This article has been accepted for the publication in the Journal of Computational Biology. Final publication is available from Mary Ann Liebert, Inc., publishers http://dx.doi.org/10.1089/cmb.2014.0064.

Stochastic Simulation Algorithm for Gene Regulatory Networks with Multiple Binding Sites

Mattia Petroni*†, Nikolaj Zimic†, Miha Mraz†

and Miha Moškon†

*To whom correspondence should be addressed.
†Faculty of Computer and Information Science, University of Ljubljana, Ljubljana, Slovenia

Abstract

Promoters with multiple binding sites present a regulatory mechanism of several natural biological systems. It has been shown that such systems reflect a higher stability in comparison to the systems with small numbers of binding sites. Regulatory mechanisms with multiple binding sites are therefore used more frequently in artificially designed biological system in recent years. While the number of possible promoter states increases exponentially with the number of binding sites, it is extremely hard to model such systems accurately. Here we present an adaptation of stochastic simulation algorithm for accurate modelling of gene regulatory networks with multiple binding sites. Small computational complexity of adapted algorithm allows us to model any feasible number of binding sites per promoter. Introduced approach is demonstrated on the model of switching mechanism in Epstein-Barr virus where 20 binding sites are observed on one of the promoters. We show that the presented approach is easy to adapt to any biological systems based on the regulatory mechanisms with multiple binding sites in order to obtain and analyse their behaviour.

Key words: multiscale stochastic simulation algorithm, multiple binding sites, gene regulatory networks, computational modelling, systems biology.

1. Introduction

Over the past decade impressive results have been achieved with synthetic biology approaches in various scientific fields and also in first commercial products, e.g. in pharmacological, environmental and fuel production applications (Weber and Fussenegger, 2012). An example of such systems are switch-operated drugs, based on the gene switching mechanism. The gene switching mechanism represents a foundation for the implementation of almost any information processing applications in biological systems (Khalil and Collins, 2010). These systems are mainly based on gene regulatory networks (GRNs) because of their programmable nature (Moškon and Mraz, 2012). However, the process of their design is rarely straightforward and often requires the trial and error strategy. Computational modelling approaches help us to simplify the design of GRNs and they also provide a basis for robustness evaluation of the system behaviour (Moškon *et al.* , 2013).

The stochastic simulation algorithm (SSA) (Gillespie, 1976, 1977) has become a very popular tool for modelling natural and engineered GRNs in the fields of systems and synthetic biology (El Samad *et al.* , 2005). The SSA provides an easy way to perform a precise molecular simulation for any chemical reaction-based system with known chemical kinetics. However, its complexity is proportional to the number of observed chemical reactions and can increase drastically with the size of the analysed system. Certain simplifications sometimes allow us to reduce the complexity of modelled systems while still capturing their precise dynamics at the same time (Cao *et al.* , 2005b; E *et al.* , 2007). Many GRNs can thus be modelled with a less than hundred reactions, which is a still manageable for the SSA. However, in some cases these reductions cause qualitative changes in

the observed system's behaviour. Therefore an alternative modelling technique is often required. A large family of such systems are GRNs regulated by promoters with multiple transcriptional control sites on which multiple transcription factors can bind competitively. In these systems the direct application of SSA requires an exponential number of binding and unbinding reactions per promoter.

Several examples of GRNs driven by transcription factors which bind competitively on multiple binding sites can be found in nature, e.g. in the switching mechanism of the Epstein-Barr virus (EBV) (Werner *et al.* , 2007b). On the other hand there has been an increasing number of synthetic GRNs comprised of a particular high number of designed transcriptional control binding sites in recent years (Boch *et al.* , 2009; Bogdanove and Voytas, 2011). A high number of binding sites can enhanced the effectiveness of gene expression, allowing more stable signals and therefore enhance the robustness of the system's behaviour. Moreover, with the advent of novel synthetic DNA binding domains, it is also possible to design more complex gene circuit with information processing capabilities (Gaber *et al.* , 2014). Modelling approaches for such systems have already been proposed (Vucko *et al.* , 2013; Werner *et al.* , 2007b). However, these approaches lack the inherent accuracy of the SSA approach and are unable to provide the detailed insights in the modelled systems on a molecular level. Ergo there is a need to provide a stochastic modelling technique for such complex system.

Here we explore the possibility to adapt the SSA to GRNs regulated by promoters with multiple DNA binding sites, in order to pertain the accuracy of original SSA and still achieve manageable computational complexity at the same time. In Section 2 we describe the multiscale SSA and its adaptation to GRNs with multiple binding sites. Section 3 demonstrates the application and verification of the introduced approach on the switching mechanism in the Epstein-Barr virus,

where 20 binding sites are observed on one of the promoters. Discussion and overview of our contributions are given in Section 4.

