Binding of Activators and Repressors to DNA

Part I: Equilibria

Peter Schuster[a]

[a] *Theoretical Biochemistry Group, Institut für Theoretische Chemie der Universität Wien, Währingerstraße 17, A-1090 Wien, Austria*
and
Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

Summary: The binding equilibria of activator and repressor to the regulatory segments of a gene are studied by conventional thermodynamics. Five different scenarios are analyzed: (i) A single regulator binds to the DNA, (ii) activator and repressor bind competitively to a single site, (iii) activator and repressor bind to two different sites, (iv) two activator molecules, and (v) four activator molecules are required for gene expression. Independent, cooperative, and anti-cooperative binding are distinguished for the cases (iii), (iv), and (v). Simplified functions with computable maximal deviations from the exact solutions are derived for the equilibrium concentrations in order to provide a rigorous basis for modelling genetic regulation.

1 Introduction and notation

The basic scenario that underlies the calculations presented here is sketched in figure 1 and has been taken from the booklet by Ptashne & Gann [1]. Two classes of molecular effectors, activators and repressors, decide on the transcriptional activity of a gene, whose activity is classified according to three states: (i) 'Naked' DNA is commonly assumed to have a low or basal transcription activity (basal state), (ii) transcription rises to the normal level when (only) the activator is bound to the regulatory region of the gene (active state), and (iii) complexes with repressor are inactive no matter whether the activator is present or not (inactive state). In this first part we consider only the binding equilibria of regulatory molecules to binding sites within the non-translated regulatory regions of genes and we shall leave all kinetic issues to the second part of this technical report. We shall discuss four different problems in separate sections: (i) single regulator binding, (ii) competitive binding of two different regulator molecules to the same site, (iii) binding of two ligands to two sites, and (iv) binding of one ligand at two or more

binding sites. At the beginning we consider a single binding site and a single regulatory molecule. This problem is identical to the well studied pre-equilibrium in Michaelis-Menten kinetics [2] (For a conventional presentation see [3]; a recent study on the quasi-steady state assumption for bimolecular reactions and further references are found in [4]). We include this example here for tutorial purposes. In addition, it allows for full analytical treatment whereas only combined numerical and analytical procedures are successful in all other cases. Chemical schemata for the mechanisms of binding two ligands are summarized in figure 2. Although activator and repressor for a given gene bind to different sites on DNA (figure 1), we consider also the case of competitive binding, because it provides a useful reference that allows for less involved calculations. As third and main example we analyze the binding of two regulators, activator and repressor, to two distinct binding sites on DNA. We shall investigate the case of (complete) independence of the sites and contrast it by cooperative and anti-cooperative binding of the two ligands. Finally, in our forth and last example we discuss binding of the same ligand to two or more binding sites. Two different mechanisms will be studied: (i) consecutive binding of ligands, one after the other, and (ii) ligand dimer or oligomer formation and binding to the site in one step (figure 24).

The notation applied throughout this part is the following: Chemical species are denoted be upper case sanserif letters, e.g the activator A or the complex $\mathsf{C} \equiv \mathsf{A} \cdot \mathsf{G}$. Equilibrium constants are either formulated as association constants and written upper case roman italics, e.g. K_j, or we shall use Greek letters, e.g. κ_j when it is advantageous to apply dissociation constants.[1] Lower case roman letters are applied for the concentration variables, e.g. the equilibrium concentration $c = [\mathsf{C}]$. To indicate total concentrations of species we shall use the subscript '$_0$' and we have, for example, $g_0 = g + c$ in the simple binding case.

2 Single regulator binding equilibrium

The simple single regulator binding equilibrium (figure 2, topmost reaction) is dealing with three variables, the concentrations of three species, $g = [\mathsf{G}]$, $a = [\mathsf{A}]$, and $c = [\mathsf{C}] = [\mathsf{A} \cdot \mathsf{G}]$, which are determined by three equations, two conservation relations and one equilibrium constant:

$$g_0 \,=\, g + c\,,\; a_0 \,=\, a + c\,, \quad \text{and} \quad K \,=\, \frac{c}{a \cdot g} \,=\, \kappa^{-1}\,. \tag{1}$$

[1] For all four equilibria in figure 2 the binding constants K_j have the dimension of a reciprocal concentration, $\mathrm{mol}^{-1}\cdot\mathrm{L}$, whereas the dimensions of the κ_j's are concentrations, $\mathrm{mol}\cdot\mathrm{L}^{-1}$.

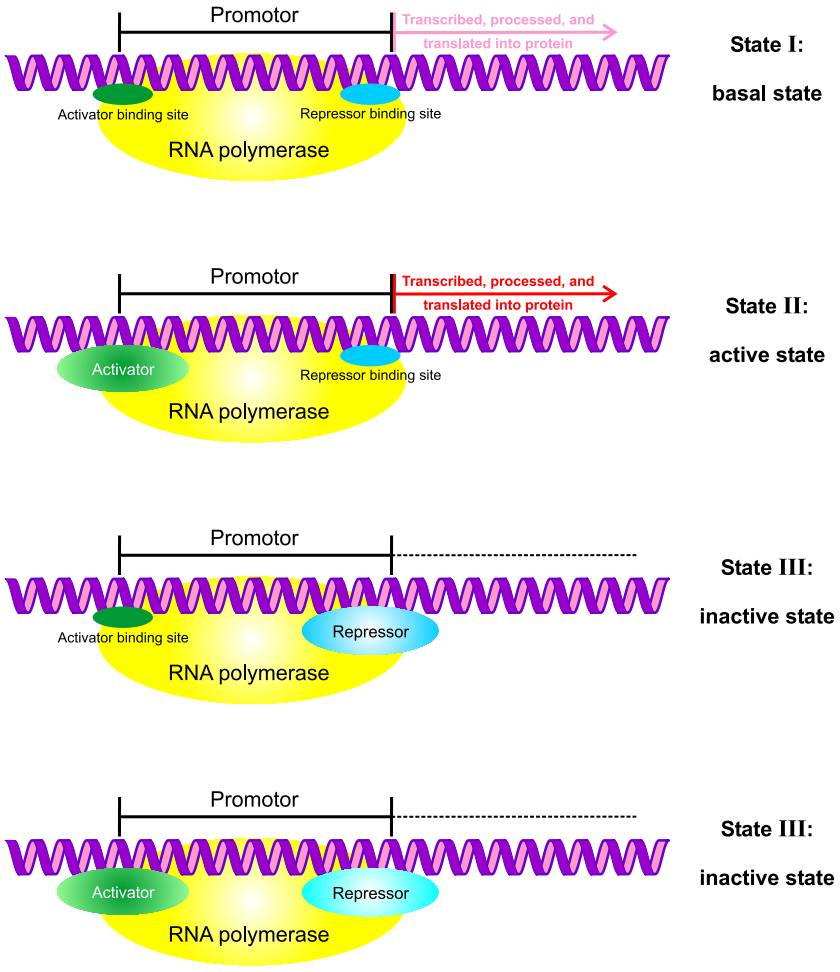


Figure 1: **Basic principle of gene regulation**. The figure sketches the *regulated recruitment* mechanism of gene activity control in prokaryote cells as discovered with the *lac* genes in *Escherichia coli* [1]. The gene has three states of activity, which are regulated by the presence or absence of glucose and lactose in the medium: State I, **basal state** occurs when both nutrients are present and it is characterized by low level transcription; neither the activator, the CAP protein, nor the *lac*-repressor protein are bound to their sites on DNA. State II, **activated state** is induced by the absence of glucose and the presence of lactose and then CAP is bound to DNA, but *lac*-repressor protein is absent. Finally, when lactose is absent the gene is in the **inactive state** no matter whether glucose is available or not. Then, the *lac*-repressor protein is bound to DNA and transcription is blocked. The promotor region of the DNA carries specific recognition sites for the RNA polymerase in addition to the binding sites for the regulatory proteins.

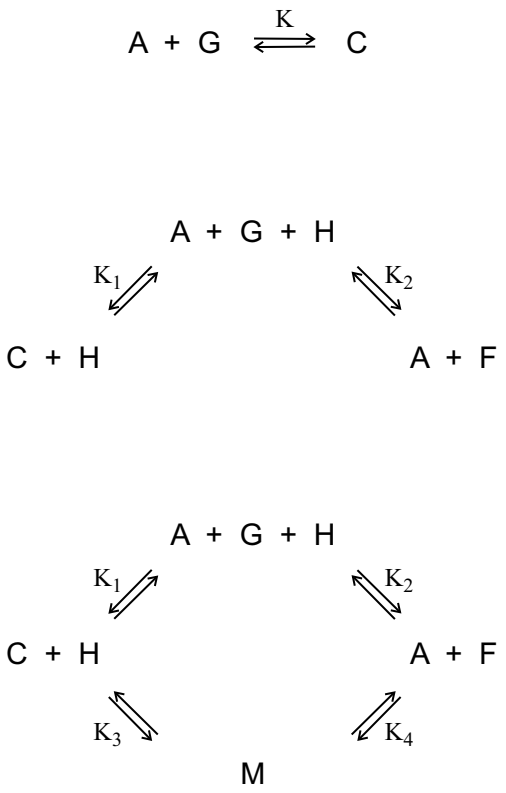


Figure 2: **Three binding mechanisms in gene regulation**. The topmost case is the simple binding equilibrium where the formation of the activator-DNA complex initiates transcription. It is identical to the pre-equilibrium in Michaelis-Menten kinetics [2, 3]. In the middle we illustrate competitive binding of two regulatory molecules, e.g. activator and inhibitor. There is only one binding site and the activity of the two complexes is determined by the nature of the bound molecules which may be either an activator or an inhibitor. The example at the bottom describes the real situation where activator and repressor are bound at different sites [1]. The four binding equilibria have to fulfil the relation: $K_1 \cdot K_3 = K_2 \cdot K_4 = K$. Lack of interaction of the two ligands in their binding sites on the DNA inplies $K_1 = K_4$ and $K_2 = K_3$. The extent of cooperative or anti-cooperative binding is measured properly in terms of a cooperativity parameter $\sigma = K_1 \cdot K_2/K$.

Calculation of the solutions for g, c, and a of (1) leads to a quadratic equation that has one positive and one negative root. Only the positive root is physically meaningful and hence the free concentration of the gene g and the

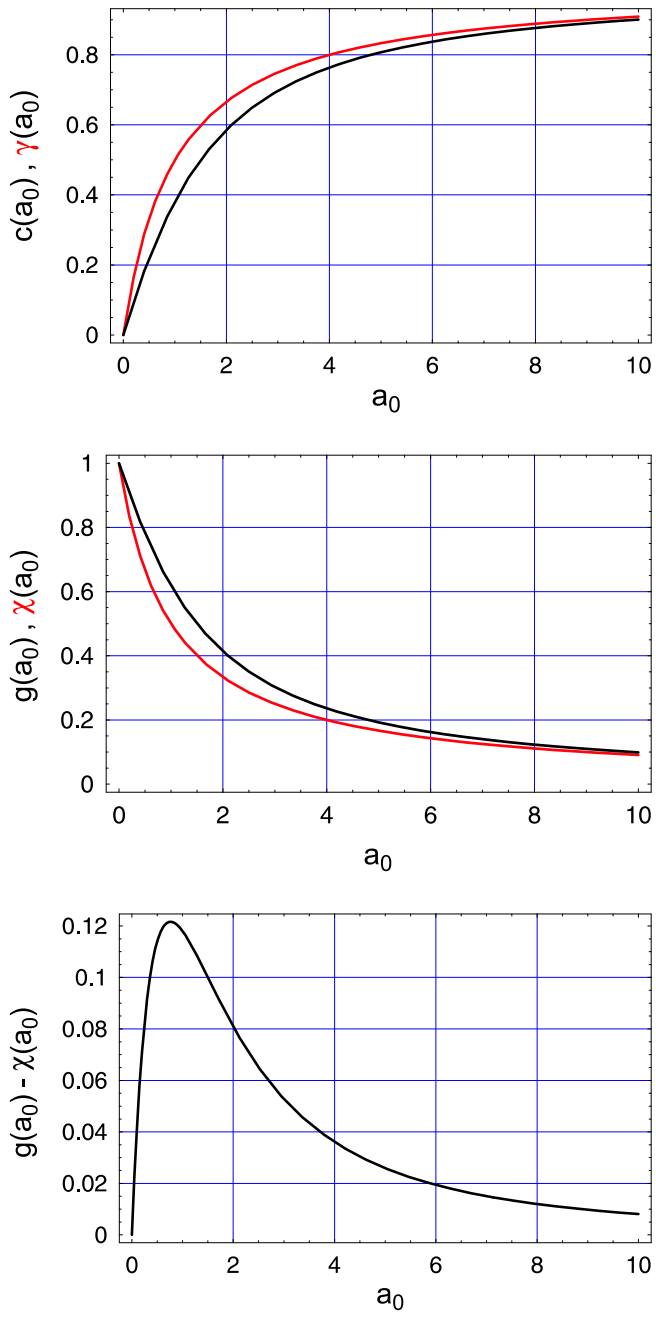


Figure 3: **Single regulator binding equilibrium.** The topmost curves represent the concentration of the complex $\mathbf{C} \equiv \mathbf{A}\cdot\mathbf{G}$, c, computed according to equation (3) [black] and its approximation γ (5) [red] as functions of the total concentration a_0. The other two parameters were chosen to be unity: $g_0 = 1$ and $K = 1$. In the middle we show the analogous curves for the concentration of free gene, $g(a_0)$ and $\chi(a_0)$, according to equations (2) and (4), respectively. The curve in the third diagram is the difference between the exact and the approximated curves: $f(\kappa, a_0, g_0) = g(\kappa, a_0, g_0) - \chi(\kappa, a_0, g_0) = -c(\kappa, a_0, g_0) + \gamma(\kappa, a_0, g_0)$ for constant κ and g_0 ($\kappa = g_0 = 1$).

gene-regulator complex c are uniquely determined by

$$g \;=\; -\frac{1}{2}\left(\kappa + a_0 - g_0 + \sqrt{(\kappa + a_0 - g_0)^2 + 4\kappa g_0}\right) \tag{2}$$

$$c \;=\; \frac{1}{2}\left(\kappa + a_0 + g_0 - \sqrt{(\kappa + a_0 + g_0)^2 - 4a_0 g_0}\right) \;. \tag{3}$$

It is useful to note that the two expressions of the square root are identical:

$$\sqrt{(\kappa + a_0 - g_0)^2 + 4\kappa g_0} \;=\; \sqrt{(\kappa + a_0 + g_0)^2 - 4a_0 g_0} \;.$$

The second expression is preferable since it makes evident the symmetry in the replacement $a_0 \Leftrightarrow g_0$ which is a consequence of the binding mechanism. Accordingly, the free concentration of **A** fulfils:

$$a \;=\; -\frac{1}{2}\left(\kappa - a_0 + g_0 + \sqrt{(\kappa + a_0 + g_0)^2 - 4a_0 g_0}\right) \;.$$

For both equations, (2) and (3), we search now for approximations that are no more complex than simple rational expressions in order to be able to use them for modelling genetic networks. The simplest approximation is to replace the free concentration of the regulator, a, by its total concentration a_0. Then we find

$$g \;\approx\; \chi \;=\; g_0\,\frac{1}{1 + K a_0} \;=\; g_0 \frac{\kappa}{\kappa + a_0} \tag{4}$$

$$c \;\approx\; \gamma \;=\; g_0\,\frac{a_0}{\kappa + a_0} \;. \tag{5}$$

The approximations become exact in the limits $\lim a_0 \to 0$ and $\lim a_0 \to \infty$. In figure 3 we show a plot of the difference between the exact solution and the approximation,

$$f(\kappa, a_0, g_0) \;=\; g(\kappa, a_0, g_0) - \chi(\kappa, a_0, g_0) \;=\; -\Big(c(\kappa, a_0, g_0) - \gamma(\kappa, a_0, g_0)\Big) \;.$$

This function exhibits a maximum at some value $[a_0]_{\max[f(\kappa,a_0,g_0)]} < g_0$ and approaches zero in the two limits mentioned above.

The position of the maximal deviation computed in the direction of the a_0-axis, $a_0 = [a_0]_{\max[f(\kappa,a_0,g_0)]} = a_0^{(\max)}$ is readily calculated from the condition

$$\max\Big\{f \;=\; \Big|\; c \;-\; \gamma \;\Big|\Big\} \;\;\Longrightarrow\;\; \frac{\partial f}{\partial a_0} \;=\; 0 \;=\; \frac{\partial c}{\partial a_0} \;-\; \frac{\partial \gamma}{\partial a_0} \;.$$

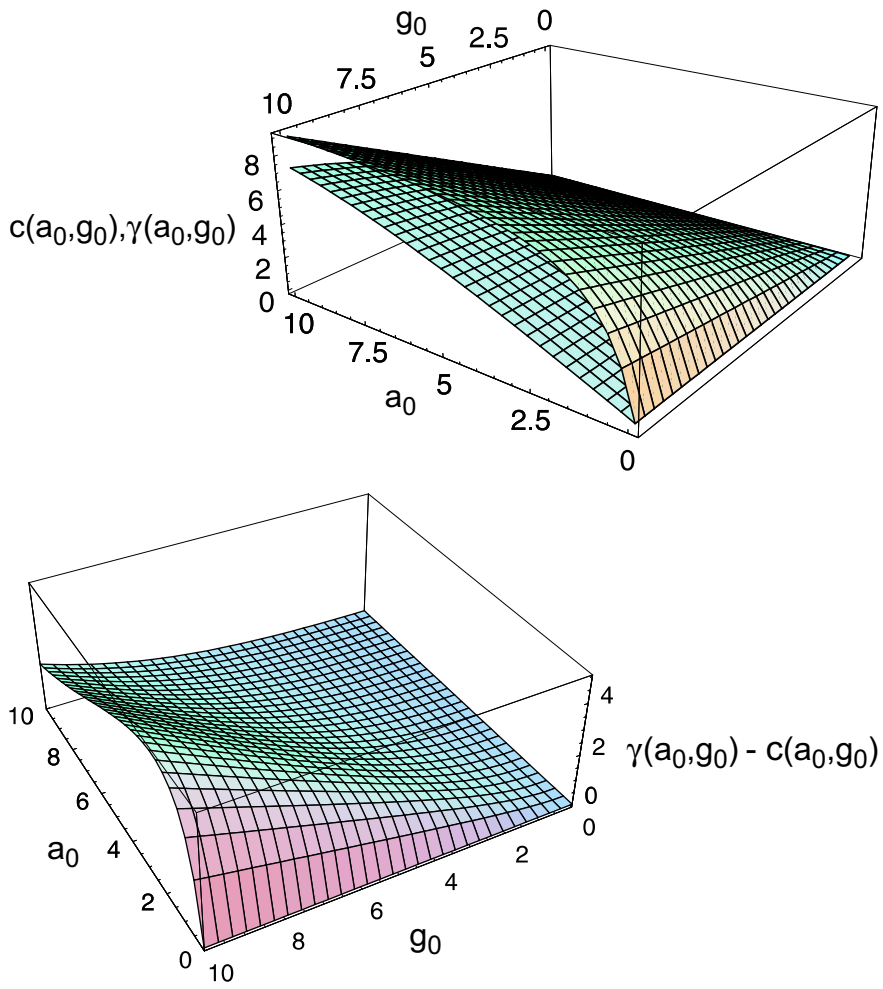


Figure 4: **Single regulator binding equilibrium.** The upper part of the figure shows the concentration of the complex **C**=**A**·**G** $c(a_0, g_0)$ and its approximation $\gamma(a_0, g_0)$ as functions of the total concentrations a_0 and g_0. The lower plot presents the (positive) difference between the two functions $f(\kappa, a_0, g_0) = |c(\kappa, a_0, g_0) - \gamma(\kappa, a_0, g_0)|$ for constant $K = \kappa = 1$.

