

## **Integração entre Biologia e Física: construindo redes de regulação gênica usando equações diferenciais**

Ricardo Z. N. Vêncio<sup>1\*</sup> e Gustavo B. de Oliveira<sup>2</sup>

<sup>1</sup>BIOINFO-USP - Núcleo de Pesquisa em Bioinformática da Universidade de São Paulo, SP, Brazil

<sup>2</sup>Instituto de Física, Universidade de São Paulo, SP, Brazil

\*rvencio@vision.ime.usp.br

### **Resumo:**

Este artigo apresenta dois elementos básicos no modelamento de redes de regulação gênica: um sistema de switch ON-OFF (chave ligado-desligado) e um sistema de oscilações periódicas. Apesar de aparentemente simples, essas duas formas básicas de controle e resposta das células a sinais externos ("circuito genético") devem ser entendidas em termos de equações químicas de equilíbrio, que por sua vez podem ser descritas em termos de Equações Diferenciais. Os dois modelos apresentados são vistos, num contexto multidisciplinar, como motivação para introdução à Biofísica de sistemas não-lineares para alunos das Ciências ditas Exatas ou motivação para o aprendizado de matemática avançada por alunos vindos das Ciências Biológicas.

**Palavras-chave:** sistemas não-lineares, equações diferenciais, biologia molecular, regulação genética

### **Abstract:**

This paper presents two basic elements used to model genetic regulatory networks: an ON-OFF switch system and a periodic oscillating system. Although apparently simple, these two basic forms of control and response to cell external signals ("genetic circuits") should be viewed as chemical equilibrium equations in the Differential Equations framework. These two models are shown in a multidisciplinary context as motivation for the introduction to non-linear Biophysics for the Physical Sciences' students, or as motivation for the studying of advanced Mathematics by the Biological Sciences' students.

**Keywords:** non-linear systems, differential equations, molecular biology, genetic regulation

## Introdução

Um dos maiores desafios enfrentados atualmente pela Biologia é entender, em nível molecular, como proteínas e genes trabalham juntos para controlar os seres vivos. A Física, e porque não dizer, as pessoas com conhecimento em Física, tem muito a contribuir para esse desafio contemporâneo da Biologia. Existe, por parte das Ciências Biológicas, uma necessidade muito grande de pessoas que consigam importar para dentro dela os métodos e as abordagens das ciências ditas Exatas. A formação em Física e Matemática é especialmente bem vinda na Biologia onde a formulação de modelos matemáticos tratáveis contribui para o avanço da área.

O objetivo deste artigo é ilustrar como alguns conceitos de Sistemas Não-Lineares podem ser apresentados, em um contexto multidisciplinar, a estudantes de graduação em Física ou também em Ciências Biológicas que possuam algum interesse ou habilidade em matemática. Para isso, vamos apresentar um modelo simples não-linear que busca descrever o comportamento de uma rede de regulação gênica.

Em Biologia, é comum nos depararmos com problemas que exigem a análise de alguma equação não-linear. Entretanto, esse tipo de análise está, em geral, fora do conhecimento de um biólogo comum. O modelo que vamos apresentar neste artigo se insere nessa classe de problemas. Problemas desse tipo despertam o interesse dos físicos por pesquisas multidisciplinares e, ao mesmo tempo, podem ser interessantes do ponto de vista do Ensino de Física e Matemática para estudantes de Biologia e áreas correlatas.

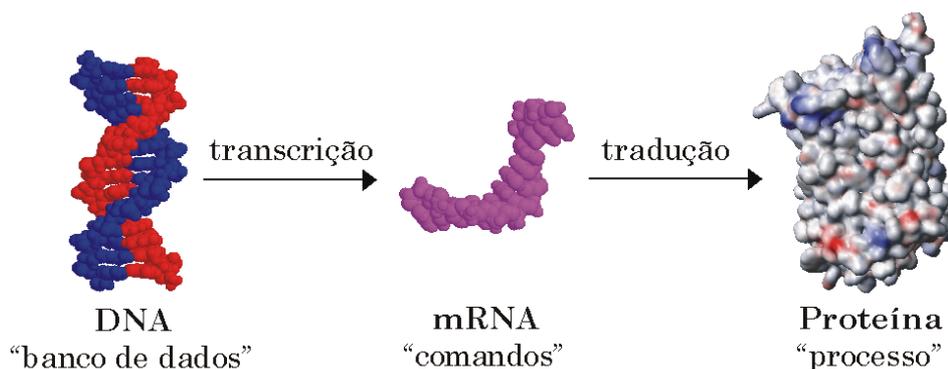
Todas as simulações deste trabalho foram feitas utilizando programas escritos na linguagem MapleV (Waterloo Maple Inc., <http://www.maplesoft.com>) e estão disponíveis no *web-site* suplementar: <http://www.vision.ime.usp.br/~rvencio/Rede>. MapleV é uma linguagem de alto nível, muito adequada para manipulação simbólica de objetos matemáticos, resolução de equações diferenciais não-lineares, e visualização de gráficos e soluções numéricas. É também uma ferramenta didática muito poderosa.

Além das simulações, estão também disponíveis, neste *web-site* suplementar, algumas figuras ilustrativas adicionais, identificadas no texto como "web".

## Redes de Regulação Gênica

O DNA é a molécula fundamental da maioria dos seres vivos. É ela que contém todas as informações genéticas de cada indivíduo e que tem a capacidade de transmiti-las aos seus descendentes. O biólogo James Watson e o físico Francis Crick publicaram uma descrição da molécula e foram os primeiros a sugerir seu mecanismo de replicação, marcando profundamente a história da Biologia [1].

Mostrou-se que o DNA tem duas funções básicas: replicação e transcrição de genes. A primeira função é responsável pela cópia do próprio DNA e é necessária para garantir a hereditariedade. A segunda função é responsável pela propagação de informação através de mensagens. A molécula de mRNA é o resultado da transcrição e é a mediadora da propagação das mensagens. Essas, por sua vez, são traduzidas em ação pelas proteínas associadas a cada uma das mensagens. Esse modelo de fluxo unidirecional de informação, ilustrado na Fig.1, foi primeiramente proposto pelo físico F. Crick e é conhecido como Dogma Central da Biologia [2].



**Figura 1:** Ilustração do Dogma Central da Biologia. Uma boa analogia para cada etapa é o funcionamento de um computador (figura fora de escala).

