Quantifying efficient information transduction of biochemical signaling cascades

Tatsuaki Tsuruyama*

Short title: Entropy production rate in cell signaling

*Department of Diagnostic Pathology, Center for Anatomical, Forensic Medical and Pathological Research, Kyoto University Graduate School of Medicine, Yoshida-Konoe-cho, Sakyo-ku, Kyoto 606-8315, Japan

*Send correspondence to: T. Tsuruyama
Department of Diagnostic Pathology, Kyoto University Graduate School of Medicine
1 Yoshida-Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan
Phone: +81-75-753-4473; Fax: +81-75-761-9591
E-mail: tsuruyam@kuhp.kyoto-u.ac.jp
Abstract

Cells can be regarded as systems that utilize changes in thermodynamic entropy as information. Therefore, they serve as useful models for investigating the relationship between entropy production and information transmission, i.e., signal transduction. Based on the hypothesis that cells apply a chemical reaction cascade for the most efficient transduction of information, a coding design was adopted that minimizes the number of bits per concentration of molecules that are employed for information transduction. As a result, the average rate of entropy production is uniform across all reactions in a cascade. Thus, the entropy production rate can be a valuable measure for the quantification of intracellular signal transduction.
Introduction

Information transmission in the cell, which is generally termed ‘signal transduction’ in the field of molecular biology, consists of biochemical reaction cascades to elicit cellular responses to stimuli from the external environment. The term transduction implies that the biological information is transmitted through the cascade of a variable number of biochemical reactions. Computational systems biology, which simulates a subset of known interactions between proteins that carry biological information (i.e., signaling molecules) using their kinetic equations, is effective in describing network of reaction cascades (1). Despite progress from computational studies, it is extremely difficult to represent entire systems in cases that include complex, non-linear interactions or to elucidate unknown interactions along with their biological functions from the overall network. The aim of this study is to provide a theory for the quantification of signal transduction that allows for the differentiation between non-specific fluctuations and a bona fide biological signaling event, and for comparing signal transduction cascades on the basis of information theory and the fluctuation theorem (FT)(2).

Specifically, signal transduction was experimentally investigated by kinetic analysis of biochemical modification/de-modification reactions, which can comprise a signaling cascade, that regulate the concentration of signaling molecules over time following extracellular stimulation. One type of protein modification/de-modification reaction is the phosphorylation/dephosphorylation cycle (3, 4). Reactions of receptors distant from the nucleus and those near the cell surface were found to occur relatively slowly (3-11). These reactions are referred to as distal reactions. In contrast, reactions occurring closer to the cell nucleus that are shared by many signaling cascades occur more rapidly, and are immediately followed by de-modification reactions to return to the pre-stimulus state. These reactions are referred to as proximal reactions. Considering that proximal reactions are generally more frequently utilized in the whole signal transduction system, reactions that are more proximal have a shorter period. This observation suggests that cellular signal transduction systems apply an effective way of coding that utilizes the signaling molecule as a symbol in reference to Brillouin’s discussion (12). Then, we made the following important assumption: “cellular signaling systems perform encoding such that the bit count per signaling molecule is at its
minimum for a given condition.” In contrast, from the perspective of robustness, biological signaling may be allowed to activate multiple signaling cascades in response to an external stimulus via a receptor on the cell surface. In an actual biological signaling cascade, the initial stimulus onto the cell triggers the receptor that subsequently activates multiple signaling cascades consisting of biochemical modifications of a signaling molecule. In particular, the signaling molecule may be an enzyme that is also a substrate in the previous reaction of the cascade. Accordingly, biological signaling cascades are a chain reaction of substrate-enzyme reactions. Herein, the specificity of the enzymatic reaction, i.e., the lock-and-key theory, is considered essential for compatibility with coding optimization and robustness employing multiple biological signaling cascades. A model signaling cascade is considered next.

