

λ 噬菌体溶源/裂解转换调控与定态熵*

冯维¹⁾ 丁辉^{1)†} 林昊¹⁾ 罗辽复²⁾

1) (电子科技大学生命科学与技术学院, 神经信息教育部重点实验室, 生物信息中心, 成都 610054)

2) (内蒙古大学物理科学与技术学院, 理论生物物理实验室, 呼和浩特 010021)

(2011年12月12日收到; 2012年2月16日收到修改稿)

被 λ 噬菌体感染的大肠杆菌是研究生物基因调控网络的重要模板之一. 根据该调控系统操纵子与调控蛋白相互作用特点, 利用热力学配分函数描述调控蛋白与操纵子的结合概率, 在给出该调控系统的动力学分岔的基础上, 计算系统各定态的热力学熵. 结果显示, 溶源态和裂解态较鞍点和分岔点具有更低的熵值, 且溶源态的熵最低, 说明处于溶源态的系统具有更高的生物序.

关键词: λ 噬菌体, 溶源/裂解转变, 分岔, 熵

PACS: 87.18.Vf, 65.40.Gd, 02.30.Oz

1 引言

基因调控网络是一系列 DNA 序列和调控蛋白构成的相互作用体系, 它用来调控网络内基因的转录和蛋白质的表达. 然而, 这些调控网络因涉及大量错综复杂的化学反应, 使得人们难以从直观上对其进行深入的了解和研究. 为了更好地认识这些网络的动力学过程, 人们构建了各种网络模型, 来了解调控网络的内部变化规律^[1-22].

λ 噬菌体感染大肠杆菌调控系统是研究基因调控网络的研究范例^[1]. λ 噬菌体感染大肠杆菌后, 病毒会进入下面两种途径之一: 1) 当细菌分裂时, 病毒会将自身 DNA 结合到宿主细菌的 DNA 中并随宿主细菌 DNA 的复制而复制, 这种途径被认为是溶源现象, 所处状态被称为溶源态; 2) 病毒通过细菌的分子结构来大量复制病毒, 并在杀死宿主细菌以后释放这些病毒, 这种状态称为裂解态. 实验研究发现, 每个细胞每一代溶源现象的固定损失率大约为 10^{-7} , 而包含在溶源现象中的每一代的基因组突变率在 10^{-6} — 10^{-7} 之间, 也就是说溶源态甚至比其基因组本身更稳定. 此外还发现, 当该系统处于某些特定条件下, 如紫外光照射, 溶源态可以

转变为裂解态. 这一系列现象是由一整套基因调控系统所控制.

过去的 20 多年里, 在获得该调控网络中的调控蛋白和操纵子相互作用的定量实验数据的基础上, 科研工作者提出许多能够模拟该系统的数学模型来解释这些现象. 早在 1982 年, Ackers 等^[2] 提出一个统计热力学模型, 预测了启动子 P_R 和 P_{RM} 的抑制曲线, 证明了抑制蛋白结合两相邻位点的协同性. 1985 年, Shea 和 Ackers^[3] 发展了他们的模型, 给出调控蛋白 CI_2 对操纵位点的协同性可使溶源态在外界微扰下保持稳定. 随后的时间里, 人们也提出了许多模型来研究该调控系统, 但这些模型大多数都是基于 Ackers 的统计热力学模型. 例如, 2002 年 Aurell 等^[5] 对该系统溶源态的稳定性进行了研究, 并根据模型建立了一些实验去验证 *cro* 和 *cI* 基因在溶源态细胞的表达. 2004 年, Santillan 和 Mackey^[6] 利用考虑了左右操纵子 (O_L 和 O_R), 共 1200 个态的统计热力学模型研究该系统的溶源/裂解转变特征, 并论证了 CI 二聚体对左右侧操纵子的协同性结合可提高溶源态的稳定性. 同年, Bakk 等^[8] 仅利用右侧操纵子的 40 个态构建的模型, 讨论了右侧操纵子的敏感性. 另一些工作则在统计热力学模型的基础上, 在调控蛋白 CI 和 Cro

* 国家自然科学基金 (批准号: 11047180, 90403010, 200408020102) 和电子科技大学科研启动经费资助的课题.

† E-mail: hding@uestc.edu.cn

的动力学方程中考虑了随机噪声,并研究了系统双稳态受随机力的影响^[9-11,14]. Lou 等^[15]和 Wang 等^[16]详细考虑了蛋白与 DNA 的特异性结合和非特异性结合,定量研究了 λ 噬菌体溶源/裂解转变机理,并利用基于布尔模型的马尔科夫链方法研究该调控网络的动力学性质. Gedeon 等^[17]利用统计热力学模型研究了蛋白与 DNA 相互作用的协同性对转换开关的影响.此外,许多工作^[14,17-19]将该模型从研究野生型调控系统推广到研究突变型调控系统,并预测了突变型调控系统的许多性质,为实验研究提供了指导.最近,Avlund 等^[20]研究了溶源/裂解选择在噪声影响下的最小调控回路,认为 CI 蛋白能有效提高系统对噪音的鲁棒性. Cao 等^[21]在统计热力学模型的基础上,将蛋白 CI 和 Cro 的变化用离散型主方程描述,讨论了系统溶源态向裂解态转化概率. Joh 和 Weitz^[22]利用随机模型研究了噬菌体侵染大肠杆菌后,系统建立溶源性或走向裂解的可能性.尽管这些工作解决了噬菌体侵染大肠杆菌调控系统中的一些重要的问题,但在当前对该系统的动力学和热力学研究中,仍然存在许多亟待解决的问题.例如,溶源态为何如此稳定,以及裂解稳态为何没有溶源稳态的稳定性高等仍缺少必要的物理解释;从噬菌体进入大肠杆菌到溶源态的建立的过程是如何发生的; N, Q 和 CII 等蛋白在双稳态调控中扮演的角色都还不清晰.本文将探求溶源态和裂解态稳定性的物理解释.