The quantity $a_0^{(\max)}$ fulfils a cubic equation, which can be simplified by a straightforward change of the variable $a_0^{(\max)} \rightarrow \alpha = \kappa + a_0^{(\max)}$:

$$2\,\alpha^3 \;-\; (3\kappa + g_0)\,\alpha^2 \;-\; 2g_0\kappa\,\alpha \;+\; g_0\kappa(4\kappa + g_0) \;=\; 0\;. \qquad (6)$$

The roots of this cubic equation provide an illustrative example for the existence of spurious solutions. Equation (6) sustains three real roots for $g_0 > 0$. At sufficiently low values of g_0 the largest root is the physically meaningful solution. The other two roots branch in the form of a parabola opening to

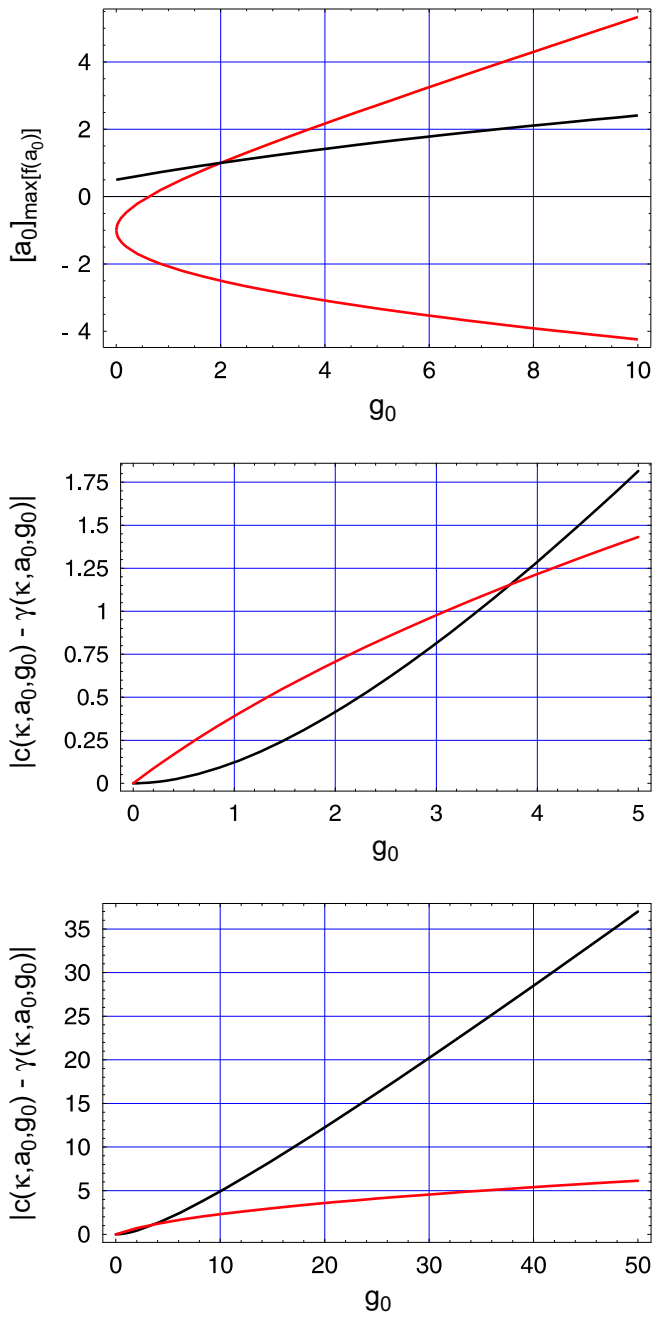


Figure 5: **Maximal deviation of the approximation $\gamma(\kappa, a_0, g_0)$ from the binding function $c(\kappa, a_0, g_0)$.** The topmost curves represent solutions of the cubic equation (6). The largest root [black] is the physically meaningful solution between $g_0 = 0$ and the crossing point of the two positive roots. From there on the smaller of the two positive roots [black] is the actual solution (see text). The two lower figures represent the maximal absolute error of the approximation $\Delta^{(\mathrm{abs})}_{\max}(\kappa, g_0) = |c(\kappa, [a_0]_{\max[f(\kappa,a_0,g_0)]}, g_0) - \gamma(\kappa, [a_0]_{\max[f(\kappa,a_0,g_0)]}, g_0)|$ [black] together with the relative error $\Delta^{(\mathrm{rel})}_{\max}(\kappa, g_0)/c(\kappa, [a_0]_{\max[f(\kappa,a_0,g_0)]}, g_0)$ [red].

the right that has the leftmost point at the position $(a_0, g_0) = (-\kappa, 0)$ (figure 5). The position on the g_0-axis at which the two positive solutions of the cubic equation cross, $g_0 = g_{0,\mathrm{cross}}$, can be calculated from the condition of a vanishing discriminant of the cubic equation (figure 6):

$$\begin{aligned} D &= -\frac{\kappa\, g_0}{432}\left(g_0^4 + 5\kappa\, g_0^3 - 5\kappa^2\, g_0^2 - 72\kappa^3\, g_0 + 108\kappa^4\right) = \\ &= -\frac{\kappa\, g_0}{432}\left(g_0 - 2\kappa\right)^2\left(g_0^3 + 7\kappa g_0^2 + 9\kappa^2 g_0 - 54\kappa^3\right) . \end{aligned} \tag{7}$$

The condition $D = 0$ is fulfilled at two values of $g^{(1)}_{0,\mathrm{cross}}(\kappa) = 0$ and at the point $g^{(2)}_{0,\mathrm{cross}}(\kappa) = 2\kappa$. Equation (6) has three real roots on both sides of $g^{(2)}_{0,\mathrm{cross}}$ implying $D < 0$ and consequently, for constant κ, $D(g^{(2)}_{0,\mathrm{cross}})$ represents a maximum of $D(g_0)$. This is straightforwardly shown by differentiation and the condition $\partial D/\partial g_0|_{\kappa=\mathrm{const}} = 0$:[2]

$$\frac{\partial D}{\partial g_0} = -\frac{\kappa}{432}\left(g_0 - 2\kappa\right)\left(5g_0^3 + 30g_0^2\kappa + 45g_0\kappa^2 - 54\kappa^3\right) ,$$

$$\left(g_0^{\mathrm{ext}}\right)_1 = 0.762793\,\kappa\,,\ \left.\frac{\partial^2 D}{\partial g_0^2}\right|_{g_0=\left(g_0^{\mathrm{ext}}\right)_1} = +11.7473\,\kappa^4 \implies \text{minimum} ,$$

$$\left(g_0^{\mathrm{ext}}\right)_2 = 2\,\kappa\,,\ \left.\frac{\partial^2 D}{\partial g_0^2}\right|_{g_0=\left(g_0^{\mathrm{ext}}\right)_2} = -196\,\kappa^4 \implies \text{maximum} .$$

The discriminant $D(g_0)$ is a fifth-order polynomial and we finalize the discussion by summarizing all five roots of $D(g_0) = 0$: $(g_0)_1 = 0$, $(g_0)_{2,3} = 2\kappa$, and two complex conjugate roots $(g_0)_{4,5} = -3\kappa(3 \pm i\sqrt{3})/2$.
For values of g_0 above the crossing point the correct solution is the second largest root of the cubic equation. It is straightforward to find out from where the spurious solution originates: At one instant in the derivation the cubic equation (6) both sides of the equation were squared in order to eliminate the square root, and this operation allows for additional solutions (which results from expressions with different signs before the operation).

From the proper solution of (6) we can easily calculate the absolute and relative maximal error in the binding function:

$$\Delta^{(\mathrm{abs})}_{\mathrm{max}}(\kappa, g_0) = \left| c(a_0^{(\mathrm{max})}) - \gamma(a_0^{(\mathrm{max})}) \right| \quad \text{and} \tag{8}$$

$$\Delta^{(\mathrm{rel})}_{\mathrm{max}}(\kappa, g_0) = \left| c(a_0^{(\mathrm{max})}) - \gamma(a_0^{(\mathrm{max})}) \right| \Big/ c(a_0^{(\mathrm{max})}) . \tag{9}$$

[2]The exact value of $(g_0^{\mathrm{ext}})_1$ is $\kappa\left[-2 + \left(\frac{5}{32+\sqrt{111}}\right)^{1/3} + \left(\frac{5}{32+\sqrt{111}}\right)^{1/3}\right]$.

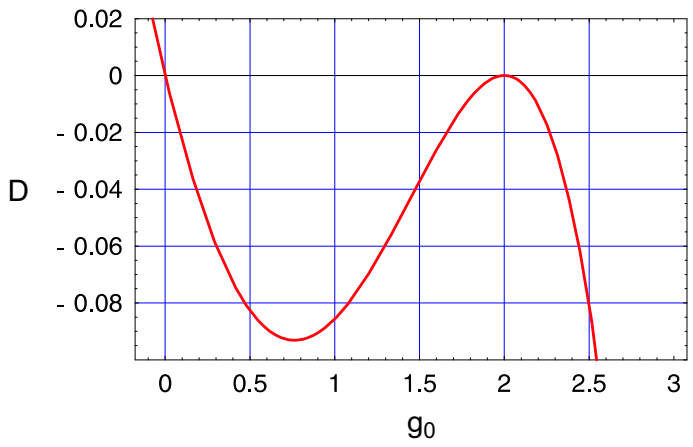


Figure 6: **Discriminant of the cubic equation** (6)**.** The figure shows the discriminant D defined in equation (7) as a function of g_0 for $\kappa = 1$. One 'double root' ($D = 0$) is always found at $g_0 = 0$ and the condition for a second 'double root' occurs at $g_0 = 2\kappa$. At this position the first derivative, dD/dg_0, vanishes as well. As shown in the text the discriminant has a maximum at $g_0 = 2\kappa$ and hence we observe three real roots of the cubic equation at both sides, $g_0 = \kappa + \varepsilon$ and $g_0 = \kappa - \varepsilon$.

Both errors are shown as functions of g_0 for constant κ in figure 5. Asymptotically for large g_0, the deviation from the exact value increases linearly with increasing total concentration of the gene. The relative error, however, shows parabolic increase corresponding to an asymptotic $\sqrt{g_0}$ proportionality.

Finally we consider changes in the free concentrations as a result of adding activator or gene. These response functions are obtained by straightforward differentiation:

$$\frac{\partial g}{\partial a_0} = -\frac{1}{2}\left(1 - \frac{\kappa + a_0 - g_0}{\sqrt{(\kappa + a_0 + g_0)^2 - 4a_0 g_0}}\right) , \tag{10}$$

$$\frac{\partial g}{\partial g_0} = \frac{1}{2}\left(1 + \frac{\kappa - a_0 + g_0}{\sqrt{(\kappa + a_0 + g_0)^2 - 4a_0 g_0}}\right) , \tag{11}$$

$$\frac{\partial c}{\partial a_0} = \frac{1}{2}\left(1 - \frac{\kappa + a_0 - g_0}{\sqrt{(\kappa + a_0 + g_0)^2 - 4a_0 g_0}}\right) , \quad \text{and} \tag{12}$$

$$\frac{\partial c}{\partial g_0} = \frac{1}{2}\left(1 - \frac{\kappa - a_0 + g_0}{\sqrt{(\kappa + a_0 + g_0)^2 - 4a_0 g_0}}\right) . \tag{13}$$

It is readily verified that the derivatives fulfil the equations:

$$\frac{\partial g}{\partial a_0} \;=\; -\frac{\partial c}{\partial a_0} \quad \text{and} \quad \frac{\partial g}{\partial g_0} + \frac{\partial c}{\partial g_0} \;=\; 1 \;.$$

The derivatives of the approximations $\chi(a_0, g_0)$ and $\gamma(a_0, g_0)$ fulfil the analogous equations,

$$\frac{\partial \chi}{\partial a_0} \;=\; -\frac{\partial \gamma}{\partial a_0} \quad \text{and} \quad \frac{\partial \chi}{\partial g_0} + \frac{\partial \gamma}{\partial g_0} \;=\; 1 \;,$$

but differ in the whole range $0 \leq a_0 < \infty$. Only in the limit $\lim a_0 \to \infty$ all derivatives approach zero as expected.

3 Binding of two ligands to one site

Binding of activator and repressor to the same regulation site of the gene on the DNA leads to the three state model shown in the middle of figure 2. Five variables, a, c, f, g, and h, are determined by three conservation relations,

$$g_0 \;=\; g \,+\, c \,+\, f \tag{14}$$

$$a_0 \;=\; a \,+\, c \tag{15}$$

$$h_0 \;=\; h \,+\, f \;, \tag{16}$$

and two equilibria,

$$K_1 \;=\; \frac{c}{a \cdot g} \quad \text{and} \quad K_2 \;=\; \frac{f}{h \cdot g} \;. \tag{17}$$

Elimination of the variables for the complexes is straightforward and introducing dissociation rather than binding constants, $\kappa_1 = K_1^{-1}$ and $\kappa_2 = K_2^{-1}$, we find:

$$g \;=\; g_0 \;\Big/\; \left(1 + \frac{a_0}{\kappa_1 + g} + \frac{h_0}{\kappa_2 + g}\right) \;,$$

$$a \;=\; \kappa_1\, a_0 \;\Big/\; (\kappa_1 + g) \;,$$

$$h \;=\; \kappa_1\, h_0 \;\Big/\; (\kappa_1 + g) \;.$$

Further, elimination of a and h yields a cubic equation for the free gene concentration g as a function of the three total concentrations, a_0, h_0, and g_0, and the two dissociation constants, κ_1 and κ_2:

$$\begin{aligned} g^3 \,+\, g^2(\kappa_1 + \kappa_2 + a_0 + h_0 - g_0) \,+ \\ +\, g\Big(\kappa_1\kappa_2 + \kappa_2 a_0 + \kappa_1 h_0 - (\kappa_1 + \kappa_2) g_0\Big) \,-\, \kappa_1\kappa_2 g_0 \;=\; 0 \;. \end{aligned} \tag{18}$$

The cubic equation has three real roots in the positive parameter orthant, $(a0 > 0, h0 > 0, g0 > 0)$, as can be easily verified by considering the discriminant of equation (18) shown in figure 7. The discriminant D is negative for all values $(a_0 > 0, h_0 > 0, g_0 \leq 0)$ and vanishes for $a_0 = h_0 = 0$ indicating a double root for this condition. Closer inspection shows that two of the three roots are negative and we are thus dealing with one and hence unique acceptable solution $g < 0$. In figure 8 we illustrate the different behavior of the three roots.

The other variables are readily obtained from equations (14-17):

$$a = \frac{\kappa_1 a_0}{\kappa_1 + g} \quad \text{and} \quad c = \frac{a_0 g}{\kappa_1 + g} , \tag{19}$$

$$h = \frac{\kappa_2 h_0}{\kappa_2 + g} \quad \text{and} \quad f = \frac{h_0 g}{\kappa_2 + g} . \tag{20}$$

In figure 9 we show the concentrations a functions of a_0 and g_0, respectively.

Applying the same strategy to find an approximation for the free concentration of the activator complex $c(a_0, h_0, g_0)$, which consisted simply in the replacement of a and h by a_0 and h_0, respectively, we obtain

$$\gamma(a_0, h_0, g_0) = \frac{\kappa_2 a_0 g_0}{\kappa_1\kappa_2 + \kappa_2 a_0 + \kappa_2 h_0} = \frac{K_1 a_0 g_0}{1 + K_1 a_0 + K_2 h_0} . \tag{21}$$

The error introduced by this approximation is shown in the plots presented in figure 10 and can be analyzed similarly as in the case of the simple binding equilibrium. The approximation is exact in the limits $a_0 \to 0$, $a_0 \to \infty$, and $h_0 \to \infty$:

$$\begin{aligned}
\lim_{a_0\to 0} c(a_0, h_0, g_0) &= \lim_{a_0\to 0} \gamma(a_0, h_0, g_0) = 0 , \\
\lim_{a_0\to \infty} c(a_0, h_0, g_0) &= \lim_{a_0\to \infty} \gamma(a_0, h_0, g_0) = g_0 , \quad \text{and} \\
\lim_{h_0\to \infty} c(a_0, h_0, g_0) &= \lim_{h_0\to \infty} \gamma(a_0, h_0, g_0) = 0 .
\end{aligned}$$

In addition, the absolute error in the competitive binding case is very similar to that in the simple binding equilibrium.

4 Binding of two ligands to two sites

In this section we shall discuss the complete binding mechanism of activator and repressor (figure 2). Binding of ligands to two different sites raises the problem of interaction between the sites. We shall distinguish here the case of independence (section 4.3) from mutual interaction between ligands (section 4.4), which can be cooperative or anti-cooperative.

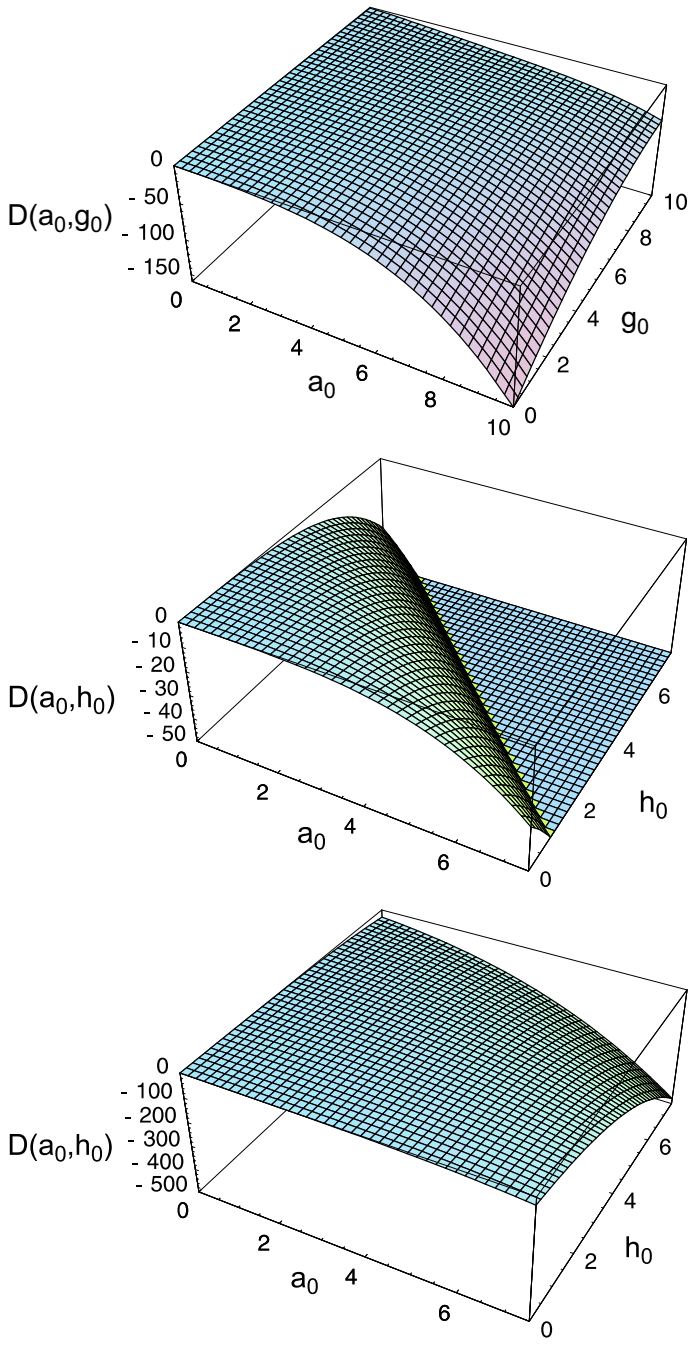


Figure 7: **Discriminant of the cubic equation** (18)**.** The topmost plot shows the discriminant D as a function of a_0 and g_0. It is negative for all values $a_0 > 0$ and $h_0 > 0$ with $g_0 \geq 0$. The two lower plots show $D(a_0, h_0)$ for smaller and larger values of $|D|$.

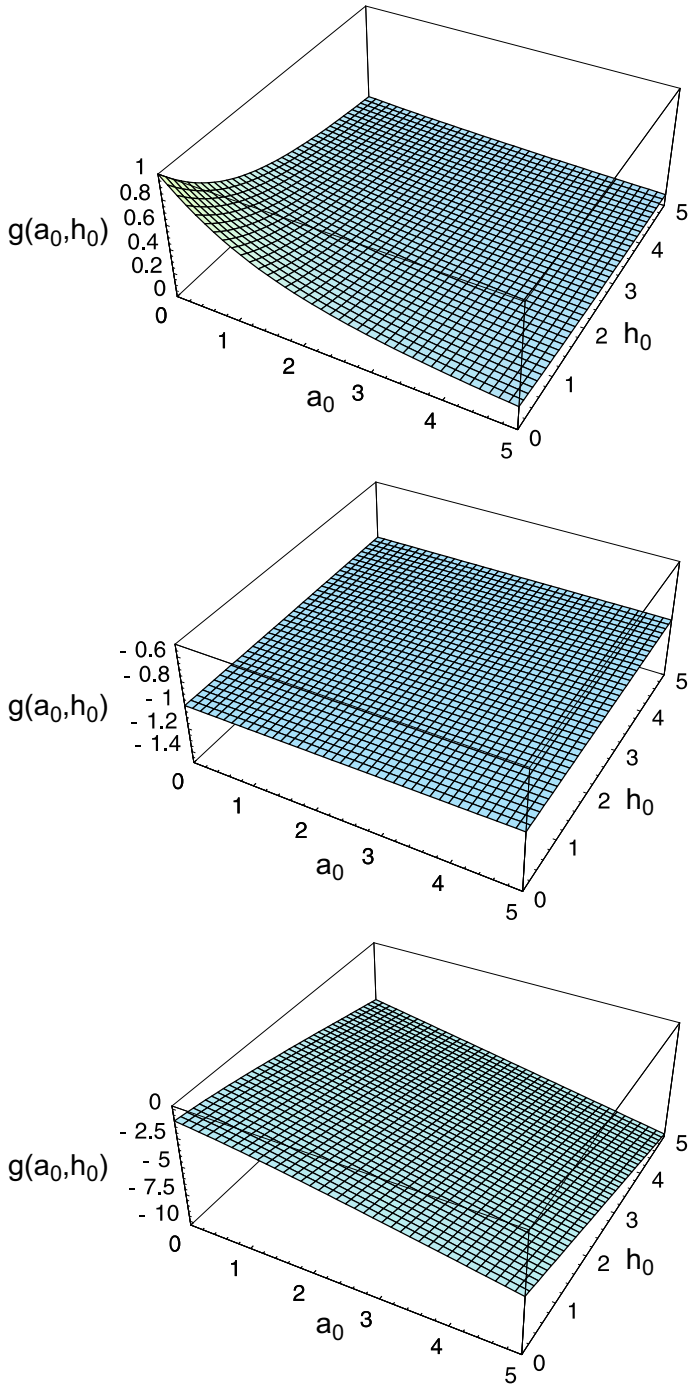


Figure 8: **Solutions of the cubic equation** (18)**.** The plots represent all three roots of the cubic equation for $K_1 = K_2 = 1$ and $g_0 = 1$: (i) The positive root corresponding to the unique physically acceptable solution in the topmost part of the figure as a function of the total concentration of activator and repressor, $g = g_3(a_0, h_0)$, (ii) the constant root $g_2(a_0, h_0) = -1$, and (iii) the second negative root $g_1(a_0, h_0)$. The analogous plots $g(a_0, g_0)$ show also one positive and two negative roots.