As redes de regulação gênica são compostas pelos grafos de interação entre as diversas espécies de mRNA e pelas regras dinâmicas associadas aos mesmos. Essas regras controlam o tráfego de mensagens no nível intermediário do Dogma Central (Fig.1). A Fig.web1 mostra um exemplo de uma rede de regulação.

A modelagem de redes de regulação gênica é uma questão de extrema relevância para aplicações em Biologia Molecular. A importância do estudo das interações entre as entidades moleculares fundamentais da Biologia tem um apelo de Ciência Básica evidente. Por outro lado, também tem um apelo aplicado, pois espera-se que o entendimento desses problemas permita prever, por exemplo, o que acontece quando se altera a disponibilidade de metabólitos no interior de uma célula ou interrompe-se o funcionamento de certos genes [3]. Com o entendimento de tais questões espera-se uma revolução tecnológica análoga àquelas ocorridas na Engenharia após os grandes avanços conceituais da Física (por exemplo, a eletrônica advinda da revolução da Mecânica Quântica). A principal aplicação está na busca de novos medicamentos, pois será muito bem-vinda a capacidade de simular o efeito de uma intervenção sobre uma via metabólica antes de se executarem caros e demorados experimentos *in vivo*.

Após o estabelecimento dos princípios básicos de regulação gênica, tornou-se evidente que para um entendimento completo dessas redes deveriam ser consideradas interações não-lineares, realimentação positiva e negativa, oligomerização de proteínas e interação entre diferentes vias metabólicas [4].

Existem três tipos de abordagem para o modelamento de redes de regulação: (i) discreta, usando estados e regras Booleanas; (ii) contínua, usando sistemas de equações diferenciais obtidas a partir da cinética bioquímica; (iii) híbridas, considerando estados discretos apenas nas interações onde a aproximação é razoável [5].

As redes de regulação foram primeiro modeladas simplesmente como redes de elementos Booleanos. Por essa abordagem assume-se que o estado de cada gene pode ser caracterizado por 1 (ligado) ou 0 (desligado), que o controle regulatório pode ser aproximado por regras de lógica Booleana, e que todos os elementos da rede têm seus estados atualizados simultaneamente [6].

Os modelos contínuos utilizam equações diferenciais para acompanhar a evolução temporal das concentrações das moléculas de interesse. Eles representam mais fielmente o fenômeno físico-químico, mas são muito mais exigentes em termos de modelagem e computação pois é preciso resolver numericamente equações que em geral não tem uma solução conhecida [3]. Ainda, nesta abordagem é preciso propor o esquema de cinética química das moléculas, o que torna necessária grande integração entre o conhecimento bioquímico e físico.

Sabe-se que os estados estacionários obtidos com modelamento discreto são qualitativamente reproduzidos com a abordagem do modelamento contínuo por equações diferenciais. Todavia, a recíproca não é necessariamente verdadeira. Além disso, soluções

periódicas obtidas com a abordagem discreta podem não ser periódicas em um modelo contínuo fisicamente mais verossímil [7]. Os modelos discretos são muito mais fáceis de serem simulados pois exigem menos tempo computacional que os modelos contínuos.

Neste trabalho revisamos dois modelos de regulação gênica construídos a partir de conhecimentos de Biologia Molecular e análise de equações diferenciais não-lineares. Algumas seções para revisar conceitos utilizados são oportunas antes do desenvolvimento dos modelos biológicos. Serão abordados dois modelos importantes de regulação gênica: sistema *switch* ON-OFF e sistema de oscilações periódicas. O modelo *switch* ON-OFF estudado mostra que é possível reproduzir estados Booleanos utilizando um esquema cinético adequado, de acordo com os modelos biológicos estabelecidos. O modelo de oscilações periódicas introduz a realimentação negativa no modelo ON-OFF, e consegue reproduzir oscilações na concentração do mRNA transcrito de forma autônoma respondendo a estímulos externos. Esses dois modelos são elementos básicos para a construção de "circuitos" genéticos que simulem redes de interação gênica.

## Equações Diferenciais

Um sistema dinâmico é composto por dois ingredientes básicos: um estado inicial, de onde o sistema parte, e uma regra de evolução. As equações diferenciais são as regras de evolução mais comuns pois uma série de leis físicas são formuladas em termos dessas. Por exemplo, a Segunda Lei de Newton em uma dimensão:

$$m\dot{x} = F(x, \dot{x}, t) \quad (1)$$

Essa é uma equação diferencial ordinária (EDO) de segunda ordem, onde a massa  $m$  é constante e  $F$  é a força que atua sobre a partícula em questão. Uma EDO de ordem  $n$  sempre pode ser escrita como um sistema de  $n$  equações de primeira ordem. No caso do exemplo anterior:

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= \frac{1}{m} F(x, v, t) \end{aligned} \quad (2)$$

Em geral:

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}) \quad \mathbf{x} \in \mathfrak{R}^n \quad (3)$$

A um sistema desse tipo associa-se um *espaço de fase* de dimensão  $n$ . A Fig.web2 mostra o espaço de fase associado às equações de um oscilador harmônico amortecido. Um caso particular de EDO ocorre quando todas as funções  $F_i$  são lineares em todas as variáveis, ou seja, quando o sistema é da forma:

$$\dot{\mathbf{x}} = \mathbf{A} \cdot \mathbf{x} \quad (4)$$

onde  $\mathbf{A}$  é uma matriz  $n \times n$  constante. O conceito de estabilidade é muito importante no estudo de equações diferenciais. Existem diferentes conceitos de estabilidade, entre eles: estabilidade do *ponto fixo* e estabilidade *estrutural*.

Os *pontos fixos* de uma equação diferencial são as soluções da equação:  $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}) = 0$ . Tais soluções definem pontos  $\bar{\mathbf{x}} = (\bar{x}_1, \dots, \bar{x}_n)$  no espaço de fase. Para fazer a análise de estabilidade, o primeiro passo é determinar os pontos fixos para em seguida investigar o que acontece na vizinhança dos mesmos. Um ponto fixo é *estável* quando todos os pontos em uma vizinhança do mesmo no espaço de fase permanecem próximos para tempos arbitrariamente longos. Mais precisamente,  $\mathbf{x}$  é estável se para toda vizinhança  $U$  de  $\mathbf{x}$ , existe uma vizinhança  $U_1$  de  $\mathbf{x}$  contida  $U$ , tal que, toda solução  $\mathbf{x}(t)$  com  $\mathbf{x}(0)$  em  $U_1$  permanece em  $U$  para todo  $t > 0$ . Um ponto fixo que não é estável é dito ser *instável*. Nesse caso, as soluções com condição inicial próxima ao ponto fixo se afastam cada vez mais do mesmo no espaço de fase [8].

