6

# Transcriptional Bursting in the Tryptophan Operon of *E. coli* and Its Effect on the System Stochastic Dynamics

Emanuel Salazar-Cavazos and Moisés Santillán
*Centro de Investigación y de Estudios Avanzados del IPN, Unidad Monterrey, Parque de Investigación e Innovación Tecnológica, Apodaca NL México*

## 1. Introduction

Transcriptional bursting, also known as transcriptional pulsing, is a fundamental property of genes from bacteria to humans (Chubb et al., 2006; Golding et al., 2005; Raj et al., 2006). Transcription of genes, the process which transforms the stable code written in DNA into the mobile RNA message can occur in "bursts" or "pulses". This phenomenon has recently come to light with the advent of new technologies, such as MS2 tagging, to detect RNA production in single cells, allowing precise measurements of RNA number, or RNA appearance at the gene. Other, more widespread techniques, such as Northern Blotting, Microarrays, RT-PCR and RNA-Seq, measure bulk RNA levels from homogeneous population extracts. These techniques lose dynamic information from individual cells, and give the impression transcription is a continuous smooth process. The reality is that transcription is irregular, with strong periods of activity, interspersed by long periods of inactivity. Averaged over millions of cells, this appears continuous. But at the individual cell level, there is considerable variability, and for most genes, very little activity at any one time.

The bursting phenomenon, as opposed to simple probabilistic models of transcription, can account for the high variability in gene expression occurring between cells in isogenic populations (Blake et al., 2003). This variability in turn can have tremendous consequences on cell behaviour, and must be mitigated or integrated. In certain contexts, such as the survival of microbes in rapidly changing stressful environments, or several types of scattered differentiation, the variability may be essential (Losick & Desplan, 2008). Variability also impacts upon the effectiveness of clinical treatment, with resistance of bacteria to antibiotics demonstrably caused by non-genetic differences (Lewis, 2010). Variability in gene expression may also contribute to resistance of sub-populations of cancer cells to chemotherapy (Sharma et al., 2010).

Bursting may result from the stochastic nature of biochemical events superimposed upon a 2 or more step fluctuation. In its most simple form, the gene can exist in 2 states, one where activity is negligible and one where there is a certain probability of activation (Raj & van Oudenaarden, 2008). Only in the second state does transcription readily occur. Whilst the nuclear and signalling landscapes of complex eukaryotic nuclei are likely to favour more than two simple states—for example, there are over twenty post-translational modifications

of nucleosomes known, this simple two step model provides a reasonable framework for understanding the changing probabilities affecting transcription. What do the restrictive and permissive states represent? An attractive idea is that the repressed state is a closed chromatin conformation whilst the permissive state is an open one. Another hypothesis is that the fluctuations reflect transition between bound pre-initiation complexes (permissive) and dissociated ones (restrictive) (Blake et al., 2003; Ross et al., 1994). Bursts may also result from bursty signalling, cell cycle effects or movement of chromatin to and from transcription factories. Nonetheless, to the best of our knowledge, there is no generally accepted explanation for this phenomenon. Transcriptional bursting in prokaryotic cell is particularly puzzling given the simplicity of the transcription initiation process, as opposed to eukaryotic cells.

There is evidence that the cooperative interaction of distant operators through a single repressor molecule has a strong influence on the transcriptional bursting observed in the *lac* operon of *E. coli* (Choi et al., 2008). Since the repression regulatory mechanism in *E. coli*'s *trp* operon involves cooperativity between two repressor molecules bound to neighbouring operators, it is interesting to investigate whether such cooperative interaction has any effect upon the system transcriptional dynamics. The present chapter is advocated to tackling this question from a mathematical modelling perspective. We also investigate the effects of the enzymatic feedback-inhibition regulatory mechanism (also present on the *trp* operon regulatory pathway) on the system dynamic behaviour.

## 2. Theory

### 2.1 The *trp* operon

The amino acid tryptophan can be synthesized by bacteria like *E. coli* through a series of catalysed reactions. The catalysing enzymes in *E. coli* are made up of the polypeptides encoded by the tryptophan operon genes: *trpE*, *trpD*, *trpC*, *trpB*, and *trpA*, which are transcribed from *trpE* to *trpA*. Transcription is initiated at promoter *trpP*, which is located upstream from gene *trpE*. The *trp* operon is regulated by three different negative-feedback mechanisms: repression, transcription attenuation, and enzyme inhibition. Below these regulatory mechanisms are briefly reviewed based on (Brown et al., 1999; Grillo et al., 1999; Jeeves et al., 1999; Xie et al., 2003; Yanofsky, 2000; Yanofsky & Crawford, 1987). The reader should consult Figure 1 for a better understanding.

**Repression** in the *trp* operon is mediated by three operators (*O1*, *O2*, and *O3*) overlapping with the operon promoter, *trpP* (see Figure 1A). When an active repressor is bound to an operator it blocks the binding of a RNA polymerase to *trpP* and prevents transcription initiation. The *trp* repressor normally exists as a dimeric protein (called the *trp* aporepressor) and may or may not be complexed with tryptophan (Trp). Each portion of the *trp* aporepressor has a binding site for tryptophan.

The *trp* aporepressor cannot bind the operators tightly when not complexed with tryptophan. However, if two tryptophan molecules bind to their respective binding sites the *trp* aporepressor is converted into a functional repressor. The resulting functional repressor complex can tightly bind to the *trp* operators and so the synthesis of the tryptophan producing enzymes is prevented. This sequence constitutes the repression negative-feedback mechanism: an increase in the concentration of tryptophan induces an increase in the concentration of the functional repressor, thus preventing the synthesis of tryptophan.