2. Materials and Methods

Any gene regulatory network can be represented with a chemical reaction network. The entire dynamic of the system can be thus fully described at a molecular level with a set of interacting chemical reactions and their kinetic rates. These define the speed of change of the molecular concentration of each chemical species that appear in the given reaction. The stochastic simulation algorithm (SSA) can be used to generate the time evolution trajectories of each chemical species by firing each reaction according to the Monte-Carlo algorithm (Gillespie, 1977). The problem of this approach is the computational complexity regarding the high number of possible reactions that have to be simulated. This is the case of many GRNs whether they are found in nature or artificially constructed. Reactions in such networks often occur in different time scales which can differ in orders of magnitude, e.g. the dissociation and association of transcription factors to DNA bindings sites occur much faster than the expression of the regulated genes. Thus the SSA may only fire the *fast* reactions for the most of the time and may not substantially affect the slow reactions. The concentrations of *slow* chemical species are thus mostly unaltered during the course of the simulation. A simple intuition for improving the simulation strategy suggests to differentiate the simulation in two parts i.e. the *slow* and the *fast*, rather than simulate the entire system at once. This intuition is the base for the stochastic multi-time-scale (or *multiscale*) modelling approach.

2.1. Multiscale stochastic simulation algorithm

A very effective approach called *multiscale SSA* reduces the computational complexity and pertains the accuracy of original SSA approach with the division

of the simulation in two parts, i.e. the slow and the fast (Cao *et al.* , 2005b). The set of reactions is divided into a *fast* and a *slow* reaction set or *slot* according to the propensity function of each reaction, which is dependant on the concentrations of reaction's reactants and its kinetic rate. The simulation is performed with two subroutines. The first one approximates the time evolution of the fast reactions to their first average moments, i.e. until a steady state is reached. The second subroutine includes the approximations of slow reactions (Cao *et al.* , 2005a,b) (see algorithm 2.1). Mathematical correctness of the approach is ensured if the sets of fast and slow chemical reactions meet certain conditions regarding their types (E *et al.* , 2007). If a reactant that is a product of a fast reaction appears in a slow reaction, then the minimal required time-scale of the multiscale SSA, used to distinguish the two sets, has to be at least the time needed for this reactant to reach the steady state in the fast reaction. The use of the multiscale SSA is thus restricted to models with a high time scale reaction differentiation. An example of clearly differentiated time scales are GRNs while the biochemical kinetics of DNA binding proteins (e.g. enzymes, protease or kinase) are substantially faster than the kinetics of protein transcription and translation. The multiscale SSA can thus drastically reduce the computational complexity of the simulations of such systems.

Algorithm 2.1 Multiscale SSA algorithm

▷ if j = 1: Reaction set of transcription, translation and degradation reactions
▷ if j = 2: Reaction set of DNA binding proteins reactions

x_j = chemical species in reaction set j
v_j = stoichiometric matrix of reaction set j
θ_j = vector of parameters of reaction set j
$\mathbf{k}$ = chemical species that are affected by fast and slow reactions
r = index of species to log in x_1
Y = reporter output vector
y_0 = initial concentration of chemical species
T = time output vector

procedure [Y,T] = MULTISCALESSA($x_1, v_1, \theta_1, x_2, v_2, \theta_2$)
    $i = 0$
    $t_1, t_2 = 0$
    $Y[0] = y_0$
    while $t_1 < T_{\mathrm{MAX}_1}$ **do**
        while $t_2 < T_{\mathrm{MAX}_2}$ **do**
            % SSA step at 2. time scale
            SSA(x_2, v_2, θ_2)
            $t_2 = t_2 + \tau_2$
        end while
        $x_1[\mathbf{k}] = x_2[\mathbf{k}]$
        % SSA step at 1. time scale
        SSA(x_1, v_1, θ_1)
        $t_1 = t_1 + \tau_1$
        $i = i + 1$
        Y[i] = $x_1[r]$
        T[i] = t_1
    end while
end procedure

2.2. Modelling the multiple promoter binding sites

Many GRNs with multiple promoter binding sites can be found in nature and may also present a suitable platform to design novel biological systems (Boch *et al.* , 2009; Bogdanove and Voytas, 2011). Reactions that change each binding site's states can be easily described as

$$\mathrm{BS} \;+\; \mathrm{P} \xrightarrow{\mathrm{k_{on}}} \mathrm{BS}^{*}, \tag{1}$$

$$\mathrm{BS}^{*} \xrightarrow{\mathrm{k_{off}}} \mathrm{BS} \;+\; \mathrm{P}, \tag{2}$$