Result

A model of biological signaling cascade

A cascade comprises a total of \( n \)-reactions and the total number of cascades is designated by \( m \). The first reaction in the cascade is the uptake or binding of the extracellular signal by a receptor \( c_1 \), leading to modification of the receptor, which then assists in the modification reaction of the signaling molecule \( c_j \) of the second step. The cascade continues in this manner, such that the \( j^{th} \)-reaction induces \( j+1^{th} \)-reaction (\( 1 \leq j \leq n \)). In the final \( n^{th} \)-reaction, the signaling molecule \( c_n \) translocates to the cell nucleus and, after binding with DNA, induces the transcription of genes into mRNA (Figure 1). Therefore, it is possible to describe a cascade of reactions from the reception of the extracellular stimulus to the induction of gene expression. As shown in Figure 1, each reaction is a per-stimulus modification/de-modification reaction that is maintained by a chemical reservoir and the de-modification reaction occurs slowly after the modification reaction. Intracellular proteins are sequentially recruited as signaling molecules in accordance with the above assumption of coding optimization. In parallel with this signaling cascade, the entire cellular system may respond to the same stimulus through multiple \( k \)-signaling cascades (\( 1 \leq k \leq m \)) (Figure 2). Given the number of signaling molecules at the \( k^{th} \)-reaction, \( A_k \), and taking the appearance frequency at the \( f^{th} \)-reaction as

```
within the $k^{th}$-signaling cascade, the combined signaling events are expressed as $\Psi_k$ and $\Psi$, respectively:

\[
\log \psi_k = \log \frac{\Lambda_k!}{\prod_{j=1}^{n} c_{jk}!} \approx -\Lambda_j \sum_{j=1}^{n} p_{jk} \log p_{jk}
\]

\[
\log \psi = \log \prod_{k=1}^{m} \psi_k \approx \log \prod_{k=1}^{m} \frac{\Lambda_k!}{\prod_{j=1}^{n} c_{jk}!} \approx -\sum_{k=1}^{m} \Lambda_k \sum_{j=1}^{n} p_{jk} \log p_{jk}
\]

using Stirling’s formula to derive [2].

Here,

\[
\sum_{k=1}^{m} \sum_{j=1}^{n} p_{jk} = 1
\]

The time required for each step is $\tau_{jk}$. Using the probability $p_{jk}$, the total duration, $\tau_k$, is obtained using the following equation:

\[
\tau_k \triangleq \Lambda_t \sum_{j=1}^{n} p_{jk} \tau_{jk}
\]

To take $\Psi$ to its maximum, [2], [3], and [4] are used for the maximum number of the $k^{th}$-cascade with the parameters, $\alpha_k$ and $\beta_k$, being independent of the step number $j$:

\[
d \log \psi = \sum_{k=1}^{m} \log \psi_k = -\sum_{k=1}^{m} \left( \alpha_k \sum_{j=1}^{n} p_{jk} - \beta_k d \tau_{jk} \right) = 0
\]

Taking the differential of [2] gives:

\[
d \log \psi = \sum_{k=1}^{m} \left( -d \Lambda_k \sum_{j=1}^{n} p_{jk} \log p_{jk} - \Lambda_k \sum_{j=1}^{n} p_{jk} (1 + \log p_{jk}) dp_{jk} \right)
\]

Substituting [5] into [6] gives:

\[
-d \Lambda_k \left[ \sum_{j=1}^{n} p_{jk} \log p_{jk} + \beta_k \sum_{j=1}^{n} p_{jk} \tau_{jk} \right] + \sum_{j=1}^{n} dp_{jk} \left[ -\alpha_k - \beta_k \Lambda_k \tau_{jk} - \Lambda_k (1 + \log p_{jk}) \right] = 0
\]

If $d \Lambda_k$ and $dp_{jk}$ are treated as independent variables:

\[
\sum_{j=1}^{n} p_{jk} \log p_{jk} + \beta_k \sum_{j=1}^{n} p_{jk} \tau_{jk} = 0
\]

\[
-\alpha_k - \beta \Lambda_k \tau_{jk} - \Lambda_k (1 + \log p_{jk}) = 0
\]
Substituting [8] into [9] gives:
\[\sum_{j=1}^{n} p_{jk} \left( -1 - \frac{\alpha_k}{\Lambda_k} \right) = 0 \]  
[10]

To satisfy [10]:
\[-1 - \frac{\alpha_k}{\Lambda_k} = 0 \]  
[11]

Thus, the following equations are obtained from [9] and [11]:
\[- \log p_{jk} = \beta_k \tau_k \]  
[12]

and