No caso da estabilidade estrutural, o interesse é estudar a robustez do espaço de fase quando o campo vetorial  $\mathbf{F}(\mathbf{x})$  é submetido a perturbações. Em outras palavras, queremos

determinar sob quais condições o espaço de fase perturbado é qualitativamente equivalente ao original. A perda de estabilidade estrutural, denominada *bifurcação*, tem um papel muito importante pois está associada a uma mudança qualitativa das propriedades do sistema [8].

Nos osciladores harmônico amortecido e não-amortecido,

$$\ddot{x} + \gamma \dot{x} + \omega_0^2 x = 0 \quad (5)$$

$$\ddot{x} + \omega_0^2 x = 0 \quad (6)$$

as diferenças qualitativas entre os espaços de fase são evidentes (vide Fig.web3). Podemos pensar que o oscilador não-amortecido (6) é um caso particular do oscilador amortecido (5) quando  $\gamma = 0$ .

Para estudar a estabilidade de sistemas  $n$ -dimensionais é interessante analisar duas-a-duas as  $n$  dimensões do espaço de fase. Essas projeções bidimensionais permitem visualizar algumas características importantes.

No caso de sistemas lineares bidimensionais,

$$\dot{\mathbf{x}} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{A} \cdot \mathbf{x} \quad (7)$$

onde  $a$ ,  $b$ ,  $c$  e  $d$  são constantes, a estabilidade estrutural e a estabilidade dos pontos fixos pode ser investigada analisando os autovalores da matriz  $\mathbf{A}$  associada à equação. Seja  $\mathbf{P}$  a matriz que diagonaliza  $\mathbf{A}$ . Fazendo a mudança de variáveis  $\mathbf{x} = \mathbf{P} \cdot \mathbf{y}$  obtemos:

$$\dot{\mathbf{y}} = \mathbf{B} \cdot \mathbf{y}$$

$$\mathbf{B} = \mathbf{P} \cdot \mathbf{A} \cdot \mathbf{P}^{-1} = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$$

onde  $\lambda$  e  $\mu$  são os autovalores de  $\mathbf{A}$ .

Se os autovalores são reais e de sinais opostos,  $\lambda\mu < 0$  e  $\text{Im}(\lambda) = \text{Im}(\mu) = 0$ , o ponto fixo é um *ponto de sela* estável em uma direção e instável em outra (as direções estão associadas aos autovetores de  $\mathbf{A}$ ). Se os autovalores são reais e de mesmo sinal,  $\lambda\mu > 0$  e  $\text{Im}(\lambda) = \text{Im}(\mu) = 0$ , o ponto fixo é um *nó* estável ou instável se os autovalores são negativos ou positivos, respectivamente. Se os autovalores são complexos com parte real negativa,  $\text{Re}(\lambda) < 0$  e  $\text{Re}(\mu) < 0$ , o ponto fixo é um *foco* estável. Se  $\text{Re}(\lambda) > 0$  ou  $\text{Re}(\mu) > 0$ , o foco é instável. Se os autovalores tem parte real zero,  $\text{Re}(\lambda) = \text{Re}(\mu) = 0$ , o ponto fixo é um *centro* estável. Finalmente, se os autovalores são iguais,  $\lambda = \mu$ , o ponto fixo é um *nó impróprio* estável se os autovalores forem negativos e instável se forem positivos [8]. A Fig.web4 ilustra os tipos de pontos fixos discutidos acima.

## Estabilidade em Sistemas Não-Lineares

Os resultados obtidos para sistemas lineares são fundamentais para o estudo de sistemas não-lineares. Podemos estudar o comportamento de um sistema não-linear através de sua *linearização* em torno dos pontos fixos. Essa é uma técnica geral e importante para o estudo de tais sistemas [8]. Como fizemos para o caso linear, vamos nos restringir a projeções bidimensionais do espaços de fase.

Considere o sistema de equações não-lineares:

$$\dot{x} = f(x, y)$$

$$\dot{y} = g(x, y)$$

para o qual existe um ponto fixo  $(\bar{x}, \bar{y})$ . Fazendo a expansão em série de Taylor em torno desse ponto, definindo novas variáveis e tomando somente os termos de 1ª ordem:

$$\xi = x - \bar{x}, \quad \eta = y - \bar{y}, \quad \dot{\xi} = \dot{x}, \quad \dot{\eta} = \dot{y}$$

obtemos o sistema linear:

$$\begin{aligned}\dot{\xi} &= a\xi + b\eta \\ \dot{\eta} &= c\xi + d\eta\end{aligned}\quad (8)$$

Com essa linearização,  $\mathbf{J}$  é a matriz jacobiana calculada no ponto fixo:

$$\mathbf{J} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

O sistema (8) é uma aproximação válida somente para pontos próximos do ponto fixo  $(\bar{x}, \bar{y})$ , ou seja, descreve o comportamento local das soluções. Assim, a análise de estabilidade dos pontos fixos reduz-se ao estudo das soluções correspondentes do sistema linearizado, onde podemos aplicar a teoria de estabilidade linear descrita na seção anterior. É possível demonstrar que a estabilidade de um ponto fixo hiperbólico (cujos autovalores associados têm parte real não nula) não é afetada pela linearização [9]. Nos casos em que os autovalores da matriz jacobiana têm parte real nula, a análise linear não permite concluir nada a respeito da estabilidade. Nesse caso, devem ser estudados termos de ordem superior [9].

A equação do pêndulo simples

$$ml^2\ddot{\theta} + mgl \sin(\theta) = 0 \quad (9)$$

serve como exemplo para a análise de estabilidade de um sistema não-linear. As curvas fechadas no espaço de fase (vide Fig.web5) representam soluções periódicas. Fontes e sorvedouros podem ser vistos como atratores ou repulsores pontuais, como os pontos fixos na Fig.2. Curvas fechadas também podem atrair ou repelir soluções próximas. Tal tipo de solução é chamada de *ciclo limite*. Esse tipo de atrator é importante para o modelamento de sistemas biológicos cujas soluções periódicas devem ser mantidas, independentemente de perturbações externas nas condições iniciais. A equação de Van der Pol

$$\begin{aligned}\dot{x} &= y \\ \dot{y} &= -\lambda(x^2 - 1)y - x\end{aligned}\quad (10)$$

é uma equação diferencial não-linear de segunda ordem que ilustra bem o conceito de ciclo limite.