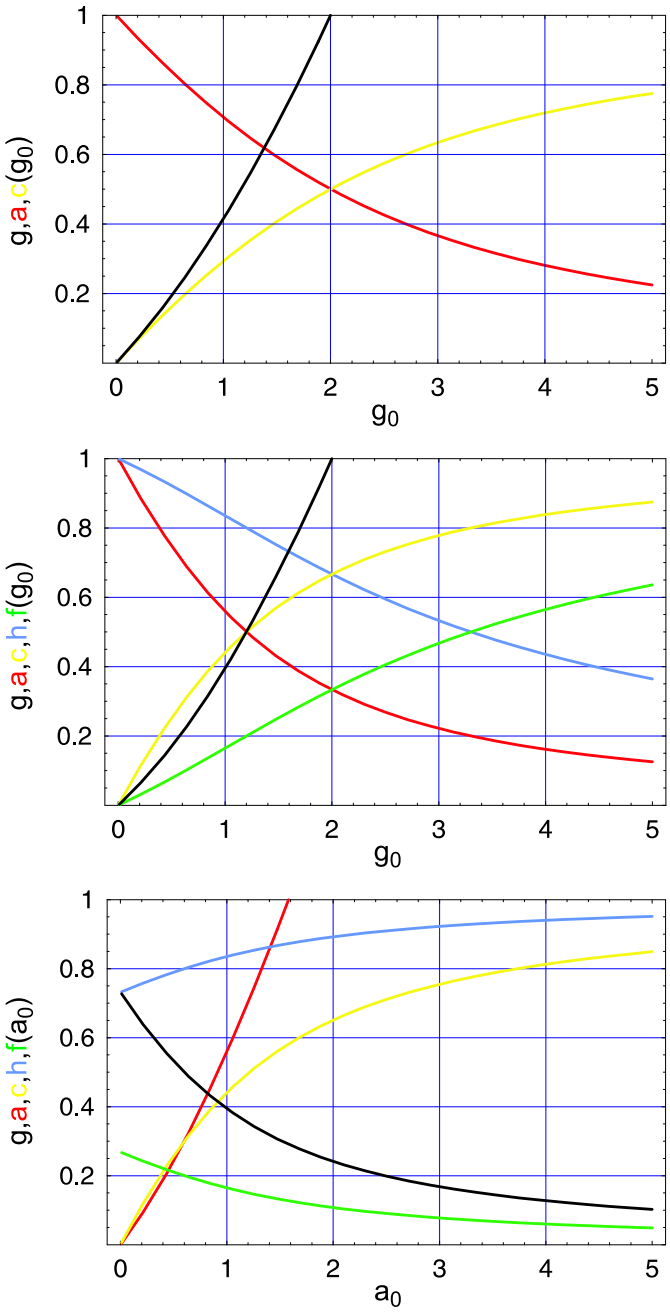
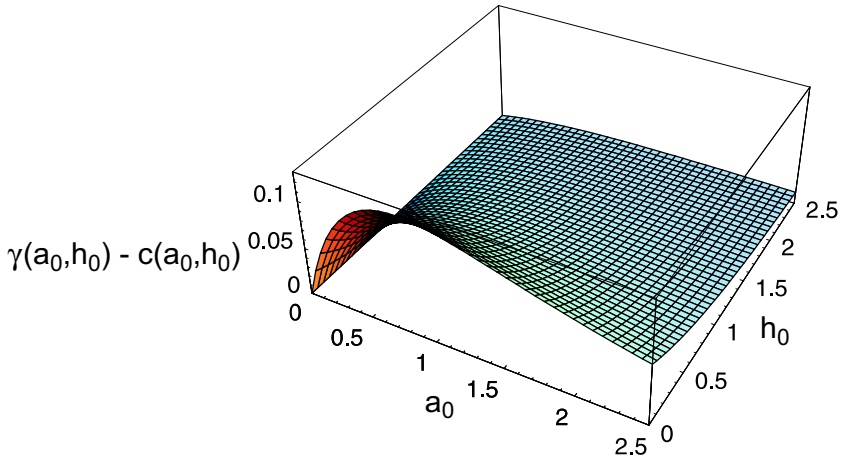
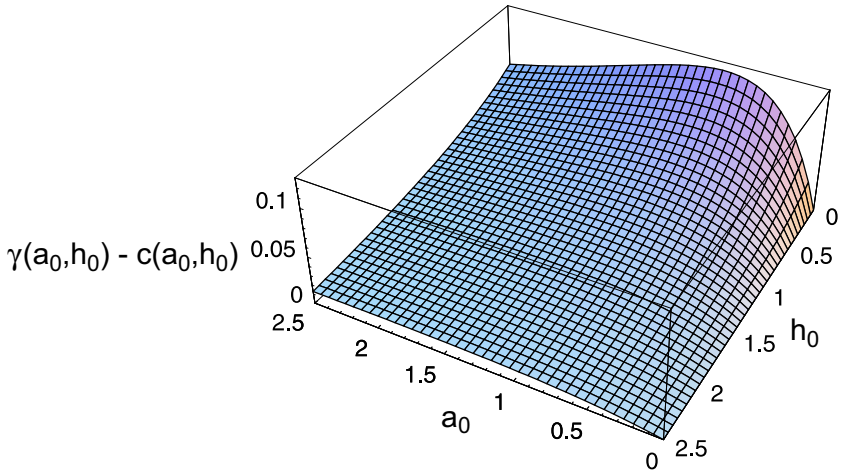
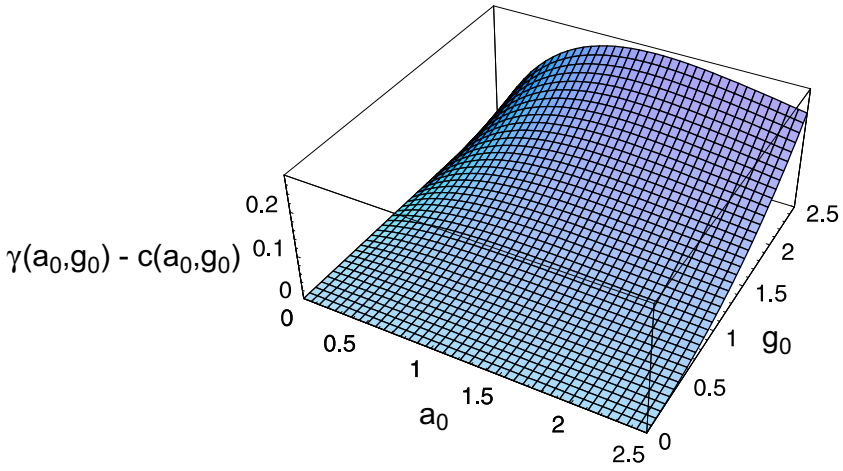


Figure 9: **Equilibrium concentrations of competitive binding.** The topmost curves present the dependence of the equilibrium concentrations on the total gene concentration g_0 for the 'symmetric' case $\kappa_1 = \kappa_2 = 1$ where we have $h(g_0) = a(g_0)$ and $f(g_0) = c(g_0)$. The two lower plots refer to an 'asymmetric' joyce of equilibrium parameters, $\kappa_1 = 0.5$ and $\kappa_2 = 2$, and show the equilibrium concentrations as functions of g_0 and a_0, respectively. Color code: g black, a red, c yellow, h blue, and f green.


γ(a0,h0) - c(a0,h0)
0.1
0.05
0
0
0.5
1
1.5
2
2.5
a0
0
0.5
1
1.5
2
2.5
h0


γ(a0,h0) - c(a0,h0)
0.1
0.05
0
2.5
2
1.5
1
0.5
0
a0
0
0.5
1
1.5
2
2.5
h0


γ(a0,g0) - c(a0,g0)
0.2
0.1
0
0
0.5
1
1.5
2
2.5
a0
0
0.5
1
1.5
2
2.5
g0

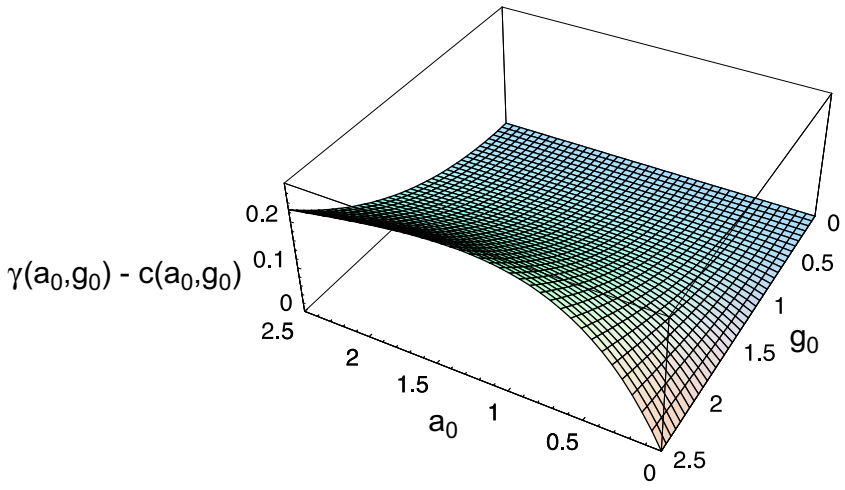


Figure 10: **Approximation of the activator complex concentration in competitive binding.** The concentration of the activator complex is approximated by the function $\gamma(a_0, h_0, g_0)$ of equation (21). The first two plots show the error $\Delta^{(\mathrm{abs})}(a_0, h_0, g_0) = \gamma(a_0, h_0, g_0) - c(a_0, h_0, g_0)$ as a function of a_0 and h_0, the other two plots as a function of a_0 and h_0.

4.1 Calculation of the equilibrium concentrations

The four state model in figure 2 with six variables, a, c, f, g, h, and m, gives rise to three conservation relations,

$$\begin{aligned} g_0 &= g + c + f + m && (22)\\ a_0 &= a + c + m && (23)\\ h_0 &= h + f + m\ , && (24) \end{aligned}$$

and three equilibria,

$$K_1 = \frac{c}{a \cdot g}\ , \quad K_2 = \frac{f}{h \cdot g} \quad \text{and} \quad K = K_1 K_3 = K_2 K_4 = \frac{m}{a \cdot h \cdot g}\ . \qquad (25)$$

First we eliminate all variables for the complexes and find for g, a, and h:

$$\begin{aligned} g &= \frac{g_0}{1 + K_1 a + K_2 h + K a h}\\ a &= \frac{1}{1 + K_1 g}\Big(a_0 - h_0 + h(1 + K_2 g)\Big)\\ h &= \frac{1}{1 + K_2 g}\Big(h_0 - a_0 + h(1 + K_1 g)\Big)\ . \end{aligned}$$

The expression for a as a function of g is the result of a quadratic equation and already fairly involved,

$$a \;=\; \frac{A(g)}{2(1+K_1 g)Kg} \quad \text{with} \quad A(g) \;=\; -(1+K_1 g)(1+K_2 g) \;-\; Kg(h_0 - a_0) \;+\; D(g) \tag{26}$$

$$D(g) \;=\; \sqrt{(1+K_1 g)^2(1+K_2 g)^2 + 2(1+K_1 g)(1+K_2 g)Kg(h_0+a_0) + K^2 g^2 (h_0 - a_0)^2}\;. \tag{27}$$

The second root of the quadratic equation leads to negative concentrations and can be ignored therefore. Defining in addition

$$H(g) \;=\; -(1+K_1 g)(1+K_2 g) \;-\; Kg(a_0 - h_0) \;+\; D(g)\;, \tag{28}$$

we obtain a final implicit equation for g:

$$\begin{aligned} 4(1+K_1 g)(1+K_2 g)Kg(g_0 - g) \;&=\; \\ =\; 2(1+K_2 g)K_1 g A(g) + 2(1+K_1 g)K_2 g H(g) + A(g)H(g)\;. \end{aligned} \tag{29}$$

The equation is well conditioned for numerical solution in the physically relevant range, $0 \le g \le g_0$, and it will be used to compute equilibrium concentrations.

In order to test the implicit equation for the free enzyme concentration, g in (29), we compute the solution for $K = 0$ and $\{K_1, K_2\} \neq 0$, which should coincide with the previous example of competitive binding. Inserting $K = 0$ into equation (27) yields

$$D \;=\; \sqrt{(1 \,+\, K_1\, g)^2 (1 \,+\, K_2\, g)^2} \;=\; (1 \,+\, K_1\, g)(1 \,+\, K_2\, g)\;.$$

Further insertion of this result into (26) results in $A(g) = 0$ and $a(g) = 0/0$. We apply now the rule of de l'Hospital,

$$a(g) \;=\; \frac{Z(g,K)}{N(g,Z)} \;=\; \frac{\lim_{K\to 0}\left(\frac{dZ}{dK}\right)}{\lim_{K\to 0}\left(\frac{dN}{dK}\right)}\;,$$

and find

$$\lim_{K\to 0} a(g) \;=\; \frac{2g\, a_0}{2g\,(1+K_1 g)} \;=\; \frac{a_0}{1+K_1 g}\;.$$

Application of the same strategy to $h(g)$ and insertion into the conservation relation (22) with $m = 0$ because of (25) with $K = 0$ yields the equation

$$g_0 \;=\; g\left(1 \,+\, \frac{K_1\, a_0}{1+K_1\, g} \,+\, \frac{K_2\, h_0}{1+K_2\, g}\right)\;,$$

which is the corresponding equation for the case of competitive binding (section 3). In addition the results derived from equation (29) are readily verified by means of integration of the kinetic equation (See Part II: Binding Kinetics).

After evaluation of the implicit equation for the free gene concentration g we are in a position to compute all other variables beginning with the concentration of the activator complex c:

$$c(a_0,h_0,g_0) \;=$$
$$\;=\; \frac{1}{2Kg}\left(-(1+K_1g+Kg(h_0-g_0+g))+\sqrt{(1+K_1g+Kg(h_0-g_0+g))^2+4KK_1a_0g^2}\right) \quad (30)$$
$$f(a_0,h_0,g_0) \;=$$
$$\;=\; \frac{1}{2Kg}\left(-(1+K_2g+Kg(a_0-g_0+g))+\sqrt{(1+K_2g+Kg(a_0-g_0+g))^2+4KK_2h_0g^2}\right) \quad (31)$$
$$m(a_0,h_0,g_0) \;=\; g_0 \,-\, g(a_0,h_0,g_0) \,-\, c(a_0,h_0,g_0) \,-\, f(a_0,h_0,g_0) \quad (32)$$
$$a(a_0,h_0,g_0) \;=\; a_0 \,-\, c(a_0,h_0,g_0) \,-\, m(a_0,h_0,g_0) \quad (33)$$
$$h(a_0,h_0,g_0) \;=\; h_0 \,-\, f(a_0,h_0,g_0) \,-\, m(a_0,h_0,g_0) \quad (34)$$

In figures 11 and 12 we show several examples of binding equilibria for the two binding sites case.

4.2 Concentrations at limits

Before we shall discuss an approximation to the equilibrium concentrations we consider the behavior of the six concentration variables in the three limits discussed in the previous section, $\lim a_0 \to 0$, $\lim a_0 \to \infty$, $\lim h_0 \to 0$, and $\lim h_0 \to \infty$ in the system with two binding sites. We start by the rather trivial limit $\lim a_0 \to 0$. The three auxiliary functions adopt the forms:

$$D(g) \;=\; (1+K_1g)(1+K_2g)+Kgh_0\,,\;\; A(g) \;=\; 0\,,\quad \text{and}\quad H(g) \;=\; 2Kgh_0\;.$$

Insertion into the implicit equation for g yields three solutions: (i) $g=0$, (ii) $g=-\kappa_1$, and (iii) the only physically acceptable solution

$$g^2 \;+\; (h_0-g_0+\kappa_2)\,g \;-\; \kappa_2\,g_0 \;\;=\;\; 0\;,$$

which is identical to the binding equilibrium $\mathbf{G}+\mathbf{H} \rightleftharpoons \mathbf{F}$. In consequence of this result we have $\lim_{a_0\to 0} c(a_0,h_0,g_0)=0$.

Next we consider the limit $\lim a_0 \to \infty$, implying $a_0 \gg \{h_0,g_0\}$ and $g \ll g_0$, that is somewhat more subtle. First we find $D(g) \approx 1+Kga_0$, $A(g) \approx 2Kga_0$, and $a \approx a_0$. All molecular species that can bind do not contain $\mathbf{A}$ will vanish in this limit, $\lim_{a_0\to\infty} g = 0$, $\lim_{a_0\to\infty} f = 0$. In particular, we have

$$\lim_{a_0\to\infty} h_0 \;=\; h+m \quad\text{and}\quad \lim_{a_0\to\infty} g_0 \;=\; c+m\;.$$

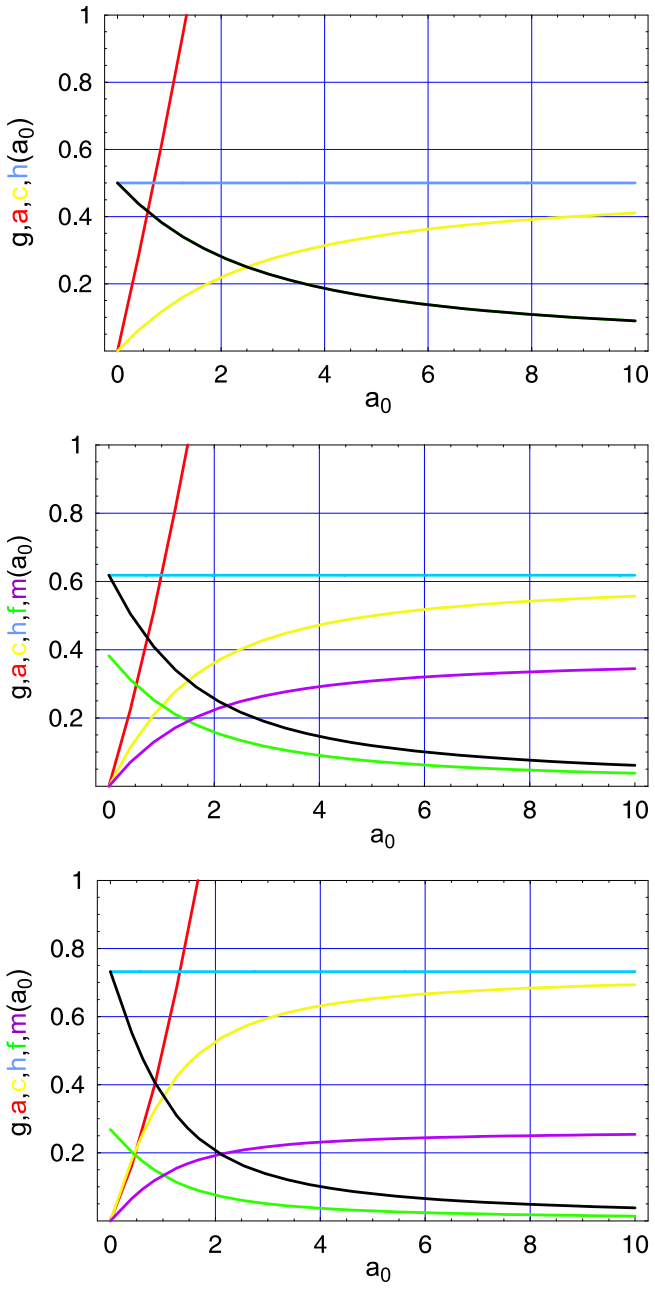


Figure 11: **Equilibrium concentrations of binding to two binding sites.** The topmost curves present the dependence of the equilibrium concentrations on the total gene concentration a_0 for an asymmetric choice of equilibrium parameters, $K = 1, K_1 = 0.5, K_2 = 2$. The 'symmetric' case $K = K_1 = K_2 = 1$ is shown in the middle and the opposite asymmetric choice, $K = 1, K_1 = 2, K_2 = 0.5$ is presented at the bottom. It is worth pointing out that we have an accidental degeneracy, $m(a_0) = c(a_0)$ and $f(g_0) = g(g_0)$ in the topmost plot. Interestingly, the free concentration of h does not depend on a_0, which is a consequence of the assumption $K = K_1 \cdot K_2$ (See section 4.3). Color code: g black, a red, c yellow, h blue, f green, and m purple.

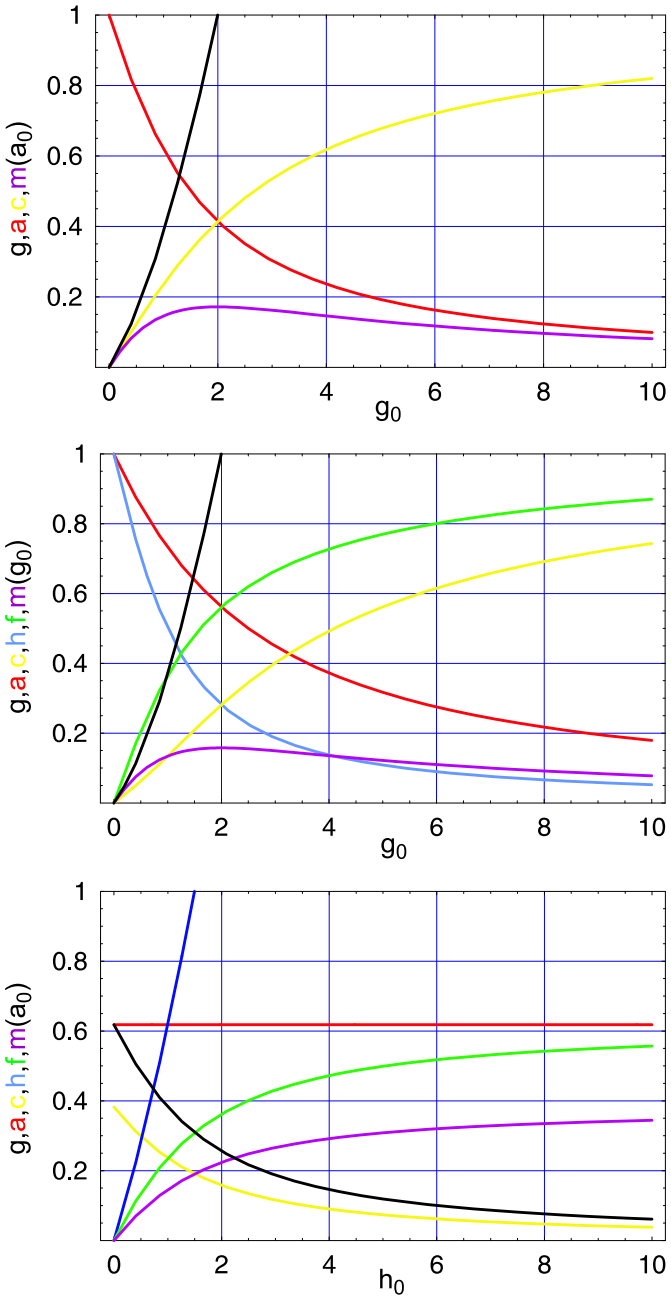


Figure 12: **Equilibrium concentrations of binding to two binding sites.** The topmost curves present the dependence of the equilibrium concentrations on the total gene concentration g_0 for the 'symmetric' case $K = K_1 = K_2 = 1$ where we have $h(g_0) = a(g_0)$ and $f(g_0) = c(g_0)$. The plot at the bottom shows the curves as a function of h_0 for the same choice of equilibrium parameters ('symmetric' case; this plot is identical to figure 11, middle plot when we exchange $\{h, f\} \leftrightarrow \{a, c\}$). The plot in the middle refer to an 'asymmetric' joyce of equilibrium parameters, $K = 1$, $K_1 = 0.5$, and $K_2 = 2$, and shows the equilibrium concentrations as a function of g_0. Color code: g black, a red, c yellow, h blue, f green, and m purple.