\[\sum_{k=1}^{n} \sum_{j=1}^{m} \exp(-\beta_k \tau_{jk}) = 1 \]  
[13]

Here, an important conclusion is obtained that \( \beta_k \) is unrelated to the step number, \( j \). This equation is equivalent to that previously derived by Brillouin (12). Consequently, it is possible to make an assessment of signal transduction from the perspective of encoding efficiency.

**Transitional probability and application of the FT**

According to the lock and key relationship in enzymatic reaction, the individual transitional probability from \( j \) to the \( j+1^{\text{th}} \)-step is equivalent to 1. Here, the transition probability is related to the total number of cascades. The total transition probability represents the probability that the signal is transmitted to the \( j^{\text{th}} \)-step and to the \( j+1^{\text{th}} \)-step using diffusion coefficients of the \( j^{\text{th}} \)-step and the \( j+1^{\text{th}} \)-step signaling molecules within the \( k^{\text{th}} \)-cascade, \( D_{jk} \) and \( D_{j+1,k} \) as follows:

\[
\rho_{j+1,j,k} = \frac{D_{jk}p_{j+1,j,k}d_{j+1,j,k}}{\sum_{j} D_{jk}p_{j+1,j,k}d_{j+1,j,k}^{-1} + \sum_{j} D_{j+1,k}p_{j+1,j,k}d_{j+1,j,k}^{-1}} , \\
\rho_{j+1,j,k} = \frac{D_{j+1,k}p_{j+1,j,k}d_{j+1,j,k}}{\sum_{j} D_{jk}p_{j+1,j,k}d_{j+1,j,k}^{-1} + \sum_{j} D_{j+1,k}p_{j+1,j,k}d_{j+1,j,k}^{-1}} 
\]  
[14]

In above, the transitional probability is given by Fick’s law. Here, the diffusion process is assumed to be the rate-limiting step of the cascade. Subsequently, using the FT (2) regarding the probability between the signal orientation and inverse orientation, the following is obtained:
\[ \rho_{j+1;j,k} = \rho_{j;j+1,k} \exp(-\bar{\sigma}_{j+1;j,k} \tau_{j+1;j,k}) \]  

[15]

with

\[ \bar{\sigma}_{j+1;j,k} = \frac{1}{\tau_{j+1;j,k}} \int_0^{\tau_{j+1;j,k}} \Delta s_{j+1;j,k} d\tau_{j+1;j,k} \]  

[16]

Here, the entropy production rate for transduction of the \( j \)-th step to the \( j+1 \)-th step represents the average rate of entropy production \( \bar{\sigma}_{j+1;j,k} \) during the interval \( \tau_{j+1;j,k} \). \( \Delta s_{j+1;j,k} \) denotes entropy production during \( \tau_{j+1;j,k} \) that encodes length (corresponding to the time required from the initiation of the signaling molecule modification to the end of the reaction) in the \( k \)-th cascade.

Here substituting [14] into [15], the following is obtained:

\[
\frac{D_{j+1;k} P_{j+1;j,k} d_{j+1;j,k}^{-1}}{\sum_j D_j P_{j+1;j,k} d_{j+1;j,k}^{-1} + \sum_j D_{j+1;k} P_{j+1;j,k} d_{j+1;j,k}^{-1}} \exp(-\bar{\sigma}_{j+1;j,k} \tau_{j+1;j,k})
\]  

[17]

Because the distance between \( j \)-th and \( j+1 \)-th steps is:

\[ d_{j+1;j,k} = d_{j,j+1} \]  

[18]

Simplifying [17] gives using [18]:

\[ D_{j+1;k} P_{j+1;j,k} = D_j P_{j+1;j,k} \exp(-\bar{\sigma}_{j+1;j,k} \tau_{j+1;j,k}) \]  

[19]

Substitution of [12] into [19] gives

\[ D_{j+1;k} \exp(-\beta \tau_{j+1;j,k}) = D_j \exp(\beta \tau_{j+1;j,k} - \bar{\sigma}_{j+1;j,k} \tau_{j+1;j,k}) \]  

[20]

and rearrangement gives:

\[ \lim_{\tau_{j+1;j,k} \to 0} \frac{1}{\log \frac{D_{j+1;k}}{D_{j+1;k}}} = \lim_{\tau_{j+1;j,k} \to 0} \frac{1}{\beta \tau_{j+1;j,k}} (\beta \tau_{j+1;j,k} - \bar{\sigma}_{j+1;j,k} \tau_{j+1;j,k} + \beta \tau_{j+1;j,k}) \]  

[21]

In an actual signaling cascade, the diffusion coefficients of signaling molecules are similar; therefore, the left side is nearly equivalent to zero when \( \tau_{j+1;j,k} \) is sufficiently longer and because \( \tau_{j+1;j,k} \ll \tau_{j+1;j,k} \),

\[ 0 = -\bar{\sigma}_{j+1;j,k} + \beta \]  

[22]
As a result, a simple result is obtained:

$$\sigma_{j+1,k} = \beta_k \hat{=} \sigma_k$$ \[23\]