A equação de Van der Pol pode ser aplicada para o modelamento de oscilações elétricas e biológicas. Ela descreve um oscilador não-linear que possui um mecanismo de controle, onde o ciclo limite representa o regime normal de trabalho. Se a amplitude é maior do que o normal, entra em ação um mecanismo dissipativo, de modo que a amplitude diminua para o valor normal. O mesmo ocorre para grandes amplitudes, onde, dessa vez, o que entra em ação é um mecanismo gerador de energia, fazendo com que a amplitude aumente. Esse tipo de mecanismo regulador é fundamental em processos biológicos, onde existe um regime normal de funcionamento (*ciclo limite*) muito bem determinado. A Fig.web6 ilustra exemplos desse processo e também o ciclo limite.

## Estabilidade Estrutural em Equações Não-Lineares

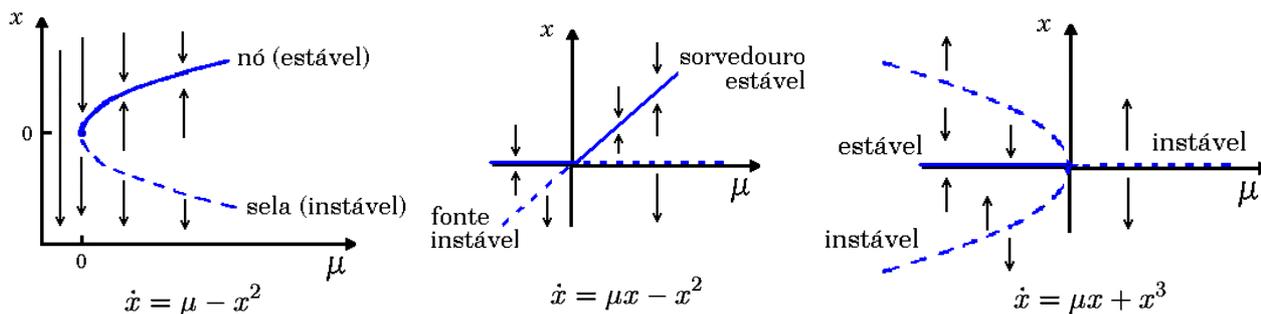
Em geral, um sistema físico depende de um ou mais parâmetros de controle. Por exemplo, no caso do oscilador harmônico simples, o parâmetro é a frequência natural  $\omega_0$ . No caso de equilíbrio químico, são as constantes de formação ou degradação das substâncias, no caso de modelos em ecologia, são as taxas de mortalidade ou natalidade e assim por diante.

Devido a essa dependência, devemos considerar tais sistemas como função de um parâmetro  $\mu$  :

$$\begin{aligned}\dot{x} &= f(x, y, \mu) = f_\mu(x, y) \\ \dot{y} &= g(x, y, \mu) = g_\mu(x, y)\end{aligned}$$

O comportamento dinâmico desse sistema pode ser bem diferente dependendo do valor do parâmetro  $\mu$ . Um sistema dinâmico é estruturalmente estável se para qualquer perturbação suficientemente pequena das equações que o descrevem, a solução resultante é

qualitativamente equivalente à solução associada ao sistema sem a perturbação.



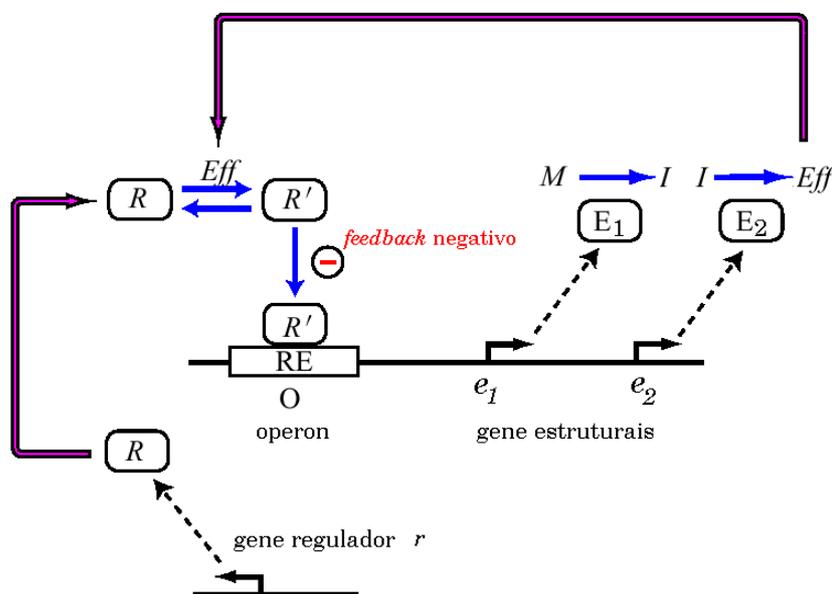
**Figura 2:** Diagramas de bifurcação para algumas equações simples. É convenção denotar pontos fixos instáveis por linhas pontilhadas. Adaptado de [8].

Nesse ponto de vista, as soluções dependem tanto do tempo quanto do parâmetro  $\mu$ , e, portanto, os autovalores e autovetores da matriz jacobiana também têm esta dependência. Assim, variando o parâmetro de controle podemos passar de um ponto fixo instável para um ponto fixo estável, por exemplo. O valor crítico  $\mu_c$  para qual essa mudança qualitativa ocorre é chamado de *ponto de bifurcação*. Nessa transição o sistema dinâmico perde a estabilidade estrutural e sofre uma *bifurcação*.

A análise de estabilidade e bifurcação é muito importante na construção de modelos para redes de regulação gênica pois os modelos devem ter certas características estruturais como robustez ou "memória", como mostraremos posteriormente.

### Biologia e Bioquímica do Modelo

Os dois modelos físicos considerados neste trabalho são baseados em dois modelos biológicos bem estabelecidos: o modelo do operon de Jacob-Monod e o modelo de cinética enzimática de Michaelis-Menten.



**Figura 3:** Modelo do operon de Jacob-Monod. Os elementos são: *r*, gene da proteína regulatória; *R*, proteínas regulatória; *R'*, repressor; *RE*, operon; *e*, genes estruturais das enzimas 1 e 2; *E*, enzimas; *M*, metabólito; *I*, intermediário; *Eff*, efetador. As setas cheias são reações bioquímicas e as pontilhadas são transcrição. Adaptado de [7].