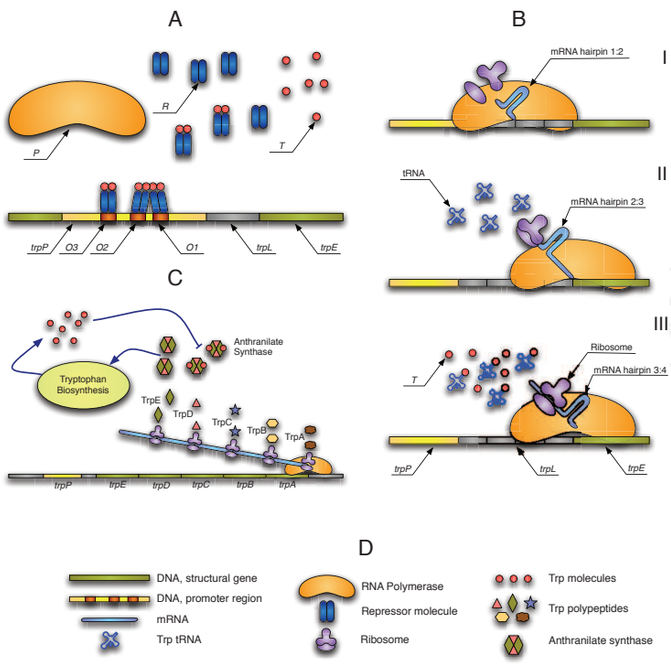



Fig. 1. Schematic representation of the three regulatory mechanisms found in the tryptophan operon: A) repression, B) transcription attenuation, and C) enzyme inhibition. A glossary with the meaning of all the geometric forms in this figure is shown in panel D. See the main text for further explanation.

**Transcription attenuation** works by promoting an early termination of mRNA transcription, see Figure 1B. The transcription starting site in the *trp* operon is separated from *trpE* by a leader region responsible for attenuation control. The transcript of this leader region consists of four segments (termed Segments 1, 2, 3, and 4) which can form three stable hairpin structures between consecutive segments. After the first two segments are transcribed they form a hairpin which stops transcription (c.f. Figure 1B-I). When a ribosome binds the nascent mRNA, it disrupts Hairpin 1:2 and transcription is re-initiated along with translation. Segment 1 contains two tryptophan codons in tandem. If there is scarcity of tryptophan, and thus of loaded $tRNA^{Trp}$, the ribosome stalls in the first segment. The development of Hairpin 2:3 (the antiterminator) is then facilitated, and transcription proceeds until the end (c.f. Figure 1B-II). However, if tryptophan is abundant, the ribosome rapidly finishes translation of Segments 1 and 2 and promotes the formation of a stable hairpin structure between Segments 3 and 4 (c.f. Figure 1B-III). RNA polymerase molecules recognize this hairpin structure as a termination signal and transcription is prematurely terminated.

**Enzyme inhibition** takes place through anthranilate synthase, the first enzyme to catalyse a reaction in the catalytic pathway that leads to the synthesis of tryptophan from chorismate. This enzyme is a hetero-tetramer consisting of two TrpE and two TrpD polypeptides. Anthranilate synthase is inhibited by tryptophan through negative feedback. This feedback inhibition is achieved when the TrpE subunits in anthranilate synthase are individually bound by a tryptophan molecule, see Figure 1C. Therefore, an excess of intracellular tryptophan inactivates most of the anthranilate synthase protein avoiding the production of more tryptophan.

## 3. Methods

### 3.1 Model development

There are three different repressor binding sites (operators) overlapping the *trp* promoter. Hence, the promoter can be in eight different states, with each operator being either free or bound by a repressor molecule. Furthermore, when two repressor molecules are bound to the first and second operators, they do it cooperatively (Grillo et al., 1999). As discussed in Appendix A, this cooperativity allows the grouping of the promoter states into two different sets that we term the permissive and the restrictive global states. The transitions within each global state and those from the permissive global state to the restrictive global state being much faster than those from the restrictive global to the permissive global states. This fact justifies the assumption that the system "instantaneously" reaches a stationary probability distribution for the states within every global state. This supposition in turn permits the derivation of the following expressions for the transition rates from the permissive into the restrictive global states ($k^+$), and vice versa ($k^-$):

$$k^+ = \frac{k_i^+ R_A/K_j + k_j^+ R_A/K_i}{1 + R_A/K_i + R_A/K_j}, \tag{1}$$

$$k^- = \frac{k_i^- + k_j^-}{k_c}, \tag{2}$$

where $k_i^+$ and $k_j^+$ are the rates for the reactions where a repressor molecule binds to the first and second operators, respectively; $k_i^-$ and $k_j^-$ are the rates for the reactions in which a repressor molecule detaches from the first and second operator; $R_A$ is the number of active repressors; $K_i = k_i^-/k_i^+$; and $K_j = k_j^-/k_j^+$.

On the other hand, when the promoter is in the permissive global state, the probability that it is not bound by any repressor and so it is free to be bound by a polymerase to start transcription is

$$P_R = \frac{1}{(1 + R_A/K_k)(1 + R_A/K_i + R_A/K_j)}. \tag{3}$$

Repressor molecules are activated when they are bound by a couple of tryptophan molecules. The kinetics of repressor activation were analysed in (Santillán & Zeron, 2004), where the number of active repressors is demonstrated to be given by

$$R_A = R_T \left( \frac{T}{T + K_T} \right)^2, \tag{4}$$

where $R_T$ stands for the total number of repressor molecules, while $T$ is the tryptophan molecule count. Substitution of Eqn. (4) into Eqns. (1) and (3) permits the calculation of the promoter inactivation rate ($k^+$) and the probability $P_R$ in terms of the tryptophan level ($T$).

Due to transcriptional attenuation, only a fraction of the polymerase molecules that initiate transcription reach the end of the *trp* genes and produce functional mRNA molecules, which in turn are translated to produce the proteins coded by the *trp* genes. Santillán and Zeron

(2004) found that the probability that transcription is not prematurely terminated due to transcriptional attenuation is:

$$P_A = \frac{1 + 2\alpha T}{(1 + \alpha T)^2}, \tag{5}$$

with $\alpha$ a parameter to be estimated.