Nicolis 和 Prigogine^[23]曾指出,可利用确定性方法、随机方法和热力学方法研究生物的自组织问题.传统的研究方法均基于前两种.最近,我们基于 Lehninger 的“最经济原理”^[24],利用动力学模型对该调控系统中的几类奇点的熵产生率进行了计算^[25,26],发现溶源态(焦点态)的熵产生率低于裂解态(焦点态)的熵产生率,且均低于鞍点态和分岔处多重奇点态的熵产生率.本文将在传统热力学模型的基础上,研究系统的分岔行为;在局域平衡假设的前提下,从热力学角度研究系统定态的信息熵,努力提供一个更加令人满意的溶源状态稳定性的物理解释.

2 网络调控模式

大量研究发现^[8,25], λ 噬菌体感染大肠杆菌调控系统主要的调控过程都发生在右侧操纵子 O_R , 该操纵子由三个操纵位点构成: O_{R1} , O_{R2} , O_{R3} .

启动子 P_{RM} 完全覆盖 O_{R3} 并部分覆盖 O_{R2} , 启动子 P_R 完全覆盖 O_{R1} 并且部分覆盖 O_{R2} . 但 RNA 聚合酶 (RNAP) 与 P_{RM} 结合时, 仅与 O_{R3} 位点结合; RNAP 与 P_R 结合时, 同时与 O_{R1} , O_{R2} 结合, 如图 1 所示.

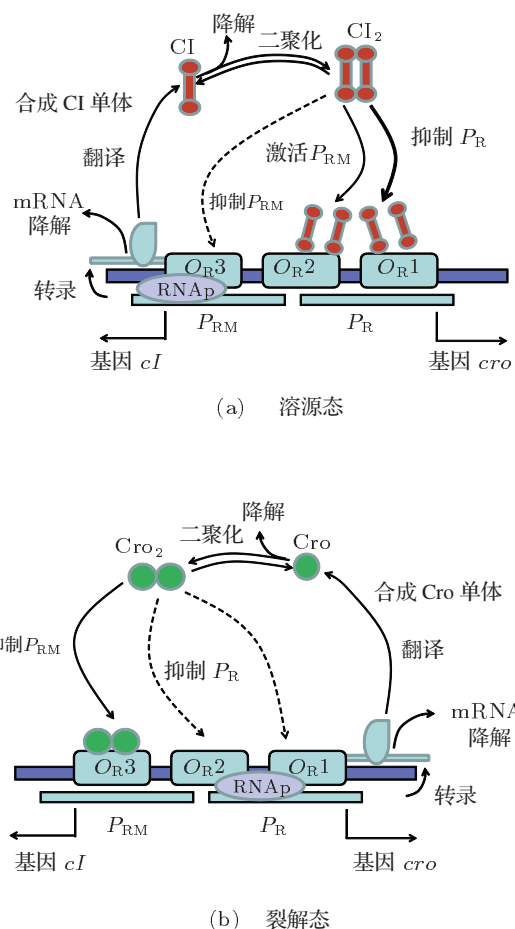


图 1 λ 噬菌体感染大肠杆菌调控网络示意图 (a) 溶源态调控; (b) 裂解态调控

CI_2 和 O_R 结合时, 与靠近转录起始的位点 O_{R1} , O_{R2} 结合较强. CI_2 首先与 O_{R1} , O_{R2} 结合, 排斥了 RNAP 与右侧早期启动子 P_R 的结合, 从而阻止了早期操纵子的转录和其他基因功能的表达. CI_2 与 O_{R1} 结合可大大增加第二个 CI_2 与 O_{R2} 结合的亲和力, 当 O_{R1} , O_{R2} 都与 CI_2 结合时, 这种相互作用一般不再延伸到 O_{R3} , 即在一个溶源菌通常的 CI_2 蛋白浓度下, 每个操纵基因上的 O_{R1} , O_{R2} 被占据后, 不再占据 O_{R3} . 若 O_{R1} 因突变而钝化, 那么 CI_2 蛋白结合到 O_{R2} , O_{R3} , 即 CI_2 与 O_{R2} 结合帮助另一个 CI_2 结合 O_{R3} . 这种协同结合的结果是在生理浓度下增加 CI_2 蛋白对操纵基因的亲和力.

Cro_2 和 O_R 结合时, 对 O_{R3} 的亲和力大于

对 O_{R2} 或 O_{R1} 的亲和力. 所以它首先结合 O_{R3} , 进而抑制 RNAP 结合于右侧维持启动子 P_{RM} , 然后 Cro_2 结合 O_{R2} , O_{R1} , 因此 Cro_2 有抑制 cI 基因转录的作用. Cro_2 与 O_R 的三个位点结合时, 也具有协同性.