where BS represents a binding site and P a binding protein, i.e. transcription factor, which can either increase, i.e. activator, or decrease, i.e. repressor, the promoter activity. While the number of possible binding sites does not affect the number of reactions that describe degradation and transcription/translation processes, the number of reactions of type 1 and 2 increase drastically with the number of binding sites. The promoter *state* can be defined with the occupancy of its binding sites with different transcription factors. Let's presume that k different types of DNA binding proteins can bind to n consecutive (equal) binding sites. Number of possible promoter states is thus

$$N = (k+1)^{n}. \tag{3}$$

The total number of binding and unbinding reactions that may occur with k-binding proteins on n available binding sites on a DNA strand is:

$$2 \times n \times \sum_{i=1}^{n} k^{i}. \tag{4}$$

It is obvious that the number of binding reactions grows exponentially with the number of available binding sites (Conzelmann *et al.* , 2008).

In order to make the system in question manageable for the multiscale SSA algorithm we propose to present the binding sites of observed GRNs by a matrix of dimensions [# of promoters in the cell] $\times$ [# of binding sites of each promoter],

which presents the entire DNA binding space for all observed promoters inside the cell. We can define the element of a *binding site matrix* M as

$$M_{ij} = \begin{cases} 0; & \text{if } BS_{ij} \textit{ is free} \\ 1; & \text{if } \textit{activator is bound to } BS_{ij} \\ 2; & \text{if } \textit{repressor is bound to } BS_{ij} \end{cases} \tag{5}$$

where BS_{ij} denotes the j-th binding site on the promoter i. An example of the binding site matrix for 10 binding sites per promoter is shown in figure 6.

$$M = \begin{bmatrix} 0\,2\,0\,0\,1\,0\,2\,0\,2\,0 \\ 1\,1\,0\,1\,1\,1\,1\,0\,0\,2 \\ 0\,2\,1\,0\,0\,1\,1\,1\,1\,1 \\ 1\,0\,0\,0\,0\,2\,0\,2\,0\,1 \\ 0\,1\,0\,2\,0\,2\,2\,2\,2\,2 \\ 0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\ \vdots \quad \vdots \\ 2\,0\,1\,1\,0\,0\,0\,0\,0\,0 \end{bmatrix} \tag{6}$$

We use the binding site matrix in the internal, i.e. *fast* subroutine of multiscale SSA. When the SSA fires a binding reaction, a random free site is chosen and a binding is performed (chosen matrix element is set to 1 for activator binding and to 2 for repressor binding). Respectively, when the SSA fires an unbinding reaction, a random occupied site is chosen and an unbinding is performed (chosen matrix element is set to 0).

The actual activity of the promoters has to be determined by specific rules that take into account the number of bound repressors and activators. The rules that determine if a specific promoter, i.e. a specific row of the binding site matrix,

is activated or inhibited, may be based on a simple *majority* rule, i.e. by simply counting the number of bounded activators and repressors. For example, in the third row of the binding site matrix of the equation (6), six repressors and one activator are bound, which makes the promoter inactive. Contrary, the promoter in the fifth row contains six activators and one repressor, so the promoter may be considered to be activated. These rules can be constructed in accordance with the experimental results, e.g. the experimental fact that the repressors proteins are usually stronger than the activators, suggests that it may be more appropriate to consider alternative rules rather than the majority one. Another realistic rule may also consider the fact that the binding sites that are closer to the actual promoter (i.e. their position is on the right side of the matrix) have stronger effect on the promoter activity.

2.3. Adaptation of multiscale SSA to GRNs with multiple binding sites

Adaptation of the multiscale SSA (see Algortihm 2.1) to an arbitrary GRN with multiple binding sites can be performed with the construction of binding matrix M for each type of the promoter in the system according to the equation (5). The pseudo-code of this adaptation can be described with the algorithm 2.2.

The algorithm 2.2 can manage the explosion of the number of promoters' states by simply covering it with the binding matrix (note that the binding matrix has to be constructed for each promoter type with multiple binding sites). Moreover, the algorithm 2.2 permits the exploitation of different rules for evaluation of the promoters' states directly. These can be inferred from the experimental data. We refer to (Werner *et al.* , 2007a) for additional details about competitive binding rules.

