B. E. Fisher, E. Batchelor, 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 (2016).
29. M. Banwarth-Kuhn, S. Sindi, How and why to build a mathematical model: A case study using prion aggregation. *J Biol Chem* **295**, 5022-5035 (2020).
30. R. Martinez-Corral, J. Garcia-Ojalvo, Modeling cellular regulation by pulsatile inputs. *Current Opinion in Systems Biology* **3**, 23-29 (2017).
31. J. Keener, J. Sneyd, *Mathematical Physiology* (Springer-Verlag, New York, 1998).
32. C. Salazar, A. Z. Politi, T. Hofer, Decoding of calcium oscillations by phosphorylation cycles: analytic results. *Biophys J* **94**, 1203-1215 (2008).
33. C. C. Friedel, L. Dölken, Z. Ruzsics, U. H. Koszinowski, R. Zimmer, Conserved principles of mammalian transcriptional regulation revealed by RNA half-life. *Nucleic Acids Res* **37**, e115 (2009).
34. R. L. Weinberg, D. B. Veprintsev, M. Bycroft, A. R. Fersht, Comparative binding of p53 to its promoter and DNA recognition elements. *J Mol Biol* **348**, 589-596 (2005).
35. L. Cai, C. K. Dalal, M. B. Elowitz, Frequency-modulated nuclear localization bursts coordinate gene regulation. *Nature* **455**, 485-490 (2008).

36. J. B. Gurdon, K. Javed, M. Vodnala, N. Garrett, 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**, 15075-15084 (2020).
37. X. Shi, A Hill type equation with the negative coefficients governs repressed gene expression induced by p53 pulsing. *J Biol Chem* **300**, s122 (2024).
38. M. B. Elowitz, A. J. Levine, E. D. Siggia, P. S. Swain, Stochastic gene expression in a single cell. *Science* **297**, 1183-1186 (2002).
39. J. M. Raser, E. K. O'Shea, Noise in gene expression: origins, consequences, and control. *Science* **309**, 2010-2013 (2005).
40. A. Raj, A. van Oudenaarden, Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**, 216-226 (2008).
41. U. Alon, *An introduction to systems biology : design principles of biological circuits*, Chapman & Hall/CRC mathematical and computational biology series (Chapman & Hall/CRC, Boca Raton, FL, 2007), pp. xvi, 301 p., 304 p. of plates.
42. J. E. Ferrell, E. M. Machleder, The biochemical basis of an all-or-none cell fate switch in Xenopus oocytes. *Science* **280**, 895-898 (1998).
43. J. M. Borghans, G. Dupont, A. Goldbeter, Complex intracellular calcium oscillations. A theoretical exploration of possible mechanisms. *Biophys Chem* **66**, 25-41 (1997).
44. D. E. Koshland, A. Goldbeter, J. B. Stock, Amplification and adaptation in regulatory and sensory systems. *Science* **217**, 220-225 (1982).
45. K. Yang *et al.*, Dynamic simulations on the arachidonic acid metabolic network. *PLoS Comput Biol* **3**, e55 (2007).
46. X. P. Zhang, F. Liu, Z. Cheng, W. Wang, Cell fate decision mediated by p53 pulses. *Proc Natl Acad Sci U S A* **106**, 12245-12250 (2009).
47. R. Gesztelyi *et al.*, *The Hill equation and the origin of quantitative pharmacology*, Archive for History of Exact Sciences (Springer, 2012), vol. 66, pp. 427-438.
48. S. Dai, *Eighteen topics on strategies and techniques in scientific investigation* (Shanghai University Press, Shanghai, 2014).
49. A. Einstein (1918) Principles of Research. Physical Society, Berlin, for Max Planck's sixtieth birtday. **https://www.site.uottawa.ca/~yymao/misc/Einstein_PlanckBirthday.html**
50. S. Dai (2011) Science Net Blog. **https://blog.sciencenet.cn/blog-330732-445344.html**.
51. J. Wu, *MIscellaneous discussions on history of mechanics* (Higher Education Press, Beijing, 2009).
52. A. V. Hill, The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. *J Physiology (London)* **40**, iv-vii (1910).
53. D. Colquhoun, The quantitative analysis of drug-receptor interactions: a short history. *Trends Pharmacol Sci* **27**, 149-157 (2006).