# Cooperative action in eukaryotic gene regulation: Physical properties of a viral example

Maria Werner, LiZhe Zhu, and Erik Aurell
*Department of Computational Biology, KTH—Royal Institute of Technology, AlbaNova University Center, SE-106 91 Stockholm, Sweden*
(Received 14 June 2007; published 18 December 2007)

The Epstein-Barr virus (EBV) infects more than 90% of the human population, and causes glandular fever as well as several more serious diseases. It is a tumor virus, and has been widely studied as a model system for cell transformation in humans. A central feature of the EBV life cycle is its ability to persist in human B cells in different latency states, denoted latency I, II, and III. In latency III the host cell is driven to cell proliferation and hence expansion of the viral population without entering the lytic pathway, while the latency I state is almost completely dormant. We here study the effective cooperativity of the viral C promoter, active in latency III EBV cell lines. We show that the unusually large number of binding sites of two competing transcription factors, one viral and one from the host, serves to make the switch sharper (higher Hill coefficient), either by cooperative binding between molecules of the same species when they bind, or by competition between the two species if there is sufficient steric hindrance.

DOI: 10.1103/PhysRevE.76.061909 PACS number(s): 87.16.Yc, 87.17.Aa, 64.60.Cn, 05.90.+m

## I. INTRODUCTION

Genetic switches, mainly in bacteria, have recently interested statistical physicists, and work in this direction has been extensively reviewed in [1,2]. The first thermodynamic model of gene regulation aimed to describe the lytic-lysogenic control switch in bacteriophage λ, and this system has been studied by several groups after the pioneering work of Shea and Ackers [3–5]. The fundamental assumption behind the thermodynamic models is that gene transcription, the copying of a stretch of DNA into mRNA, is either "on" or "off". This state of transcription depends on whether certain gene specific DNA binding proteins—transcription factors—are bound, or not, to the promoter region of the gene. A gene may be controlled by one or more transcription factors, each having a varying number of binding sites in the promoter region. The action of the transcription factor may in turn be either inhibitory or excitatory. Inhibition can arise from blocking access of the RNA polymerase to the transcription start site, while a stimulating effect is obtained if the bound factor stabilizes the polymerase-DNA complex. DNA looping, where distantly bound transcription factors interact and affect transcription, is also possible.

At a given transcription factor concentration, each possible state of promoter bound factors occurs with a probability given by a grand canonical ensemble formula. The promoter region with the binding sites (with or without transcription factors) corresponds to the small system, and the cytoplasm, with a large number of transcription factors moving around, serves as the reservoir. Quite often transcription factors bind in dimer (or multimer) form, in which case the relevant concentration is determined by balance from the total concentration. In summary, the rate of transcription is a nonlinear, sometimes quite complicated, function of the concentrations of the transcription factors regulating the gene.

One important property in gene regulation is cooperativity. If a single copy of a protein molecule in monomer form were to (positively) regulate a certain gene, the activity of that gene would follow the well-known Michaelis-Menten curve. The transcription rate would then be proportional to the concentration of the regulating molecule, up to a threshold above which it would level off. In other words, there would be appreciably high transcription even at very low concentrations of the regulating protein. The rationale for transcription factor often binding in multimer form, and of multiple DNA binding sites enabling cooperative interactions, is therefore assumed to be that it results in a sharper, more "all-or-nothing" switch [6].

Multiple binding sites for one and the same transcription factor are common in promoter regions. The object of this paper is one particular viral example of no less than 20 binding sites for a viral factor, where transcriptional activity has been observed to require eight bound molecules [7,8] (see Sec. II below). In addition, these sites are interleaved with an equal number of binding sites of a host transcription factor, imposing the opposite effect. In a previous contribution [9], we introduced, for reasons of computational simplicity, a thermodynamic model of this promoter switch ignoring eventual cooperative bindings and allowing only some steric hindrance. Although direct experimental evidence is lacking, cooperative bindings of the viral transcription factor at this promoter is likely to be present, as well as more extensive blocking scenarios due the closely spaced sites. Both these mechanisms are likely to affect the sharpness of a regulatory element [10,11].

We show in this paper that while cooperative protein interactions is one way to achieve effective cooperativity of the switch, accounting for full steric hindrance (blocking) of one species of molecules on the other is an equally effective one. Therefore, a possible functional role of the alternating pattern of binding sites could be to increase effective cooperativity when the promoter architecture does not allow for cooperative molecular interactions.

The paper is organized as follows: in Sec. II we describe our example, and in Sec. III we describe our model of cooperativity and competition in this example. In Sec. IV we summarize and discuss our results.