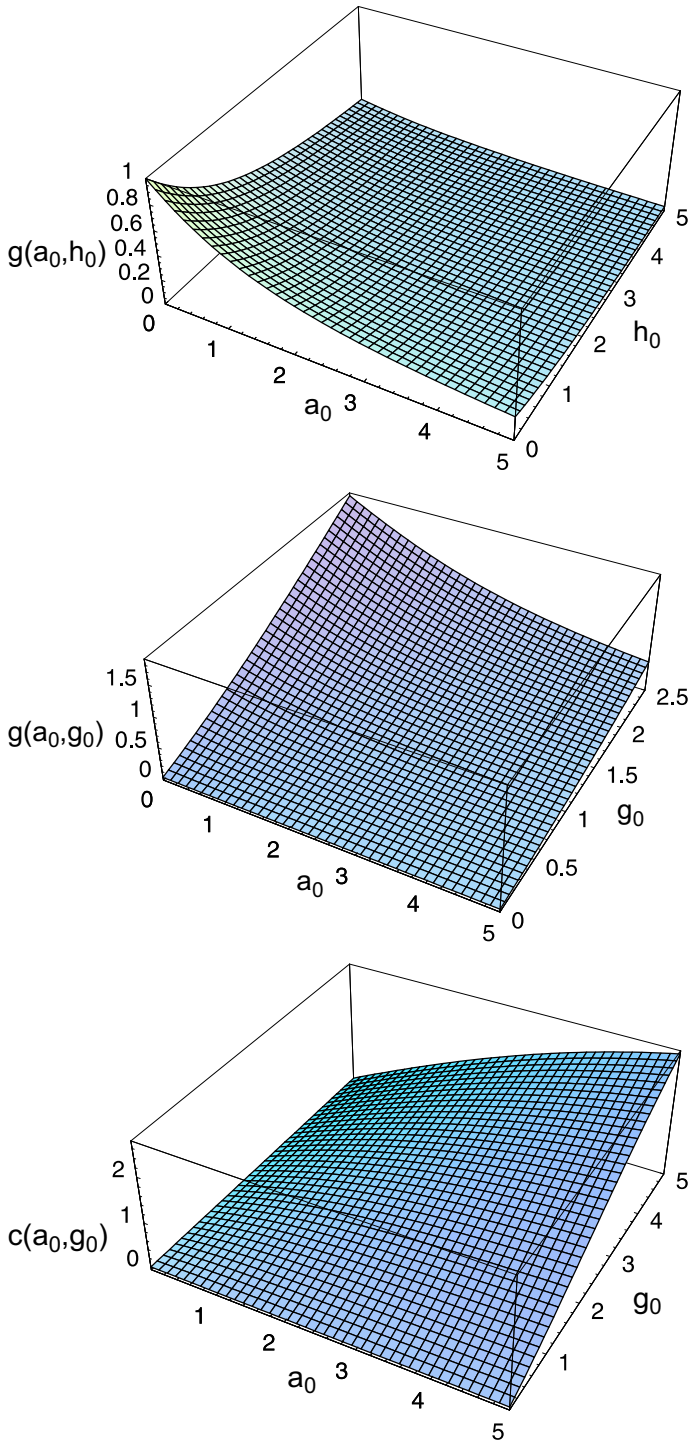


Figure 13: **Solutions of the implicit equation** (29) **for the free gene and the activator complex concentration.** The topmost and the middle plot show g as functions of (a_0, h_0) and (a_0, g_0) according to equation (29). The plot at the bottom presents the concentration of the activator complex $c(a_0, g_0)$ as expressed by equation (30).

Straightforward calculation for the limit lim $a_0 \to \infty$ yields

$$\lim_{a_0\to\infty} c \;=\; \frac{1}{2K}\left(-(K_1 + K(h_0 - g_0) + \sqrt{K_1^2 + K(h_0 - g_0)^2 + 2K_1K(h_0 + g_0)}\right) \quad (35)$$

$$\lim_{a_0\to\infty} h \;=\; \frac{1}{2K}\left(-(K_1 - K(h_0 - g_0) + \sqrt{K_1^2 + K(h_0 - g_0)^2 + 2K_1K(h_0 + g_0)}\right) \quad (36)$$

$$\lim_{a_0\to\infty} m \;=\; g_0 - \lim_{a_0\to\infty} c \;=\; h_0 - \lim_{a_0\to\infty} h \quad (37)$$

$$\lim_{a_0\to\infty} f \;=\; \frac{\lim_{a_0\to\infty} m}{K\, a_0} \;=\; 0. \quad (38)$$

Accordingly $\lim_{a_0\to\infty} f(a_0, h_0, g_0) = 0$ as argued above.

The limit lim $h_0 \to \infty$, implying $h_0 \gg \{a_0, g_0\}$ and $g \ll g_0$ can be handled by symmetry arguments and accordingly we have

$$\lim_{h_0\to\infty} c(a_0, h_0, g_0) \;=\; 0\ . \quad (39)$$

Similarly, the limit lim $h_0 \to 0$ leads to a simple equilibrium $\mathbf{G} + \mathbf{A} \rightleftharpoons \mathbf{C}$ with the consequence

$$\lim_{h_0\to 0} c(a_0, h_0, g_0) \;=\; \frac{1}{2}\left(\kappa_1 + a_0 + g_0 - \sqrt{(\kappa_1 + a_0 + g_0)^2 - 4a_0g_0}\right)\ . \quad (40)$$

The behavior of approximations in relation to the exact expression is discussed in the next section 4.3.

4.3 Independent binding

In case the two ligands bind independently or, in other words, the binding equilibrium of one effector is not influenced by absence or presence of the second effector the over-all binding constant fulfils the relation $K = K_1 \cdot K_2$ or $K_1 = K_4$ and $K_2 = K_3$, respectively. The whole system is described now by two rather than three equilibrium parameters. With respect to the calculations of free concentrations the reduction of parameters provides only a minor simplification. It is straightforward to show, however, that fulfilling this relation is responsible for the fact that h is independent of the input concentration a_0 as seen, for example, in figure 11. The same is true, of course, when we consider a as a function of h_0 (figure 12). As we model binding to two independent sites, we may argue, why should a change in the total concentration of one effector influence the second effector that binds to the other site. Thus the result is not unexpected, but we shall, nevertheless, derive it from conservation relations and equilibrium parameters. Partial

differentiation of equations (22-25) yields:

$$\frac{\partial g}{\partial a_0} + \frac{\partial c}{\partial a_0} + \frac{\partial f}{\partial a_0} + \frac{\partial m}{\partial a_0} = 0 \tag{41}$$

$$\frac{\partial h}{\partial a_0} + \frac{\partial f}{\partial a_0} + \frac{\partial m}{\partial a_0} = 0 \tag{42}$$

$$\frac{\partial a}{\partial a_0} + \frac{\partial g}{\partial a_0} + \frac{\partial m}{\partial a_0} = 1 \tag{43}$$

$$\frac{\partial c}{\partial a_0} = K_1 \left(a \frac{\partial g}{\partial a_0} + g \frac{\partial a}{\partial a_0} \right) \tag{44}$$

$$\frac{\partial f}{\partial a_0} = K_2 \left(h \frac{\partial g}{\partial a_0} + g \frac{\partial h}{\partial a_0} \right) \tag{45}$$

$$\frac{\partial m}{\partial a_0} = K_3 \left(h \frac{\partial c}{\partial a_0} + c \frac{\partial h}{\partial a_0} \right) = K_4 \left(a \frac{\partial f}{\partial a_0} + f \frac{\partial a}{\partial a_0} \right) . \tag{46}$$

From $K_3\, c \cdot f = K_2\, g \cdot m$ we derive an equation that contains exclusively the derivatives of c and g:

$$(1 + K_2 g + K_3 c + K_3 h) \frac{\partial c}{\partial a_0} = -(1 + K_2 g + K_3 c + K_2 h) \frac{\partial g}{\partial a_0} .$$

Now we introduce the condition for independent binding sites, $K_2 = K_3$ and obtain the following three results:

$$\frac{\partial c}{\partial a_0} = -\frac{\partial g}{\partial a_0} , \tag{47}$$

$$\frac{\partial f}{\partial a_0} = -\frac{\partial m}{\partial a_0} , \quad \text{and} \tag{48}$$

$$\frac{\partial h}{\partial a_0} = 0 . \tag{49}$$

As expected and as seen from a comparison of figures 11 and 17 the independence of h caused by $(\partial h/\partial a_0) = 0$ is found only in the case of independent binding sites. The derivatives presented in figure 14 are further confirmation of all three relations (47-49).

Next we try to find an approximation for the case of independent binding sites. In the equation

$$c = K_1\, a \cdot g = g_0 \frac{K_1\, a}{1 + K_1\, a + K_2\, h + K\, ah}$$

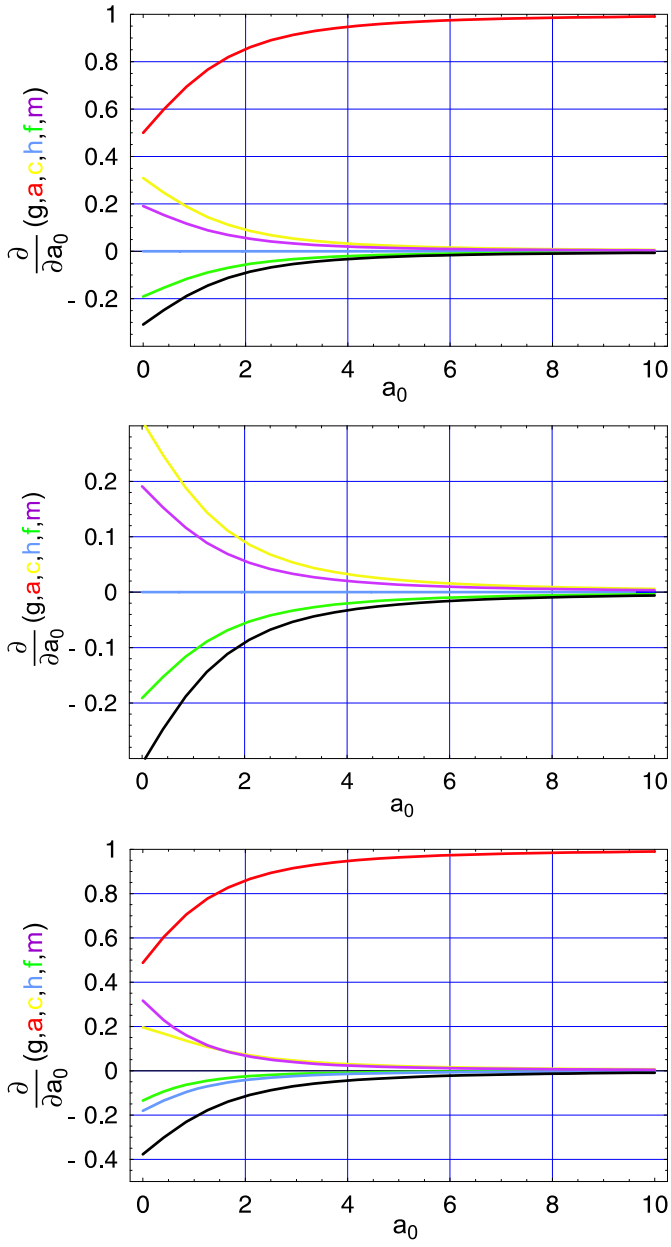


Figure 14: **Derivative of concentrations.** The topmost plot shows the first partial derivatives of the concentrations g, a, h, c, f, and m with respect to a_0 for the parameters $K = K_1 = K_2 = 1$ and the total concentrations $h_0 = g_0 = 1$. The plot in the middle is an enlargement in order to demonstrate the symmetry in the derivatives: $(\partial c/\partial a_0) = -(\partial g/\partial a_0)$ and $(\partial m/\partial a_0) = -(\partial f/\partial a_0)$. The plot at the bottom was computed with $K = 1$, $K_1 = 0.5$, $K_2 = 0.3$, and $h_0 = g_0 = 1$. Color code: $(\partial g/\partial a_0)$ black, $(\partial a/\partial a_0)$ red, $(\partial c/\partial a_0)$ yellow, $(\partial h/\partial a_0)$ blue, $(\partial f/\partial a_0)$ green, and $(\partial m/\partial a_0)$ purple.

the free concentrations of activator and repressor are replaced by the total concentrations, $a \approx a_0$ and $h \approx h_0$. Introduction of the independent binding sites approximation, $K = K_1 K_2$, allows for factorization of the denominator, $1 + K_1 a_0 + K_2 h_0 + K_1 K_2 a_0 h_0 = (1 + K_1 a_0)(1 + K_2 h_0)$, and eventually we find:

$$\gamma(g_0, a_0, h_0) \;=\; g_0 \cdot \frac{K_1 a_0}{1 + K_1 a_0} \cdot \frac{1}{1 + K_2 h_0} \;. \tag{50}$$

Comparisons of the functions $c(a_0, g_0)$ and $\gamma(a_0, g_0)$ and their difference are presented in terms of plots in figures 15 and 16.

It is worth to consider the behavior of $c(a_0, h_0)$ and its approximation $\gamma(a_0, h_0)$ in the limits of vanishing and large values of a_0 and h_0. The exact limits were already presented in equation (35-40). For the approximation we derive easily the cases where the asymptotic behavior is correct:

$$\lim_{a_0 \to 0} \gamma(a_0, h_0, g_0) \;=\; 0 \quad \text{and} \tag{51}$$

$$\lim_{h_0 \to \infty} \gamma(a_0, h_0, g_0) \;=\; 0 \;. \tag{52}$$

The other two limits where γ does not vanish require more care, since we observe differences between the exact analysis and the approximation there. In particular, we find

$$\lim_{a_0 \to \infty} \gamma(a_0, h_0, g_0) \;=\; \frac{g_0}{1 + K_2 h_0} \quad \text{and} \tag{53}$$

$$\lim_{h_0 \to 0} \gamma(a_0, h_0, g_0) \;=\; g_0 \, \frac{K_1 a_0}{1 + K_1 a_0} \;. \tag{54}$$

In figure 16 the error in the limit $a_0 \to \infty$ is easily recognized, since the tilt in the surface indicates a change in the sign if the deviation. The interpretation of the differences between c and γ is straightforward: The limits in both cases, (53) and (54), are represented by the equilibrium of complex formation, $\mathbf{A} + \mathbf{G} \rightleftharpoons \mathbf{C}$ and an error results from the replacement $a \approx a_0$ and $h \approx h_0$.

4.4 Cooperative binding of two different ligands

In the previous we assumed that the two binding sites are independent. In other words binding of the second ligand was neither facilitated nor impeded by the presence of the first one, and hence the equilibrium constant were identical: $K_1 = K_4$, $K_2 = K_3$, and $K = K_1 K_2$. Now we shall introduce a cooperativity parameter that measures the deviation from independent

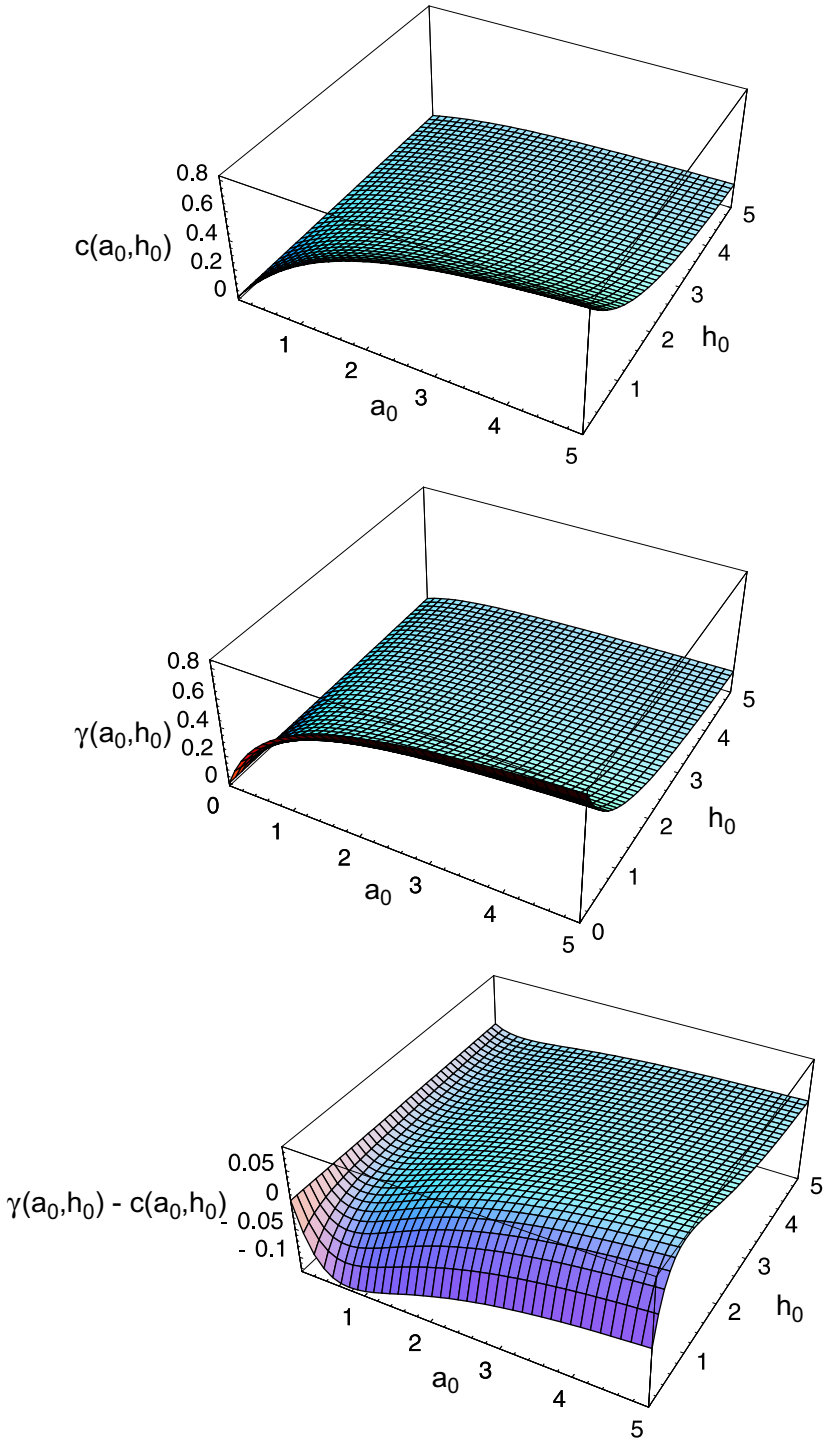


Figure 15: **Concentration of the activator complex** (30) **and the independent binding approximation.** The topmost plot shows $c(a_0, h_0)$ according to equation (30). In the middle we show the approximation $\gamma(a_0, h_0)$ according to equation (50). The plot at the bottom presents the difference $\gamma(a_0, h_0) - c(a_0, h_0)$. Choice of parameters: $K = K_1 = K_2 = 1$ and $g_0 = 1$.

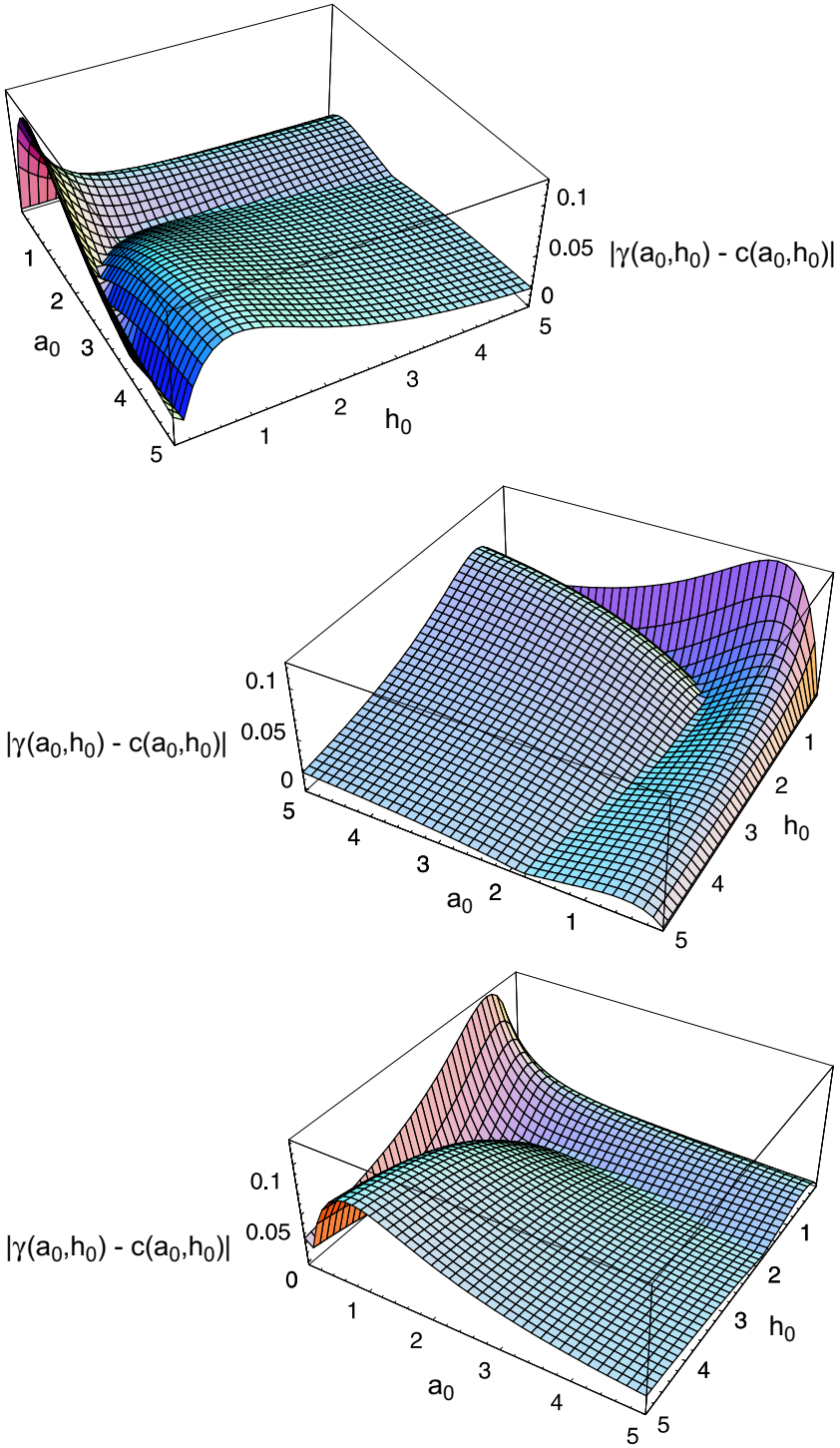


Figure 16: **Absolute value of the error in the independent binding approximation** (50)**.** All three plots show the absolute value of $\gamma(a_0, h_0) - c(a_0, h_0)$ from three different view points. Choice of parameters: $K = K_1 = K_2 = 1$ and $g_0 = 1$.