It follows from the considerations in the previous paragraphs that a promoter in the restrictive global state is completely incapable of being expressed, but its activity level when it is in the permissive global state is a function of the tryptophan level $T$ and is given by the product of $P_R$ and $P_A$. Therefore, if $k_E$ denotes the rate of enzyme synthesis by a fully active promoter, the enzyme synthesis rate when the promoter is in the permissive state at a given tryptophan level turns out to be:

$$k_E P_R(T) P_A(T). \tag{6}$$

Tryptophan is synthesized by proteins which are assembled from the polypeptides coded by the *trp* genes. Conversely, tryptophan is mainly consumed in the synthesis of all kinds of proteins in *E. coli*. Thus, the equation governing the tryptophan-level dynamics is:

$$\frac{dT}{dt} = k_T E P_I(T) - \rho(T) - \mu T,$$

where $\mu$ is the bacterial growth rate, $k_T$ is the tryptophan rate of synthesis per enzyme molecule

$$P_I(T) = \frac{K_I^n}{T^n + K_I^n}$$

is the probability that an enzyme molecule is not feedback inhibited by tryptophan, and

$$\rho(T) = \rho_{max} \frac{T}{T + K_\rho}$$

is the rate of tryptophan consumption associated to protein synthesis. If we assume that these processes are much faster than those associated to gene expression and protein degradation, then we can make the following quasi-steady state approximation: $dT/dt = 0$, and the tryptophan molecule count can be uniquely calculated in terms of the enzyme molecule count as the root of the following algebraic equation:

$$k_T E P_I(T) - \rho(T) - \mu T = 0. \tag{7}$$

Following previous modelling studies we assume that the enzyme degradation rate is negligible as compared with the bacterial growth rate, $\mu$. On the other hand, instead of considering a cell that grows exponentially, we assume that we have a constant-volume cell and that the effective enzyme degradation rate is $\mu$.

The facts previously discussed in the present section provide enough information to develop a model for the *trp* operon regulatory pathway. This model consists of four chemical reactions: promoter activation, promoter inactivation, enzyme synthesis, and enzyme degradation, whose rates are $k^-$, $k^+(T(E))R_A(T(E))$, $k_E P_R(T(E))P_A(T(E))$, and $\mu E$, respectively. Figure 2 provides a schematic representation of such a model. It is worth emphasizing that the repression regulatory feedback loop is implicitly accounted for by functions $k^+(T(E))$ and $P_R(T(E))$, that function $P_A(T(E))$ corresponds to the attenuation feedback regulatory mechanism, and the feedback enzyme inhibition is implicit in the function $T(E)$, obtained after solving Eqn. (7).

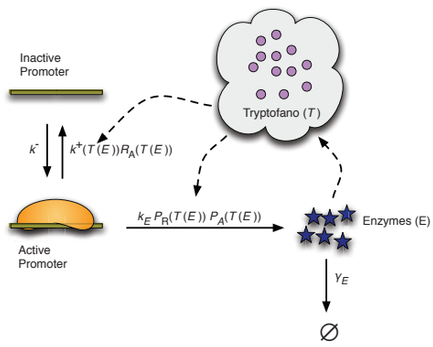



Fig. 2. Schematic representation of the mathematical model here developed for tryptophan operon gene regulatory circuit.

### 3.2 Parameter estimation

We paid special attention to the estimation of all the model parameters from reported experimental data. The parameter values we employ in the present work and the detailed procedure to estimate them are presented in Appendix B.

### 3.3 Numerical methods

The time evolution of the reaction network that models the tryptophan operon regulatory pathway was simulated by means of the Gillespie algorithm, which we implemented in `Python`.

## 4. Results

We carried out stochastic simulations with the model described in the previous section. As formerly stated, we made use of Gillespie's algorithm to mimic the system dynamic evolution for 200,000 min. In the first simulation we employed the parameter values estimated in Appendix B, which correspond to a wild-type bacterial strain. The results are summarized in Figure 3. In Figure 3A the cumulative sum of the promoter activity is plotted vs. time. We can appreciate there the existence of alternated activity and inactivity periods, just like it has been observed in transcriptional bursting. To further investigate this phenomenon we calculated the histograms of the permissive and restrictive period lengths. The results are shown in Figures 3B and 3C, respectively. Observe that both histograms are well fitted by exponential distributions, in agreement with the reported experimental data on transcriptional bursting. Finally, we present in Figures 3D and 3E the histograms for the enzyme and tryptophan molecule counts, respectively.

One feature worth noticing is that the histogram for the enzyme abundance is well fitted by a gamma distribution with parameters $k = 32.5$ and $\theta = 63$. This last fact is in agreement with the existence on transcriptional bursting in the *trp* operon of *E. coli*. It has been proved that in such a case, the protein count obeys a gamma distribution, with parameters $k$ and $\theta$ respectively interpreted as the average number of transcriptional bursts occurring during an average protein lifetime and the mean number of proteins produced per burst (Shahrezaei & Swain, 2008). It is also interesting to point out that the coefficient of variation in the tryptophan molecule count is similar to that of the enzyme molecule count.