Thus, an important conclusion is reached, whereby the average entropy production $\sigma_{j+1,k}$ is fixed and is independent of the step number, $j$. The entropy production rate can be a valuable measure for the quantification of intracellular signal transduction. When the entropy production rate is independent of the step number, the cascade is the most effective signaling cascade among the observed cascades.

**The average information entropy**

The average information entropy of the $k$th-cascade, the Shannon’s entropy $i_k$, is then given using [4]:

$$i_k = -\Lambda_k \sum_j p_j \log p_j = \Lambda_k \sum_j p_j \sigma_j \tau_j = \hat{=} \sigma_k \tau_k$$ \[24\]

Where the maximum amount of transmitted information, i.e., the channel capacity, $I_k$, is:

$$I_k \hat{=} \lim_{\tau_i \rightarrow \tau_k} \frac{i_k}{\tau_k} = \sigma_k$$ \[25\]

In this way, in non-equilibrium states that are not differentiable from the equilibrium state as described by the FT, a signaling cascade between signaling molecules can be established that satisfies equations [23] to [25] by measuring the duration of the reactions that are sequentially altered within a complex network.

**Discussion**

According to Sagawa et al. (13-17), when the feedback system of the biochemical reaction is given, the work that can be retrieved, $W_k$, is limited by the mutual information or the channel capacity, $I_k$, and the free energy change $\Delta F_k$ in the $k$th-cascade as described below.

$$W_k \leq \Delta F_k + k_B T I_k = \Delta F_k + k_B T \sigma_k$$ \[26\]

Here, $T$ is the temperature of the cellular system and $k_B$ is the Boltzmann constant. If we assume that the signaling cascade consists of modification/de-modification reactions of the signaling molecule, $\Delta F_k$ in [26] is equivalent to zero, because the initial and final status of the signaling cascade is identical. As a result, the work done by the external chemical reservoir of the mediator, $W_k$, is comprised of $k_B T \sigma_k$. 
In a Szilard engine, \( W_j \) is equivalent to \( k_b T \sigma_j = k_b T \sigma_j \) (13-17). Therefore, the conclusion from equation [26] suggests that the Szilard engine may be used as a model of biological signaling cascades. Further discussion and validation will be required in the future.

**ACKNOWLEDGMENTS**

This work was supported by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (Synergy of Fluctuation and Structure: Quest for Universal Laws in Non-Equilibrium Systems, P2013-201 Grant-in-Aid for Scientific Research on Innovative Areas, MEXT, Japan). There are no conflicts of interest.

**FIGURE LEGENDS**

**Figure 1.** Scheme for intracellular signal transduction. In the reaction, \( \Delta F_j = 0 \), and the work done to the external environment, \( W_{jk} \), is composed of \( \Delta F_{jk} + k_b T \sigma_k \tau_{jk} \).

**Figure 2.** Schematic of the reaction cascade in cell signal transduction. \( L \) is a ligand and \( R \) is a receptor that mediates cellular responses to external environmental changes. Individual signaling molecules \( S_j \) \( (S = W, Z, X, Y, V; 1 \leq j \leq n) \) relay the activation of individual cascades and the last species \( X_{n+1} \) is translocated to the nucleus, where it controls gene expression. We hypothesized that the multiple responses to the same ligand reflect the robustness of the cellular system that contains the reaction cascade.
REFERENCES

1. Cheong, R., A. Rhee, C. J. Wang, I. Nemenman, and A. Levchenko. 2011. Information transduction capacity of noisy biochemical signaling networks. Science 334:354-358.