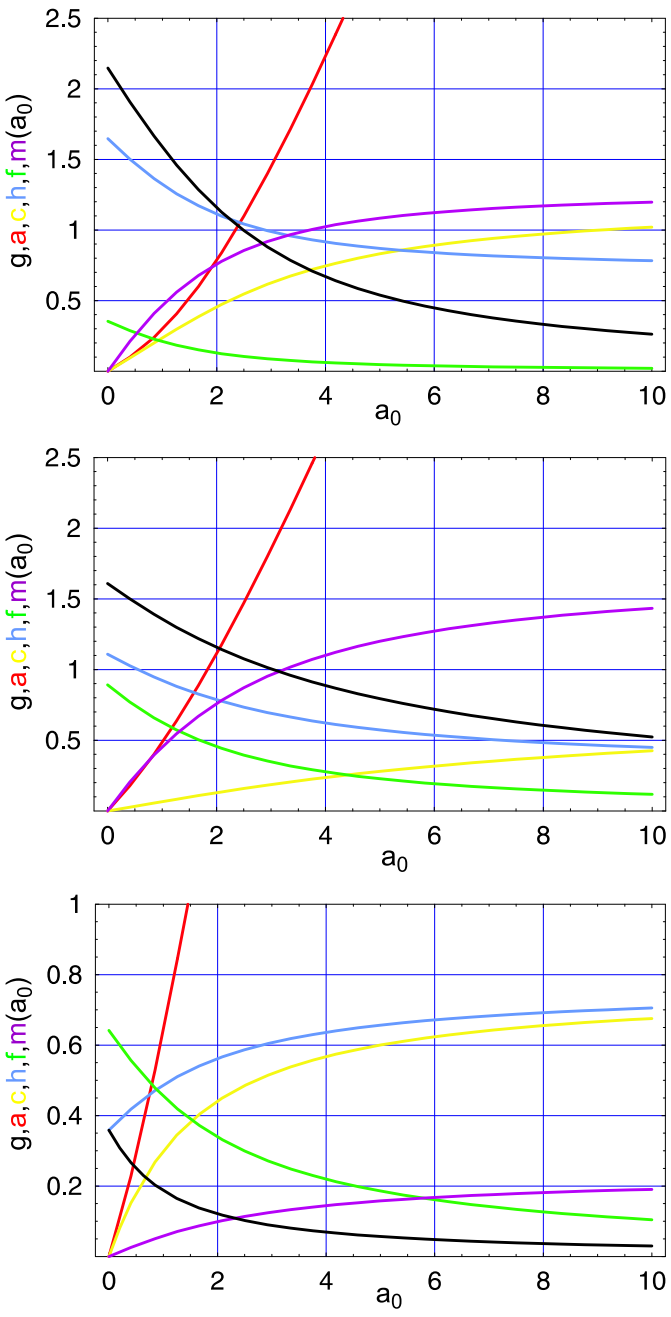


Figure 17: **Cooperative binding to two different sites.** The plots show the six concentrations, g, a, c, h, f, and m, as a function of a_0 The choice of parameters is: $K = 0.75$, $K_1 = 0.5$, $K_2 = 0.1$, $h_0 = 2$, and $g_0 = 2.5$ for the topmost plot, $K = 0.75$, $K_1 = 0.1$, $K_2 = 0.5$, $h_0 = 2$, and $g_0 = 2.5$ for the plot in the middle, and $K = 1$, $K_1 = 2.5$, $K_2 = 5$, $h_0 = g_0 = 1$ for the plot at the bottom. The upper two examples represent cases of cooperative binding, whereas the plot at the bottom is an anti-cooperative case. Color code: g black, a red, c yellow, h blue, f green, and m purple.

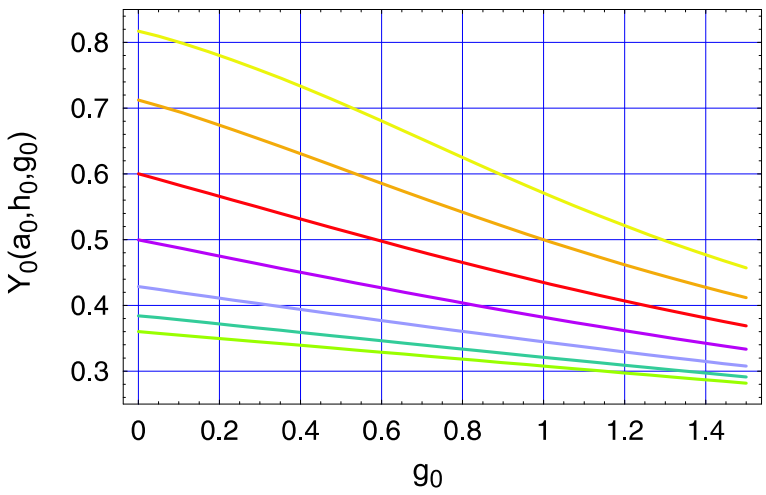


Figure 18: **The fraction of occupied sites Y_0 as a function of total concentration g_0.** The figures refers to a series of equilibrium constants, $K_1 = K_2 = 1$ and $K = 0.125, 0.25, 0.5, 1, 2, 4, 8$ (Series I). The total concentrations of the ligands were constant, $a_0 = h_0 = 1$. Color code: K =0.125 green, 0.25 cyan, 0.5 blue, 1 violet, 2 red, 4 orange, and 8 yellow.

binding:

$$\sigma = \frac{K_1 K_2}{K} \quad \text{and} \quad K_3 = \frac{K_2}{\sigma}, \; K_4 = \frac{K_1}{\sigma} \; . \tag{55}$$

The value $\sigma = 1$ is tantamount to no interaction and was discussed above (section 4.3, $0 < \sigma < 1$ is representative for the cooperative case in which the second ligand has higher affinity to the target than the first, and $\sigma > 1$ finally characterizes anti-cooperativity where binding of the second ligand is made more difficult by the presence of the first ligand. All three cases find straightforward physical interpretation: (i) no interaction is typical when the two binding sites are far apart and binding of the first ligand does not change the conformation of the macromolecule, (ii) cooperativity may have different causes – the first ligand, for example, provides a binding site for the second one, which is then bound by the target as well as by the first ligand, or a conformational change induces by the first ligand may increase the affinity for ligand number two – and (iii) anti-cooperativity may result from steric hindrance between the ligands to close-by binding sites or from conformational changes. Three typical examples of binding equilibria are shown in figure 17. Condition (49) is no longer valid and accordingly h varies with $a0$.

4.4.1 The binding function Y

The consequences of cooperative binding are illustrated best in terms of the binding function

$$Y(a,h) \;=\; \frac{[\text{number of occupied sites}]}{[\text{total number of sites}]} \;=\; \frac{c+f+2m}{2g_0} \;= \\ = \;\frac{1}{2}\cdot\frac{K_1\,a \;+\; K_2\,h \;+\; 2\,K\,a\,h}{1 \;+\; K_1\,a \;+\; K_2\,h \;+\; K\,a\,h}\;, \tag{56}$$

which describes the fraction of occupied binding sites as a function of the ligand concentrations, and which is normalized, $0 \leq Y \leq 1$. The first definition of the binding function allows also for usage of total concentrations, $Y_0(a_0, h_0, g_0)$ since we have the concentration of the complexes, c, f, and m available form equations (30-32). For the sake of simplicity, however, the binding function is often discussed in terms of the free ligand concentrations, a and h, only. We shall consider here both cases, the fraction of occupied sites as a function of total as well as free concentrations.

Comparison of $Y_0(a_0, h_0, g_0)$ and $Y(a, h)$ shows a difference already at the first glance: Y_0 depends on the total concentration of the gene, g_0, whereas $Y(a, h)$ does not. The dependence of Y_0 is shown for constant ligand total concentrations in figure 18. We notice almost linear relations and as expected

$$\lim_{g_0 \to 0} Y_0(a_0, h_0, g_0) \;=\; Y(a_0, h_0, 0) \;=\; Y(a, h)\;, \tag{57}$$

since in the absence of the gene we have $a = a_0$ and $h = h_0$.

In figure 19 we show the results of numerical computations of the binding function Y_0 from c, f, m, and g_0, and plot Y as a function of the total concentrations a_0 and h_0 for constant g_0 and different values of the equilibrium constants. The ratio of these parameters is color coded in the figure: yellow indicates the largest value of ratio K/K_1 or K/K_2, green the smallest value, violet represents the non-cooperative case $K = K_1 = K_2 = 1$, which fulfils the condition for independent binding of the two effectors, $K = K_1 K_2$ (section 4.3), and thus separates the curves for cooperative binding from those showing anti-cooperativity. Cooperative binding is often associated with a *sigmoid* binding function.[3] An inspection of figure 19, middle plot, shows, however, that the curve for $K/K_1 = K/K_2 = 2$ is not sigmoid and only the two curves for $K/K_1 = K/K_2 = 4$ and 8 fulfil this criterion. Hence, the notions of cooperativity and sigmoid binding curves are not equivalent. We

[3]The notion *sigmoid* indicates an 'S'-shape of the binding curve and is understood in contrast to a *hyperbolic* curve, which results, for example, for binding of a single ligand or for independent binding.

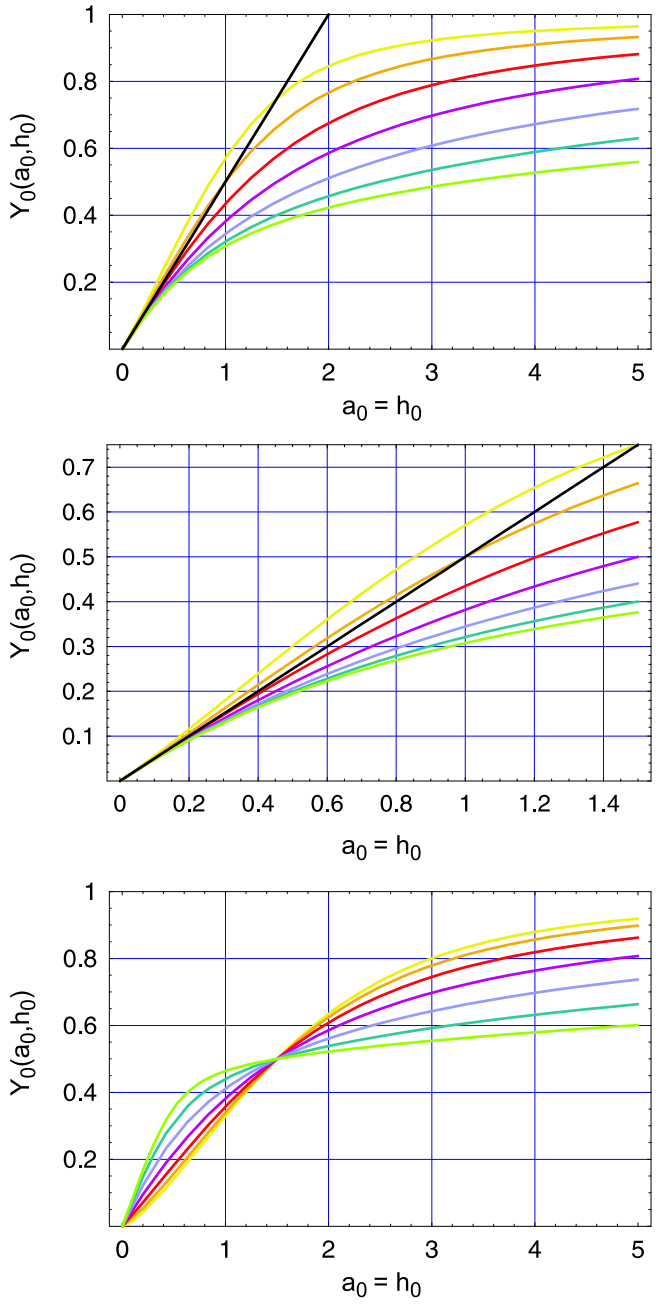


Figure 19: **The fraction of occupied sites Y_0 as a function of total concentrations a_0 and h_0.** The upper two figures refer to one series of equilibrium constants, $K_1 = K_2 = 1$ and $K = 0.125, 0.25, 0.5, 1, 2, 4, 8$ (Series I) whereas the bottom figure shows the series $K = 1$ and $K_1 = K_2 = 8, 4, 2, 1, 0.5, 0.25, 0.125$ (Series II). The total concentration of the gene was constant, $g_0 = 1$. Color code: (0.125;8) green, (0.25;4) cyan, (0.5;2) blue, (1;1) violet, (2;0.5) red, (4;0.25) orange, and (8;0.125) yellow, where the first and the second number in the parentheses refer to the varied parameters of series I and series II, respectively; upper two figures $a_0/2$ black.

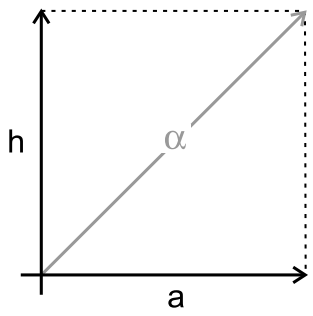


Figure 20: **Sketch of the definition of the direction $\boldsymbol{\alpha}$ for the variation of $Y(a,h)$.** The direction $\boldsymbol{\alpha}$ in the (a,h) plane of concentration space for the differentiation of $Y(a,h)$ is chosen to fulfil $\boldsymbol{\alpha} = \mathbf{a} + \mathbf{h}$. An identical sketch holds for α_0 in (a_0, h_0).

repeat: cooperative binding implies $K > K_1 K_2$ and anti-cooperativity means $K < K_1 K_2$. In the remaining part of this paragraph and in section 5.1.2 we shall analyze the relation between cooperativity and sigmoid binding curves in more detail.

4.4.2 Derivatives of the binding function

.Sigmoid curves have an inflection point where the sign of the second derivative changes sign.[4] Therefore we shall consider now first and second derivatives of the binding function $Y_0(a_0, h_0, g_0)$. Numerical differentiation yields the two partial derivatives $\partial Y_0/\partial a_0$ and $\partial Y_0/\partial h_0$. The directions of these two partial derivatives correspond to two orthogonal vectors in the (a_0, h_0)-plane of concentration space (figure 20). The actual differential dY_0 is a linear combination of the two partial derivatives depending on the applied changes in a_0 and h_0 corresponding to direction in the plane of the concentrations:

$$dY_0 \;=\; \left(\frac{\partial Y_0}{\partial a_0}\right) da_0 \;+\; \left(\frac{\partial Y_0}{\partial h_0}\right) dh_0 \;.$$

In order to be able to compare with the analogous system with two identical ligands (section 5.1 we choose the direction $\boldsymbol{\alpha} = \mathbf{a}+\mathbf{h}$ as sketched in figure 20 corresponding to $d\alpha_0 = da_0 + dh_0$. In other words, the change in total concentrations a_0 and h_0 is the same. This yields for the first derivative

$$\frac{dY_0}{d\alpha_0} \;=\; \left(\frac{\partial Y_0}{\partial a_0}\right) + \left(\frac{\partial Y_0}{\partial h_0}\right) \;.$$

[4]For the relation between the second derivative and the curvature of a function $y = f(x)$ see section 5.1.2.

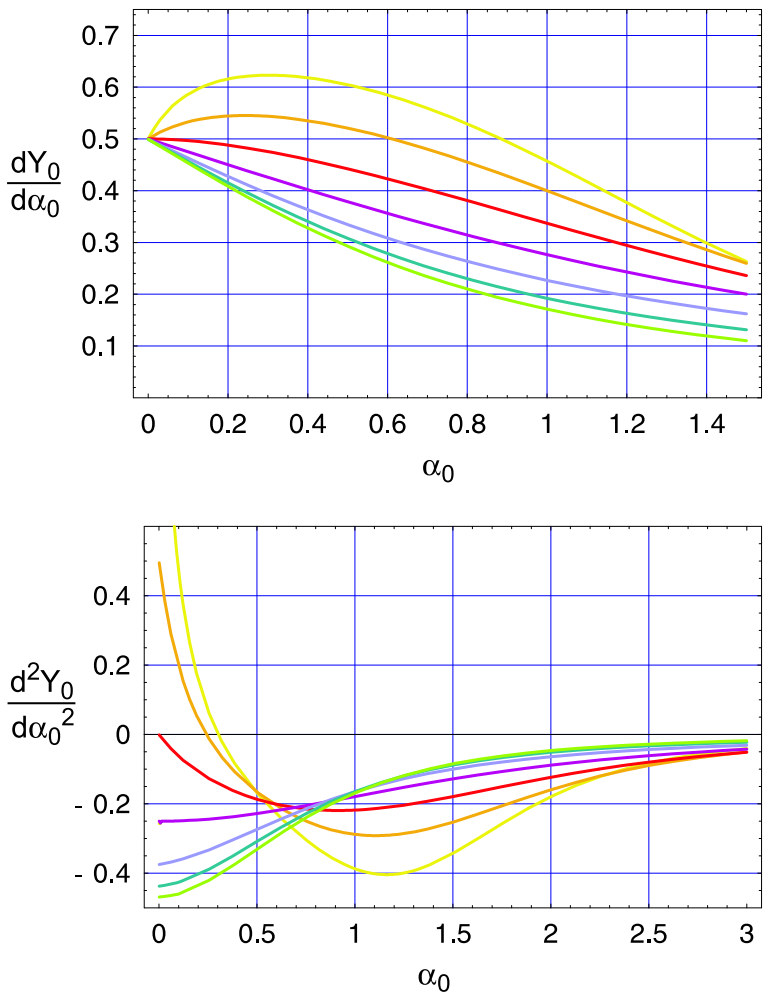


Figure 21: **First and second derivatives of the binding function $Y_0(a_0, h_0, g_0)$ plotted against total concentrations.** The derivatives were computed numerically by using the functions c, f, and m according to equation (56). The curves refer to series I of figure 19, K_1, K_2, and g_0 remain unchanged ($K_1 = K_2 = 1$, $g_0 = 1$) and $K = 0.125, 0.25, 0.5, 1, 2, 4, 8$. Color code: $K = 0.125$ green, 0.25 cyan, 0.5 blue, 1 violet, 2 red, 4 orange, and 8 yellow.

The the second differential d^2Y_0 is of the general form

$$d^2Y_0 \;=\; \left(\frac{\partial^2 Y_0}{\partial a_0^2}\right) da_0^2 \;+\; 2\left(\frac{\partial^2 Y_0}{\partial a_0 \partial h_0}\right) da_0\, dh_0 \;+\; \left(\frac{\partial^2 Y_0}{\partial h_0^2}\right) dh_0^2 \;.$$

Repeating now the consideration concerning the direction $\boldsymbol{\alpha}$ we find for the second derivative

$$\frac{d^2Y_0}{d\alpha_0^2} \;=\; \left(\frac{\partial^2 Y_0}{\partial a_0^2}\right) \,+\, 2\left(\frac{\partial^2 Y_0}{\partial a_0 \partial h_0}\right) \,+\, \left(\frac{\partial^2 Y_0}{\partial h_0^2}\right) \;.$$

The results of the numerical computations are shown in figure 21. The first derivatives are always positive and approach zero in the limit of large total concentrations of the ligands, $a_0 = h_0$. Our choice of variation in K at constant values of K_1 and K_2 leads to the same tangent in the limit $\alpha_0 \to 0$ which happens to be 0.5 in our example. The behavior of the second derivative is more interesting for the discussion of sigmoid curves. As expected $d^2Y_0/d\alpha_0^2$ is negative everywhere in the anti-cooperative case and for independent binding sites, $K \le K_1K_2$. A positive second derivative for small values of a_0 occurs only for $K > 2K_1K_2$. Then, the binding function Y_0 has an inflection point at some value $a_0 > 0$ and is 'S'-shaped. In the range $K_1K_2 < K < 2K_1K_2$ we observe a hyperbolic shape of $Y_0(\alpha_0)$ despite cooperativity in the sense of the definition applied here stating that the second ligand is bound stronger than the first one.