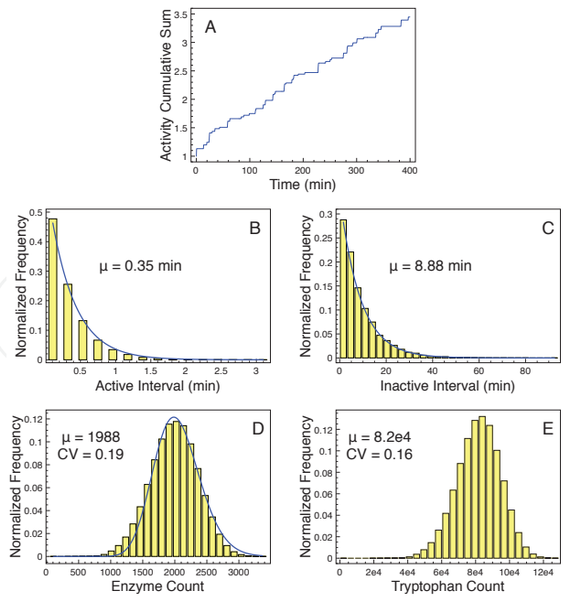



Fig. 3. Statistical analysis of the simulation corresponding to the wild-type strain. A) Plot of the cumulative promoter activity vs. time. B) Histogram of the promoter permissive time intervals and best fit to an exponential distribution. C) Histogram of the promoter restrictive time intervals and best fit to an exponential distribution. D) Histogram of the enzyme molecule count and best fit to a gamma distribution. E) Histogram of the tryptophan molecule count.

In order to understand the influence of the slow promoter gating between the global permissive and restrictive states on the operon dynamics, we increased the value of parameters $k_x^+$ and $k_x^-$ $(x = I, j, k)$ by a factor of 100. In this way, the promoter switching rate among all its available states gets faster, without altering each state's stationary probability. We repeated the simulation described in the previous paragraph with this new parameter set and the results are condensed in Figure 4.

We observe by comparing Figures 3A and 4A that there are many more alternated activity and inactivity periods in the fast-switching model than in that corresponding to the wild type strain. Concomitantly, the periods are shorter in the former case. Interestingly, the accumulated promoter activity is quite similar for both models. This last result comes from our increasing parameters $k_x^+$ and $k_x^-$ by the same factor, and is in agreement with the fact that the mean enzyme and tryptophan counts are quite similar in both models (see below).

In Figures 4B and 4C we present the histograms for the activity and inactivity periods, and the corresponding fits to exponential distributions. By comparing with the wild-type period distributions we can see that the mean values of both the activity and inactivity periods decrease. However the decrease of the activity-period average is about twice as large as that of the average of the inactivity period.

Finally, the histograms for the enzyme and tryptophan molecule counts are plotted in Figures 4D and 4E. Notice that the histogram for the enzyme count is well approximated by a gamma distributions with parameters $k = 2079$ and $\theta = 1$. Therefore, by making the promoter switching rate faster we increased the frequency of bursting, but decreased in the same

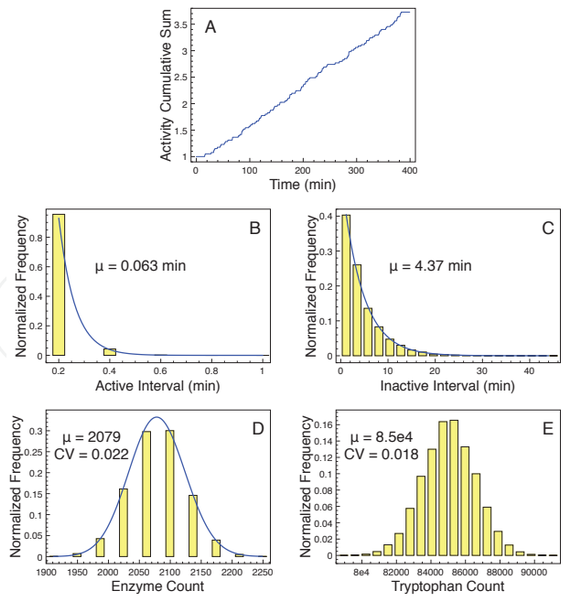



Fig. 4. Statistical analysis of the simulation corresponding to the strain with fast promoter-transition rates. A) Plot of the cumulative promoter activity vs. time. B) Histogram of the promoter permissive time intervals and best fit to an exponential distribution. C) Histogram of the promoter restrictive time intervals and best fit to an exponential distribution. D) Histogram of the enzyme molecule count and best fit to a gamma distribution. E) Histogram of the tryptophan molecule count.

amount the number of proteins synthesized per burst. In that way, the average enzyme count remains the same (compare with Figure 3D). However, the variation coefficient is much smaller in the fast promoter switching model than in the wild-type one. Recall that, in the gamma distribution, the mean and the standard deviation are: $\mu = \theta k$ and $\sigma = \theta\sqrt{k}$, while the variation coefficient is given by $CV = \sigma/\mu = 1/\sqrt{k}$.

We further simulated a bacterial mutant strain lacking the feedback inhibition regulatory mechanism. To mimic this mutation we increased the value of parameter $K_I$ by two orders of magnitude, up to $K_I = 500,000$ molecules. We analysed this last simulation in a similar way than the previous ones and present the results in Figure 5.

A comparison of Figures 3A and 5A reveals that the promoter level of activation is generally smaller in the inhibition-less mutant strain than in the wild-type strain, because the accumulated activity is about three times smaller in the former case. Nonetheless the length of the activity and inactivity periods seem to be similar. This last assertion is corroborated by the plots in Figures 5B and 5C, where we can see the the activity and inactivity period histograms are well fitted by exponential distributions, and that the corresponding mean values are similar to the corresponding ones in the wild-type strain.