图 1(a) 所示, 在溶源态, cI 基因开启, cro 基因关闭. cI 产物 CI 蛋白自发地聚合成二聚体 CI_2 . 同时, cro 产物 Cro 蛋白自发地聚合成二聚体 Cro_2 . 由于协同性, 在一般的溶源态, 大多数 O_{R1} , O_{R2} 位点被 CI_2 占据. 这有两个作用: 其一是启动子 P_R 被抑制; 其二是在启动子 P_{RM} 上的转录被启动. 结合到 O_{R2} 上的 CI_2 不影响 RNAP 与启动子 P_{RM} 的结合. 另一方面, 二聚体 CI_2 抑制 cro 转录、翻译成 Cro, 提高 CI 的产生. 然而如果 CI_2 的浓度非常高时, CI_2 与 O_{R3} 结合的概率将增加, 这将抑制 RNAP 与 P_{RM} 的结合. 这样, 在 CI_2 浓度不太高的时候, CI_2 通过增加 cI 转录、翻译成 CI 蛋白来调节自身的浓度, 否则可通过抑制基因 cI 的转录、翻译来调节自身的浓度.

如果 CI_2 的浓度降低, O_{R1} 和 O_{R2} 不被 CI_2 占据的概率就会增加. 这很可能使 RNAP 结合到 P_R 上, 并且开始 cro 基因的转录, 从而导致 Cro_2 浓度的增加. 当 Cro_2 浓度足够高时, Cro_2 与 O_{R3} 结合,

抑制 CI 的产生, 这样就转换到裂解态, 如图 1(b) 所示. 在这个态, cro 基因开启, cI 基因关闭. 当 Cro_2 浓度太高时, Cro_2 与 O_{R2} 甚至 O_{R1} 结合, 抑制 Cro 的产生.

3 调控网络的熵

Boltzmann 在研究热力学系统分子运动统计现象的基础上, 提出该系统的热力学熵 S , 表示为

$$S = k \ln W, \quad (1)$$

其中, k 为玻尔兹曼常数, W 为系统的微观状态数 (或系统分子的状态数).

Shannon 将热力学熵引入到信息论, 用来衡量一个随机变量出现的信息量的期望值, 定义为

$$H = - \sum P_i \ln P_i, \quad (2)$$

其中, P_i 为系统某一状态出现的概率.

由于热力学熵 S 的微观态密度遵守 Liouville 方程, 信息熵 H 的概率分布遵守 Kolmogorov 方程, 二者涉及的自由度不同, 是投影关系 [27]. 因此, 可通过计算信息熵来研究 λ 噬菌体感染大肠杆菌调控网络的热力学性质.

表 1 λ 噬菌体调控网络的右操纵子被占据的 40 种状态的能量值

结合态	$E/\text{kcal}\cdot\text{mol}^{-1}$	结合态	$E/\text{kcal}\cdot\text{mol}^{-1}$	结合态	$E/\text{kcal}\cdot\text{mol}^{-1}$	结合态	$E/\text{kcal}\cdot\text{mol}^{-1}$
000	0	001	-12.5	010	-10.5	100	-9.5
011	-25.7	101	-22.0	110	-22.9	111	-38.1
002	-12.0	020	-10.8	200	-13.4	022	-23.8
202	-25.4	220	-24.8	222	-37.1	021	-23.3
012	-22.5	102	-21.5	201	-25.9	210	-23.9
120	-20.3	112	-34.9	121	-32.8	211	-39.1
221	-37.3	212	-35.9	122	-33.3	033	-12.5
300	-11.5	133	-22.0	301	-24.0	310	-22.0
233	-25.9	302	-23.5	320	-22.3	333	-24.0
312	-34.0	321	-34.8	311	-37.2	322	-35.3

依据 λ 噬菌体感染大肠杆菌调控网络中三种调控蛋白 (CI_2 , Cro_2 , RNAP) 与右侧操纵子三个位点的结合情况, 可出现 40 种不同的结合状态, 如表 1 所示. 表 1 中从左到右三个数值分别代表右侧操纵子三个位点 O_{R3} , O_{R2} , O_{R1} 被占据的情况, 这里 0 代表该位置为空置状态, 即无蛋白结合到该位置, 1 代表调控蛋白 CI_2 结合到该位点, 2 代表调控

蛋白 Cro_2 结合到该位点, 3 代表 RNAP 结合到该位点. 这三种蛋白与右侧操纵子位点间的相互作用能量已由实验测得 [6], 并在表 1 中给出.

对于上述 40 种状态, 每一个状态发生的概率 P_i 为右侧操纵子 O_R 和调控蛋白 CI_2 , Cro_2 , RNAP 分子接近到一定距离的概率 P_{ai} 和它们之间产生结合的概率 P_{bi} 的乘积, 即

$$P_i = P_{ai}P_{bi}, \quad (3)$$

其中, $P_{ai} \sim [\text{Cl}_2]^{\alpha_i} [\text{Cro}_2]^{\beta_i} [\text{RNAP}]^{\gamma_i}$, $P_{bi} \sim \exp(-E_i/RT)$, E_i 是第 i 状态的相互作用能, R 是理想气体常数, T 是绝对温度. 在这里我们取 $T = 37^\circ\text{C}$, 所以 $RT \simeq 0.617 \text{ kcal/mol}$. $[\text{Cl}_2]$, $[\text{Cro}_2]$, $[\text{RNAP}]$ 分别代表 Cl_2 , Cro_2 , RNAP 分子的浓度. $\alpha_i, \beta_i, \gamma_i$ 分别代表 $\text{Cl}_2, \text{Cro}_2, \text{RNAP}$ 蛋白结合于右侧操纵子第 i 状态下的分子个数.