 

## II. EPSTEIN-BARR VIRUS, THE EBNA-1 PROTEIN, AND THE C PROMOTER

The Epstein-Barr virus (EBV) belongs to the gammaherpes virus family, with relatives among other primate lymphocryptoviruses, and has likely coevolved with man for a very long time [12]. Although not discovered until the 1960s, it is now known to infect more than 90% of the human population. The infection is asymptomatic if it occurs early in life, while later infection may result in glandular fever, also known as infectious mononucleosis or "the kissing disease." The virus infects new hosts by virus particles shed from epithelial cells in the throat, and can persist in the host blood B cells for long times, in at least three distinct latent states known as latency I, II, and III. EBV is medically important primarily because some cancer forms are invariably associated with the viral infection [13].

The most vital EBV protein is EBNA-1, a transcription factor involved in replication, episome partitioning, as well as gene regulation [14]. In latency I, EBNA-1 is produced from RNA transcripts originating from the Q promoter on the EBV genome. EBNA-1 downregulates transcription from Qp by binding to sites downstream of the transcription start site [15]. In latency III, on the other hand, EBNA-1 is produced together with five other proteins by alternative splicing of a longer RNA transcribed from the EBV C promoter (Cp) [16]. EBNA-1 positively regulates Cp activity by binding to the "family-of-repeats" (FR) region, positioned upstream of the start site [17]. The physical description of this regulatory element is the topic of the present paper.

The FR region consists of 20 consecutive binding sites for EBNA-1 [18]. There are minor variations in the DNA sequence among these sites, but they are all experimentally verified, and approximately equally strong, binding sites [19]. Comparing promoter activity, from constructs with a varying number of binding sites in FR, revealed that at least eight sites are necessary to have full transcriptional activation [7,8] (see Table I). Recent studies have identified an equal number of octamer binding sites at FR, juxtaposed with the EBNA-1 sites [20]. The action of the human transcription factor Oct-2, identified as binding to these octamer sites, can, when bound alone, activate Cp transcription [21]. However, when bound in complex with the cofactor Groucho/TLE, it has been shown to downregulate Cp activity [22]. Supported by displayed altered levels of Oct-2 in the two latency forms, Oct-2 together with Groucho/TLE is believed to mainly act as an inhibitor *in vivo*.

In summary, the Cp activity is largely regulated by the binding of two species of molecules, EBNA-1 and Oct-2. They each can bind to 20 sites, and have antagonistic effects when bound. Due to the closely spaced binding sites, Oct-2 and EBNA-1 compete for binding to FR. It is, however, not experimentally known if one bound Oct-2 blocks out one or both of the neighboring sites for EBNA-1, and vice versa. The other unknown aspect is whether there exists cooperative binding between EBNA-1 proteins at FR, and if so, the strength of these interactions. We have therefore explored the effects of cooperative binding and blocking, with emphasis on how the effective cooperativity of the promoter switch is affected, i.e., the sharpness of the switch.

TABLE I. Activity of the C promoter in EBV mutants with different numbers of binding sites for EBNA-1 in the family of repeats site, adapted after [8]. The activity level is relative to control.

| Number of sites | Activity |
|---|---|
| 20 | 280 |
| 19 | 229 |
| 17 | 226 |
| 14 | 169 |
| 12 | 206 |
| 11 | 169 |
| 8 | 87 |
| 6 | 19 |
| 5 | 19 |
| 4 | 11 |
| 3 | 3.3 |
| 2 | 2.1 |
| 1 | 1.2 |
| 0 | 3.3 |

## III. COOPERATIVE BINDING AND COMPETITION

The general thermodynamic framework is the following. Suppose a number of transcription factors $TF_1, TF_2, \ldots, TF_m$ can bind in different states indexed by $s$ around the start of a gene. The number of transcription factors of type $TF_i$ bound in state $s$ is $n_i(s)$, the association free energy is $\Delta G_s$, and the rate of transcription of the gene is $R_s$. Suppose further $[TF_i]$ is the concentration of transcription factor $TF_i$ in the surrounding cytoplasm, in the form in which this transcription factor binds. Then the binding sites, with or without bound transcription factors, can be considered a small system, exchanging particles (transcription factors) and energy with the larger reservoir. The probability of the small system being in state $s$ is

$$P_s \propto [TF_1]^{n_1(s)} \cdots [TF_m]^{n_m(s)} \exp\left(-\frac{\Delta G_s}{RT}\right), \tag{1}$$

and the net average rate of transcription is

$$R([TF_1], \ldots, [TF_m]) = \sum_s R_s P_s. \tag{2}$$

The key assumption behind Eq. (2) is that the time scale at which the probabilities in Eq. (2) equilibrate is much faster than the time scales at which the concentrations $[TF_1], [TF_2], \ldots, [TF_m]$ change appreciably. The transcriptional activity also depends on the polymerase attachment to the core promoter site, aggregated into the effective parameter $R_s$ in Eq. (2). Whether to explicitly include the polymerase binding in the state description depends somewhat on the promoter architecture and the complexity of the system. Here, we have a complex eukaryotic system with the operator sites distant from the polymerase attachment site, not allowing direct competition. Hence, the polymerase binding is not explicitly included in our description.