2. Crooks, G. E. 1999. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics 60:2721-2726.

3. Kim, D. S., E. S. Hwang, J. E. Lee, S. Y. Kim, S. B. Kwon, and K. C. Park. 2003. Sphingosine-1-phosphate decreases melanin synthesis via sustained ERK activation and subsequent MITF degradation. J Cell Sci 116:1699-1706.

4. Kodama, H., K. Fukuda, J. Pan, M. Sano, T. Takahashi, T. Kato, S. Makino, T. Manabe, M. Murata, and S. Ogawa. 2000. Significance of ERK cascade compared with JAK/STAT and PI3-K pathway in gp130-mediated cardiac hypertrophy. Am J Physiol Heart Circ Physiol 279:H1635-1644.

5. Lee, C. S., M. Park, J. Han, J. H. Lee, I. H. Bae, H. Choi, E. D. Son, Y. H. Park, and K. M. Lim. 2013. Liver X receptor activation inhibits melanogenesis through the acceleration of ERK-mediated MITF degradation. J Invest Dermatol 133:1063-1071.

6. Mackeigan, J. P., L. O. Murphy, C. A. Dimitri, and J. Blenis. 2005. Graded mitogen-activated protein kinase activity precedes switch-like c-Fos induction in mammalian cells. Mol Cell Biol 25:4676-4682.

7. Newman, D. R., C. M. Li, R. Simmons, J. Khosla, and P. L. Sannes. 2004. Heparin affects signaling pathways stimulated by fibroblast growth factor-1 and -2 in type II cells. Am J Physiol Lung Cell Mol Physiol 287:L191-200.

8. Tao, R., H. E. Hoover, N. Honbo, M. Kalinowski, C. C. Alano, J. S. Karliner, and R. Raffai. 2010. High-density lipoprotein determines adult mouse cardiomyocyte fate after hypoxia-reoxygenation through lipoprotein-associated sphingosine 1-phosphate. Am J Physiol Heart Circ Physiol 298:H1022-1028.

9. Wang, H., J. J. Ubl, R. Stricker, and G. Reiser. 2002. Thrombin (PAR-1)-induced proliferation in astrocytes via MAPK involves multiple signaling pathways. Am J Physiol Cell Physiol 283:C1351-1364.

10. Yeung, K., T. Seitz, S. Li, P. Janosch, B. McFerran, C. Kaiser, F. Fee, K. D.
Katsanakis, D. W. Rose, H. Mischak, J. M. Sedivy, and W. Kolch. 1999. Suppression of Raf-1 kinase activity and MAP kinase signalling by RKIP. Nature 401:173-177.

11. Zhang, W., N. Yano, M. Deng, Q. Mao, S. K. Shaw, and Y. T. Tseng. 2011. beta-Adrenergic receptor-PI3K signaling crosstalk in mouse heart: elucidation of immediate downstream signaling cascades. PLoS One 6:e26581.

12. Brillouin, L. Science and Information Theory. Dover Publication Press, Inc., Mineola, New York, 2nd ed. :pp42.

13. Ito, S., and T. Sagawa. 2015. Maxwell's demon in biochemical signal transduction with feedback loop. Nat Commun 6:7498.

14. Ito, S., and T. Sagawa. 2013. Information thermodynamics on causal networks. Phys Rev Lett 111:180603.

15. Sagawa, T., and M. Ueda. 2010. Generalized Jarzynski equality under nonequilibrium feedback control. Phys Rev Lett 104:090602.

16. Sagawa, T., and M. Ueda. 2012. Nonequilibrium thermodynamics of feedback control. Phys Rev E Stat Nonlin Soft Matter Phys 85:021104.

17. Shiraishi, N., and T. Sagawa. 2015. Fluctuation theorem for partially masked nonequilibrium dynamics. Phys Rev E Stat Nonlin Soft Matter Phys 91:012130.
Figure 1

Information (entropy production rate $\sigma$)

chemical reservoir

$\Delta F_{jk} = k_B T \sigma_{jk}$

$\Delta F_{j+1, k} = k_B T \sigma_{j+1, k}$

Gene expression

Figure 2

Ligand

Receptor

DNA

nucleus

$W_{jk}$

$Y_{jm}$

$Z_{jk}$

$X_{jk}$

$Y_{nm}$

$Z_{n1}$

$X_{nk}$

$Y_{nm}$

$Z_{m1}$

$X_{nk}$

$Y_{nm}$

$Z_{n1}$

$X_{nk}$

$Y_{nm}$

$Z_{m1}$

$X_{nk}$

$Y_{nm}$

$Z_{n1}$

$X_{nk}$

$Y_{nm}$

$Z_{m1}$

$X_{nk}$

$Y_{nm}$

$Z_{n1}$

$X_{nk}$

$Y_{nm}$

$Z_{m1}$

$X_{nk}$

$Y_{nm}$

$Z_{n1}$

$X_{nk}$

$Y_{nm}$

$Z_{m1}$

$X_{nk}$

$Y_{nm}$

$Z_{n1}$

$X_{nk}$

$Y_{nm}$

$Z_{m1}$

$X_{nk}$

$Y_{nm}$