Em 1961, F. Jacob e J. Monod publicaram um trabalho propondo um mecanismo para explicar a indução ou repressão de enzimas [10]. Esse modelo foi bem consolidado experimentalmente e é conhecido como o modelo do operon de Jacob-Monod. Nesse modelo,

cada gene estrutural que codifica uma enzima ou proteína está ligado a um gene "operador" que regula sua transcrição, chamado *operon*. Cada operon está sujeito a repressão por uma molécula repressora. Cada repressor, por sua vez, tem sua afinidade de ligação com o gene operon modulada por outra molécula chamada efetador. Na Fig.3 está ilustrado um exemplo de uma regulação por operon. Nesse caso existe um *feedback* negativo envolvendo duas enzimas  $E_1$  e  $E_2$ , que reprimem sua própria transcrição na presença de um metabólito  $M$ . O metabólito é convertido pelas enzimas no efetador  $Eff$ , que interage com uma proteína regulatória  $R$ . A interação converte  $R$  em  $R'$ , que reprime a transcrição de  $E_1$  e  $E_2$ , fechando o ciclo de *feedback* negativo.

As reações bioquímicas são governadas pela Lei cinética da ação das massas:



onde os termos entre colchetes representam concentrações,  $k$  é uma constante cinética e,  $\alpha$  e  $\beta$  são coeficientes estequiométricos da reação.

Mas a dinâmica de reações bioquímicas enzimáticas não é simples como as reações químicas. Em 1913, L. Michaelis e M. Menten propuseram uma série de mecanismos e hipóteses para reações enzimáticas que levaram a um modelo de cinética bem estabelecido experimentalmente [11]. O modelo sugere que a reação enzimática é composta por duas reações elementares intermediárias, onde o substrato forma um complexo com a enzima e depois se decompõem em produto e enzima:



Para este mecanismo cinético:

$$[\dot{P}] = k_2[ES]$$

$$[\dot{ES}] = k_1[E][S] - k_{-1}[ES] - k_2[ES] \quad (12)$$

onde a constante com sub-índice negativo denota a reação no sentido oposto.

Esse sistema não pode ser integrado explicitamente. Algumas aproximações biologicamente razoáveis podem ser feitas para resolver (12): assumir o equilíbrio da primeira etapa da reação, e assumir atingido o estado estacionário na formação do complexo intermediário.

A hipótese de equilíbrio não é sempre verdadeira mas é uma boa aproximação. Biologicamente significa que a transformação do complexo em produto e recuperação da enzima é muito mais lenta que a primeira etapa de encontro entre substrato e enzima. Matematicamente  $k_{-1} \gg k_2$ , levando à constante de dissociação da enzima, que pode ser medida experimentalmente.

A hipótese de estado estacionário é uma boa aproximação pois o transiente, quando enzima e substrato são colocados em contato, dura relativamente pouco tempo. Biologicamente, isso significa que praticamente toda molécula de enzima está "ocupada" em reação, isto é, a taxa de síntese de  $ES$  é igual ao consumo de  $ES$ . Matematicamente:

$$[\dot{ES}] = 0$$

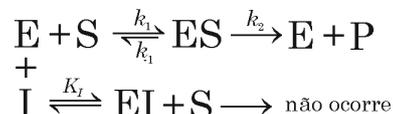
Assim pode-se manipular (12) adequadamente com auxílio da equação de conservação  $[E_{tot}] = [E] + [ES]$  e obter finalmente:

$$[\dot{P}] = k \frac{[S]}{K_M + [S]} \quad (13)$$

onde  $k = k_2[E_{tot}]$  e  $K_M = (k_{-1} + k_2)/k_1$  podem ser obtidos experimentalmente e  $[E_{tot}]$  é suposto ser constante (todas as passagens dessa dedução podem ser encontradas no *web-site* suplementar).

O modelo pode ser adaptado para lidar com um inibidor da enzima [11]. Um inibidor é uma molécula, sintética ou natural, que também liga-se à enzima como o substrato. Esse inibidor pode ser competitivo, isto é, ligando-se à enzima da mesma forma que o substrato, competindo pelo acesso ao sítio ativo da enzima; ou pode ser um inibidor não competitivo, que tem outros mecanismos para interferir na ligação do substrato com a enzima.

Para o inibidor competitivo,

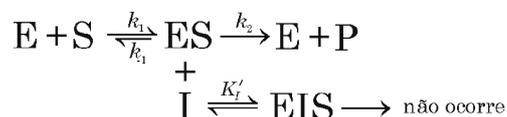


a equação de conservação muda para  $[E_{tot}] = [E] + [ES] + [EI]$ . Resolvendo (12) e tendo  $[E]$  no estado estacionário (derivação no *web-site* suplementar):

$$[\dot{P}] = k \frac{[S]}{K_M (1 + [I]/K_I) + [S]} \quad (14)$$

onde  $K_I = [E][I]/[EI]$  é observável.

Para o inibidor não competitivo,

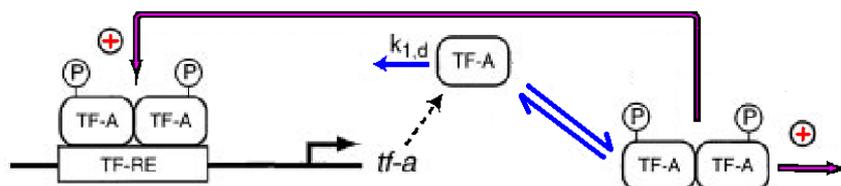


a equação de conservação é:  $[E_{tot}] = [E] + [ES] + [ESI]$ . Analogamente ao caso competitivo, a taxa de produção é dada por (14) mas agora  $K_I = [ES][I]/[ESI]$ , também observável.

Uma molécula repressora, no modelo do operon, pode ser entendida como um competidor da proteína regulatória pelo sítio de ligação do operon, e por isso a dinâmica de Michaelis-Menten é relevante no modelamento da regulação via operon.

## Modelo Switch ON-OFF

O modelo que será considerado aqui está proposto e resolvido em [7]. Trata-se de um exemplo de autoregulação com fatores de transcrição. O modelo consegue reproduzir o comportamento de um sistema de dois níveis para a expressão de um certo gene, variando do estado ligado ON para o estado desligado OFF, de acordo com estímulos externos.



**Figura 4:** Modelo *switch* ON-OFF. TF-A, proteína fator de transcrição; TF-RE, promotor; TF-A/TF-A, dímero do fator de transcrição; *tf-a*, gene para TF-A;  $k_{1,d}$ , constante de degradação; P, fosforilação. As setas cheias são reações bioquímicas e a pontilhada é a transcrição. Os dímeros TF-A também promovem transcrição de outros genes não mostrados. Adaptado de [7].