In agreement with the fact that the promoter level of activity is smaller in the inhibition-less than in the wild-type strain, the mean protein count is smaller in the first case (compare Figures 3D and 5D). On the other hand, the coefficient of variation (CV) is similar in both cases. To understand why this happens when one would expect a larger CV in the inhibition-less

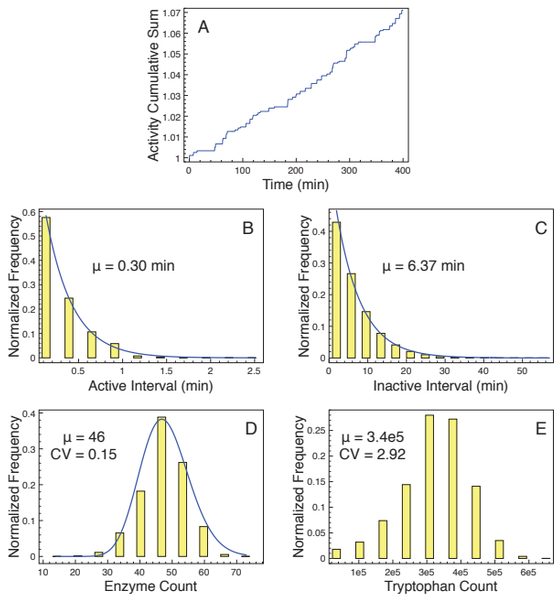



Fig. 5. Statistical analysis of the simulation corresponding to the inhibition-less strain. A) Plot of the cumulative promoter activity vs. time. B) Histogram of the promoter permissive time intervals and best fit to an exponential distribution. C) Histogram of the promoter restrictive time intervals and best fit to an exponential distribution. D) Histogram of the enzyme molecule count and best fit to a gamma distribution. E) Histogram of the tryptophan molecule count.

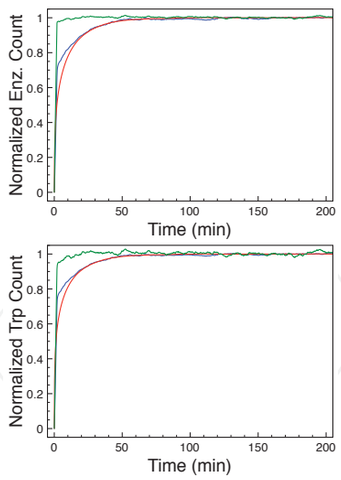



Fig. 6. Plots of the normalized enzyme count and the normalized tryptophan count, averaged over 1,000 independent simulations, vs. time for the wild-type (blue line), the fast promoter-transition (red line), and the inhibition-less (green line) *E. coli* strains.

mutant because of the reduced enzyme count, we fitted the histogram in Figure 5D and found that the best fit is obtained with parameters $k = 40$ and $\theta = 1.2$. Since the values of $k$ for the inhibition-less and the wild-type strains are similar, we conclude that the burst frequency is

comparable in both cases. However, the number of proteins produced per burst is notably smaller in the inhibition-less strain due to the reduced promoter activity, as well as to the increased level of transcriptional attenuation.

We note that, in contrast with the two previous simulations, the variation coefficient of the tryptophan molecular count in the inhibition-less strain is much larger than the coefficient of variation for the enzyme count. To our understanding, this happens because, being enzyme inhibition absent, there is a highly non-linear relation between the enzyme and tryptophan molecule counts.

In order to investigate the effect of the inhibition-less mutation and of the increased promoter switching rate on the *trp* operon response time, we carried out 1,000 simulations with each bacterial strain (including the wild type), starting with the promoter in the restrictive state and zero enzyme and tryptophan molecules as initial conditions. Then, we averaged the enzyme and the tryptophan counts over all the simulations, and plotted the results in Figure 6 to compare how fast each strain approaches the steady state. We can see there that the inhibition-less strain is the one with the shortest response time, followed by the wild type and the fast promoter switching strains, respectively.

## 5. Concluding remarks

In this work we have introduced a stochastic mathematical model for the tryptophan operon. Our objectives were twofold: 1) to investigate whether the reported reaction rates of the interaction between repressor and the operators can give rise to transcriptional bursting; and 2) to study the dynamic effects of transcriptional bursting, if it exists, and of the feedback enzyme-inhibition regulatory mechanism.

Regarding the first objective our results indicate that, indeed, the reported reaction rates make the promoter switching between its available states slow enough so as to give rise to transcriptional bursting. As previously discussed, this assertion is supported by the agreement between our model results and a number of reported experimental facts. Interestingly, experiments on the *lac* operon also suggest that transcriptional bursting has its origin in the kinetics of the repressor-operator interaction. As a matter of fact, Choi et al. (2008) demonstrated that the cooperativity present when a single repressor molecule is bound to two distant operators is responsible of generating two different types of bursts in the expression of *lac* operon.

In our model we have assumed that the trp promoter activation rate $k^-$ is independent of the tryptophan concentration. This assumption is supported by direct and indirect experimental measurements of the promoter activation and deactivation rates on the *lac* and several other promoters (Choi et al., 2008; So et al., 2011). Those reports demonstrate that modulation of gene expression is mainly achieved by changing the promoter deactivation rate.

Regarding our second objective, our results allow us to put forward the following conclusions:

- Transcriptional bursting increases the noise level and decreases the system response time after a nutritional shift. To the best of our understanding, the noise level is increased because the promoter-transition events become less frequent as the promoter-repressor interactions slow down, thus enhancing the concomitant stochastic effects. On the other hand, the faster system response can be explained by a single burst of intense transcriptional activity, occurring during the first couple of minutes after the nutritional

shift. This burst allows most of the cells to reach tryptophan levels superior to two thirds of the steady-state level.

- Enzyme inhibition also has important dynamic effects. It increases the noise level in the enzyme count, but decreases the noise level in the number of tryptophan molecules. Furthermore, this regulatory mechanism also increases the system response time. Knowing that the presence of a strong negative feedback is capable of reducing the noise in a biological system (Austin et al., 2006; Becskei & Serrano, 2000; Dublanche et al., 2006), we can explain the above observations as follows: the fact that the wild-type *E. coli* strain has lower tryptophan levels than the inhibition-free strain means that the transcriptional-attenuation and the repression feedback loops are weaker in the first strain; this weakening of both negative feedback loops is responsible for the increment of the noise level in the enzyme count. For the same reasons, the presence of the enzyme-inhibition feedback loop reduces the noise level in the tryptophan count, but makes it necessary to produce much more enzymes to fulfil the required tryptophan production. This increased enzyme synthesis requirement lengthens the system response time.