由概率归一化, 每一个状态发生的概率 P_i 可通过下式计算:

$$P_i = \frac{\exp(-E_i/RT)[\text{Cl}_2]^{\alpha_i} [\text{Cro}_2]^{\beta_i} [\text{RNAP}]^{\gamma_i}}{Z}, \quad (4)$$

其中,

$$Z = \sum_i \exp(-E_i/RT)[\text{Cl}_2]^{\alpha_i} [\text{Cro}_2]^{\beta_i} [\text{RNAP}]^{\gamma_i},$$

称为配分函数. 因此, 系统的信息熵可表示为

$$H = - \sum_{i=1}^{40} \frac{\exp(-E_i/RT)[\text{Cl}_2]^{\alpha_i} [\text{Cro}_2]^{\beta_i} [\text{RNAP}]^{\gamma_i}}{Z} \times \ln \frac{\exp(-E_i/RT)[\text{Cl}_2]^{\alpha_i} [\text{Cro}_2]^{\beta_i} [\text{RNAP}]^{\gamma_i}}{Z}. \quad (5)$$

由于当大肠杆菌的生长速率 $\mu \simeq 2.0 \times 10^{-2} \text{ min}^{-1}$ 时, 在每个宿主细胞中约有 1500 个 RNAP , 因此, RNAP 的浓度 $[\text{RNAP}] \simeq 3.0 \mu\text{mol/L}$ ^[6]. 在公式 (5) 中, 除 $[\text{Cl}_2]$ 和 $[\text{Cro}_2]$ 外, 其余均为常数, 因此, 熵 H 为调控蛋白 $\text{Cl}_2, \text{Cro}_2$ 浓度的函数, 表示为 $H([\text{Cl}_2], [\text{Cro}_2])$.

单体 CI 和 Cro 在形成对应的二聚物时, 是一个动态平衡过程. 由文献 [6, 25, 26], 二聚物浓度 $[\text{Cl}_T]$ 和 $[\text{Cro}_T]$ 与总体浓度 $[\text{Cl}_2]$ 和 $[\text{Cro}_2]$ 的关系为

$$[\text{Cl}_2] = \frac{1}{2}[\text{Cl}_T] - \frac{K_D^{\text{Cl}}}{8} \left[\sqrt{1 + 8 \frac{[\text{Cl}_T]}{K_D^{\text{Cl}}}} - 1 \right], \quad (6)$$

$$[\text{Cro}_2] = \frac{1}{2}[\text{Cro}_T] - \frac{K_D^{\text{Cro}}}{8} \left[\sqrt{1 + 8 \frac{[\text{Cro}_T]}{K_D^{\text{Cro}}}} - 1 \right], \quad (7)$$

等式中 K_D^{Cl} 和 K_D^{Cro} 分别表示 CI 和 Cro 二聚作用的解离常数, 由表 2 给出. 进而熵 H 为调控蛋白 CI 和 Cro 总浓度的函数, 表示为 $H([\text{Cl}_T], [\text{Cro}_T])$.

接下来考虑 $[\text{Cl}_T]$ 和 $[\text{Cro}_T]$ 的演化形式. 令 $[\text{M}_{cI}]$ 和 $[\text{M}_{cro}]$ 分别代表 cI 和 cro 两者的 mRNA 分子的浓度. 则 $[\text{M}_{cI}], [\text{M}_{cro}], [\text{Cl}_T]$

和 $[\text{Cro}_T]$ 的动态演化方程可表述为^[6,25]

$$\begin{aligned} \frac{d[\text{M}_{cI}]}{dt} = & k_{cI}^q [\text{O}_R] f_{\text{RM}}^q([\text{Cl}_2]_{\tau_m}, [\text{Cro}_2]_{\tau_m}) \\ & + k_{cI}^s [\text{O}_R] f_{\text{RM}}^s([\text{Cl}_2]_{\tau_m}, [\text{Cro}_2]_{\tau_m}) \\ & - (\gamma_m + \mu)[\text{M}_{cI}], \end{aligned} \quad (8)$$

$$\begin{aligned} \frac{d[\text{M}_{cro}]}{dt} = & k_{cro} [\text{O}_R] f_{\text{R}}([\text{Cl}_2]_{\tau_m}, [\text{Cro}_2]_{\tau_m}) \\ & - (\gamma_m + \mu)[\text{M}_{cro}], \end{aligned} \quad (9)$$

$$\frac{d[\text{Cl}_T]}{dt} = v_{cI}[\text{M}_{cI}]_{\tau_{cI}} - (\gamma_{\text{Cl}} + \mu)[\text{Cl}_T], \quad (10)$$

$$\frac{d[\text{Cro}_T]}{dt} = v_{cro}[\text{M}_{cro}]_{\tau_{cro}} - (\gamma_{\text{Cro}} + \mu)[\text{Cro}_T], \quad (11)$$

其中, γ_m 是 mRNA 的降解速率; γ_{Cro} 是 Cro_T 的降解速率; γ_{Cl} 是 Cl_T 的降解速率; τ_m 是从转录起始到有一个 mRNA 开始与核糖体结合所用的时间; τ_{cI} 是 mRNA 翻译单体 CI 所需要的时间; τ_{Cro} 是 mRNA 翻译单体 Cro 的时间; v_{cI} 和 v_{cro} 分别是 cI 和 cro 翻译起始速率; μ 是细菌的降解速率; $[\text{M}_{cI}]_{\tau_{cI}} = [\text{M}_{cI}](t - \tau_{cI})$, $[\text{M}_{cro}]_{\tau_{cro}} = [\text{M}_{cro}](t - \tau_{cro})$, $[\text{Cl}_2]_{\tau_m} = [\text{Cl}_2](t - \tau_m)$, $[\text{Cro}_2]_{\tau_m} = [\text{Cro}_2](t - \tau_m)$, 代表时间延迟; $[\text{O}_R]$ 表示总的操纵子浓度; $f_{\text{RM}}^q([\text{Cl}_2]_{\tau_m}, [\text{Cro}_2]_{\tau_m})$ 代表 Cl_2 结合于 O_R2 且 RNAP 结合于 O_R3 时的所有情况; $f_{\text{RM}}^s([\text{Cl}_2]_{\tau_m}, [\text{Cro}_2]_{\tau_m})$ 代表的是 Cl_2 不结合 O_R2 且 RNAP 结合于 O_R3 时的所有情况; $f_{\text{R}}([\text{Cl}_2]_{\tau_m}, [\text{Cro}_2]_{\tau_m})$ 代表的是 RNAP 结合于 O_R1 和 O_R2 时的所有情况. 表 2 中给出所需实验参数^[6].