Vamos considerar uma via de transmissão de sinal na qual estímulos levam a produção de um segundo mensageiro e a fosforilação (aquisição de uma molécula de fósforo) de fatores de transcrição (TF), que por sua vez ligam-se ao DNA em seqüências conhecidas como promotores (RE) e controlam a transcrição de genes específicos. A atividade regulatória dos TF sempre é modulada pela fosforilação ou interações intermoleculares. Os TF sempre ligam-se aos promotores na forma de homodímeros ou heterodímeros. Este modelo trata de TF que tem dependência de dimerização e fosforilação para regulação da transcrição. Um único fator de transcrição, TF-A, é considerado como parte de uma via mediando uma resposta celular a um

estímulo. O TF-A forma um dímero que se liga no promotor TF-RE. No operon de TF-RE está o gene *tf-a* que auto regula a transcrição de TF-A, fechando o *feedback* positivo. Somente dímeros fosforilados podem atuar sobre TF-RE, o que torna a regulação dependente de proteínas quinases e fosfatases. A atividade destas enzimas é regulada por sinais externos, o que fornece o canal para interação do ambiente com o modelo pela mudança da taxa de produção de TF-A em função da atividade daquelas. A Fig. 4 ilustra o modelo.

A expressão para a dinâmica de TF-A é :

$$\dot{A} = k_{1,f} \frac{A^2}{K_{1,d} + A^2} - k_{1,d} A + r_{1,bas} \quad (15)$$

onde:  $A$  é a concentração de TF-A,  $k_{1,f}$  é a constante da reação de formação de TF-A a partir do dímero fosforilado,  $K_{1,d}$  é constante de Michaelis da reação,  $k_{1,d}$  é a constante de degradação de TF-A, e  $r_{1,bas}$  é a taxa basal de transcrição de TF-A a partir de *tf-a*.

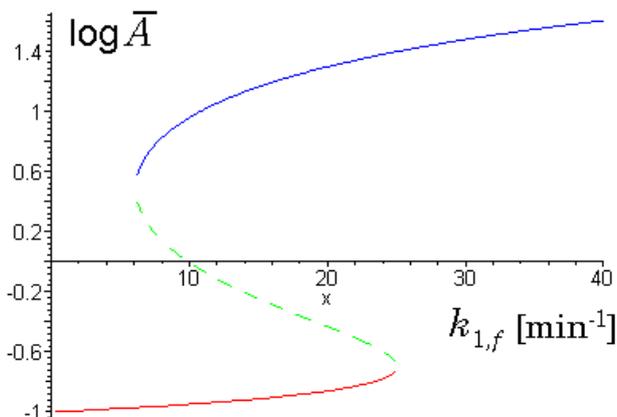
O primeiro termo da equação diferencial nada mais é que uma adaptação de (13) as condições do modelo. Como o substrato neste caso são duas moléculas idênticas (reação dímero  $\rightarrow$  monômero), a concentração é quadrática, conforme (11). Esse expoente é importante para as propriedades não lineares interessantes do modelo. Segundo [7], sem a dimerização o sistema perde a capacidade de sustentar 2 estados estacionários, ON e OFF, tornando-se inútil. O segundo termo dá conta do decaimento exponencial na quantidade de TF-A devido a degradação natural. O terceiro termo modela a taxa basal, independente de estímulos, na qual a proteína TF-A, como a maioria das proteínas, é transcrita a partir do seu gene *tf-a*.

A interação com estímulos do ambiente se dá pela alteração da eficiência de fosforilação do dímero TF-A e portanto influi diretamente na constante  $k_{1,f}$ . Este é considerado o parâmetro de controle do sistema, conforme discutido na introdução à análise de bifurcação. Os outros parâmetros de (15) foram escolhidos conforme [7] e são todos valores biologicamente plausíveis no contexto do modelo:  $K_{1,d} = 10 \text{ min}^{-1}$ ,  $k_{1,d} = 1 \text{ min}^{-1}$  e  $r_{1,bas} = 0.1 \text{ min}^{-1}$ .

Utilizando a teoria de Sistemas Não-Lineares desenvolvida anteriormente, é possível identificar o intervalo de valores para os pontos fixos de (15) para os quais existem 2 soluções estacionárias estáveis. Para encontrar os pontos fixos  $\bar{A}$ , resolvemos (15) para  $A$ , para cada um dos valores de  $k_{1,f}$ :

$$k_{1,f} \frac{A^2}{K_{1,d} + A^2} - k_{1,d} A + r_{1,bas} = 0$$

Como esta é uma equação do 3º grau temos 3 soluções para cada escolha do parâmetro de controle, mas apenas os pontos fixos reais de fato existem. A Fig.5 mostra o diagrama de bifurcação deste sistema.



**Figura 5:** Diagrama de bifurcação para o modelo *switch* ON-OFF. No intervalo  $6.1 < k_{1,f} < 25.1$  existem 3 *pontos fixos*. As duas soluções reais são dois níveis estáveis possíveis.

Para descobrir a estabilidade dos pontos fixos, usamos a linearização, como em (10), e verificamos os autovalores da matriz Jacobiana, como desenvolvido anteriormente. Como neste caso não temos um sistema de equações, mas apenas uma, o cálculo torna-se mais simples:

$$\lambda = \frac{\partial}{\partial A} \left[ k_{1,f} \frac{A^2}{K_{1,d} + A^2} - k_{1,d}A + r_{1,bas} \right] (\bar{A})$$

O sinal do autovalor real para cada  $k_{1,f}$  determina se o equilíbrio é estável ou instável. Na 1ª região, com  $0 < k_{1,f} < 6.1$ , os autovalores são sempre negativos e portanto o único ponto fixo é estável. Na 2ª região, com  $6.1 < k_{1,f} < 25.1$ , os autovalores são negativos para 2 pontos fixos e positivos para o outro. Isso significa que existem dois estados estacionários coexistentes, característica mais importante do modelo. Na 3ª região novamente, com  $25.1 < k_{1,f}$ , os autovalores voltam a ser negativos, implicando que a única solução é estável.

Resolvendo a equação (15) com valores de  $k_{1,f}$  representantes de cada uma dessas regiões, podemos verificar explicitamente as soluções. Algumas figuras suplementares mostram bem as diferentes soluções estacionárias alcançadas partindo-se de diversas condições iniciais. Apesar de possuírem condições iniciais de magnitude muito diferente, todas as soluções convergem para o mesmo e único estado estacionário nas Fig.web7 e Fig.web8. Na Fig.web9 observamos que dependendo da condição inicial o sistema evolui para o estado estacionário mais baixo (OFF) ou mais alto (ON).