Finally, if we assume that having a tryptophan operon with short response times and low noise levels in the tryptophan molecular count are beneficial traits for *E. coli* then we can speculate from the previously discussed facts that evolution has bestowed this system with an optimal trade-off between short response times and low tryptophan noise.

## 6. Acknowledgements

This research was partially supported by Consejo Nacional de Ciencia y Tecnología (CONACyT, MEXICO) under Grant: 55228.

## 7. Appendix

### A. Promoter dynamics modelling

Three different operators are overlap with the *trp* promoter and each of them can be bound by a repressor molecule. Therefore, the eight promoter states can be denoted as $\{i, j, k\}$, where $i, j, k = 0, 1$ represent the binding state of the first, second, and third repressors, respectively. A zero (one) value means that the corresponding operator is free from (bound by) a repressor molecule.

Let $k_i^+$, $k_j^+$, and $k_k^+$ respectively denote the rates of binding of a repressor molecule to the first, second, and third operators, when the other two are free. Similarly, let $k_i^-$, $k_j^-$, and $k_k^-$ respectively represent the dissociation rate of a repressor molecule solely bound to the first, second, and third operators.

It is known that the first and second operators are bound by repressor molecules cooperatively and the cooperativity constant $k_c > 1$ has been measured. Here we assume that this cooperativity means that the rate of dissociation for a repressor molecule bound to either the first or the second operator is respectively given by $k_i^-/k_c$ or $k_j^-/k_c$, when both operators are bound by repressor molecules. Under this assumption, the eight different promoter states can be grouped into two sets that we call the permissive and the restrictive global states. The permissive global state consists of states $(0,0,0)$, $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, $(1,0,1)$, and $(0,1,1)$, while the restrictive global state consists of $(1,1,0)$, and $(1,1,1)$. From the

way the were constructed, the transitions within each global state and the transitions from the permissive to the restrictive global state are much faster than the transitions from the restrictive to the permissive global state. From the above considerations we make an adiabatic approximations for all the transitions within the permissive state, and thus:

$$\begin{aligned} P(0,0,0)\tfrac{R_A}{K_i} &= P(1,0,0), \\ P(0,0,0)\tfrac{R_A}{K_j} &= P(0,1,0), \\ P(0,0,0)\tfrac{R_A}{K_k} &= P(0,0,1), \\ P(0,0,1)\tfrac{R_A}{K_i} &= P(1,0,1), \\ P(0,0,1)\tfrac{R_A}{K_j} &= P(0,1,1), \end{aligned}$$

where $P(i,j,k)$ stands for the probability of state $i,j,k$ within the restrictive global state, $R_A$ denotes the number of active repressors, and $K_x = k_x^-/k_x^+$ $(x = i,j,k)$. It follow from the equations above and the constraint $P(0,0,0) + P(1,0,0) + P(0,1,0) + P(0,0,1) + P(1,0,1) + P(0,1,1) = 1$ that

$$\begin{aligned} P(0,0,0) &= \tfrac{1}{(1+R_A/K_k)(1+R_A/K_i+R_A/K_j)}, \\ P(1,0,0) &= \tfrac{R_A/K_i}{(1+R_A/K_k)(1+R_A/K_i+R_A/K_j)}, \\ P(0,1,0) &= \tfrac{R_A/K_j}{(1+R_A/K_k)(1+R_A/K_i+R_A/K_j)}, \\ P(0,0,1) &= \tfrac{R_A/K_k}{(1+R_A/K_k)(1+R_A/K_i+R_A/K_j)}, \\ P(1,0,1) &= \tfrac{R_A/K_i \times R_A/K_k}{(1+R_A/K_k)(1+R_A/K_i+R_A/K_j)}, \\ P(0,1,1) &= \tfrac{R_A/K_j \times R_A/K_k}{(1+R_A/K_k)(1+R_A/K_i+R_A/K_j)}. \end{aligned}$$

Let $k^+$ $(k^-)$ be the transition rate from each of the states in the permissive (restrictive) global state to each if the states in the restrictive (permissive) global state. From Zeron and Santillán (2010), these rates are given by

$$k^+ = k_i^+(P(0,1,0) + P(0,1,1)) + k_j^+(P(1,0,0) + P(1,0,1)) = \frac{k_i^+ R_A/K_j + k_j^+ R_A/K_i}{1 + R_A/K_i + R_A/K_j},$$

and

$$k^- = \frac{k_i^- + k_j^-}{k_c}(P(1,1,0) + P(1,1,1)) = \frac{k_i^- + k_j^-}{k_c},$$

since $P(1,1,0) + P(1,1,1) = 1$.

Finally, when the promoter is in the permissive global state, the probability that it is not bound by any repressor and so it is free to be bound by a polymerase to start transcription is

$$P_R = P(0,0,0) = \frac{1}{(1 + R_A/K_k)(1 + R_A/K_i + R_A/K_j)}.$$

Thus $P_R$ can be interpreted as the operator activity level.

### B. Estimation of the model parameters

In this work we consider a bacterial doubling time of 40 min, and thus

$$\mu \simeq 0.017 \text{ min}^{-1}.$$

From the website `E. coli Statistics`, the total number of proteins in *E. coli* is about 3.6 millions, while the average protein size is 360 residues. On the other hand, we have from the website `B1ONUMB3RS` (`http://bionumbers.hms.harvard.edu/`) that the abundance of tryptophan in the *E. coli* proteins is around 1.1%. The data above imply that there are of the order of 14.256 million tryptophan molecules assembles in the *E. coli* proteins at any given time. If we further consider that all the proteins in a bacterium have to be doubled before it duplicates (every 40 min), then the average tryptophan consumption rate is

$$\rho_{max} \simeq 360,000 \text{ molecules/min}.$$

Since this consumption rate cannot be maintained when the tryptophan level is too low, we assumed that the consumption rate for this amino acid is given by

$$\rho(T) = \rho_{max} \frac{T}{T + K_\rho},$$

with

$$K_\rho = 1,000 \text{ molecules}.$$

This choice for $K_\rho$ guarantees that the tryptophan consumption rate is most of the time very close to $\rho_{max}$, except when there are of the order of a few thousand molecules of the amino acid.