Algorithm 2.2 The adapted multiscale stochastic simulation algorithm

x_i = chemical species in reaction set i
v_i = stoichiometric matrix of reaction set i
θ_i = vector of parameters of reaction set i
k = chemical species that are affected by fast and slow reactions
p = promoters species used in fast reactions
s = promoters species used in slow reactions
r = index of species to log in x_1
Y = reporter output vector
y_0 = initial concentration of chemical species
T = time output vector
A = average number of plasmids in a cell
n = number of binding sites per promoter type
for each type of promoter in the system **do**
  $M_p(j)$ = *empty matrix* of size: A $\times$ n
end for

procedure [Y,T] = MULTISCALESSA($x_1, v_1, \theta_1, x_2, v_2, \theta_2$)
  $i = 0$
  $t_1, t_2 = 0$
  $Y[0] = y_0$
  while $t_1 < T_{\text{MAX}_1}$ **do**
    // Quantify the number of activated, repressed
    // and free binding sites for each type of promoters
    $x_2[p]$ = EvaluateBindingSites($M_p(j)$)
    while $t_2 < T_{\text{MAX}_2}$ **do**
      // Perform an SSA step in for fast reactions without updating the state
      //vector x_2 and retrieve the next reaction c that will be fired
      c = SSA(x_2, v_2, θ_2)
      j = randomly choose a j promoter for which the reaction c can be fired
      // Set the value of the chosen promoter binding site as in equation (5)
      $M_p(j)$ = c $\rightarrow [0, 1, 2]$
      $t_2 = t_2 + \tau_2$
    end while
    for for each promoter j **do**
      $x_1[s]$ = EvaluatePromotorState($M_p(j)$)
    end for
    // Update the outermost state vector with the remaining common species
    $x_1[\mathbf{k}] = x_2[\mathbf{k}]$
    SSA(x_1, v_1, θ_1) // perform a SSA step for slow reactions
    $t_1 = t_1 + \tau_1$
    Y[i] = $x_1[r]$
    T[i] = t_1
  end while
end procedure

3. A case study: the Epstein-Barr virus genetic switch

The Epstein-Barr virus (EBV) affects a very high percentage of the entire adult population (Young and Rickinson, 2004). It is widely confirmed that the infection of this virus can increase the risk of cancer development and recently a notable effort in the way to prevent this infection has been achieved (Cohen *et al.* , 2011). The EBV is one of the most studied viruses. Majority of the virus-related diseases have already been discovered (Maeda *et al.* , 2009). The genome of the virus comprises 172000 base pairs. Its location is usually the nucleus of the B-lymphocytes. Like the majority of other viruses the EBV exhibits two different states, a *proliferating* (active) and a *latent* (resting) state. Roughly speaking, when the virus is in latent state, the cell – the B-lymphocyte – behaves normally and the virus genome duplicates together with the cell in a lysogenic cycle. When the virus is in the proliferating state (or lytic state), the cell membrane decays, the virus spreads outside the cell, it begins to duplicate in a very high number and it becomes highly infective for neighbour cells. The switch between these two states depends on the promoter activity of two promoters (*Qp* and *Cp*) on the virus genome. A high Cp promoter activity indicates a lytic state, while a high Qp promoter activity indicates a resting state. These two promoters are believed to be mutually exclusive, although they are responsible of a more complex gene regulation (Robertson, 2010). A key component in this regulation is the Cp promoter transcriptional regulator site, which is composed of 20 consecutive binding sites known as *Family of Repeats* (FR). On these binding sites two different transcription factors that may inhibit or activate the Cp promoter can simultaneously bind competitively. A very simplified scheme of the regulatory network controlling the Cp and Qp promoter activity is depicted in figure 1. We refer to (Robertson, 2010;

Werner *et al.* , 2007b; Young and Rickinson, 2004) for the precise details about the complex gene regulation in the Epstein-Barr virus.

[Figure 1 about here.]

The exact switching mechanism between the proliferating and resting state of the virus is still not perfectly understood, but a reasonable simplified mechanism can be explained as follows. Protein complex Oct2/Grg/TLE can bind to the 20 available binding sites in the FR region, which causes the repression of the Cp promoter. The dimer of Epstein-Barr Antigen-1 (EBNA-1) can on the other hand bind to FR region binding sites, which causes the activation of the Cp promoter. Moreover, EBNA1 can also bind to two binding sites in the Qp promoter region, which causes the repression of the Qp promoter. Active Cp promoter provides a high gene expression of all the nuclear proteins EBNA-1-6 encoded in the EBNA-1-6 gene. The transcription of the EBNA-1-6 proteins codes a long mRNA molecule which provides a translation of the 619 amino acid long EBNA-1 protein (Goldsmith *et al.* , 1993; Ohara *et al.* , 2000). While the EBNA-1 binding requires a dimeric structure, the activation rate of the bound Cp promoter is highly increased. This gene regulation results in a positive feedback, which drastically increases the gene expression of the EBNA-1-6 gene during the lytic state. On the other hand active Qp promoter provides a feeble gene expression of the the EBNA-1 gene, which keeps the EBNA-1 protein concentrations low, resulting in a weakly negative feedback in the two binding sites in the Qp promoter downstream region. The trigger to initiate the switch between Cp and Qp promoter activity is represented by the external signal provided by the Oct-2 protein complex.