The dependence of the binding function $Y(a,h)$ on the free concentrations a and h allows for straightforward analysis since the derivatives are obtained by analytical differentiation. We begin with the first partial derivatives of $Y(a,h)$:

$$\frac{\partial Y}{\partial a} \;=\; \frac{K_1 + 2K\,h + KK_2\,h^2}{2(1 + K_1a + K_2h + K\,ah)^2} \tag{58}$$

$$\frac{\partial Y}{\partial h} \;=\; \frac{K_2 + 2K\,a + KK_1\,a^2}{2(1 + K_1a + K_2h + K\,ah)^2} \;. \tag{59}$$

As required by the symmetry of the kinetic system (figure 2) the two equation are interchangeable under the substitutions $a \leftrightarrow h$ and $K_1 \leftrightarrow K_2$. The change of Y with α is obtained by summation of the two partial derivatives:

$$\frac{dY}{d\alpha} \;=\; \frac{K_1 + K_2 + 2K\,(a+h) + K\,(K_1\,a^2 + K_2\,h^2)}{2(1 + K_1a + K_2h + K\,ah)^2} \;. \tag{60}$$

For the partial second derivatives we obtain:

$$\frac{\partial^2 Y}{\partial a^2} \;=\; -\frac{(K_1 + K\,h)(K_1 + 2K\,h + KK_2\,h^2)}{(1 + K_1a + K_2h + K\,ah)^3} \;, \tag{61}$$

$$\frac{\partial^2 Y}{\partial a \partial h} \;=\; -\frac{(-1 + K\,ah)(K - K_1K_2)}{2(1 + K_1a + K_2h + K\,ah)^3} \;, \tag{62}$$

$$\frac{\partial^2 Y}{\partial h^2} \;=\; -\frac{(K_2 + K\,a)(K_2 + 2K\,a + KK_2\,a^2)}{(1 + K_1a + K_2h + K\,ah)^3} \;. \tag{63}$$

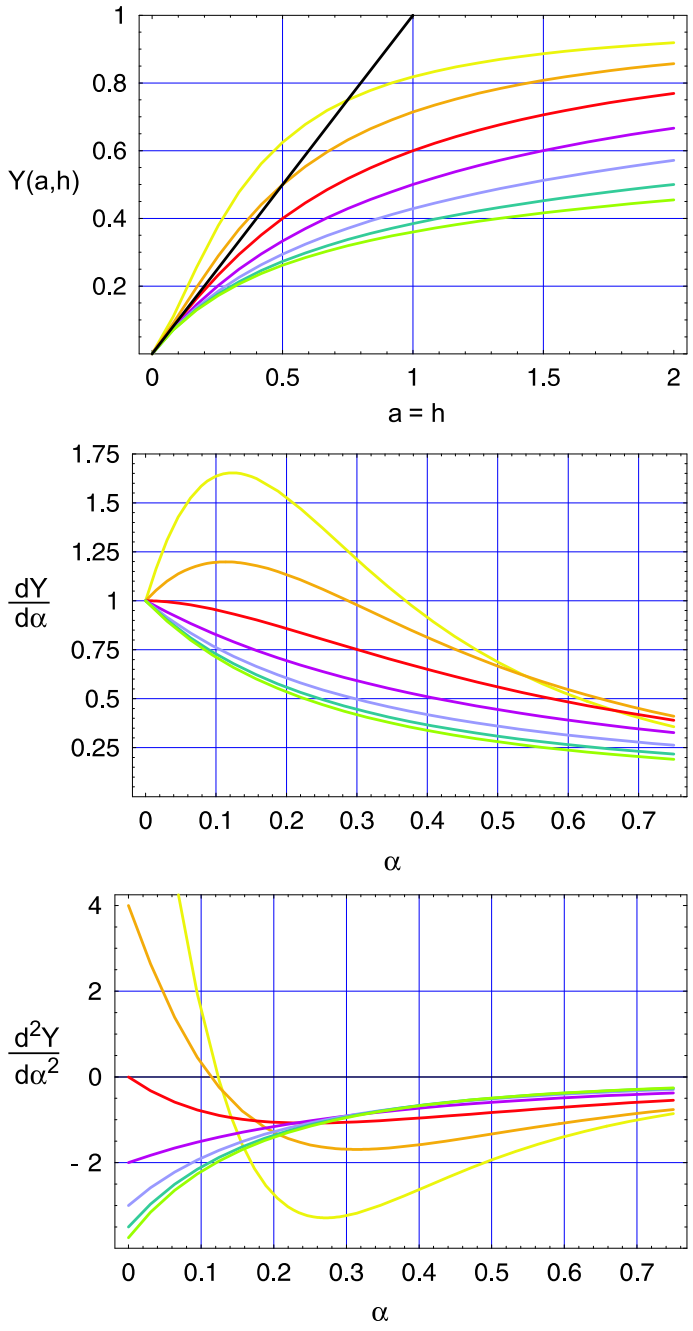


Figure 22: **The fraction of occupied sites Y and their derivatives as functions of the free concentrations a and h.** The series of equilibrium constants was chosen to be $K_1 = K_2 = 1$ and $K = 0.125, 0.25, 0.5, 1, 2, 4, 8$ (Series I). The total concentration of the gene was constant, $g_0 = 1$. Color code: 0.125 green, 0.25 cyan, 0.5 blue, 1 violet, 2 red, 4 orange, and 8 yellow, and in the topmost figure: a_0 black.

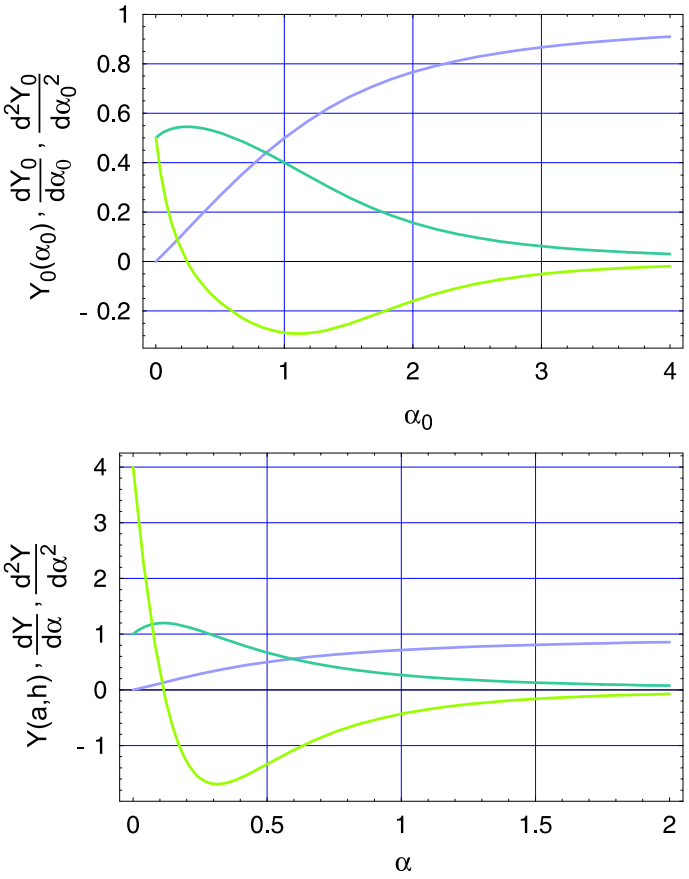


Figure 23: **Comparison of the binding functions $Y_0(a_0, h_0, g_0)$ and $Y(a, h)$ and their derivatives.** The two binding functions are compared for the equilibrium parameters $K_1 = K_2 = 1$ and $K = 4$, and $g_0 = 1$ for Y_0. This parameter choice belongs to the cooperative binding regime. Color code: Y blue, $dY/d\alpha$ cyan, $d^2Y/d\alpha^2$ green.

The second derivative in α is computed again by summation

$$\begin{aligned}\frac{\partial^2 Y}{\partial \alpha^2} &= -\frac{1}{(1+K_1a+K_2h+K\,ah)^3}\cdot \\ &\cdot\Big\{(K_1+K_2)^2-2K+3K(K_2a+K_1h)+ \\ &+K\Big((2K+K_1K_2)(a^2+h^2)+(K-K_1K_2)ah\Big)+K^2(K_1a^3+K_2h^3)\Big\}\,.\end{aligned} \tag{64}$$

Plots of $Y(a, h)$ as well as its derivatives for different values of the equilibrium constants are summarized in figure 22. We observe striking similarity between the two binding functions Y_0 and Y and their first and second derivatives.

Similar arguments as used in the case of equation (57) for Y_0 can be used to show that the derivatives of the two functions converge in the limit

lim $g_0 \to 0$, and we find

$$\lim_{g_0 \to 0} \frac{dY_0(\alpha_0, g_0)}{d\alpha_0} = \frac{dY(\alpha)}{d\alpha} , \tag{65}$$

$$\lim_{g_0 \to 0} \frac{d^2Y_0(\alpha_0, g_0)}{d\alpha_0^2} = \frac{d^2Y(\alpha)}{d\alpha^2} . \tag{66}$$

We remark that it is somewhat tricky to reach numerical convergence of the second derivative $d^2Y_0(\alpha_0, g_0)/d\alpha_0^2$ in the limit lim $g_0 \to 0$. In order to see how the two series of curves differ for nonzero values of total DNA concentration we compare at the value $g_0 = 1$ that has also been used before (figures 19-22). The qualitative agreement is complete and the observed quantitative differences are readily explained by the inequalities $a_0 \geq a$ and $h_0 \geq h$. In other words and somewhat sloppily phrased: More a_0 than a or more h_0 than h, respectively, is needed to cause the same effect on the binding curves. Therefore the function Y_0 and its derivatives appear stretched in comparison to Y and its derivatives.

5 Binding of a single ligand to several binding sites

In this section we study binding equilibria of a single ligand to the regulatory region of a gene. In particular we shall consider two and four binding sites. There is ample experimental evidence [1] that effectors may bind to DNA as two or more molecules, for example, in the form of dimers or higher oligomers. We shall analyze two different situations: (i) consecutive binding of single ligand molecules and (ii) dimerization or oligomerization of the ligand and binding to DNA as oligomer.

5.1 Two binding sites

In this section we present the analysis of two identical ligands binding cooperatively to two (neighboring) binding sites:

$$\mathsf{G} + 2\,\mathsf{A} \overset{K_1}{\rightleftharpoons} \mathsf{C} + \mathsf{A} \overset{K_2}{\rightleftharpoons} \mathsf{M} .$$

Herein, C and M represent the complexes $\mathsf{G}{\cdot}\mathsf{A}$ and $\mathsf{A}{\cdot}\mathsf{G}{\cdot}\mathsf{A}$, respectively. The non-cooperative case is characterized by independence of the binding equilibria for the two activator molecules. Because of statistical factors this case does not coincide with $K_1 = K_2$. In an alternative pathway the activator A forms first a dimer before it binds to the gene G:

$$\mathsf{G} + 2\,\mathsf{A} \overset{K_D}{\rightleftharpoons} \mathsf{G} + \mathsf{D} \overset{K_B}{\rightleftharpoons} \mathsf{M} .$$

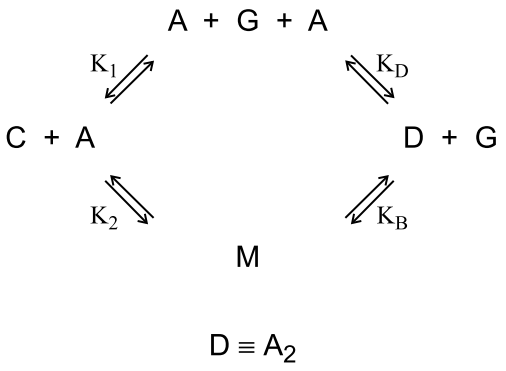


Figure 24: **Gene activation by two identical activator molecules**. The sketch shows a situation where gene activation occurs through binding of two activator molecules, which bind either consecutively or after dimer formation. The four binding constants are not independent and fulfil: $K_1 \cdot K_2 = K_D \cdot K_3 = K$. Accordingly we may use K_1, K_D, and K as three independent equilibrium constants.

If both pathways are relevant, as indicated in figure 24, we have the thermodynamic condition $K_1 \cdot K_2 = K_D \cdot K_B$. We shall start with the complete mechanism and discuss the two limiting cases thereafter.

5.1.1 Dimerization and consecutive binding

The full reaction scheme of figure 24 is characterized by two conservation relations

$$g_0 \;=\; g \,+\, c \,+\, m \quad \text{and} \tag{67}$$

$$a_0 \;=\; a \,+\, c \,+\, 2d \,+\, 2m \;, \tag{68}$$

and three equilibria

$$K_1 \;=\; \frac{c}{g \cdot a}\,, \quad K_D \;=\; \frac{c}{a^2}\,, \quad \text{and} \quad K \;=\; \frac{m}{g \cdot a^2}\;. \tag{69}$$

Conventional calculations yield a forth order equation in the free concentration of **A**, whereby we use again the dissociation constants rather than binding constants, $\kappa_1 = K_1^{-1}$, $\kappa_D = K_D^{-1}$, and $\kappa = K^{-1}$:

$$a^4 \,+\, \frac{\kappa_1\kappa_D + 2\kappa}{2\kappa_1}\, a^3 \,+\, \frac{(2\kappa_1 + \kappa_D)\kappa + \kappa_1\kappa_D(2g_0 - a_0)}{2\kappa_1}\, a^2 \,+$$

$$+\, \frac{\kappa_d\kappa(\kappa_1 + g_0 - a_0)}{2\kappa_1}\, a \,+\, \frac{\kappa_D\kappa\, a_0}{2} \;=\; 0\;. \tag{70}$$

The other concentrations are readily obtained from a:

$$g \;=\; \frac{\kappa_1 \kappa\, g_0}{\kappa_1\kappa + \kappa a + \kappa_1 a^2}\,,\; c \;=\; \frac{g\,a}{\kappa_1}\,,\; d \;=\; \frac{a^2}{\kappa_D}\,,\;\; \text{and}\;\; m \;=\; \frac{g\,a^2}{\kappa}\;. \tag{71}$$

The binding function in this case has the form

$$Y \;=\; \frac{c \,+\, 2m}{2g_0} \;=\; \frac{\kappa\, a \,+\, 2\kappa_1\, a^2}{2(\kappa_1\kappa \,+\, \kappa\, a + \kappa_1\, a^2)}\;. \tag{72}$$

Two properties of the binding function Y are of interest: (i) As in the previous case of activator-repressor binding, $Y(a)$ does not depend on g_0, and (ii) it does not depend explicitly on dimer formation. Indeed, the expression (72) is identical with the one we shall obtain for the binding mechanism without the dimer pathway. Consecutive binding of two identical molecules requires a special consideration of binding equilibria: Since **G** has two binding sites for the activator and **M** has two equivalent activator molecules bound in the complex, a statistical factor of 2 or $\frac{1}{2}$ is required to account for the difference between the macroscopic constants, K_1 and K, respectively, and the local or *microscopic* equilibrium parameters as outlined in section 5.1.2.

In figure 25 we show three examples for the dependence of equilibrium concentrations on a_0 with different combinations of equilibrium constants K_1, K, and K_D. A comparison of the lower two plots is of interest, because only K_D has been changed in this case and hence the binding function $Y(a)$ remains formally unchanged. The difference between the two plots concerns indeed dimer formation which dominates in the middle plot. Thus, although the actual shape of $Y(a_0)$ may be very different, the relation between Y, m, and c seems to only slightly effected.

5.1.2 Consecutive binding

Now we restrict the mechanism in figure 24 by eliminating the dimer and have as introduced at the beginning of this chapter

$$\mathbf{G} \,+\, 2\,\mathbf{A} \;\overset{K_1}{\rightleftharpoons}\; \mathbf{C} \,+\, \mathbf{A} \;\overset{K_2}{\rightleftharpoons}\; \mathbf{M}\;.$$

The system is determined by two total concentrations and two equilibrium constants:

$$g_0 \;=\; g \,+\, c \,+\, m \;\;\text{and}\;\; a_0 \;=\; a \,+\, c \,+\, 2m\;, \tag{73}$$

$$K_1 \;=\; \frac{c}{g\cdot a} \;\;\text{and}\;\; K \;=\; \frac{m}{g\cdot a^2} \;=\; K_1\, K_2\;. \tag{74}$$

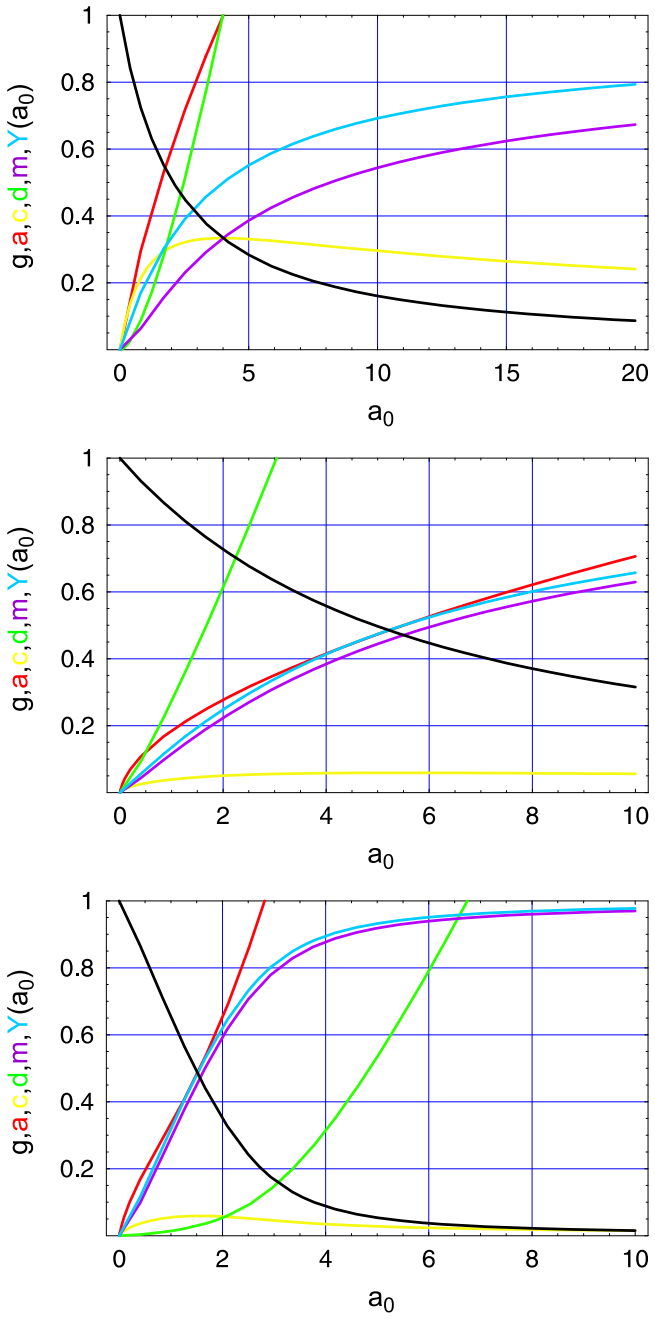


Figure 25: **Binding of two activator molecules to the gene with and without dimer formation.** The plots show the five concentrations, g, a, c, d, and m, and the binding function Y as a function of a_0. The choice of parameters is: $K = K_1 = K_D = 1$, $K = 4, K_1 = 0.25, K_D = 8$, and $K = 4, K_1 = 0.25, K_D = 0.125$, for the upper, middle and lower plot, respectively. The total gene concentration was constant: $g_0 = 1$. Color code: g black, a red, c yellow, d green, m purple, and Y blue.

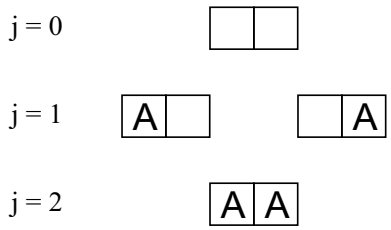


Figure 26: **Degeneracy of ligand binding to two sites**. A regulatory unit on DNA has two identical binding sites leading to a twofold degeneracy of the state with single occupancy ($j = 1$), which results in differences between the local or microscopic and the measurable or macroscopic binding constants.

For the purpose of comparison we shall also use the reciprocal equilibrium parameters: $\kappa = K^{-1}$ and $\kappa_1 = K_1^{-1}$.

Standard calculations yield the expressions for the free concentrations of the molecules and their complexes:

$$a^3 + \frac{K_1 + K(2g_0 - a_0)}{K}\, a^2 \;+\; \frac{1 + K_1(g_0 - a_0)}{K}\, a \;-\; \frac{a_0}{K} \;=\; 0\;, \tag{75}$$

$$g \;=\; \frac{g_0}{1 \;+\; K_1 a \;+\; K a^2}\;, \tag{76}$$

$$c \;=\; 2\,(g_0 - g) \;-\; (a_0 - a)\,, \quad \text{and} \quad m \;=\; (a_0 - a) \;+\; (g_0 - g)\;. \tag{77}$$

The binding function is the same as in the previous example:

$$Y \;=\; \frac{c \;+\; 2m}{2g_0} \;=\; \frac{K_1 a \;+\; 2K a^2}{2(1 \;+\; K_1 a \;+\; K a^2)} \;=\; \frac{\kappa\, a \;+\; 2\kappa_1\, a^2}{2(\kappa_1 \kappa \;+\; \kappa\, a + \kappa_1\, a^2)}\;. \tag{72}$$

Before we shall analyze this binding function we make a comparison with the binding function of the activator-inhibitor binding which we recall from section 4.4.1:[5]

$$\tilde{Y}(a,h) \;=\; \frac{1}{2} \cdot \frac{\tilde{K_1}\, a \;+\; \tilde{K_2}\, h \;+\; 2\,\tilde{K}\, a\, h}{1 \;+\; \tilde{K_1}\, a \;+\; \tilde{K_2}\, h \;+\; \tilde{K}\, a\, h}\;. \tag{56}$$

Now we assume the two ligands, the activator and the inhibitor to be identical, $\mathbf{H} \equiv \mathbf{A}$, $h = a$, and $\tilde{K_1} = \tilde{K_2}$ and find:

$$\tilde{Y}(a,h) \;=\; \frac{1}{2} \cdot \frac{2\,\tilde{K_1}\, a \;+\; 2\,\tilde{K}\, a^2}{1 \;+\; 2\,\tilde{K_1}\, a \;+\; \tilde{K}\, a^2}\;.$$

[5]In order to be able to distinguish between the two cases we marked the functions and constants of section 4.4.1 by tilde ($\tilde{\cdot}$).