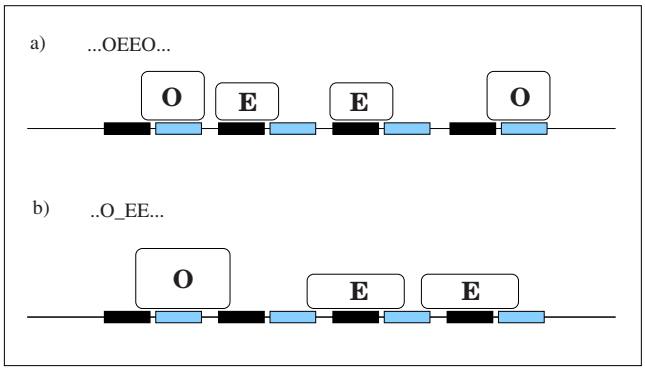


FIG. 1. (Color online) Illustration of the two blocking scenarios. EBNA-1 ($E$) binds black sites while Oct-2 ($O$) binds blue sites. (a) The single-side blocking model where a bound $E$ blocks binding of $O$ to the closest site to the right, and a bound $O$ blocks $E$ binding to the closest site to the left. (b) The double-side blocking model where one bound $E$ or $O$ blocks the opposite molecular species binding on both neighboring sites.

In the present example, states can be labeled by $n$, the number of EBNA-1 molecules bound, $k$, the number of Oct-2 molecules bound, $n_1$, the number of cooperative bindings between bound EBNA-1 molecules, and $k_1$, the number of cooperative bindings between bound Oct-2 molecules. Every such state has a binding free energy of

$$\Delta G_{n,k,n_1,k_1} = nE_E + kE_O + n_1E_{E1} + k_1E_{O1}, \tag{3}$$

where $E_E$=−15.45 kcal/mol [19] and $E_O$=−12.28 kcal/mol [23] are the known binding free energies of EBNA-1 and Oct-2 to binding sites in FR, and $E_{E1}$ and $E_{O1}$ are the unknown cooperative binding energies. In the numerical experiments described in this paper we only examine EBNA-1 cooperativity. $E_{E1}$ is proportional to $E_E$ in the range from 0% (no cooperativity) up to 40%. The total probability of the states with given values of $n$, $k$, $n_1$, and $k_1$ is hence

$$P_{n,k,n_1,k_1} \propto \xi(n,k,n_1,k_1)[E]^n[O]^k \exp\left(-\frac{\Delta G_{n,k,n_1,k_1}}{k_BT}\right), \tag{4}$$

where $\xi(n,k,n_1,k_1)$ is the number of such states, and the overall probability of transcription is

$$P = \sum_{n=8}^{N} \sum_{k=0}^{N-n} \sum_{n_1=0}^{n-1} \sum_{k_1=0}^{N-n-1} P_{n,k,n_1,k_1}, \tag{5}$$

where $N$ is the number of binding sites. Transcription is defined active only if eight or more EBNA-1 are bound, with a constant rate of transcription.

As described briefly in the Introduction, one can imagine two plausible blocking scenarios at FR. The first and simplest, is that each molecule bound hinders binding of the competitive species to the closest neighboring site on one side. This is referred to as single-side blocking [Fig. 1(a)]. The other scenario is that each bound molecule sterically hinders both neighboring sites for the other molecule; a double-sided blocking [Fig. 1(b)]. The blocking method naturally affects the number of possible bound configurations, as seen in Eq. (5). The upper bound in the sum over $k$ is $N-n$ in the single-side blocking model, but at most $N-n-1$ in the double-side blocking model for all $n$ greater than zero. Similarly, the sums over $n_1$ and $k_1$ may effectively go over smaller ranges, e.g., in the double blocking scenario with both molecules bound and $n+k=N-1$ all EBNA-1 and Oct-2 molecules bind together in two groups, hence $n_1=n-1$ and $k_1=k-1$.

Brute-force counting of $\xi(n,k,n_1,k_1)$ is not feasible as the number of states in this model is up to $3^{20} \approx 3.4 \times 10^9$ (in the model with single-side blocking only). Efficient calculation of $\xi(n,k,n_1,k_1)$ involves two aspects. First, elementary combinatorics is used to build up a paradigm "balls-baskets" problem. It counts, under different constraints, the number of ways that one can put a certain number of balls into another number of baskets. Second, we find a way that can describe efficiently all effects including double-side blocking, cooperativity, and a combination of both in a three-step algorithm:

(1) Construct a backbone sequence ($S0$) made up by two types of baskets ($b_E$, $b_O$), the two types of molecules.

(2) Distribute $n$ Es and $k$ Os among these baskets, forming a sequence ($S1$) consisting only of $E$ and $O$.

(3) Consider the front, end, and the $n+k-1$ in-between positions of $S1$ as baskets ($b_\phi$) for empty binding sites $\phi$. Insert $N-(n+k)$ empty sites into these positions and get the final pattern ($S2$).

By setting $N$=20, the actual number of sites is reduced by half, and the single-sided blocking model is the default. The double-side blocking is realized by setting the $b_\phi$ between an "OE" segment in $S2$ as must-be-filled baskets [Fig. 1(b)]. The number of cooperative units, $n_1$, are counted by recording number of "EE" in $S2$, minus the number of $b_\phi$ that have been filled with $\phi$.

To examine the effective cooperativity in the transition from $P \approx 0$ to $P \approx 1$ we compute the Hill coefficient. This is the logarithmic derivative of the ratio of probability of transcription to the probability of no transcription, with respect to the logarithm of the free ligand concentration. The Hill coefficient is a function of the ligand concentration, but the effective Hill coefficient is customarily taken at half saturation as follows:

$$\frac{d \lg \frac{P}{1-P}}{d \lg[E_{free}]} \quad \text{at} \quad P = 0.5. \tag{6}$$

In this paper we explore the Hill coefficient functions to see how blocking and cooperative binding influence the effective cooperativity of the switch. There are three cases studied: (1) cooperative binding of EBNA-1 and no competing molecular species, (2) cooperative binding of EBNA-1 with single-side blocking between the competing species, and (3) cooperative binding of EBNA-1 with double-side blocking between the competing species.

## IV. EFFECTIVE COOPERATIVITY OF THE SWITCH

One convenient way to visualize the cooperativity of the switch is as the ratio $\frac{P}{1-P}$ vs the local Hill coefficient given as $\frac{d \lg \frac{P}{1-P}}{d \lg[E_{free}]}$. For very high and very low concentrations of

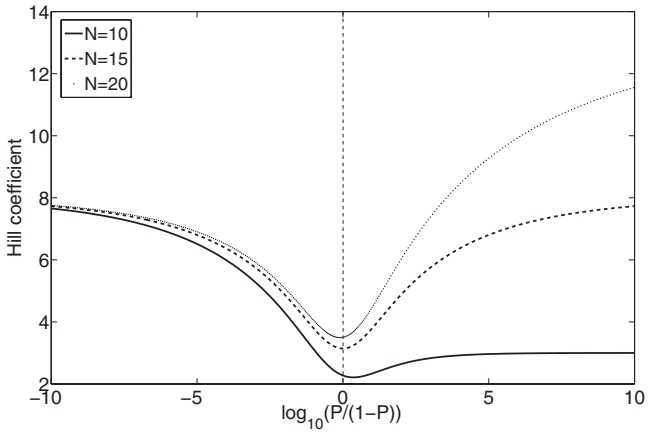


FIG. 2. Hill coefficient curves for varying numbers of binding sites $N$ in the promoter. The promoter is defined as active when more than eight EBNA-1 are bound. Hence, the Hill coefficient at the limit of low EBNA-1 concentrations (low $P$) is 8, independent of $N$. At high $P$, the Hill coefficient approaches $N-7$, yielding higher effective cooperativity for longer promoters.

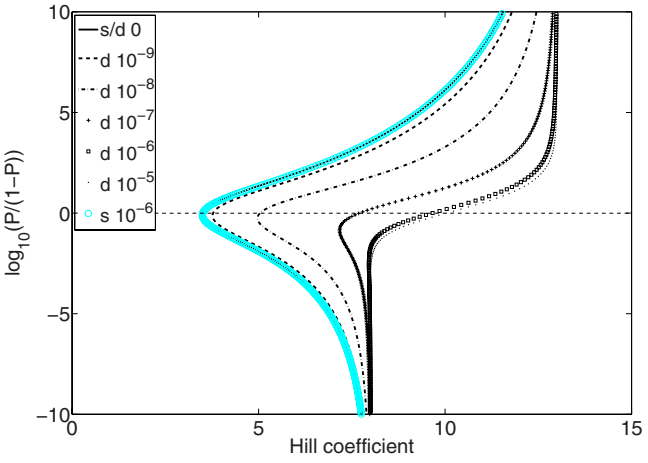


FIG. 3. (Color online) Hill coefficient curves for both the single- and double-sided blocking models, for various concentrations of Oct-2 (legends show Oct-2 concentration). Without Oct-2, both models have the same coefficient curve (solid line). For the single-side blocking, an increase in Oct-2 does not lead to any effect on the Hill coefficient (cyan circles). On the other hand, in the double-sided blocking model, an increase in Oct-2 concentration dramatically shifts the curve (five black marker lines). For saturating levels of Oct-2, the effective Hill coefficient approaches 10.5.

EBNA-1, corresponding to very large and very small values of $P$, it is easy to see that in our model $1-P \sim A[E_{free}]^{7-N}$, respectively, $P \sim B[E_{free}]^8$. $A$ and $B$ are constants, and $N$ is the total number of binding sites in FR. Accordingly, the extreme local Hill coefficients are $N-7$ and 8. Figure 2 illustrates this limit behavior for three values of $N$.

In the region of main interest, where $P \sim \frac{1}{2}$, the Hill coefficient curves show very different behavior for the three models. Without any cooperative interactions, and without competition, the effective Hill coefficient is substantially lower than both its limits. This baseline function for the system has an effective Hill coefficient of 3.5 (Fig. 3, circled lines). This low Hill coefficient remains even with competition from Oct-2 binding, for the single-side blocking, the effective cooperativity practically insensitive to Oct-2 levels. On the contrary, competition with double-side blocking dramatically alters the shape of the Hill coefficient curve, to a sigmoidal interpolation between the limits 8 and $N-7$. The effective Hill coefficient then changes from 3.5 up to 10.5, for saturating amounts of Oct-2 (Fig. 3, dotted lines).

From a theoretical point of view, the thermodynamic model of the switch is a (finite, one-dimensional) Ising-type model with three states at each site: bound by EBNA-1, bound by Oct-2, or free. The only complication in computing the "ON" probability ($P$) is that only states with enough bound EBNA-1 count, which mixes in a global variable in the elementary statistical mechanical model. The single-blocking results can, however, be readily understood. With no cooperative binding and only single blocking, one can sum over $k$ in Eq. (5) to obtain the model studied in [9], that is,

$$P_n \propto \binom{N}{n} [E]^n e^{-nE_E/k_BT} (1 + [O]e^{-E_O/k_BT})^{N-n}. \tag{7}$$

Including the normalization this means

$$P_n = \binom{N}{n} \frac{z^n}{(1+z)^N}, \quad z = \frac{[E]e^{E_E/k_BT}}{(1+[O]e^{-E_O/k_BT})}, \tag{8}$$

and the ratio between ON and OFF probabilities is therefore a function of the variable $z$ only.

$$\frac{P}{1-P} = \frac{\sum_{n=8}^{20} \binom{N}{n} z^n}{\sum_{n=0}^{7} \binom{N}{n} z^n} = f(z). \tag{9}$$

The local Hill coefficients are

$$\frac{d \lg \frac{P}{1-P}}{d \lg [E_{free}]} = \frac{d \lg f(z)}{d \lg z}, \tag{10}$$

which like the ratio $\frac{P}{1-P}$ depends on the concentration of the second molecule $[O]$ only through $z$. The effective cooperativity in the model without cooperative binding and only single blocking hence does not depend on $[O]$, as shown in the curves in Fig. 3. The Hill coefficient at $P \approx \frac{1}{2}$ can be estimated by approximating the binomials with a Gaussian distribution, i.e.,

$$P \approx C^{-1} \int_{\bar{x}+c\sigma}^{\infty} \exp\left( -\frac{1}{2\sigma^2}(x-\bar{x})^2 + x \log z \right) dx, \tag{11}$$

where, in the case at hand, $C = e^{\sigma^2/2 \log^2 z} \sqrt{\frac{2\pi}{\sigma^2}}$, $\bar{x} = \frac{N}{2} = 10$, $\sigma = \sqrt{\frac{N}{4}} = \sqrt{5}$, and $c = -\frac{2}{\sigma}$. Half-filling is achieved at $z^* = e^{c/\sigma}$, and the Hill coefficient is $\sqrt{8\sigma^2/\pi} \approx 3.6$, which accords quite well with the minimum value in Fig. 3. The switch is therefore much less sharp than the limits of 8 and $N-7=13$, at, respectively, $P \approx 0$ and $P \approx 1$ could have led one to believe.

We note that the sharpness increases with $N$ (as long as the threshold stays around $N/2$), but only as the square root of $N$: more than a hundred consecutive binding sites are necessary to reach a Hill coefficient of about ten in a model of this kind.

In the model with double blocking on the other hand, clearly the effective cooperativity can be much larger, and also depend on $[O]$. That is easy to understand in the limit where $[O]$ is large; if so EBNA-1 and Oct-2 compete for binding sites, and the possibility that a site is left free can be disregarded. For every state with a certain number of bound EBNA-1, the most probable configuration will be the one with the maximum bound Oct-2, since all other configurations are dampened by at least a factor $[O]e^{-E_O/Tk_B}$. Therefore, if $n$ copies of EBNA-1 are bound, then also $N-n$ copies of Oct-2 are bound, altogether in the pattern $EEEE\cdots OOOO$ with statistical weight

$$P_n = \frac{x^n}{1 + x + x^2 + \cdots x^N}, \quad x = \frac{[E]e^{-E_E/k_BT}}{[O]e^{-E_O/k_BT}}. \tag{12}$$

The Hill coefficient is then only a function of $x$, such that the curve in Fig. 3 has a limit when $[O]$ becomes large, and the value of the Hill coefficient at, e.g., $x=1$ then lies between the limits of 8 and 13. Competition with a second molecule therefore makes the switch sharper for double-sided blocking, in contrast to the situation in single-sided blocking.

The case with cooperativity can be understood qualitatively, using the helix-coil model of protein physics. Without Oct-2, the statistical model can be written as a factor $s = [E]e^{(-E_E+E_{coop})/k_BT}$ for each letter $E$, and a penalty $\sigma = e^{E_{coop}/k_BT}$ for every start letter of a string of $E$'s. In an infinitely long string, the *fraction* of letters $E$ as well as the *frequency of initiation* of a string of $E$'s are calculated from the leading eigenvalue of the transfer matrix $T=\begin{pmatrix} 1 & 1 \\ \sigma s & s \end{pmatrix}$ [24]. In our case, the interesting region is obviously when that fraction is around 40%, as eight sites out of 20 need to be filled to have transcription from Cp. If $\sigma$ is close to one, cooperative binding is weak, and the switch is similar to the single-blocking case discussed above. If, on the other hand, $\sigma$ is much less than one, the expected fraction of letters $E$ can be larger than 40%, while the expected frequency of initialization of a string of $E$'s is less than once in 20 sites. Eventually, we would expect that either all 20 sites are bound, or no sites in FR be bound. This describes a situation where all 20 molecules bind in simultaneously, in which case the Hill coefficient is 20.

Even though the effects are similar, the increase in Hill coefficient from cooperative binding and from double-sided blocking hence do not come about in exactly the same way. Figure 4 displays the curves for five different cooperative binding strengths, when no Oct-2 is competing for the FR sites. The range of cooperative strength here is from 0% up to 40% of the DNA affinity, i.e., ≈6.2 kcal/mol, resulting in an increase in effective Hill coefficient from 3.5 to 16, which is larger than can be achieved in the double-blocking scenario. In cooperative binding the increased effective cooperativity is an energetic effect, while in double blocking it is an entropic effect. In terms of the helix-coil model, double

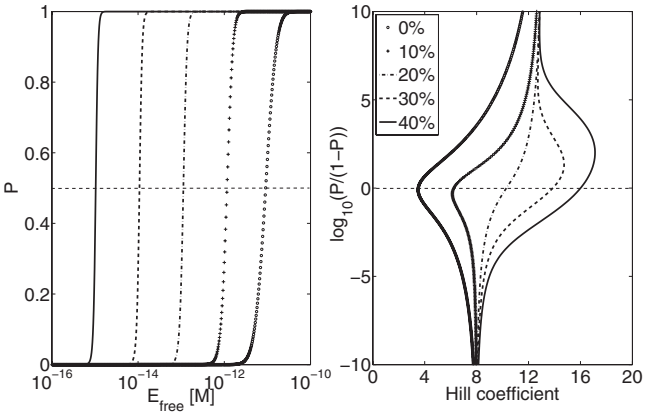


FIG. 4. Hill coefficient curves with no Oct-2 and increasing strength of cooperative binding between EBNA-1. The cooperative binding is varied from 0%–40% of the DNA binding strength of EBNA-1 (0–6.2 kcal/mol). With no competition and only added cooperative interaction, the effective Hill coefficient changes dramatically from 3.5 up to 16 for the largest added cooperative binding energy.

blocking at high Oct-2 concentration has, relative to a background of all sites bound by $O$'s, a factor $s=\frac{[E]e^{-E_E/k_BT}}{[0]e^{-E_O/k_BT}}$ for each letter $E$, and a penalty $\sigma=\frac{1}{[0]e^{-E_O/k_BT}}$ for a start letter of a string of $E$'s (since sequence $OE$ is forbidden). This is not precisely the same as in cooperative binding, because there is an exception if the string of $E$'s starts with the first site, in which case there is no penalty. Therefore, the mixed sequences of $E$'s and $O$'s, which begin with $E$'s, carry less penalty in the double-blocking scenario than in cooperative binding.

Single blocking with cooperative binding and Oct-2 present can again be understood qualitatively from a transfer matrix

$$T = \begin{pmatrix} 1 & 1 & 1 \\ s & \sigma s & s \\ \tau & \tau & \tau \end{pmatrix},$$

where $\tau=[O]e^{-E_O/k_BT}$ is the factor for each letter $O$. Qualitatively the results are similar to single blocking without $O$, but the fraction of letters $E$ and the frequency of initiation of a string of $E$'s now depend on $\tau$. As a result, an additional cooperative binding of EBNA-1 does not have the same impact when Oct-2 levels are high. Instead of a fourfold change, the effective Hill coefficient is now only doubled, from 3.5 to 7 (compare Figs. 4 and 5, solid lines). This is to be compared with the double-sided blocking model, where even no cooperative bindings have a relatively high effective cooperativity. Adding up to 40% cooperative binding strength, the Hill coefficient is almost doubled, from 10.5 to 18 (Fig. 3).

A conclusion to draw from this investigation is that to create an effective switch for genetic control, this type of architecture, with alternating binding sites for two antagonistic factors, can be one approach. An experimental study by Rossi *et al.* on a promoter with seven repetitive sites, each

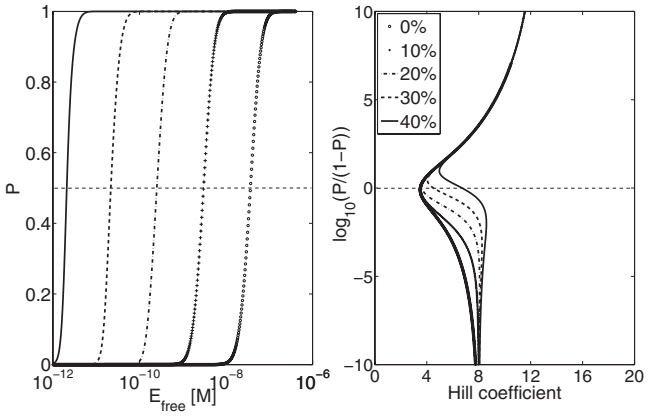


FIG. 5. Hill coefficient curves for the single-sided blocking model at a high concentration of Oct-2, and different strength of cooperative binding between EBNA-1. The cooperative binding is varied from 0%–40% of the DNA binding strength of EBNA-1. The effect of adding cooperative binding energy for EBNA-1 doubles the effective Hill coefficient from 3.5 to 7.

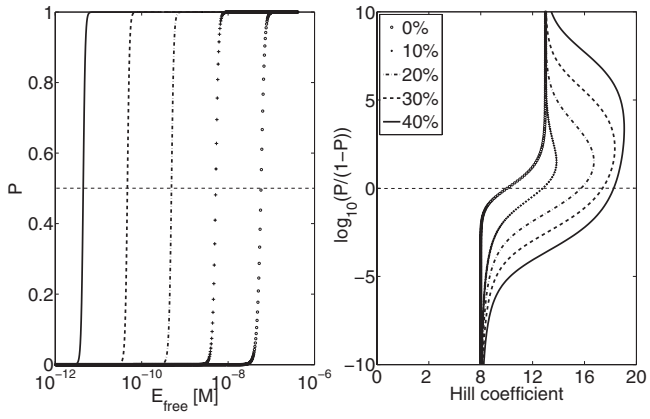


FIG. 6. Hill coefficient curves for the double-sided blocking model at high concentration of Oct-2, and different strength of cooperative binding between EBNA-1. The cooperative binding is varied from 0%–40% of the DNA binding strength of EBNA-1. The double blocking in itself gives a relatively high effective Hill coefficient, and the extra cooperative interactions almost double this coefficient from 10.5 up to 18.

binding an activator or repressor, showed that the competition doubled the effective Hill coefficient [11]. A direct quantitative comparison between theirs and our study is not straightforward since the conditions for activation and repression are not known in the same detail in their system as for the Epstein-Barr C promoter. Qualitatively, the effect observed in [11] could correspond to a single-sided blocking model with cooperative binding between the activating or repressing factors (or both), or a double-sided blocking with or without cooperative binding. For EBV, the FR region is known for its enhancer function, as well as forming a looped structure with another EBNA-1 binding region on the viral genome; the dyad symmetry (DS) [25,26]. This structure is involved in replication initiation control. If the EBNA-1 binding sites in FR were to be arranged in the same manner as in DS, i.e., much closer in space, there might be cooperative bindings forming even at FR. However, since FR also seem to play an important role in forming a looped structure, there might be a structural reason behind these more sparsely placed sites, not enabling the same type of tight interactions. And, as we show here, there is no need for cooperative interactions to get a sharp switch of Cp activity, as long as there is efficient steric hindrance (Figs. 3 and 6).