It is obvious that the ordinary SSA approach would fail on the described system (only the FR region would give us 3^{20} different promoter states which have to be modelled individually). We have established an adapted multiscale

SSA model with the reactions described in tables 1 (fast reactions) and 2 (slow reactions). These reactions represent a minimum-reaction model of the EBV switch. This model comprises DNA binding, transcription, translation and degradation reactions, which are still sufficient to capture a relatively accurate behaviour of the observed system.

[Table 1 about here.]

We performed several simulations with the adapted SSA algorithm (see algorithm 2.2) varying initial conditions, i.e. number of promoters per cell and other variable parameters given in table 3. We presented the FR region with the binding matrix M with the reference size $2x20$ (for 2 promoters with 20 binding sites). The simulation time was set to 20 days, in order to perform a direct comparison with the model proposed in (Werner *et al.* , 2007b). We inferred the reaction rates of both fast and slow reactions from the selected literature 3. We adopted the weak majority binding rule for the repressors, which states that the promoter is be inhibited when at least one repressor is bound in the multiple binding site region. Results of the reference simulations performed are presented in figure 2.

[Table 2 about here.]

[Table 3 about here.]

Once we established the reference model, which is defined with the reference parameter values, we performed several simulations by changing these parameters within the feasible intervals. In accordance with the previous modelling technique presented in (Werner *et al.* , 2007b) we tested the Epstein-Barr genetic switch, by altering the level of the external signal of the protein complex Oct-2/Grg/TLE. We found out that the model depicted in figure 1, behaves as a poor genetic switch

when altered with an insufficient amount of Oct-2/Grg/TLE. We discovered a successful switch between the two viral states with an amount of Oct-2/Grg/TLE of 500000 molecules inserted after a simulation time of 10 days (plot d in figure 2).

[Figure 2 about here.]

Results of stochastic simulations performed with the adapted SSA approach were in accordance with the experimental results of previous works. Correctness of the introduced approach was therefore successfully confirmed. Although there are several different genetic model motifs with multiple binding sites to which the algorithm 2.2 can be applied, the EBV genetic switch was chosen for the demonstration because of its high importance in systems biology.

4. Conclusion

The Epstein-Barr virus genetic switch is able to achieve remarkable resistance to the external signals and robustness with its positive feedback structure. Our stochastic simulations and analysis confirmed these statements. Although our reaction model is quite simple for describing the complex dynamics that occur in a complex environment such as the virus responsible cell lysate, where it might also be important to consider additional diffusion properties and delays at the molecular-reaction level, our simulations are still in accordance with previously reported experimental work.

We applied the introduced adapted SSA approach to Epstein-Barr genetic switch mechanism because of its importance in systems biology. However, the application of our approach to any biological system that can be presented by a GRN with multiple promoter binding sites would be straightforward. With the recent advances in synthetic biology, which also justify the applications of GRNs with very high numbers of promoter binding sites, the necessity of stochastic modelling of such systems is even more important. We believe that the adapted SSA approach presented here should find many interesting applications by allowing manageable computational complexity of stochastic simulations of these networks.

Acknowledgements

We would like to thank the colleagues of Computer Structures and Systems Laboratory for their support and motivation during the preparations of this paper. The research was supported by the national postgraduate programme Higher Education National Scheme (*Inovativna shema za sofinanciranje doktorskega šudija za spodbujanje sodelovanja z gospodarstvom in reševanja aktualnih družbenih izzivov — generacija 2011 Univerza v Ljubljani*), financed by the European Union (EU), University of Ljubljana and Slovenian Ministry of Higher Education, Science and Technology.

Author disclosure statement

The authors declare that no competing financial interests exist.

References

Ambinder, R.F., Shah, W.A., Rawlins, D.R., et al. 1990. Definition of the sequence requirements for binding of the EBNA-1 protein to its palindromic target sites in Epstein-Barr virus DNA. *J. Virol.*, **64**(5), 2369–2379.

Boch, J., Scholze, H., Schornack, S., et al. 2009. Breaking the code of dna binding specificity of tal-type iii effectors. *Sci.*, **326**(5959), 1509–1512.

Bogdanove, A.J., and Voytas, D.F. 2011. Tal effectors: Customizable proteins for dna targeting. *Sci.*, **333**(6051), 1843–1846.

Cao, Y., Gillespie, D., and Petzold, L. 2005a. Multiscale stochastic simulation algorithm with stochastic partial equilibrium assumption for chemically reacting systems. *J. Comput. Phys.*, **206**, 395–411.

Cao, Y., Gillespie, D.T, and Petzold, L.R. 2005b. The slow-scale stochastic simulation algorithm. *J. Chem. Phys.*, **122**(1), 014116(1–18).