表 2 方程 (6)—(11) 所需参数值^[6]

$\mu \approx 2.0 \times 10^{-2} \text{ min}^{-1}$	$k_{cro} \approx 2.76 \text{ min}^{-1}$
$k_{cI}^q \approx 4.29 \text{ min}^{-1}$	$k_{cI}^s \approx 0.35 \text{ min}^{-1}$
$\gamma_m \approx 0.12 \text{ min}^{-1}$	$v_{cI} \approx 0.09 \text{ min}^{-1}$
$v_{cro} \approx 3.2 \text{ min}^{-1}$	$\tau_{cI} \approx 0.24 \text{ min}^{-1}$
$\tau_{cro} \approx 6.6 \times 10^{-2} \text{ min}^{-1}$	$\tau_m \approx 5.1 \times 10^{-3} \text{ min}^{-1}$
$K_D^{\text{Cl}} \approx 5.56 \times 10^{-3} \mu\text{mol/L}$	$K_D^{\text{Cro}} \approx 3.26 \times 10^{-1} \mu\text{mol/L}$
$[\text{O}_R] \approx 5.0 \times 10^{-3} \mu\text{mol/L}$	$[\text{RNAP}] \approx 3.0 \mu\text{mol/L}$

对于微分方程 (8)—(11), 参照表 2 中数据, 时间延迟 τ_m 比 τ_{cI} 小两个数量级, 比 τ_{cro} 小一个数量级, 因此可简化计算, 忽略时间延迟 τ_m . 对 $[\text{M}_{cI}]_{\tau_{cI}}$ 和 $[\text{M}_{cro}]_{\tau_{cro}}$, 采用泰勒展开, 一级近似, 可忽略 τ_{cI} 和 τ_m , 这样即化为无时间延迟的微分方程组, 便于

求解. 该简化并不影响对该系统的定性分析.

当系统达到定态时, $d[M_{CI}]/dt = 0$, $d[M_{Cro}]/dt = 0$, $d[CI_T]/dt = 0$ 和 $d[Cro_T]/dt = 0$. 联立方程 (8)—(11), 则 $[CI_T]$ 和 $[Cro_T]$ 满足以下方程:

$$\begin{aligned} & \Phi([CI_T], [Cro_T], \gamma_{CI}) \\ &= \frac{v_{CI}}{\gamma_m + \mu} [O_R] (k_{CI}^a f_{RM}^a + k_{CI}^s f_{RM}^s) \\ & - (\gamma_{CI} + \mu) [CI_T] = 0, \end{aligned} \quad (12)$$

$$\begin{aligned} & \Theta([CI_T], [Cro_T]) \\ &= \frac{v_{Cro}}{\gamma_m + \mu} [O_R] k_{Cro} f_R - (\gamma_{Cro} + \mu) [Cro_T] \\ &= 0, \end{aligned} \quad (13)$$

通过求解 $[CI_T]$ 和 $[Cro_T]$, 便可计算信息熵 $H([CI_T], [Cro_T])$.

4 结果与讨论

4.1 系统的动力学行为

方程 (12) 和 (13) 是 $[CI_T]$ 与 $[Cro_T]$ 的相方程. 通过求解方程 (12) 和 (13), 可获得该调控系统的动力学行为. Cro_T 的降解速率 γ_{Cro} 和 CI_T 的降解速率 γ_{CI} 受环境的影响, 可视为变化参数. 在本研究中, 仅考虑 $\gamma_{Cro} = 0$ 时的情况, 即只研究 γ_{CI} 对系统的影响. 对于 $\gamma_{Cro} \neq 0$ 的情况, 已在文献 [26] 进行过讨论, 并不影响系统的分岔行为.

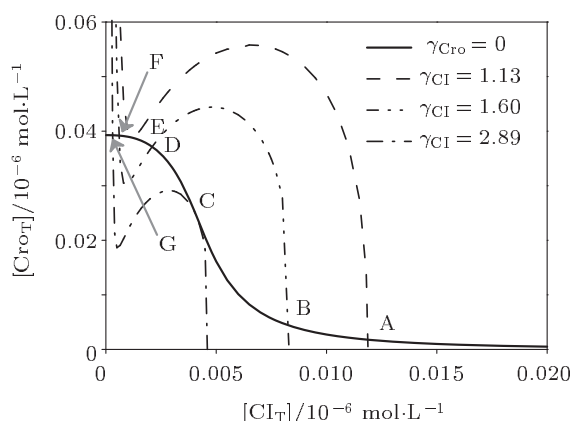


图2 系统在 $[CI_T]/[Cro_T]$ 相空间中随 γ_{CI} 改变的演化行为

图2描述了系统随 CI 蛋白降解速率 γ_{CI} 改变, 在 $[CI_T]/[Cro_T]$ 相空间的演化行为. 当 $\gamma_{CI} < 1.13 \text{ min}^{-1}$ 时, 系统仅存在一个稳态 (实线 A 点右侧), 即系统处于溶源态. 随着 γ_{CI} 的增大,