According to the website `E. coli Statistics` (`http://redpoll.pharmacy.ualberta.ca/CCDB/cgi-bin/STAT_NEW.cgi`), the average tryptophan molecule count in this bacterium is

$$T^* \simeq 80,000 \text{ molecules}.$$

From Morse et al. (1968), the average number of anthranilate synthase enzymes in *E. coli* is

$$E^* \simeq 2,000 \text{ molecules}.$$

Caligiuri & Bauerle (1991) found from their experimental data that the probability that an anthranilate synthase enzyme is not feedback inhibited by tryptophan can be approximated by the following function:

$$P_I(T) = \frac{K_I^n}{K_I^n + T^n},$$

with

$$K_I \simeq 2,500 \text{ molecules}, \quad \text{and} \quad n \simeq 1.2.$$

Let $k_T$ be the tryptophan synthesis rate per anthranilate synthase molecule. Given, that the average tryptophan synthesis rate must equal the consumption rate for this amino acid, we can solve for $k_T$ from the following equation $k_T E^* P_I(T*) = \rho_{max}$. After doing the math we obtain

$$k_T \simeq 12,000 \text{ molecules/min}.$$

Consider the promoter model discussed in Section 3. Let $P_{act}$ be the probability that the promoter is in the global permissive state. We have thus that, in the stationary state, $k^+ R_A P_{act} = k^-(1 - P_{act})$. Then, by solving for $P_{act}$ in the previous equation and taking into consideration that, when the promoter is in the global permissive state, the probability that it is ready to be bound by a polymerase is $P_R$. The stationary promoter activity level as a function of the tryptophan molecule count is

$$\frac{1}{1 + k^+(T)R_A(T)/k^-} P_R(T).$$

It is straightforward to test that the promoter activity level equals one when $T = 0$ molecules. On the other hand, Yanofsky & Horn (1994) measured that the operon expression level is maximal under conditions of tryptophan starvation, and that the repression regulatory mechanism decreases the promoter activity level by 60 times when the tryptophan level reaches its normal value. Thus, we must have that

$$\frac{1}{1 + k^+(T^*)R_A(T^*)/k^-} P_R(T^*) = 1/60.$$

This last result can then be used to estimate parameter $K_T$—see Eqn. (4). Thus, from Eqn. (4) and given that (Gunsalus et al., 1986)

$$R_T \simeq 400 \text{ molecules},$$

we obtain after some algebra that

$$K_T \simeq 1.7 \times 10^6 \text{ molecules}.$$

Grillo et al. (1999) estimated the following values for the promoter-repressor interaction rates:

$$\begin{aligned}
k_i^+ &\simeq 8.1 \text{ molecules}^{-1}\text{min}^{-1}, \\
k_i^- &\simeq 6.0 \text{ min}^{-1}, \\
k_j^+ &\simeq 0.312 \text{ molecules}^{-1}\text{min}^{-1}, \\
k_j^- &\simeq 0.198 \text{ min}^{-1}, \\
k_k^+ &\simeq 0.3 \text{ molecules}^{-1}\text{min}^{-1}, \\
k_k^- &\simeq 36.0 \text{ min}^{-1}, \\
k_c &\simeq 40.0
\end{aligned}$$

The probability that transcription is not prematurely terminated due to transcriptional attenuation is given by Eqn. (5). On the other hand, Yanofsky & Horn (1994) measured that one of every ten polymerases that have initiated transcription finish transcribing the operon genes when the tryptophan level is at its normal value. This means that $P_A(T) = 0.1$. We obtain from this that

$$\alpha \simeq 2.3 \times 10^{-4} \text{ molecules}^{-1}.$$

Finally, the value of parameter $k_E$ is chosen so that, when $T = T^*$, the average enzyme molecule count is $E^*$. We found by inspection that

$$k_E \simeq 30,000 \text{ molecules/min}.$$

complies with this requirement.

## 8. References

Austin, D., Allen, M., McCollum, J., Dar, R., Wilgus, J., Sayler, G., Samatova, N., Cox, C. & Simpson, M. (2006). Gene network shaping of inherent noise spectra, *Nature* 439(7076): 608–611.

Becskei, A. & Serrano, L. (2000). Engineering stability in gene networks by autoregulation, *Nature* 405(6786): 590–593.

Blake, W., Kaern, M., Cantor, C. & Collins, J. (2003). Noise in eukaryotic gene expression, *Nature* 422(6932): 633–637.

Brown, M. P., Grillo, A. O., Boyer, M. & Royer, C. A. (1999). Probing the role of water in the tryptophan repressor-operator complex, *Protein Sci* 8(6): 1276–85.

Caligiuri, M. G. & Bauerle, R. (1991). Identification of amino acid residues involved in feedback regulation of the anthranilate synthase complex from *Salmonella typhimurium*, *J. Biol. Chem.* 266: 8328–8335.

Choi, P. J., Cai, L., Frieda, K. & Xie, S. (2008). A stochastic single-molecule event triggers phenotype switching of a bacterial cell, *Science* 322(5900): 442–446.

Chubb, J., Trcek, T., Shenoy, S. & Singer, R. (2006). Transcriptional pulsing of a developmental gene, *Current Biology* 16(10): 1018–1025.