Ceccarelli, D.F.J., and Frappier, L. 2000. Functional Analyses of the EBNA1 Origin DNA Binding Protein of Epstein-Barr Virus. *J. Virol.*, **74**(11), 4939–4948.

Cohen, J.I., Fauci, A.S., Varmus, H., et al. 2011. Epstein-Barr Virus: An Important Vaccine Target for Cancer Prevention. *Sci. Transl. Medicine*, **3**(107), 107fs7.

Conzelmann, H., Fey, D., and Gilles, E. 2008. Exact model reduction of combinatorial reaction networks. *BMC Syst. Biol.*, **2**(1), 78.

Davenport, M.G., and Pagano, J.S. 1999. Expression of EBNA-1 mRNA Is Regulated by Cell Cycle during Epstein-Barr Virus Type I Latency. *J. Virol.*, **73**(4), 3154–3161.

E, W., Liu, D., and Vanden-Eijnden, E. 2007. Nested stochastic simulation algorithms for chemical kinetic systems with multiple time scales. *J. Comput. Phys.*, **221**(1), 158–180.

El Samad, H., Khammash, M., Petzold, L., et al. 2005. Stochastic modelling of gene regulatory networks. *Int. J. Robust Nonlinear Control.*, **15**, 691–711.

Gaber, R., Lebar, T., Majerle, A., Šter, B., et al. 2014. Designable DNA-binding domains enable construction of logic circuits in mammalian cells. *Nat. Chem. Biol.*, **10**(3), 203–208.

Gillespie, D.T. 1976. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, **22**, 403–434.

Gillespie, D.T. 1977. Exact stochastic simulation of coupled chemical reactions. *The J. Phys. Chem.*, **81**, 2340–2361.

Goldsmith, K., Bendell, L., and Frappier, L. 1993. Identification of EBNA1 amino acid sequences required for the interaction of the functional elements of the Epstein-Barr virus latent origin of DNA replication. *J. Virol.*, **67**(6), 3418–3426.

Khalil, A.S., and Collins, J.J. 2010. Synthetic biology: applications come of age. *Nat. Rev. Genet.*, **11**(5), 367–379.

Levitskaya, J., Sharipo, A., Leonchiks, A., et al. 1997. Inhibition of ubiquitin proteasome-dependent protein degradation by the Gly-Ala repeat domain of the Epstein-Barr virus nuclear antigen 1. *Proc. Natl. Acad. Sci.*, **94**(23), 12616–12621.

Maeda, E., Akahane, M., Kiryu, S., et al. 2009. Spectrum of Epstein-Barr virus-related diseases: a pictorial review. *Jpn. J. Radiol.*, **27**(1), 4–19.

Moškon, M., and Mraz, M. 2012. Modelling and analysing the information processing capabilities of simple biological systems. *Math. Model. Analysis*, **17**, 467–484.

Moškon, M., Bordon, J., Mraz, M., et al. 2013. *Recent advances in system biology*. Hauppauge NY, USA: Nova Science Publishers Inc. Chap. Computational approaches in synthetic and systems biology.

Ohara, N., Hayashi, K., Teramoto, N., et al. 2000. Sequence Analysis and Variation of EBNA-1 in Epstein-Barr Virus-Related Herpesvirus of Cynomolgus Monkey. *Intervirology*, **43**, 102–106.

Robertson, E.S. 2010. *Epstein-Barr Virus: Latency and Transformation*. Caister Academic Press.

Tsurumi, T. 1991. Primer terminus recognition and highly processive replication by Epstein-Barr virus DNA polymerase. *Biochem. J.*, **280(Pt 3)**, 703–708.

Vucko, D., Mraz, M., Zimic, N., et al. 2013. Hybrid discrete algorithm for the modelling of gene regulatory networks. *Biosyst. Inf. Technol.*, **2(2)**, 32–36.

Weber, W., and Fussenegger, M. 2012. Emerging biomedical applications of synthetic biology. *Nat. reviews. Genet.*, **13**(1), 21–35.

Werner, M., Zhu, L., and Aurell, E. 2007a. Cooperative action in eukaryotic gene regulation: Physical properties of a viral example. *Phys. Rev. E*, **76**(Dec), 061909.

Werner, M., Ernberg, I., Zou, J., et al. 2007b. Epstein-Barr virus latency switch in human B-cells: a physico-chemical model. *BMC Syst. Biol.*, **1**(1), 40.

Yang, E., van Nimwegen, E., Zavolan, M., et al. 2003. Decay Rates of Human

mRNAs: Correlation With Functional Characteristics and Sequence Attributes. *Genome Res.*, **13**(8), 1863–1872.

Young, L.S., and Rickinson, A.B. 2004. Epstein-Barr virus: 40 years on. *Nat. Rev. Cancer*, **4**(10), 757–768.