当 $\gamma_{CI} = 1.13 \text{ min}^{-1}$ 时, 出现一个新的临界态 (E 点), 为多重奇点, 原先稳定的溶源态 (A 点) 保持不变. 当 γ_{CI} 稍有增大, 发生第一次分岔, 多重奇点分成两个定态: 一个是稳定焦点的裂解态 (F 点); 另一个是不稳定的鞍点 (D 点). 先前的溶源态 (A 点随实线演化为 B 点) 仍然存在, 所以第一次分岔为单点态向三点态的转变. 随着 γ_{CI} 进一步增加, 虚线继续下降, 鞍点 (D 点) 和溶源稳态 (B 点) 沿实线逐渐靠近. 直到 $\gamma_{CI} = 2.89 \text{ min}^{-1}$, 鞍点和溶源稳态彼此碰撞湮灭产生第二个多重奇点 (C 点), 但裂解态 (F 点随实线演化为 G 点) 仍然存在. 当 γ_{CI} 进一步增大, 分岔再一次产生, 系统从三点态转变为单点态 (裂解态) (G 点随实线左移). 当 $\gamma_{CI} > 2.89 \text{ min}^{-1}$ 时, 系统仅存在裂解态.

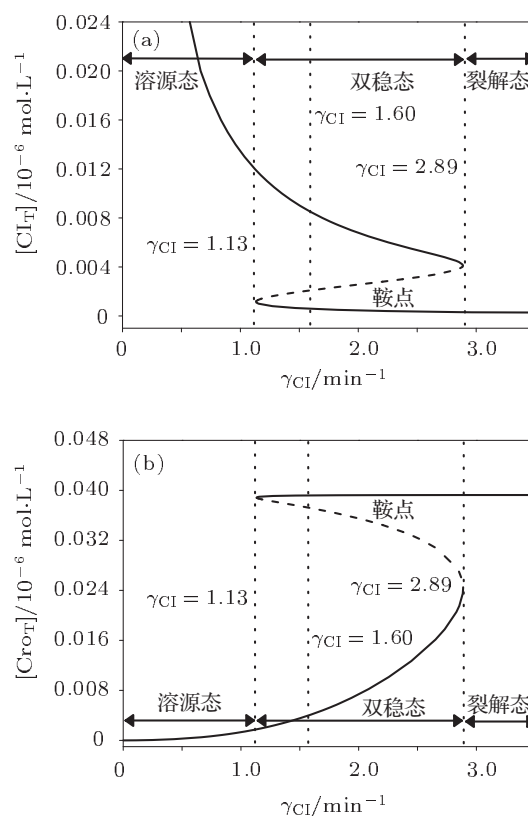


图3 $[CI_T]$ 和 $[Cro_T]$ 随降解速率 γ_{CI} 变化的分岔图 (a) $[CI_T]$ 的分岔图; (b) $[Cro_T]$ 的分岔图

图3描述了调控蛋白 CI 和 Cro 单体总浓度 $[CI_T]$ 和 $[Cro_T]$ 随 CI 降解速率 γ_{CI} 变化的分岔. 在溶源态, 系统具有较大的 $[CI_T]$ 和较小的 $[Cro_T]$; 而在裂解态, 系统则具有较小的 $[CI_T]$ 和较大的 $[Cro_T]$; 在中间的双稳态区, 出现三个定态, 分别是两个稳定焦点对应溶源态、裂解态和不稳定的鞍点. λ 噬菌体侵染大肠杆菌基因调控网络是

典型的双稳态系统. 通过对该网络的动力学模拟, 发现仅当 CI 蛋白降解速率 γ_{CI} 增大到某一临界值 (第二次分岔点) 时, 系统通过两次分岔, 经历一次三点态, 才会从溶源态转变为裂解态, 该结果与实验结果以及其他理论结果一致 [6,17]. 然而, 系统并不能从裂解态再回到溶源态, 因为在裂解态, 宿主细胞已被产生的大量噬菌体破坏. 在双稳态区域, 任一微扰都会使系统从不稳定的鞍点落入稳定的溶源态或裂解态. 需要注意的是该调控系统的溶源态和裂解态均为非平衡定态, 平衡态须遵从爱因斯坦细致平衡原理, 而方程 (12) 和 (13) 的解却并不满足该原理, 这意味着在各个定态 λ 噬菌体侵染大肠杆菌调控系统还在不停地进行运动. 因此溶源态和裂解态并非平衡态.

4.2 系统定态的熵

在生物体中, 能量转换和耗散是不可避免的. Lehninger^[24] 则认为生物体应符合最经济原则. 因此, 基于热力学的观点, 可计算该调控系统的热力学熵. 热力学熵描述了系统的混乱程度. 对于稳定的溶源态, 具有较高的生物序, 因此, 我们猜测溶源态的熵最低, 裂解态的熵其次, 而不稳定鞍点的熵应该最高. 这里可通过计算系统各定态的信息熵来研究系统的热力学性质.