Dublanche, Y., Michalodimitrakis, K., Kuemmerer, N., Foglierini, M. & Serrano, L. (2006). Noise in transcription negative feedback loops: simulation and experimental analysis, *Molecular Systems Biology* 2: 41.

Golding, I., Paulsson, J., Zawilski, S. & Cox, E. (2005). Real-time kinetics of gene activity in individual bacteria, *Cell* 123(6): 1025–1036.

Grillo, A. O., Brown, M. P. & Royer, C. A. (1999). Probing the physical basis for *trp* repressor-operator recognition, *J. Mol. Biol.* 287: 539–554.

Gunsalus, R. P., Miguel, A. G. & Gunsalus, G. L. (1986). Intracellular *trp* repressor levels in *Escherichia coli*, *J Bacteriol* 167(1): 272–8.

Jeeves, M., Evans, P. D., Parslow, R. A., Jaseja, M. & Hyde, E. I. (1999). Studies of the *Escherichia coli trp* repressor binding to its five operators and to variant operator sequences, *Eur J Biochem* 265(3): 919–28.

Lewis, K. (2010). Persister cells, *Annual Review of Microbiology, Vol 64, 2010* 64: 357–372.

Losick, R. & Desplan, C. (2008). Stochasticity and cell fate, *Science* 320(5872): 65–68.

Morse, D. E., Baker, R. F. & Yanofsky, C. (1968). Translation of the tryptophan messenger RNA of *Escherichia coli*, *Proc. Natl. Acad. Sci.* 60(4): 1428–1435.

Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y. & Tyagi, S. (2006). Stochastic mRNA synthesis in mammalian cells, *Plos Biology* 4(10): 1707–1719.

Raj, A. & van Oudenaarden, A. (2008). Nature, nurture, or chance: Stochastic gene expression and its consequences, *Cell* 135(2): 216–226.

Ross, I. L., Browne, C. M. & Hume, D. A. (1994). Transcription of individual genes in eukaryotic cells occurs randomly and infrequently, *Immunol Cell Biol* 72(2): 177–85.

Santillán, M. & Zeron, E. S. (2004). Dynamic influence of feedback enzyme inhibition and transcription attenuation on the tryptophan operon response to nutritional shifts, *J. Theor. Biol.* 231: 287–298.

Shahrezaei, V. & Swain, P. S. (2008). Analytical distributions for stochastic gene expression, *Proceedings of the National Academy of Sciences of the United States of America* 105(45): 17256–17261.

Sharma, S. V., Lee, D. Y., Li, B., Quinlan, M. P., Takahashi, F., Maheswaran, S., McDermott, U., Azizian, N., Zou, L., Fischbach, M. A., Wong, K.-K., Brandstetter, K., Wittner, B.,

Ramaswamy, S., Classon, M. & Settleman, J. (2010). A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations, *Cell* 141(1): 69–80.

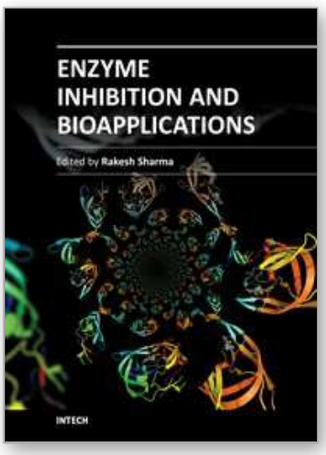
So, L.-H., Ghosh, A., Zong, C., Sepulveda, L. A., Segev, R. & Golding, I. (2011). General properties of transcriptional time series in escherichia coli, *Nature Genetics* 43(6): 554–U84.

Xie, G., Keyhani, N. O., Bonner, C. A. & Jensen, R. A. (2003). Ancient origin of the tryptophan operon and the dynamics of evolutionary change, *Microbiol. Mol. Biol. Rev.* 67: 303–342.

Yanofsky, C. (2000). Transcription attenuation, once viewed as a novel regulatory strategy, *J. Bacteriol.* 182: 1–8.

Yanofsky, C. & Crawford, I. P. (1987). The tryptophan operon, *in* F. C. Neidhart, J. L. Ingraham, K. B. Low, B. Magasanik & H. E. Umbarger (eds), *Escherichia coli and Salmonella thyphymurium: Cellular and Molecular Biology, Vol. 2*, Am. Soc. Microbiol., Washington, DC, pp. 1453–1472.

Yanofsky, C. & Horn, V. (1994). Role of regulatory features of the *trp* operon of *Escherichia coli* in mediating a response to a nutritional shift, *J. Bacteriol.* 176: 6245–6254.




**Enzyme Inhibition and Bioapplications**

Edited by Prof. Rakesh Sharma

ISBN 978-953-51-0585-5
Hard cover, 314 pages
**Publisher** InTech
**Published online** 09, May, 2012
**Published in print edition** May, 2012

Enzyme Inhibition and Bioapplications is a concise book on applied methods of enzymes used in drug testing. The present volume will serve the purpose of applied drug evaluation methods in research projects, as well as relatively experienced enzyme scientists who might wish to develop their experiments further. Chapters are arranged in the order of basic concepts of enzyme inhibition and physiological basis of cytochromes followed by new concepts of applied drug therapy; reliability analysis; and new enzyme applications from mechanistic point of view.

**How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Emanuel Salazar-Cavazos and Moisés Santillán (2012). Transcriptional Bursting in the Tryptophan Operon of E. coli and Its Effect on the System Stochastic Dynamics, Enzyme Inhibition and Bioapplications, Prof. Rakesh Sharma (Ed.), ISBN: 978-953-51-0585-5, InTech, Available from: http://www.intechopen.com/books/enzyme-inhibition-and-bioapplications/transcriptional-bursting-in-the-tryptophan-operon-of-e-coli-and-its-effect-on-the-system-stochastic-

INTECH
open science | open minds

**InTech Europe**
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

**InTech China**
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen