Supplement Material

Supplement A: Modeling transcription activity of *PahdICR* promoter

We here present a quantitative model of *PahdICR* transcription control by C.AhdI. The model is based on the following set of reaction:

$$M + M \underset{K_1}{\rightleftarrows} D \quad (1.1)$$

$$DNA + RNAP \underset{K_2}{\rightleftarrows} RNAP - DNA \quad (1.2)$$

$$D + DNA \underset{K_3}{\rightleftarrows} D - DNA \quad (1.3)$$

$$D - DNA + D \underset{K_4}{\rightleftarrows} T - DNA \quad (1.4)$$

$$D - DNA + RNAP \underset{K_5}{\rightleftarrows} D - DNA - RNAP \quad (1.5)$$

In the above equations, we introduced the following notation. M and D denote C.AhdI monomers and dimers respectively; D-DNA denotes C.AhdI dimer bound to the promoter distal (high affinity) C box; RNAP-DNA and T-DNA denote, respectively, complex of operator DNA with RNAP and C.AhdI tetramer; D-DNA-RNAP is the complex consisting of C.AhdI dimer, DNA and RNAP.

The physical meaning of the equations is the following. Equation (1.1) presents dimerization of C.AhdI. Equation (1.2) presents reversible binding of RNA polymerase to the core promoter in the absence of C.AhdI, which is needed to reproduce the experimentally observed small basal transcription rate of C-R genes. Equation (1.3) presents binding of C.AhdI dimer to the promoter distal (high affinity) C box. Equation (1.4) presents recruitment of C.AhdI dimer to the promoter proximal (low affinity) C box, by the dimer that is bound at the promoter distal (high affinity) C box (1.4). Finally, Eq. (1.5) presents the competing reaction of RNAP recruitment to the promoter by C.AhdI dimer that is bound at the promoter distal C box. One should note that the reaction which presents binding of a dimer to the promoter proximal box is neglected, since the promoter proximal box has a much lower binding affinity compared to the promoter distal box.

The equilibrium condition for Eqs. (1.1)-(1.5) leads to:

$$K_1 = \frac{[M]^2}{[D]} \tag{1.6}$$

$$K_2 = \frac{[DNA][RNAP]}{[RNAP-DNA]} \tag{1.7}$$

$$K_3 = \frac{[D][DNA]}{[D-DNA]} \tag{1.8}$$

$$K_4 = \frac{[D-DNA][D]}{[T-DNA]} \tag{1.9}$$

$$K_5 = \frac{[D-DNA][RNAP]}{[D-DNA-RNAP]} \tag{1.10}$$

Similarly as in the previous subsection, we assume that transcription activity φ is proportional to occupancy of the promoter by RNA polymerase:

$$\varphi \propto \frac{[RNAP-DNA]+[D-DNA-RNAP]}{[DNA]+[RNAP-DNA]+[D-DNA-RNAP]+[T-DNA]} \tag{1.11}$$

In simplifying the above expression, we assumed that the concentration of $[D-DNA]$ is much smaller compared to the concentration of $[T-DNA]$, as indicated by the *in-vitro* binding studies. Next, by using Eqs. (1.6)-(1.10), from Eq. (1.11) follows:

$$\varphi([M]) \propto \frac{a+b[M]^2}{1+a+b[M]^2+c[M]^4} \tag{1.12}$$

,where

$$a = \frac{[RNAP]}{K_2},\ b = \frac{[RNAP]}{K_1 K_3 K_5},\ c = \frac{1}{K_1^2 K_3 K_4}.$$

Supplement B: Changes of the model constants with the mutations

We here address how constants *b* and *c* in the model given by Eq. (1.12) change with introducing mutations in the two C boxes. To derive this, we use the relationship between equilibrium binding constant *K* and the free energy of the reaction products. That is, if one starts from the following reaction:

$$A + B \underset{K}{\rightleftarrows} C + D, \qquad (1.13)$$

the equilibrium binding constant is given by:

$$K \sim \exp\left(\Delta G_C + \Delta G_D - \Delta G_A - \Delta G_B\right). \qquad (1.14)$$

In the above expression, ΔG_A, ΔG_B, ΔG_C and ΔG_D correspond, respectively, to the free energies of the complexes A, B, C and D in the units of $k_B T$ (where k_B is the Boltzmann constant and T is temperature). Equation (1.13) and the related expression given by Eq. (1.14) can be straightforwardly generalized to the cases where different number of reactants and/or reaction products is involved. Further, let we denote by $\Delta G_d^{(w)}$ and $\Delta G_p^{(w)}$ the interaction energy of a C.AhdI dimer with, respectively, the wild type promoter distal and promoter proximal C-box. Similarly, let we denote by $\Delta G_d^{(m)}$ and $\Delta G_p^{(m)}$ the interaction energy of a C.AhdI dimer with the operator sequences in which, respectively, the promoter distal and the promoter proximal C-boxes are mutated.

We first start from the case in which the mutations are introduced in the promoter proximal C-box. In that case, the general expression given by Eq. (1.14) leads to the following change in the equilibrium constant K_4 for Eq. (1.4):

$$\frac{K_4^{(m)}}{K_4^{(w)}} = \exp\left(\Delta G_p^{(m)} - \Delta G_p^{(w)}\right) \qquad (1.15)$$

In the above expression, $K_4^{(w)}$ and $K_4^{(m)}$ are the equilibrium binding constants that correspond, respectively, to the wild type operator sequence and to the sequence in which the promoter proximal C-box is mutated.

Since all other equilibrium constants do not change by the mutation in the promoter proximal site, from Eqs (1.12) and (1.15) follows that the change in the constants *b* and *c* is given by the following expression:

$$\frac{b_p^{(m)}}{b_p^{(w)}} = 1 \qquad (1.16)$$

$$\frac{c_p^{(m)}}{c_p^{(w)}} = \exp\left(\Delta G_p^{(w)} - \Delta G_p^{(m)}\right) \qquad (1.17)$$

In the expressions above, the superscripts *(m)* and *(w)* indicate, respectively, mutant and wild-type operator sequences, while the subscript *p* denotes that the mutation is exhibited in the promoter proximal C-box. We therefore conclude that constant *b* does not change

due to the mutations, while constant *c* should decrease. (Note that $\Delta G_p^{(w)}$ is more negative compared to $\Delta G_p^{(m)}$ since the mutations reduce the binding affinity to the promoter proximal C-box).

Similarly to the analysis given above, in the case of the mutation in the promoter distal C-box, the following relations for the changes in the constants are obtained:

$$\frac{b_d^{(m)}}{b_d^{(w)}} = \exp\left(\Delta G_d^{(w)} - \Delta G_d^{(m)}\right) \tag{1.18}$$

$$\frac{c_d^{(m)}}{c_d^{(w)}} = \exp\left(\Delta G_d^{(w)} - \Delta G_d^{(m)}\right) \tag{1.19}$$

From the above two expressions, it follows that both constants decrease due to the mutations in the promoter distal C-box.

Finally, in the case in which mutations are introduced in both the promoter proximal and promoter distal C-box, the following relations for the changes in *b* and *c* hold:

$$\frac{b_{dp}^{(m)}}{b_{dp}^{(w)}} = \exp\left(\Delta G_d^{(w)} - \Delta G_d^{(m)}\right) \tag{1.20}$$

$$\frac{c_{dp}^{(m)}}{c_{dp}^{(w)}} = \exp\left(\Delta G_d^{(w)} + \Delta G_p^{(w)} - \Delta G_d^{(m)} - \Delta G_p^{(m)}\right), \tag{1.21}$$

where the subscript indicates that the mutation is exhibited in both of the two C-boxes. The interpretation of this result is analogous as in the previous two cases.

Supplement C: Modeling the in-vivo dynamics of C-R and M-S loops

We start from the following equations, which are explained in the main text:

$$\frac{dc(t)}{dt} = \varphi\left(C(t)\right) - \lambda c(t) \tag{1.22}$$

$$\frac{dC(t)}{dt} = \alpha c(t) - \beta C(t) \tag{1.23}$$

The equilibrium for the two equations is given by the condition $d\left(C(t)\right)/dt = 0$ and $dc(t)/dt = 0$, which leads to the following expression:

$$\varphi\left(C_{eq}\right) = \frac{\lambda\beta}{\alpha} C_{eq} \tag{1.24}$$

From Eq. (1.24), we see that the value of the equilibrium is determined by the intersection of $\varphi(C)$ curve and the line with the slope $\lambda\beta/\alpha$, as stated in the main text.

Further, we are interested under which conditions is the equilibrium state determined by Eq. (1.24) stable. For the equilibrium state to be stable, the system has to return to the equilibrium when perturbed for a small value. We, therefore, substitute $C(t) \to C_{eq} + \delta C(t)$ and $c(t) \to c_{eq} + \delta c(t)$ in Eqs. (1.22) and (1.23), where $\delta c(t)$ and $\delta C(t)$ are small perturbations of the transcript and protein concentrations respectively. If we use Eq. (1.24) together with $\varphi\left(C_{eq} + \delta C(t)\right) = \varphi\left(C_{eq}\right) + \left.\frac{d\varphi(C)}{dC}\right|_{C_{eq}} \delta C(t)$, the substitution leads to the following two equations:

$$\frac{d\delta c(t)}{dt} = \left.\frac{d\varphi(C)}{dC}\right|_{C_{eq}} \delta C - \lambda \delta c(t) \tag{1.25}$$

$$\frac{d\delta C(t)}{dt} = \alpha \delta c(t) - \beta \delta C(t) \tag{1.26}$$

We can now eliminate $\delta c(t)$ from Eq. (1.26), which leads to:

$$\frac{d^2\delta C(t)}{dt^2} = -(\lambda + \beta)\frac{d\delta C(T)}{dt} + \left(\alpha \left.\frac{d\varphi(C)}{dC}\right|_{C_{eq}} - \lambda\beta \right) \delta C(t) \tag{1.27}$$

Equation (1.27) is an equation of a dumped charmonic oscillator, which exhibits bound motion, under the condition:

$$\frac{\lambda\beta}{\alpha} > \left.\frac{d\varphi(C)}{dC}\right|_{C_{eq}} \tag{1.28}$$

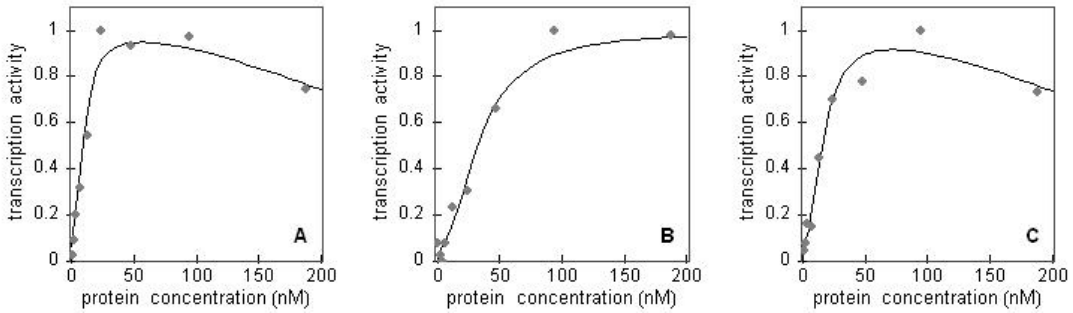
One should note that the equation with the same form as Eq. (1.27), and with the same condition for the bound motion (Eq. (1.28)), is obtained when $\delta C(t)$ (instead of $\delta c(t)$)

is eliminated from Eq. (1.26). Therefore, as stated in the main text, the existence of the stable equilibrium, given by Eq. (1.28), is geometrically equivalent to the condition that the slope of the linear line is larger than the slope of the $\varphi(C)$ curve, at the point of their intersection.

Finally, we want to find how $C(t)$, determined by Eqs. (1.22) and (1.23), changes with time, with the initial conditions $C(0)=0$ and $\frac{dC(0)}{dt}=0$. By eliminating $c(t)$ from Eq. (1.23) and substituting in Eq. (1.22), we obtain:

$$\frac{d^2C(t)}{dt^2}+(\lambda+\beta)\frac{dC(t)}{dt}=(\alpha\varphi(C)-\lambda\beta)C(t) \qquad (1.29)$$

Equation (1.29) is a second order non-linear differential equation, which we numerically solve by using a Runge-Kutta method (MATLAB, MathWorks), to obtain the solutions shown in Fig. 9B**.**


transcription activity
protein concentration (nM)
A
B
C
0
0.2
0.4
0.6
0.8
1
0
50
100
150
200