由 (5)—(7) 式知信息熵 H 是 $[CI_T]$, $[Cro_T]$ 的函数, 由相方程 (12) 和 (13) 得 $[CI_T]$, $[Cro_T]$ 是 γ_{CI} 的函数 ($\gamma_{Cro} = 0$), 因此信息熵 H 是 γ_{CI} 的函数, 表示为 $H(\gamma_{CI})$. 图 4 描述了信息熵 $H(\gamma_{CI})$ 随 γ_{CI} 的变化. 对于 40 个状态的平均信息量应在 5.32 (40 个态等概率出现) 到 0 (只出现一个态) 之间. 结果显示, 稳定的溶源态和裂解态比不稳定的鞍点和分岔点具有更低的熵, 并且溶源态的熵比裂解态的熵更低. 特别是在溶源态, 其熵值已接近 0, 即某些状态以很高的概率出现, 表现出极高的生物组织有序性. 可以发现, 尽管鞍点的熵最高 (约为 1.5), 但远小于随机状态的熵 5.32, 仍表现出有序性. 因此, 我们计算了在溶源态、裂解态和鞍点中 O_R 的 40 个状态的发生概率. 结果显示, 在溶源态中, CI_2 结合 O_{R2} 和 O_{R1} , 同时 RNAP 结合 P_{RM} (表 1 中 311 态) 的概率达到约 90%; 在裂解态中, Cro_2 结合 O_{R2} 和 O_{R1} , 同时 RNAP 结合 P_{RM} (表 1 中 322 态) 的概率达到约 70%; 而在鞍点, 主要以 311 态和 322 态共存, 两态出现的概率都在 30%—50% 之间, 而其他态出现的概率极低. 我们还注意到, 到 CI 降解速率很小

时 (约为 0), 熵又发生了增加. 因此, 通过计算此时各态的概率, 发现 311 态的概率下降到约 75% 左右, 新出现了 111 态, 其概率约 25% 左右, 这主要是由于在 γ_{CI} 较小时, CI 的浓度很高, CI_2 与 RNAP 竞争占据 O_{R3} , 进而 O_R 的三个位点可能都被 CI_2 占据.

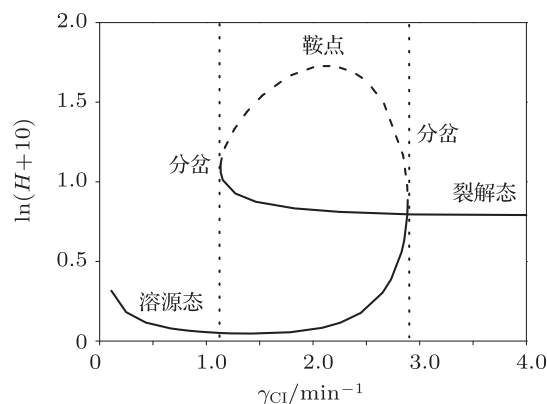


图 4 定态熵 H 随 CI 蛋白降解速率 γ_{CI} 的变化

熵值直接反映了系统所处状态的均匀程度, 系统的熵值越小, 其所处的状态越是有序, 越不均匀; 系统的熵值越大, 它所处的状态越是无序, 越均匀. 系统总是力图自发地从熵值较小的状态向熵值较大 (即从有序走向无序) 的状态转变. 在 λ 噬菌体侵染大肠杆菌基因调控系统中, 熵最大的情况是鞍点, 因此, 无论溶源态还是裂解态都应尽可能地向鞍点态或随机分布状态转变. 但是, 实验观察发现, 溶源态和裂解态都是十分稳定的, 溶源态的稳定性甚至超过其本身基因组的稳定性. 裂解态之所以比溶源态的熵高, 是因为在裂解态, λ 噬菌体不断复制进而破坏了宿主大肠杆菌细胞, 使得整个系统的混乱度增加. 但复制出的子代噬菌体仍是高度自组织的生命体, 具有较高的生物序, 所以裂解态的熵又低于鞍点和分岔点的熵. 因此, 该系统在两个稳定态, 特别是溶源态, 必定在不停地从环境吸取负熵, 以维持生命的有序性.

λ 噬菌体侵染大肠杆菌调控网络作为生物网络的研究范例, 已进行了许多工作. 这些工作利用调控蛋白与基因左右操纵子的协同作用, 导出了网络的双稳态特性 [6,15,17], 并进一步分析了突变性网络的双稳态结构 [17—19]. 另一些工作通过在 CI 和 Cro 蛋白演化的相方程中加入高斯白噪声, 研究了随机信号对该网络的影响 [9—11,14]. 本文的工作首先确定了调控蛋白与基因操纵子结合的概率的基本含义 (公式 (3)), 在此基础上研究了右侧操纵

子对该调控网络的性质,发现仅需右侧操纵子系统即可建立双稳态结构,与已有研究结论一致.熵作为基本物理量,已被应用于生物调控网络研究中,Stoll等^[28]计算了酵母细胞周期网络的熵,发现该网络定态的熵始终远小于随机最大熵.然而,据我们所知,并没有文献对噬菌体感染大肠杆菌调控网络的定态熵进行研究.通过对各定态的熵的计算,我们发现各定态的熵(无论是溶源态、裂解态还是鞍点)都远小于随机系统产生的最大熵,且溶源态的熵最低.该结果与酵母细胞周期网络的熵的结果类似,暗示生命系统拥有相同的物理学理论基础.该结果有助于阐明 λ 噬菌体感染大肠杆菌调控系统的物理机理.

5 结论

在充分考虑调控蛋白与基因组右侧操纵子结合的各种情况的基础上,计算了系统的动力学分岔;并讨论了系统定态的熵,建立了定态稳定性与熵的关系,发现稳定的定态具有较低的熵.该研究成果将为调控网络的物理解释提供依据.该研究同时给出了薛定谔关于“生命以负熵为生”^[29]的一个例子.我们曾利用化学动力学方程研究了该调控系统的动力学行为^[25,26],并计算了系统各定态的熵产生率,发现溶源态的熵产生率最低,而鞍点的熵产生率最高.两者结合可得:“生命以负熵为生,生命以熵产生最小而活”,构成了比较完整的热力学图画^[30].