## ACKNOWLEDGMENTS

We thank Ingemar Ernberg for sharing his knowledge of the Epstein-Barr virus and many discussions on mechanisms and modeling of the EBV lat I/lat III switch. This work was supported by the Swedish Science Council (M.W. and E.A.).

[1] L. Bintu, N. E. Buchler, H. G. Garcia, U. Gerland, T. Hwa, J. Kondev, T. Kuhlman, and R. Phillips, Curr. Opin. Genet. Dev. **15**, 116 (2005).
[2] L. Bintu, N. E. Buchler, H. G. Garcia, U. Gerland, T. Hwa, J. Kondev, T. Kuhlman, and R. Phillips, Curr. Opin. Genet. Dev. **15**, 125 (2005).
[3] M. Shea and G. Ackers, J. Mol. Biol. **181**, 211 (1985).
[4] J. Reinitz and J. R. Vaisnys, J. Theor. Biol. **145**, 295 (1990).
[5] E. Aurell and K. Sneppen, Phys. Rev. Lett. **88**, 048101 (2002).
[6] N. E. Buchler, U. Gerland, and T. Hwa, Proc. Natl. Acad. Sci. U.S.A. **100**, 5136 (2003).
[7] D. A. Wysokenski and J. L. Yates, J. Virol. **63**, 2657 (1989).
[8] H. Zetterberg, C. Borestrom, T. Nilsson, and L. Rymo, Int. J. Oncol. **25**, 693 (2004).
[9] M. Werner, I. Ernberg, J. Zou, J. Almqvist, and E. Aurell, BMC Syst. Biol. **1**, 40 (2007).
[10] R. Hermsen, S. Tans, and P. R. ten Wolde, PLOS Comput. Biol. **2**, e164 (2006).
[11] F. M. V. Rossi, A. M. Kringstein, A. Spicher, O. Guicherit, and H. M. Blau, Mol. Cell **6**, 723 (2000).
[12] C. S. Gerner, A. Dolan, and D. J. McGeoch, Virus Res. **99**, 187 (2004).
[13] L. S. Young and A. B. Rickinson, Nat. Rev. Cancer **4**, 757 (2004).
[14] E. R. Leight and B. Sugden, Rev. Med. Virol. **10**, 83 (2000).
[15] J. Sample, E. B. Henson, and C. Sample, J. Virol. **66**, 4654 (1992).
[16] M. Bodescot, M. Perricaudet, and P. J. Farrell, J. Virol. **61**, 3424 (1987).
[17] D. Reisman and B. Sugden, Mol. Cell. Biol. **6**, 3838 (1986).
[18] T. Nilsson, H. Zetterberg, Y. C. Wang, and L. Rymo, J. Virol. **75**, 5796 (2001).
[19] R. F. Ambinder, W. A. Shah, D. R. Rawlins, G. S. Hayward, and S. D. Hayward, J. Virol. **64**, 2369 (1990).

[20] J. Almqvist, J. Zou, C. Borestrom, S. Petterson, L. Rymo, and I. Ernberg, J. Gen. Virol. **86**, 1261 (2005).
[21] S. Malin, Y. Linderson, J. Almqvist, I. Ernberg, T. Tallone, and S. Petterson, Nucleic Acids Res. **33**, 4618 (2005).
[22] J. Almqvist, J. Zou, C. Borestrom, L. Rymo, and I. Ernberg, J. Gen. Virol. (to be published).
[23] P. C. Shah, E. Bertolino, and H. Singh, EMBO J. **16**, 7105 (1997).
[24] K. Sneppen and G. Zocchi, *Physics in Molecular Biology* (Cambridge University Press, Cambridge, England, 2005).
[25] L. Frappier and M. O'Donnell, Proc. Natl. Acad. Sci. U.S.A. **88**, 10875 (1991).
[26] W. Su, T. Middelton, B. Sudgen, and H. Echols, Proc. Natl. Acad. Sci. U.S.A. **88**, 10870 (1991).