As expected the binding function $\tilde{Y}$ becomes identical with Y of equation (72) if we set $K_1 = 2\tilde{K_1}$ and $K = \tilde{K}$. For the second sequential binding constants follows $K_2 = K/K_1 = \frac{1}{2}(\tilde{K}/\tilde{K_1})$. This apparent difference is readily interpreted by the statistical factors originating form the degeneracy of single ligand binding to two identical sites (figure 26). In this situation we have to distinguish between *macroscopic* or global binding constants and *microscopic* or local binding constants. The global constants are measured by experimental determination. In other words the observed difference in the comparison performed above results from the fact that the molecules **A** and **H** giving rise to equation (56) are distinguishable whereas the two molecules **A** considered in equation (72) are not. As show in figure 26 this gives rise two the factors 2 and $\frac{1}{2}$ between the global and the local binding constants for the first and the second step, respectively.

In order to derive general binding functions for local constants, γ, we introduce the following notation formulated for dissociation constants:

$$K_1^{-1} \,=\, \kappa_1 \,=\, \frac{1}{2}\,\gamma_1\,,\;\; K_2^{-1} \,=\, \kappa_2 \,=\, 2\,\gamma_2\,,\quad \text{and}\quad K^{-1} \,=\, \kappa \,=\, \gamma_1 \cdot \gamma_2\;. \tag{78}$$

The binding function rewritten in terms of local dissociation constants reads:

$$Y(a) \;=\; \frac{(\gamma_2 \,+\, a)\,a}{\gamma_1\gamma_2 \,+\, 2\,\gamma_2\,a \,+\, a^2}\;. \tag{79}$$

Calculation of the derivatives of the binding function is straightforward:

$$\frac{dY}{da} \;=\; \gamma_2\,\frac{\gamma_1\gamma_2 + 2\gamma_1 a + a^2}{(\gamma_1\gamma_2 + 2\gamma_1 a + a^2)^2} \quad \text{and} \tag{80}$$

$$\frac{d^2Y}{da^2} \;=\; 2\gamma_2\,\frac{\gamma_1\gamma_2(\gamma_1 - 2\gamma_2) - 3\gamma_1\gamma_2 a - 3\gamma_1 a^2 - a^3}{(\gamma_1\gamma_2 + 2\gamma_1 a + a^2)^3}\;. \tag{81}$$

The non-cooperative case is now easy to analyze by putting $\gamma_1 = \gamma_2 = \gamma$, and we find:

$$Y(a) \;=\; \frac{a}{\gamma + a}\,,\quad \frac{dY}{da} \;=\; \frac{\gamma}{(\gamma + a)^2}\,,\quad \text{and}\quad \frac{d^2Y}{da^2} \;=\; -\,\frac{2\gamma}{(\gamma + a)^3}\;.$$

As mentioned before the non-cooperative binding function is a hyperbola[6] with the second derivative $d^2Y/da^2 = -2/\gamma^2$ in the limit $\lim a \to 0$ exactly as found for independent binding of to different regulator proteins.

[6]This can be verified easily since $y = x/(1+x) \implies (x+1)(y-1)+1 = 0$ is indeed the conventional hyperbola, $x^2 - y^2 = 2$ rotated by $\varphi = -\pi/4$ and shifted such that the center coincides with the point $(x, y) = (-1, 1)$.

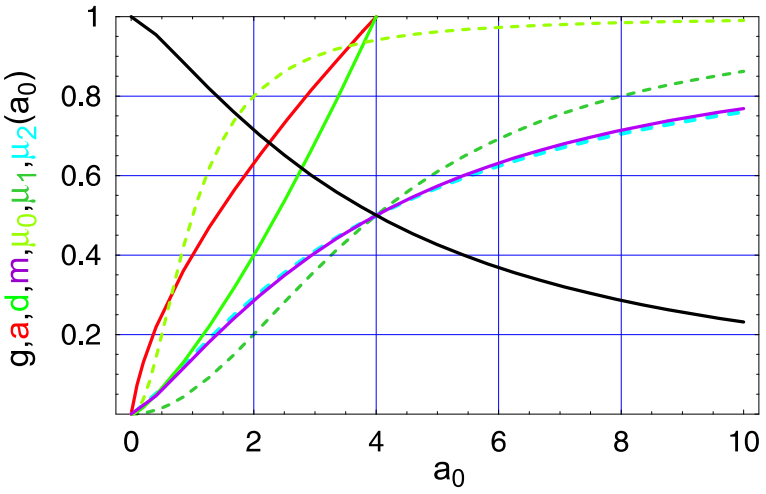


Figure 27: **Binding of two activator molecules as dimer to the gene.** Equilibrium concentrations as a function of the total activator concentration a_0. The equilibrium constants are $K_D = \kappa_D = 1$ and $K = \kappa = 1$. The total gene concentration was constant: $g_0 = 1$. Three approximations to $m(a_0)$ are shown: μ_0, μ_1, and μ_2 (dotted lines; definitions see in the text). Color code: g black, a red, d green, m purple, μ_0 yellow-green, μ_1 blue-green, and μ_2 cyan.

Finally, we make a remark concerning the curvature of a function and its second derivative that is relevant for the diagnosis of a sigmoid curve. The (extrinsic) curvature of a function $y(x)$ in two dimensions is defined as

$$\gamma(x) \;=\; \frac{d^2y}{dx^2} \Bigg/ \left(1 + \left(\frac{dy}{dx}\right)^2\right)^{3/2} . \tag{82}$$

To be sigmoid implies that the curve starts with positive curvature at $a = a_0 = 0$, with increasing a or a_0 the curvature decreases, goes to zero and finally becomes negative. From equation (82) follows that the second derivative is equally well suited for the qualitative discussion since the denominator is always positive and thus leaves the sign of the expression unchanged.

5.1.3 Dimer binding

The alternative pathway to consecutive binding is dimer formation and binding of the dimer

$$\mathbf{G} + 2\,\mathbf{A} \overset{K_D}{\rightleftharpoons} \mathbf{G} + \mathbf{D} \overset{K_B}{\rightleftharpoons} \mathbf{M} .$$

Again we are dealing with two conservation relations and two equilibrium constants

$$g_0 \;=\; g + m \;\text{ and }\; a_0 \;=\; a + 2d + 2m \;, \tag{83}$$

$$K_D \;=\; \frac{d}{a^2} \;\text{ and }\; K \;=\; \frac{m}{g \cdot a^2} \;=\; K_B\, K_D \;. \tag{84}$$

In this case the use of dissociation rather than biding constants, $\kappa_D = K_D^{-1}$ and $\kappa = K^{-1}$, is of advantage. Computation of equilibrium concentrations requires the (numerical) solution of the quartic equation:

$$a^4 \;+\; \frac{\kappa_D}{2}\, a^3 \;+\; \frac{2\kappa + \kappa_D(2g_0 - a_0)}{2}\, a^2 \;+\; \frac{\kappa_D \kappa}{2}\, a \;-\; \frac{\kappa_D \kappa\, a_0}{2} \;=\; 0 \;. \tag{85}$$

All other equilibrium concentrations follow straightforwardly from the free concentration of **A**

$$g \;=\; g_0\, \frac{\kappa}{\kappa + a^2} \;,\; d \;=\; \frac{a^2}{\kappa_D} \;,\;\; \text{and }\; m \;=\; g_0\, \frac{a^2}{\kappa + a^2} \;. \tag{86}$$

The binding function has a particularly simple form,

$$Y(a) \;=\; \frac{2m}{2g_0} \;=\; \frac{a^2}{\kappa + a^2} \;, \tag{87}$$

and for $g_0 = 1$ it is identical with the concentration of the activated complex **M**. A plot of the equilibrium concentration versus a_0 is shown in figure 27.

In order to derive suitable approximations for the concentration of the active activator-gene complex, m, we start by approximating a by a_0 in equation (88) as we did before and obtain:

$$\mu_0(\kappa, a_0) \;=\; g_0\, \frac{a_0^2}{\kappa + a_0^2} \tag{88}$$

Although μ_0 becomes exact in the limits $\lim a_0 \to 0$ and $\lim a_0 \to \infty$, it is a rather bad approximation as we can see in figure 27. Without invoking an additional parameter we can find a much better approximation $\mu_1(a_0, \tilde{\kappa})$ by applying a constant $\tilde{\kappa}$ that is determined by the condition that the approximation $\mu_1(\tilde{\kappa}, a_0) =$ is exact not only in above mentioned the limits but coincides with m also at the point $a_0 = \tilde{a}_0$ with $m(\tilde{a}_0) = g_0/2$. Using equations (85) and (86) we find:

$$\tilde{a}_0 \;=\; \sqrt{\tilde{\kappa}} \;=\; 2\kappa_B + \sqrt{\kappa} + g_0 \;\text{ and }\; \mu_1 \;=\; g_0\, \frac{a_0^2}{(2\kappa_B + \sqrt{\kappa} + g_0)^2 + a_0^2} \tag{89}$$

The improvement of the approximation is evident in figure 27. Still further improvement is possible through adjustment of the exponent $n = 2 \Rightarrow \tilde{n}$. The additional condition to be fulfilled is an identical tangent of $m(a_0)$ and $\mu_2(a_0, \tilde{\kappa}, \tilde{n})$ at the point $a_0 = \tilde{a}_0$:

$$\left.\frac{\partial m}{\partial a_0}\right|_{a_0=\tilde{a}_0} = \left.\frac{\partial \mu_2}{\partial a_0}\right|_{a_0=\tilde{a}_0} = \frac{\tilde{n} \cdot g_0}{4(2\kappa_B + \sqrt{\kappa} + g_0)} \ ,$$

$$\tilde{n} = \frac{4(2\kappa_B + \sqrt{\kappa} + g_0)}{g_0} \left.\frac{\partial m}{\partial a_0}\right|_{a_0=\tilde{a}_0} \quad \text{and}$$

$$\mu_2(a_0, \tilde{\kappa}, \tilde{n}) = g_0 \frac{a_0^{\tilde{n}}}{(2\kappa_B + \sqrt{\kappa} + g_0)^2 + a_0^{\tilde{n}}} \ . \tag{90}$$

The adjusted value for $\kappa = \kappa_D = g_0 = 1$ is $\tilde{n} = 4/3$. The approximation has now become excellent but we were using the exact tangent to the equilibrium concentration of **M** that, of course is not available without solving equation (85).

5.1.4 Properties of the functions $a^n/(\kappa^n + a^n)$

In the previous section 5.1.3 we used a binding function of the general class:

$$Y(a, \kappa, n) = \frac{a^n}{\kappa^n + a^n} \ . \tag{91}$$

Here we shall analyze now some properties of these two-parameter functions and relate them to the conventional notions of cooperativity. Clearly, the general properties of a binding function are fulfilled by equation (91):
(i) $\lim_{a\to 0} Y(a, \kappa, n) = 0$, (ii) $\lim_{a\to\infty} Y(a, \kappa, n) = 0$, and (iii) $Y(a, \kappa, n)$ is monotonously increasing with a. Next we shall study first and second derivatives of $Y(a, \kappa, n)$ and consider their behavior in the limit $\lim a \to 0$.

Differentiation is straightforward and yields:

$$\frac{\partial Y}{\partial a} = \frac{n\,\kappa^n\,a^{n-1}}{(\kappa^n + a^n)^2} \quad \text{and} \tag{92}$$

$$\frac{\partial^2 Y}{\partial a^2} = -\frac{n\,\kappa^n\,a^{n-2}\ (a^n(n+1) + k^n(1-n))}{(\kappa^n + a^n)^3} \ . \tag{93}$$

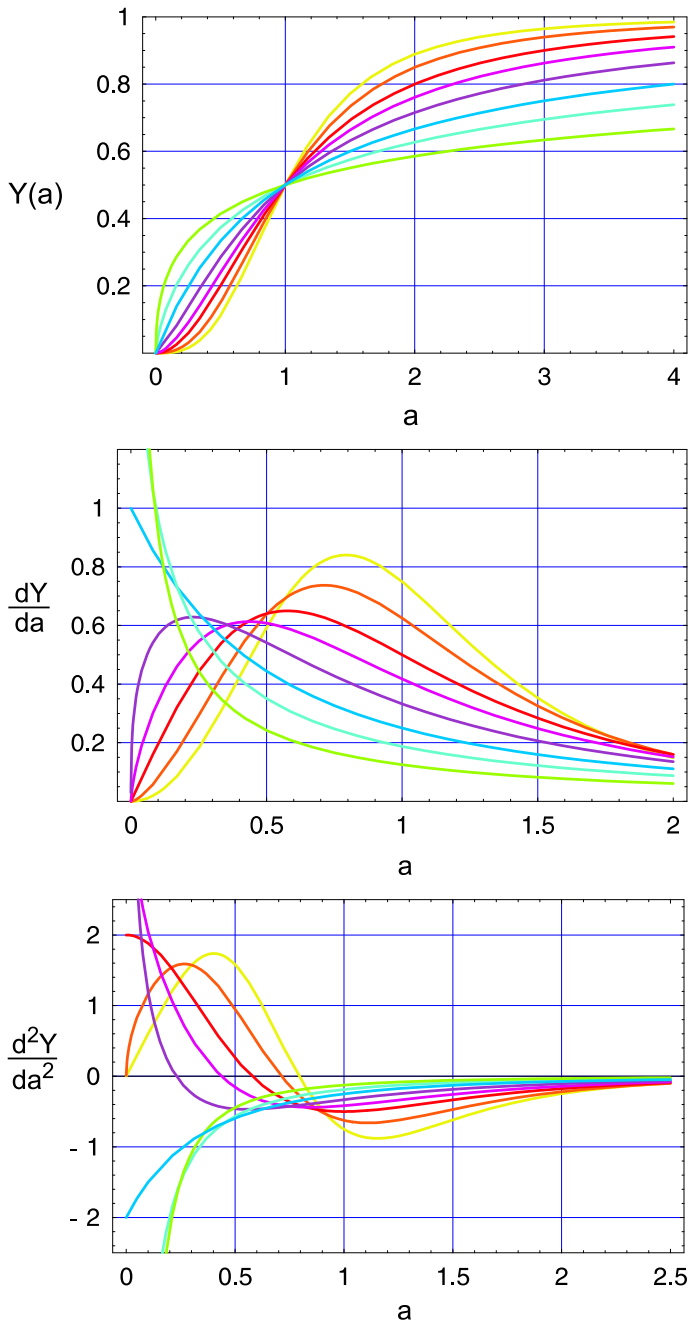


Figure 28: **The binding function $Y(a) = a^n/(\kappa^n + a^n)$ and its derivatives.** The binding functions and the first and second derivative are shown as functions of a for $\kappa = 1$ and different values of n. Color code: $n = 0.5$ yellow-green, 0.75 blue-green, 1 blue, 1.33 blue-violet, 1.67 red-violet, 2 red, 2.5 orange, and 3 yellow.

In the limit $\lim a \to 0$ the functions Y show a variety of different behavior:

$$\lim_{a\to 0} \frac{\partial Y}{\partial a} = \begin{cases} 0 & \text{if } n > 1\,, \\ \kappa^{-1} & \text{if } n = 1\,, \quad \text{and} \\ \infty & \text{if } n < 1\,. \end{cases} ,$$

$$\lim_{a\to 0} \frac{\partial^2 Y}{\partial a^2} = \begin{cases} 0 & \text{if } n > 2\,, \\ +2\kappa^{-2} & \text{if } n = 2\,, \\ +\infty & \text{if } 2 > n > 1\,, \\ -2\kappa^{-2} & \text{if } n = 1\,, \quad \text{and} \\ -\infty & \text{if } n < 1\,. \end{cases} .$$

These results are easily interpreted by considering the analytical continuation of Y-functions to negative a-values where they are available. For $n = 2$, for example, $Y(a)$ is well approximated by $Y(a) \approx \kappa^{-2} \cdot a^2$ around $a = 0$ and thus behaves like a parabola with the minimum at the origin and a second derivative of $2\kappa^{-2}$ there. By the same token the origin is an inflection point for $n = 3$ since $Y(a) \approx \kappa - 3 \cdot a^3$ has a vanishing derivative at $a = 0$. For the non-cooperative case $n = 1$ we find

$$\lim_{a\to 0} \frac{\partial Y}{\partial a} = \frac{1}{\kappa} \quad \text{and} \quad \lim_{a\to 0} \frac{\partial^2 Y}{\partial a^2} = -\frac{2}{\kappa^2}$$

in agreement with our previous observations for several special cases. At $a = 0$ the curve for non-cooperative binding starts with positive slope and negative second derivative. Examples of binding curves $Y(a, \kappa, n)$ and their derivatives are shown in figure 28.

In summary, the binding functions $Y(a, \kappa, n)$ show highly different behavior for different values of n with a few important features:

(i) Negative values of n describe inhibition whereas, as seen in the previous section 5.1.3, the positive values correspond to activation.

(ii) The functions have an analytical continuation into negative values of a only for integer n.

(iii) The non-cooperative case has a negative second derivative at the origin $a = 0$ and in this aspect the simplified functions are very similar to the exact curves discussed before.

5.2 Four binding sites

Finally, we shell consider two cases of binding four identical molecules to a target: consecutive binding and binding as a tetramer that is formed by association of two dimers.

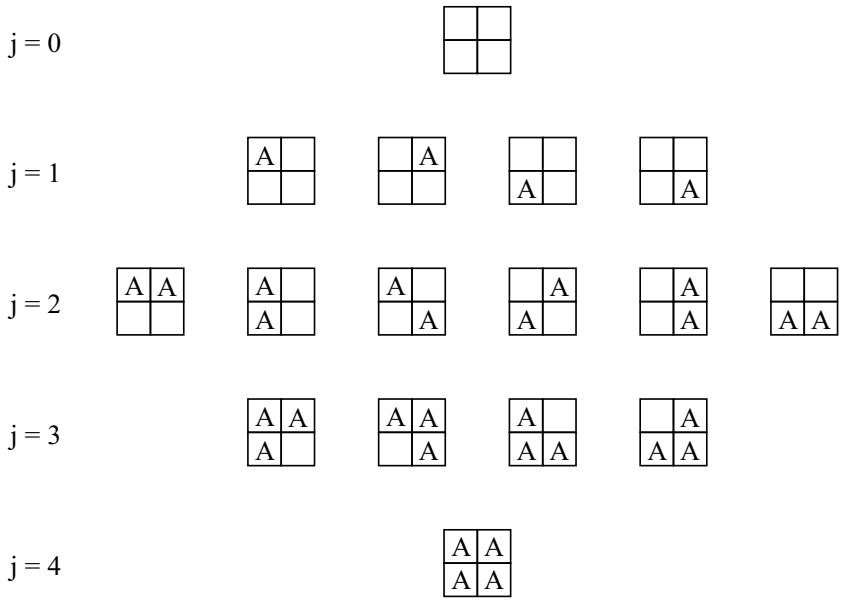


Figure 29: **Degeneracy of ligand binding to four sites**. A regulatory unit on DNA has four identical binding sites leading to fourfold, sixfold, and fourfold degeneracy of the state with single, double, and triple occupancy ($j = 1, 2, 3$), respectively. This degeneracy results in differences between the local or microscopic and the measurable or macroscopic binding constants.

5.2.1 Consecutive binding

In consecutive binding the ligands associate with the target one at a time:

$$\mathbf{G} + 4\,\mathbf{A} \overset{K_1}{\rightleftharpoons} \mathbf{C} + 3\,\mathbf{A} \overset{K_2}{\rightleftharpoons} \mathbf{B} + 2\,\mathbf{A} \overset{K_3}{\rightleftharpoons} \mathbf{F} + \mathbf{A} \overset{K_4}{\rightleftharpoons} \mathbf{M}\,.$$

We have two conservation rules and four equilibrium constants to determine the six unknown concentrations:

$$g_0 \,=\, g + c + b + f + m \quad \text{and} \quad a_0 \,=\, a + c + 2b + 3f + 4m\;, \tag{94}$$

$$K_1 \,=\, \frac{c}{g\cdot a}\,,\; K_2 \,=\, \frac{b}{c\cdot a}\,,\; K_3 \,=\, \frac{f}{b\cdot a}\,,\;\; \text{and}\;\; K_4 \,=\, \frac{m}{f\cdot a}\,. \tag{95}$$

As in most of the previous sections we shall use dissociation rather than binding constants, $\kappa_j = K_j^{-1}$; $(j = 1, 2, 3, 4)$. The computation of equilibrium concentrations requires solution of a quintic equation:

$$\begin{aligned} &a^5 + (\kappa_4 + 4g_0 - a_0)\,a^4 + \kappa_4(\kappa_3 + 3g_0 - a_0)\,a^3 + \\ &\quad + \kappa_4\kappa_3(\kappa_2 + 2g_0 - a_0)\,a^2 + \kappa_4\kappa_3\kappa 2(\kappa_1 + g_0 - a_0\,)a - \\ &\quad - \kappa_4\kappa_3\kappa_2\kappa_1\,a_0 = 0\;. \end{aligned} \tag{96}$$

It is not difficult to recognize the built-up principle of this equation that allows for straightforward extension to an arbitrary number of consecutively bound ligands. The equilibrium concentrations of the other species are readily obtained from known a:

$$g \;=\; \frac{\kappa_4\kappa_3\kappa_2\kappa_1\, g_0}{\kappa_4\kappa_3\kappa_2\kappa_1 \,+\kappa_4\kappa_3\kappa_2\, a \,+\, \kappa_4\kappa_3\, a^2 \,+\, \kappa_4\, a^3 \,+\, a^4} \;, \tag{97}$$

$$c \;=\; \frac{a\cdot g}{\kappa_1}\,,\; b \;=\; \frac{a\cdot c}{\kappa_2}\,,\; f \;=\; \frac{a\cdot b}{\kappa_2}\,, \quad \text{and} \tag{98}$$

$$m \;=\; g_0\,\frac{a^4}{\kappa_4\kappa_3\kappa_2\kappa_1 \,+\kappa_4\kappa_3\kappa_2\, a \,+\, \kappa_4\kappa_3\, a^2 \,+\, \kappa_4\, a^3 \,+\, a^4} \;. \tag{99}$$

Then we find for the binding function

$$\begin{aligned} Y(a) \;&=\; \frac{c \,+\, 2\,b \,+\, 3\,f \,+\, 4\,m}{4\,g_0} \;= \\ &=\; \frac{\kappa_4\kappa_3\kappa_2\, a \,+\, 2\,\kappa_4\kappa_3\, a^2 \,+\, 3\,\kappa_4\, a^3 \,+ 4\, a^4}{4\,(\kappa_4\kappa_3\kappa_2\kappa_1 \,+\kappa_4\kappa_3\kappa_2\, a \,+\, \kappa_4\kappa_3\, a^2 \,+\, \kappa_4\, a^3 \,+\, a^4)} \;. \end{aligned} \tag{100}$$

Examples of plots of the equilibrium concentrations as functions of a_0 are shown in figures 30 and 31.