- [1] Arkin A, Ross J, McAdams H H 1998 *Genetics* **149** 1633
- [2] Ackers G K, Johnson A D, Shea M A 1982 *Proc. Natl. Acad. Sci. U. S. A.* **79** 1129
- [3] Shea M A, Ackers G K 1985 *J. Mol. Biol.* **181** 211
- [4] Yi M, Jia Y, Liu Q, Zhan X 2008 *Acta Phys. Sin.* **57** 621 (in Chinese) [易鸣, 贾亚, 刘泉, 詹璇 2008 物理学报 **57** 621]
- [5] Aurell E, Brown S, Johanson J, Sneppen K 2002 *Phys. Rev. E* **65** 051914
- [6] Santillan M, Mackey M C 2004 *Biophys. J.* **86** 75
- [7] Wan X, Zhou J, Liu Z R 2012 *Acta Phys. Sin.* **61** 010203 (in Chinese) [万茜, 周进, 刘曾荣 2012 物理学报 **61** 010203]
- [8] Bakk A, Metzler R, Sneppen K 2004 *Biophys. J.* **86** 58
- [9] McAdams H H, Shapiro L 1995 *Science* **269** 650
- [10] Aurell E, Sneppen K 2002 *Phys. Rev. Lett.* **88** 048101
- [11] Tian T, Burrage K 2004 *J. Theor. Biol.* **227** 229
- [12] Liu S B, Wu Y, Hao Z W, Li Y J, Jia N 2012 *Acta Phys. Sin.* **61** 020503 (in Chinese) [刘少宝, 吴莹, 郝忠文, 李银军, 贾宁 2012 物理学报 **61** 020503]
- [13] Li R, Yan P L, Chen J, Li J, Li J, Zhang K W, Zhong J X 2009 *Acta Phys. Sin.* **58** 6703 (in Chinese) [李蓉, 颜平兰, 陈健, 李俊, 李金, 张凯旺, 钟建新 2009 物理学报 **58** 6703]
- [14] Zhu X M, Yin L, Hood L, Ao P 2004 *Funct. Integr. Genomics* **4** 188
- [15] Lou C B, Yang X J, Liu X L, He B, Ouyang Q 2007 *Biophys. J.* **92** 2685
- [16] Wang S Y, Zhang Y P, Ouyang Q 2006 *Phys. Rev. E* **73** 041922
- [17] Gedeon T, Mischaikow K, Patterson K, Traldi E 2008 *Biophys. J.* **94** 3384
- [18] Morelli M J, Wolde P R, Allen R J 2009 *Proc. Natl. Acad. Sci. U. S. A.* **106** 8101
- [19] Werner M, Aurell E 2009 *Phys. Biol.* **6** 046007
- [20] Avlund M, Krishna S, Semsey S, Dodd I B, Sneppen K 2010 *PLoS One* **5** e15037
- [21] Cao Y, Lu H M, Liang J 2010 *Proc. Natl. Acad. Sci. U. S. A.* **107** 18445
- [22] Joh R I, Weitz J S 2011 *PLOS Comput. Biol.* **7** e1002006
- [23] Nicolis G, Prigogine I 1977 *Self-Organization in Nonequilibrium Systems* (New York: Wiley Publishing)
- [24] Lehninger A L 1982 *Principles of Biochemistry* (Delhi: CBS Publishers and Distributors)
- [25] Ding H, Luo L 2009 *Chin. Phys. Lett.* **26** 098701
- [26] Ding H, Luo L F, Lin H 2011 *Commun. Theor. Phys.* **55** 371
- [27] Luo L F 2009 *Front. Phys. China* **4** 122
- [28] Stoll G, Rougemont J, Naef F 2006 *Bioinformatics* **22** 2539
- [29] Schrödinger E 1945 *What is Life?* (Cambridge: Cambridge University Press) p24
- [30] Luo L F 2010 *Sci. Tech. Rev.* **28** 106 (in Chinese) [罗辽复 2010 科技导报 **28** 106]

The lysogeny/lysis switch and entropies of stationary states in λ phage*

Feng Wei¹⁾ Ding Hui^{1)†} Lin Hao¹⁾ Luo Liao-Fu²⁾

1) (Key Laboratory for Neuro Information of Ministry of Education, Center of Bioinformatics, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China)

2) (Laboratory of Theoretical Biophysics, School of Physical Science and Technology, Inner Mongolia University, Hohhot 010021, China)

(Received 12 December 2011; revised manuscript received 16 February 2012)

Abstract

The system of *Escherichia coli* infected by λ phage is one of the templates for quantitative study of regulated networks. In this article, according to the characteristics of the interactions between operators and regulators, the probabilities of regulators binding to operators are described by thermodynamic partition function. On the basis of bifurcation analysis of this regulated network, the entropies of stationary states are calculated. Results show that the entropies of both lysogenic state and lytic state are lower than those of saddle-point and bifurcation point states. Moreover, we find that the lysogenic state has lower entropy than lytic state, which proves that the lysogenic state has a higher biological order.

Keywords: λ phage, lysogenic/lytic switch, bifurcation, entropy

PACS: 87.18.Vf, 65.40.Gd, 02.30.Oz

* Project supported by the National Natural Science Foundation of China (Grant Nos. 11047180, 90403010, 200408020102) and the Scientific Research Startup Foundation of University of Electronic Science and Technology of China.

† E-mail: hding@uestc.edu.cn