Address correspondence to:
Mattia Petroni
Faculty of Computer and Information Science,
University of Ljubljana,
Tržaška 25, 1000 Ljubljana
Slovenia
E-mail: mattia.petroni@fri.uni-lj.si

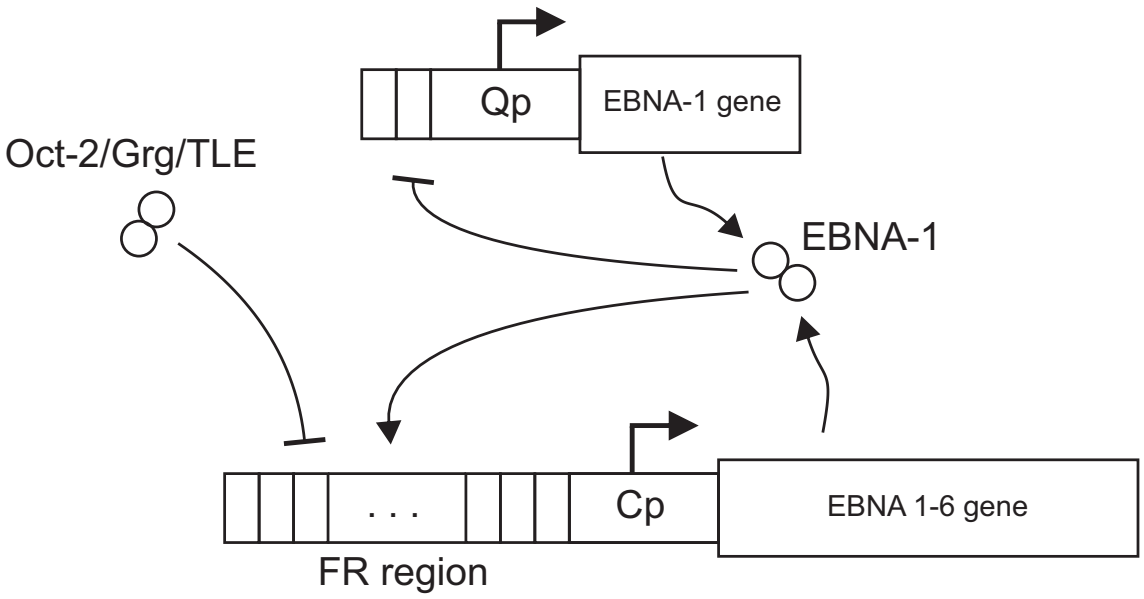

Figure 1: A simple scheme of the gene regulatory network responsible for the regulation of the switching mechanism in the Epstein-Barr virus infected B-lymphocytes.

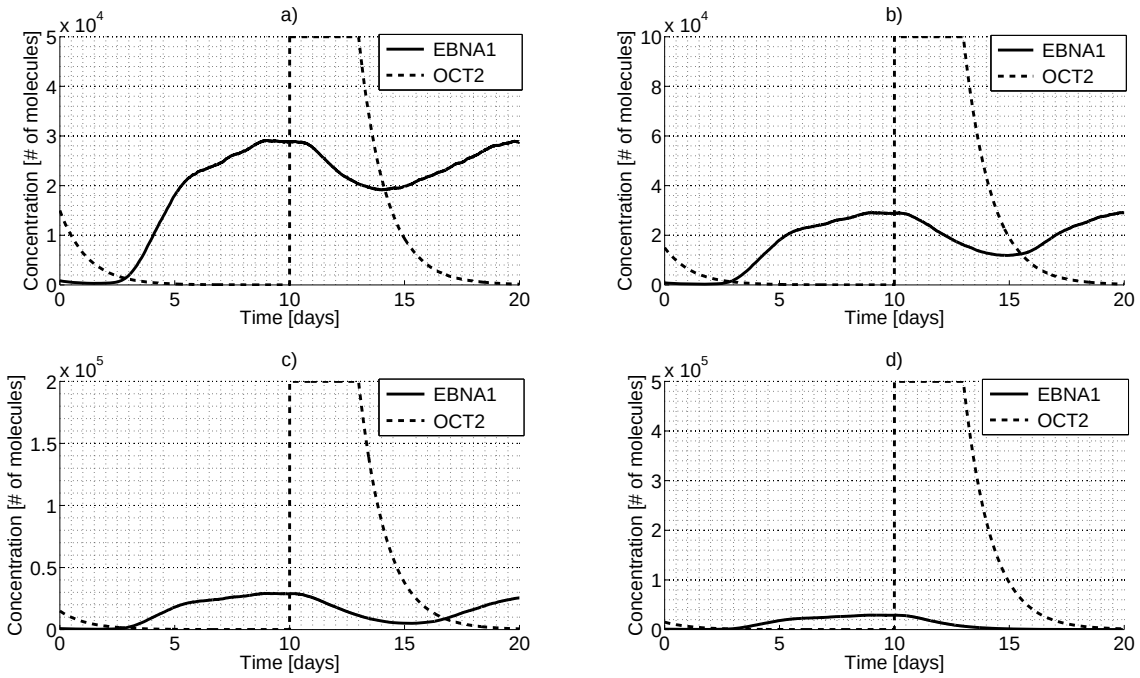

Figure 2: The Epstein-Barr virus genetic switch triggered by different Oct-2 concentrations a) 50000 molecules, b) 100000 molecules, c) 200000 molecules and d) 500000 molecules.