In the case of binding identical ligands at equivalent sites requires to consider statistical factors as we did already in section 5.1.2. The situation with four ligands is sketched in figure 29. Here, we derive now the expressions for the general case of j ligands binding to n equivalent sites. If all ligand combinations for given n and j are equivalent the statistical factor $\omega_j^{(n)}$ is the number of combinations of j objects, the ligands, chosen from a set of n objects, the sites, and this is simply the binomial coefficient $\binom{n}{j}$. Next we consider the binding equilibrium between the occupation numbers $j-1$ and j. The state final state 'j' has degeneracy $\binom{n}{j}$ and the initial state is $\binom{n}{j-1}$ fold degenerate and this yields for the macroscopic equilibrium constant

$$K_j \;=\; \frac{\omega_j^{(n)}}{\omega_{j-1}^{(n)}}\cdot\chi_j \;=\; \frac{\binom{n}{j}}{\binom{n}{j-1}}\cdot\chi_j \;=\; \frac{n-j+1}{j}\cdot\chi_j \;,$$

where χ_j is the microscopic equilibrium parameter corresponding to K_j. In terms of dissociation equilibria, which we are using preferentially here, the relation between microscopic constants, $\gamma_j = \chi_j^{-1}$, and macroscopic constants, $\kappa_j = K_j^{-1}$ is of the form

$$\kappa_j \;=\; g_j^{(n)}\,\gamma_j \;=\; \frac{\omega_{j-1}(n)}{\omega_j^{(n)}} \;\cdot\; \gamma_j \;=\; \frac{j}{n-j+1} \;\cdot\; \gamma_j \;. \tag{101}$$

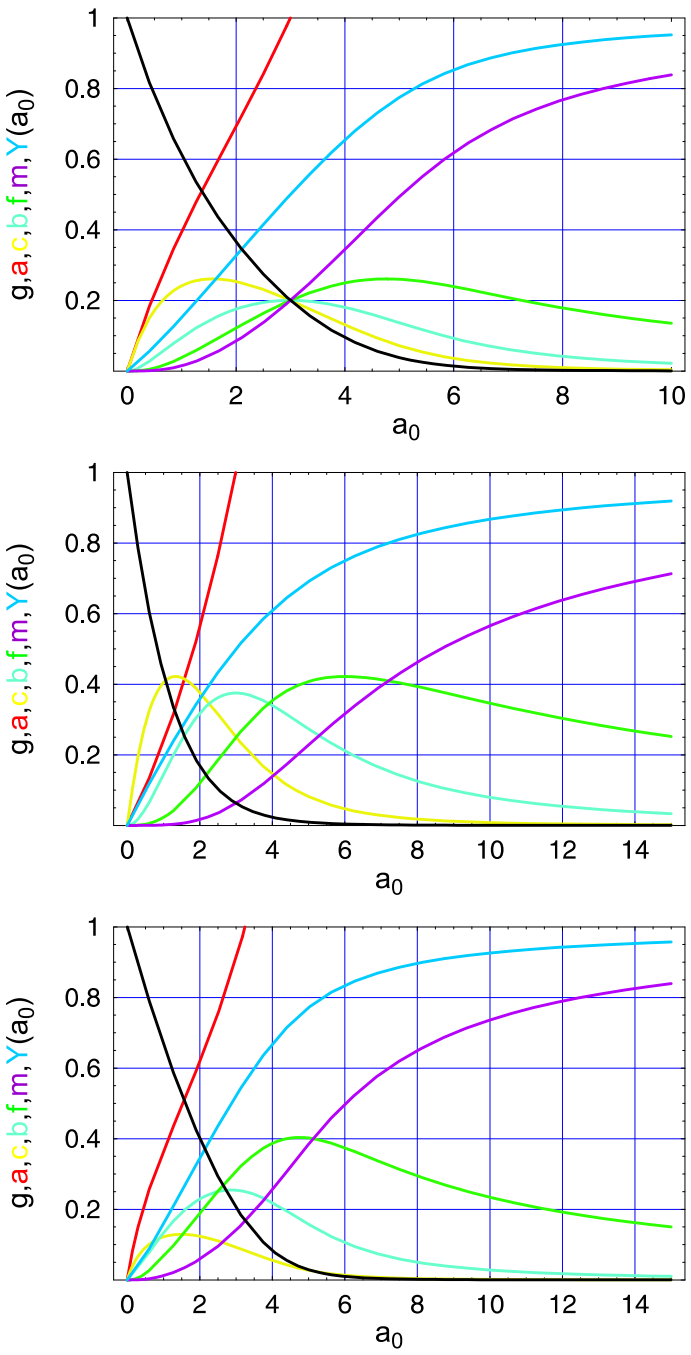


Figure 30: **Consecutive binding of four ligands.** The plots show the the equilibrium concentrations and the binding ration Y as functions a_0 at $g_0 = 1$. The topmost plot corresponds to $\kappa_1 = \kappa_2 = \kappa_3 = \kappa_4 = 1$. The middle and lower plots are examples for the non-cooperative and the cooperative case and correspond to the middle and lower plot in figure 31, respectively. The values of the equilibrium constants are given in the caption of figure 31. Color code: g black, a red, c yellow, b blue-green, f green, m purple, and Y blue.

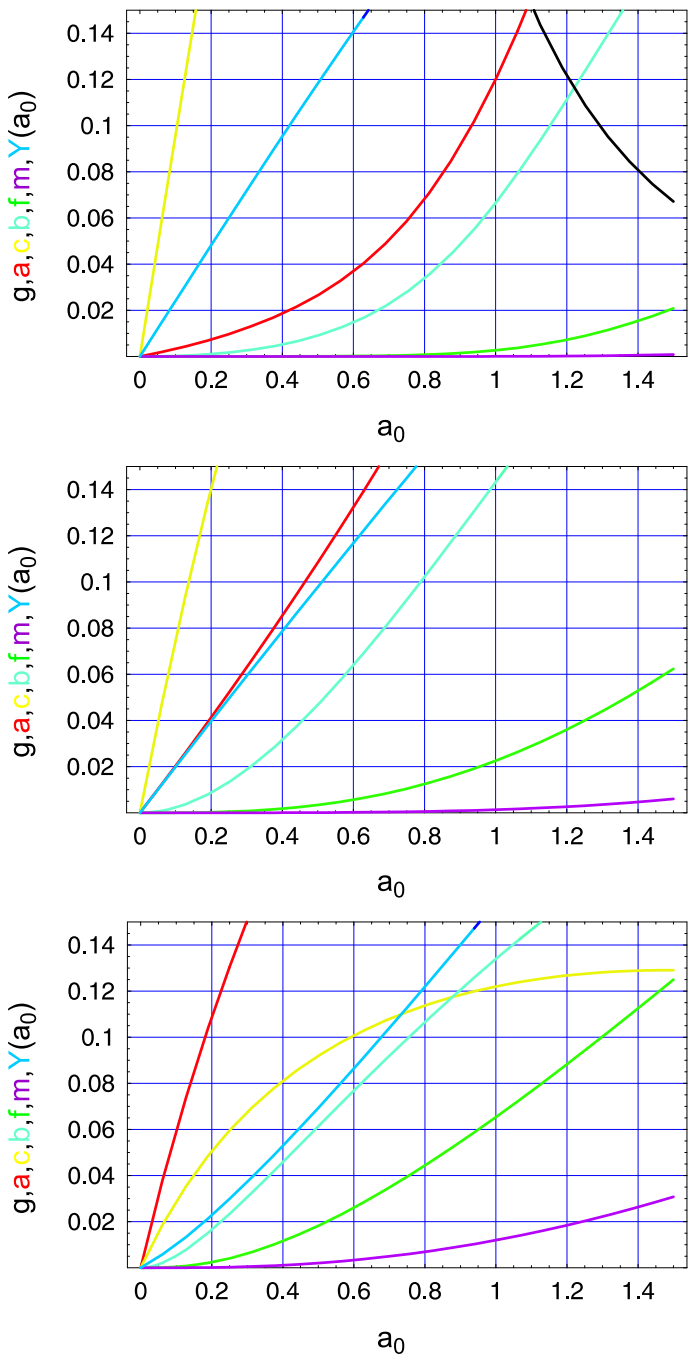


Figure 31: **Initial phase of consecutive binding of four ligands.** The plots show the the equilibrium concentrations and the binding function at low values of a_0. The upper, middle, and lower plot represent examples for anti-, non-cooperative and cooperative binding. The equilibrium constants were: $\kappa_1 = \rho\gamma/4$ and $3\kappa_2/2 = 2\kappa_3/3 = \kappa_4/4 = \gamma$ with $\rho = 0.0625, 1, 16$ and $\gamma = 2, 1, 0.5$, respectively, and $g_0 = 1$. Color code: g black, a red, c yellow, b blue-green, f green, m purple, and Y blue.

For our special case with $n = 4$ we have accordingly

$$\kappa_1 \;=\; \frac{1}{4}\,\gamma_1\,,\;\; \kappa_2 \;=\; \frac{2}{3}\,\gamma_2\,,\;\; \kappa_3 \;=\; \frac{3}{2}\,\gamma_3\,,\quad \text{and}\quad \kappa_4 \;=\; 4\,\gamma_4\;.$$

In general, and as we have seen in the case of two binding sites, we find that the product of all κ's is equal to the product of all γ's:[7] $\kappa_1\kappa_2\kappa_3\kappa_4 = \gamma_1\gamma_2\gamma_3\gamma_4$. It is straightforward now to compute the binding function in terms of microscopic equilibrium constants:

$$Y(a) \;=\; \frac{a\,(\gamma_2\gamma_3\gamma_4 \;+\; 3\,\gamma_3\gamma_4\,a \;+\; 3\,\gamma_4\,a^2 \;+\; a^3)}{\gamma_1\gamma_2\gamma_3\gamma_4 \;+\; 4\,\gamma_2\gamma_3\gamma_4\,a \;+\; 6\,\gamma_3\gamma_4\,a^2 \;+\; 4\,\gamma_4\,a^3 \;+\; a^4}\;. \tag{102}$$

All γ's are equal in the non-cooperative case, $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = \gamma$ and then we obtain the hyperbolic binding curve $Y(a) = a/(\gamma + a)$ as expected. Figures 30 and 31 show several examples of equilibrium concentrations and binding curves. For the anti-cooperative and the cooperative cases we assumed that only the first microscopic dissociation constant differs from the others: $\gamma_1 = \rho \cdot \gamma$ and $\gamma_2 = \gamma_3 = \gamma_4 = \gamma$.

Two results derived from first and second derivatives of the binding function are of general interest:

$$\lim_{a\to 0}\frac{\partial Y}{\partial a} \;=\; \frac{1}{\gamma_1} \qquad\qquad \text{and} \tag{103}$$

$$\lim_{a\to 0}\frac{\partial^2 Y}{\partial a^2} \;=\; \frac{2(3\gamma_1 - 4\gamma_2)}{\gamma_1^2\,\gamma_2}\;. \tag{104}$$

As in the dimer binding case the binding curve $Y(a)$ thus starts with a negative curvature at $a = 0$ and not with curvature zero. The curves for the cooperative cases $\gamma_1 > \gamma_2 \geq 3\gamma_1/4$ are not sigmoid.

It is possible to generalize the results shown above to arbitrarily long chains of consecutive binding of n ligands: Equation (103) is valid for all integer $n \geq 1$, and for the second derivative (104) we find:

$$\lim_{a\to 0}\frac{\partial^2 Y}{\partial a^2} \;=\; \frac{2\big((n-1)\,\gamma_1 \;-\; n\,\gamma_2\big)}{\gamma_1^2\,\gamma_2}\;. \tag{105}$$

In the limit $n \to \infty$ the curvature of the binding function $Y(a)$ vanishes indeed for $\gamma_1 = \gamma_2$ and then the non-cooperative case separates sigmoid and hyperbolic binding functions.

[7] The same holds, of course, for the product of all K's and all χ's.

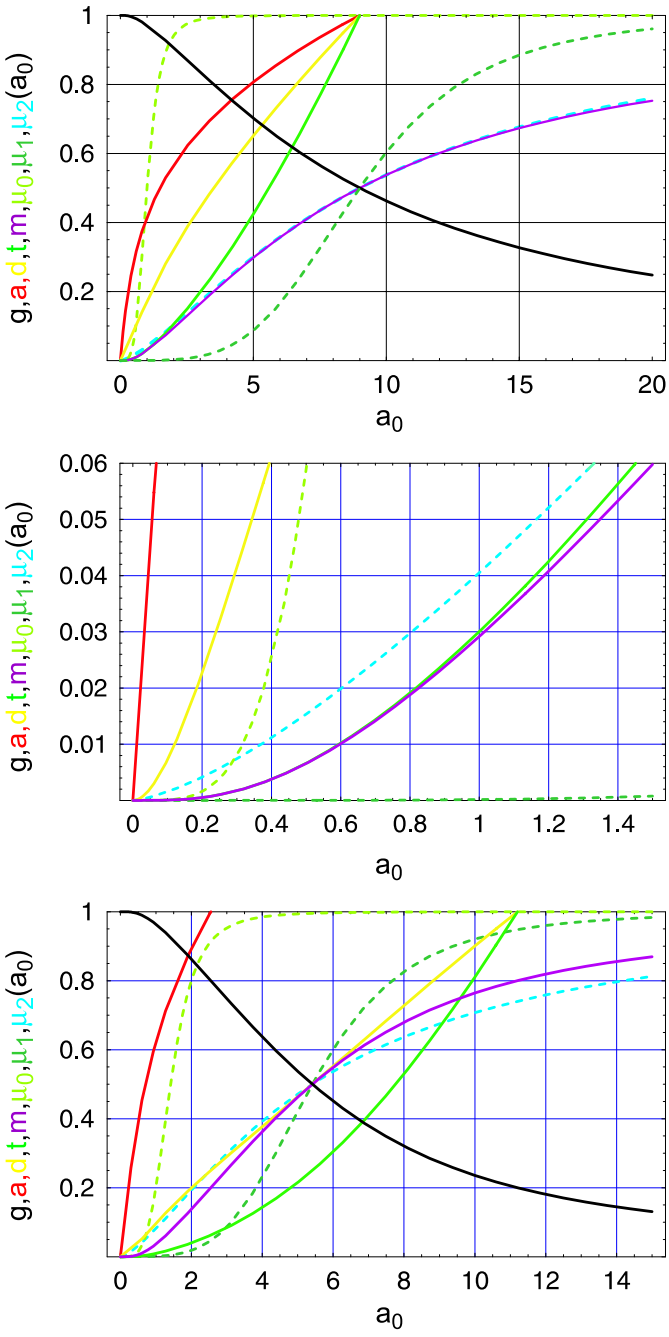


Figure 32: **Binding of four activator molecules as tetramer to the gene.** Equilibrium concentrations as a function of the total activator concentration a_0. The total gene concentration was constant: $g_0 = 1$. The equilibrium constants were $\kappa_D = \kappa_T = \kappa_B = 1$ for the upper and middle plot, and $\kappa_D = 4$, $\kappa_T = 1$, and $\kappa_B = 0.25$ for the plot at the bottom. Three approximations to $m(a_0)$ are shown: μ_0, μ_1, and μ_2 (dotted lines; definitions see in the text). Color code: g black, a red, d yellow, t green, m purple, μ_0 yellow-green, μ_1 blue-green, and μ_2 cyan.

5.2.2 Tetramer binding

Our last example of ligand binding involves dimerization and tetramerization prior to binding to the DNA:

$$\mathsf{G} + 4\,\mathsf{A} \overset{K_D}{\rightleftharpoons} \mathsf{G} + 2\,\mathsf{D} \overset{K_T}{\rightleftharpoons} \mathsf{G} + \mathsf{T} \overset{K_B}{\rightleftharpoons} \mathsf{M}\,.$$

Two conservation relation and three binding equilibria determine the five unknown equilibrium concentrations:

$$g_0 \;=\; g + m \quad \text{and} \quad a_0 \;=\; a + 2d + 4t + 4m\;, \tag{106}$$

$$K_D \;=\; \frac{d}{a^2}\,,\; K_T \;=\; \frac{t}{d^2}\,,\;\; \text{and} \;\; K_B \;=\; \frac{m}{g\cdot t}\,. \tag{107}$$

For convenience we us dissociation constants, $\kappa_D = K_D^{-1}$, $\kappa_T = K_T^{-1}$, and $\kappa_B = K_B^{-1}$. The equilibrium concentration of A is obtained from the equation

$$\begin{aligned} a^8 \;+\; \frac{\kappa_D\kappa_T}{2}\,a^6 \;+\; \frac{\kappa_D^2\kappa_T}{4}\,a^5 \;+\; \frac{4\kappa_D^2\kappa_T\kappa_B + \kappa_D^2\kappa_T(4g_0 - a_0)}{4}\,a^4 \;+ \\ +\; \frac{\kappa_D^3\kappa_T^2\kappa_b}{2}\,a^2 \;+\; \frac{\kappa_D^4\kappa_T^2\kappa_B}{4}\,a \;+\; \frac{\kappa_D^4\kappa_T^2\kappa_B a_0}{4} \;\;=\;\; 0\;. \end{aligned} \tag{108}$$

The other equilibrium concentrations follow from a in the following equations:

$$g \;=\; g_0\,\frac{\kappa_D^2\kappa_T\kappa_B}{\kappa_D^2\kappa_T\kappa_B \;+\; a^4}\;, \tag{109}$$

$$d \;=\; \frac{a^2}{\kappa_D}\,, \quad t \;=\; \frac{a^4}{\kappa_D^2\kappa_T}\,, \quad \text{and} \tag{110}$$

$$m \;=\; g_0\,\frac{a^4}{\kappa_D^2\kappa_T\kappa_B \;+\; a^4}\;. \tag{111}$$

Since we have only one complex containing the gene the binding function is simply $Y(a) = m/g_0$.

In complete analogy to section 5.1.3 we derive the same hierarchy of approximations. In the zeroth order approximation, $\mu_0(a_0)$, we replace a by a_0, in first order, $\mu_1(a_0)$, the equilibrium constants in the denominator are chosen such that the approximation and $m(a_0)$ coincide at the point $m(\tilde{a}_0) = g_0/2$, and finally in $\mu_2(a_0)$ the tangent at this point is adjusted to

coincide with the tangent of $m(a_0)$:

$$\mu_0(\kappa_D, \kappa_T, \kappa_B, a_0) \;=\; g_0 \, \frac{a_0^4}{\kappa_D^2 \kappa_T \kappa_B \;+\; a_0^4} \;, \tag{112}$$

$$\mu_1(\kappa_D, \kappa_T, \kappa_B, a_0) \;=\; g_0 \, \frac{a_0^4}{\tilde{a}_0^4 + a_0^4} \quad \text{with} \tag{113}$$

$$\tilde{a}_0 \;=\; 4\kappa_B + 2\sqrt{\kappa_T \kappa_B} + \sqrt{\kappa_D \sqrt{\kappa_T \kappa_B}} + 2g_0 \;,$$

$$\mu_2(\kappa_D, \kappa_T, \kappa_B, \tilde{n}, a_0) \;=\; g_0 \, \frac{a_0^{\tilde{n}}}{\tilde{a}_0^{\tilde{n}} + a_0^{\tilde{n}}} \quad \text{with} \tag{114}$$

$$\tilde{n} \;=\; \frac{4\,\tilde{a}_0}{g_0} \, \frac{\partial m}{\partial a_0} \bigg|_{a_0 = \tilde{a}_0} \;.$$

The quality of the improvements is shown in figure 32. For equal equilibrium constants, $\kappa_D = \kappa_T = \kappa_B = 1$, the optimally adjusted value is $\tilde{n} = 13/9$ and the approximated curve $\mu_2(a_0)$ is almost indistinguishable from the exact curve $m(a_0)$, the middle plot enlarges the behavior of the curves at small a_0-values where deviations are recognizable. The approximation is somewhat less good for different values of the equilibrium constants (bottom plot in figure 32).

6 Concluding remarks

A few results of this analysis of biologically relevant binding equilibria seem to go beyond the conventional approach. First, present day computational facilities allow for a straight forward approach combining mathematical analysis with numerical calculations. Neither higher order equations nor quantities that are obtainable only in implicit form are an obstacle for rigorous investigations of the rather complicated expressions.

Second, it was possible to derive justifiable approximations that are suitable for the application in modelling genetic regulatory netwoks and metabolic networks. In addition, the approximations can be analyzed in detail and improvements of the conventionally used approaches are possible without a large increase in the numerical efforts.

Third, we presented a lengthy analysis of the relation between sigmoid binding functions and cooperative binding, and we were able to show that these two notions are not equivalent. There exist a domain of cooperative binding next to the non-cooperative case where the binding functions are not sigmoid.

References

[1] M. Ptashne and A. Gann. *Genes & Signals*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2002.

[2] L. Michalies and M. L. Menten. Kinetik der Invertinwirkung. *Biochemische Zeitschrift*, 49:333–369, 1913.

[3] A. Cornish-Bowden. *Fundamentals of Enzyme Kinetics*. Portland Press, London, third edition, 2004.

[4] A. R. Tzafriri and E. R. Edelman. On the validity of the quasi-steady state approximation of bimolecular reactions in solution. *J. Theor. Biol.*, 233:343–350, 2005.

[5] C. R. Cantor and P. R. Schimmel. *Biophysical Chemistry, Vol.I – III.* W. H. Freeman and Co., San Francisco, CA, 1980.

Contents