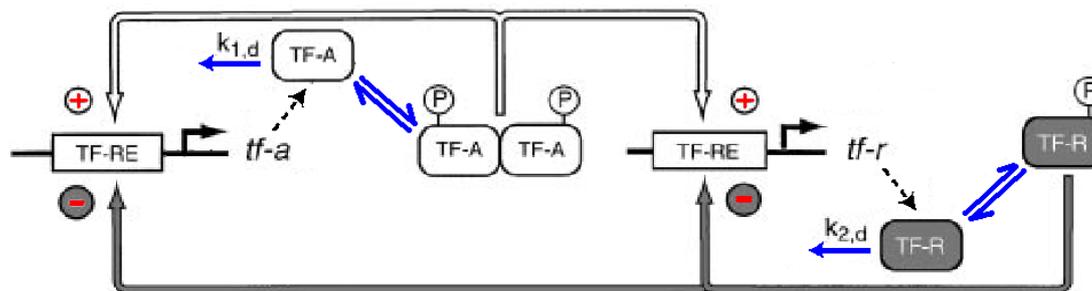
A característica mais importante do modelo é manter uma bi-estabilidade, pois verifica-se que com breves variações no parâmetro de controle  $k_{1,f}$  o sistema passa de um estado estacionário para outro. Biologicamente, estímulos externos, representados por breves mudanças em  $k_{1,f}$ , mudam o estado do sistema para dois patamares de concentração do fator TF-A, como um sistema de dois níveis ON-OFF. As figuras suplementares Fig.web10, Fig.web11 e Fig.web12 mostram a bi-estabilidade para algumas perturbações no parâmetro de controle  $k_{1,f}$ . É assim que se acredita que as células "percebem" seu ambiente e respondem às mudanças externas.

Dependendo das constantes do modelo os níveis reais de transcrição, isto é [TF-A], pode ser ajustados aos resultados experimentais reproduzindo o comportamento de sistemas reais. Exemplos de sistemas reais que podem ser descritos por estes modelos são os conhecidos fatores de transcrição  $\text{Ca}^{+2}/\text{cAMP}$  ou CREB/ATF, e os AP-1 das famílias de proteínas Fos e Jun [7].

## Modelo de Oscilações Periódicas

O modelo que será considerado aqui está proposto e resolvido em [7]. Trata-se de um exemplo de autoregulação com fatores de transcrição e repressores. O modelo consegue reproduzir oscilação autônoma de um fator de transcrição. É possível mostrar que modelos que incorporaram somente um tipo de *feedback*, positivo ou negativo, não podem reproduzir oscilações estáveis.

Uma adaptação do modelo anterior, incorporando uma molécula repressora (TF-R), adiciona o *feedback* negativo ao modelo anterior, permitindo a observação de oscilações. O repressor TF-R funciona como um inibidor competitivo do TF-A na ligação com os sítios TF-RE. A Fig.6 mostra um esquema do modelo.



**Figura 6:** Modelo de oscilações periódicas. TF-A, proteína fator de transcrição; TF-R, repressor; TF-RE, promotor; TF-A/TF-A, dímero do fator de transcrição; *tf-a*, gene para TF-A; ; *tf-r*, gene para TF-R;  $k_{1,d}$ , constante de degradação de TF-A;  $k_{2,d}$ , constante de degradação de TF-R; P, fosforilação. As setas cheias são reações bioquímicas e as pontilhadas são processos de transcrição. Adaptado de [7].

Esse modelo passa a exigir 2 equações diferenciais, uma para TF-A e outra para TF-R, e a cinética leva em conta a competição por TF-RE como deduzido em (14):

$$\begin{aligned} \dot{A} &= k_{1,f} \frac{A^2}{K_{1,d}(1 + R/K_{R,d}) + A^2} - k_{1,d}A + r_{1,bas} \\ \dot{R} &= k_{2,f} \frac{A^2}{K_{2,d}(1 + R/K_{R,d}) + A^2} - k_{2,d}R \end{aligned} \quad (16)$$

onde  $A$  é a concentração de TF-A,  $R$  é a concentração de TF-R,  $k_{1,f}$  é a constante da reação de formação de TF-A a partir do dímero fosforidado,  $K_{1,d}$  é constante de Michaelis desta reação,  $k_{1,d}$  é a constante de degradação de TF-A,  $r_{1,bas}$  é a taxa basal de transcrição de TF-A a partir de *tf-a*,  $k_{2,f}$  é a constante da reação de formação de TF-R,  $K_{2,d}$  é constante de Michaelis desta reação e  $k_{2,d}$  é a constante de degradação de TF-R.

Analogamente ao modelo anterior é possível analisar a estabilidade dos pontos fixos e observar o ciclo limite da oscilação. Novamente o parâmetro de controle é  $k_{1,f}$ . Os outros parâmetros foram escolhidos conforme [7]:  $K_{1,d} = 10.0 \text{ min}^{-1}$ ,  $K_{2,d} = 10.0 \text{ min}^{-1}$ ,  $K_{R,d} = 0.2$ ,  $k_{1,d} = 1.0 \text{ min}^{-1}$ ,  $k_{2,f} = 1.0 \text{ min}^{-1}$ ,  $k_{2,d} = 0.2 \text{ min}^{-1}$ ,  $r_{1,bas} = 0.4 \text{ min}^{-1}$ . Para encontrar os pontos fixos  $(\bar{A}, \bar{R})$ , resolvemos (16) para  $A$  e  $R$  em cada um dos valores de  $k_{1,f}$ :