n.	**fast reactions**	**type**	k_i
(1)	$Cp + EBNA1 \xrightarrow{k_1} Cp^*$	Cp promoter activation	$9.2462 * 10^{-12} s^{-1}$
(2)	$Cp^* \xrightarrow{k_2} Cp + EBNA1$	Cp - EBNA1 dissociation	$1.5 * 10^{-11} s^{-1}$
(3)	$Cp + Oct2 \xrightarrow{k_3} Cp'$	Cp promoter inhibition	$9.2462 * 10^{-12} s^{-1}$
(4)	$Cp' \xrightarrow{k_4} Cp + Oct2$	Cp - Oct2 dissociation	$2.5 * 10^{-9} s^{-1}$
(5)	$Qp + EBNA1 \xrightarrow{k_5} Qp'$	Qp promoter inhibition	$9.2462 * 10^{-12} s^{-1}$
(6)	$Qp' \xrightarrow{k_6} Qp + EBNA1$	Qp - EBNA1 dissociation	$2.1 * 10^{-10} s^{-1}$

Table 1: Fast reactions in the Epstein-Barr virus genetic switch model.

n.	slow reactions	type	k_i
(1)	$Cp^* \xrightarrow{k_{trsc16}} Cp^* + mRNA16$	Cp transcription	$0.002069\ s^{-1}$
(2)	$Qp^* \xrightarrow{k_{trsc1}} Qp^* + mRNA1$	Qp transcription	$0.006240\ s^{-1}$
(3)	$mRNA16 \xrightarrow{k_{trsl16}} mRNA16 + EBNA1$	mRNA16 translation	$0.005172\ s^{-1}$
(4)	$mRNA1 \xrightarrow{k_{trsl1}} mRNA1 + EBNA1$	mRNA1 translation	$0.0156\ s^{-1}$
(5)	$EBNA1 + EBNA1 \xrightarrow{k_{dim}} dimEBNA1$	EBNA-1 dimerization	$1.8492 * 10^{-10}\ s^{-1}$
(6)	$mRNA1 \xrightarrow{k_{dis}} mRNA1 + EBNA1$	EBNA-1 dimer dissociation	$10^{-8}\ s^{-1}$
(7)	$mRNA16 \xrightarrow{k_{degRNA}} \emptyset$	mRNA16 decay	$1.9254 * 10^{-5}\ s^{-1}$
(8)	$mRNA1 \xrightarrow{k_{degRNA}} \emptyset$	mRNA1 decay	$1.9254 * 10^{-5}\ s^{-1}$
(9)	$EBNA1 \xrightarrow{k_{degE}} \emptyset$	EBNA-1 decay	$9.627 * 10^{-6}\ s^{-1}$
(10)	$dimEBNA1 \xrightarrow{k_{degE}} \emptyset$	EBNA-1 dimer decay	$9.627 * 10^{-6}\ s^{-1}$
(11)	$Oct2 \xrightarrow{k_{degO}} \emptyset$	Oct-2 decay	$9.627 * 10^{-6}\ s^{-1}$

Table 2: Fast reactions in the Epstein-Barr virus genetic switch model.

parameter	**value**	**reference**
Nuclear volume V	1.796 $*10^{-13}$ lt.	(Werner *et al.* , 2007b)
mRNA half-life $\tau_{\frac{1}{2}mRNA}$	10 hr	(Yang *et al.* , 2003)
EBNA-1 amino acid length	619 - 641	(Ceccarelli and Frappier, 2000; Goldsmith *et al.* , 1993; Ohara *et al.* , 2000)
EBNA-1 - FR binding dissociation constant (K_{dEFR})	$15 * 10^{-12} M$	(Ambinder *et al.* , 1990; Werner *et al.* , 2007b)
EBNA-1 transcription rate	12 b.p. s^{-1}	(Tsurumi, 1991)
EBNA-1 half-life $\tau_{\frac{1}{2}mRNA}$	$20 - 48$ hr	(Davenport and Pagano, 1999; Levitskaya *et al.* , 1997)

Table 3: Parameters used in the Epstein-Barr virus genetic switch model.