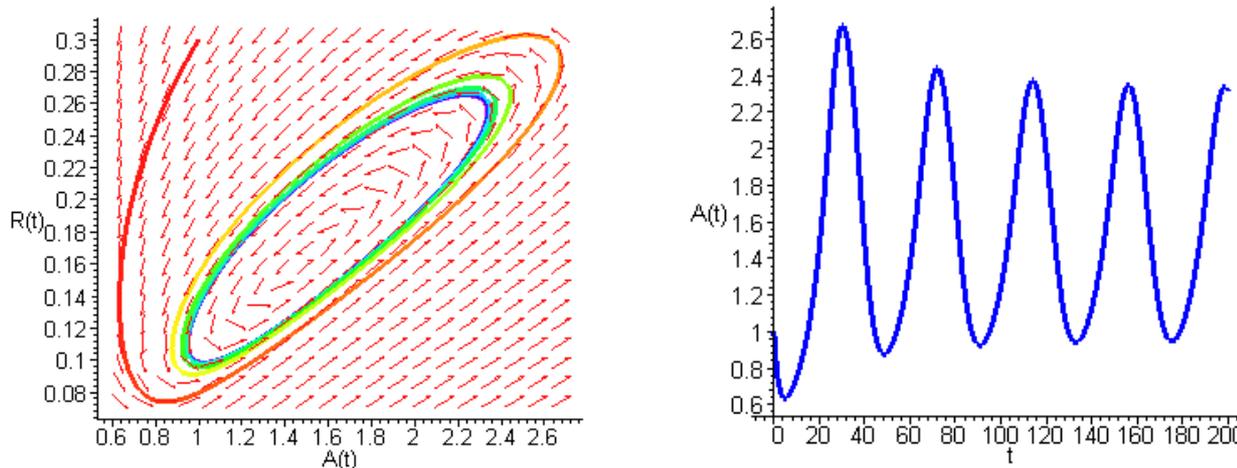
$$\begin{cases} \dot{A} = 0 \\ \dot{R} = 0 \end{cases}$$

Para analisar a estabilidade de cada um dos pontos fixos aplicamos (10) às equações anteriores e determinamos os autovalores de  $\mathbf{J}$ :

$$\mathbf{J} = \begin{bmatrix} \frac{\partial}{\partial A}[\dot{A}](\bar{A}, \bar{R}) & \frac{\partial}{\partial R}[\dot{A}](\bar{A}, \bar{R}) \\ \frac{\partial}{\partial A}[\dot{R}](\bar{A}, \bar{R}) & \frac{\partial}{\partial R}[\dot{R}](\bar{A}, \bar{R}) \end{bmatrix}$$

Analisando cada um dos autovalores, identificamos a região em que  $9.9 < k_{1,f} < 10.8$  onde existe solução periódica conforme as regras discutidas anteriormente. Nesta região temos duas bifurcações de Hopf, uma abrindo em 9.9 e outra colapsando em 10.8.

Modificando o parâmetro de controle, o que corresponde fisicamente ao recebimento de um sinal externo pelo sistema, é possível modular a oscilação na transcrição de TF-A. A Fig.7 mostra o ciclo limite obtido e as oscilações da solução, e a Fig.web13 mostra uma solução não-periódica que pode ocorrer dependendo do parâmetro de controle escolhido.



**Figura 7:** Soluções do modelo de oscilações para  $k_{1f} = 10.0 \text{ min}^{-1}$ . A esquerda é possível ver nitidamente o ciclo limite no espaço de fase das soluções para uma condição inicial arbitrária. A direita a integração do sistema mostra as oscilações.

A capacidade de modulação de um sinal periódico é uma tarefa não trivial quando se imagina fisicamente sistemas biológicos. Entretanto, este modelo relativamente simples contém todas as características necessárias para descrever sistemas reais.

Oscilações biológicas com períodos da ordem de horas podem ser explicadas por este modelo. Exemplos de aplicação são a secreção periódica de hormônios como o hormônio do crescimento ou o hormônio luteinizante LH, e transcrição de genes relacionados a ciclo circadiano [7].

## Conclusões

Neste artigo apresentamos uma ilustração de como um assunto relativamente avançado para estudantes de graduação pode ser introduzido em um contexto multidisciplinar. A aplicação em Biologia Molecular pode servir tanto para motivar o interesse dos físicos por essa área quanto para introduzir alguns conceitos de Sistemas Não-Lineares aos alunos da área de Ciências Biológicas com conhecimentos de cálculo.

A decisão do metabolismo de entrar em um estado ou outro em resposta à estímulos externos pode parecer um problema simples, mas devemos lembrar que todos esses processos são realizados em um contexto de equilíbrio químico das concentração de moléculas no interior da célula. Montar e modelar sistemas químicos que possam funcionar como verdadeiros circuitos genéticos de controle é uma atividade de pesquisa não-trivial.

Os dois principais tipos de controle metabólico, ligar-desligar e ciclar, foram revisados neste trabalho.

## Agradecimentos

Este trabalho foi realizado com o suporte da FAPESP. Agradecemos aos professores Eduardo J. Neves (IME-USP) e Nelson Fidler-Ferrara (IF-USP) pela orientação durante a elaboração deste trabalho.

## Referências Bibliográficas

- [1] T.H.G Oliveira, N.F. dos Santos, L. M. Beltramini (2004) O DNA: uma sinopse histórica, Rev. Bras. Ens. Bioq. Biol. Mol. 1, A1-A16.

- [2] D. Voet, J. Voet (1995) *Biochemistry*, Ed. Wiley & Sons Inc., New York, U.S.A.
- [3] M. Tomita, K. Hashimoto, K. Takahashi, T. S. Shimizu, Y. Matsuzaki, F. Miyoshi, K. Saito, S. Tanida, K. Yugi, J.C. Venter, C. A. Hutchison (1999) E-CELL: software environment for whole-cell simulation, *Bioinformatics* 15, 72-84.
- [4] P. Smolen, D. A. Baxter, J. H. Byrne (1998) Frequency selectivity, multistability, and oscillations emerge from models of genetic regulatory systems, *Am J Physiol* 274, C531-C542.
- [5] H. McAdams, L. Shapiro (1995) Circuit simulation of genetic networks, *Science* 269, 650-656.
- [6] T. Genoud, M. T. Santa Cruz e JP. Métraux (2001) Numeric simulation of plant signaling networks, *Plant Physiol* 126, 1430-1437.
- [7] P. Smolen, D. A. Baxter e J. H. Byrne (2000) Modeling transcriptional control in gene networks-- methods, recent results, and future directions, *Bulletin of Mathematical Biology* 62, 247-292.
- [8] N. Fidler-Ferrara e C.P.C. do Prado (1994), *Caos: uma Introdução*, Ed. Edgard Blücher, São Paulo.
- [9] P. Glendinning (1994) *Stability, Instability and Chaos: an introduction to the theory of nonlinear differential equations*, Ed. Cambridge Press.
- [10] F. Jacob e J. Monod (1961) Genetic regulatory mechanisms in the synthesis of proteins, *J. Mol. Biol.* 3, 318-356
- [11] A. Marzzoco e B.B. Torres (1999) *Bioquímica Básica*, Ed. Guanabara Koogan, Rio de Janeiro