生物物理复习重点整理笔记

HW Zhong Phys.1701.Hust

2020年8月23日

1 Chap1

重点建立整个框架的思路,最重要的地方是 P24 页的表格,里面列出了生物学估算的经验数据。 其中重点关注几个数据:

- 大肠杆菌: 细胞体积 $1\mu m^3$,细胞质量 1pg,细胞表面积 $6\mu m^2$
- **酵母 (真菌):** 细胞体积 60μm³, 细胞质量 60pg, 细胞直径 5μm
- **DNA**: 碱基对长度 0.34nm, 碱基对体积 1nm³, 驻留长度 50nm
- **蛋白质:** 平均半径 2nm, 平均质量 30000Da, 平均体积 25nm³, 平均长度 300 个氨基酸

2 Chap2

重点在于估算! 估算! 估算!

2.1 模式生物

给出了几种研究中常用的模式生物:大肠杆菌,黑腹果蝇,拟南芥,斑马鱼,非洲爪蟾。 主要针对大肠杆菌进行了细胞内的分子普查,就是利用一些比例系数进行估算。有一个有意思的参数值得一 记:每个大肠杆菌中一个分子对应大约 2nmol/L 的浓度 (1 分子/1fL=2nmol/L)。【P30 大肠杆菌具体估算】

2.2 细胞内部高度拥挤

有几下 4 个方面的影响: 改变反应的速率平衡常数,增强 DNA-蛋白质相互作用,影响蛋白质折叠,增加晶状体蛋白质浓度。

2.3 研究细胞内结构的实验技术 (P36)

荧光显微镜; 原子力显微镜; 电子显微镜; 冷冻电镜

2.4 细胞及其内部结构

根据核糖体 RNA 序列的相似性和差异性把细胞分为三类:细菌、古核生物和真核生物。

P40 给出了酵母菌估算

实验测定细胞中蛋白质数目的方法: 1. 在细胞的每一个蛋白质上附加一个可以被<mark>抗体</mark>识别的氨基酸片段,就可能产生一个菌株。2. 构造<mark>蛋白质融合体</mark>,即在要研究的蛋白质上附加一个荧光蛋白质,比如 GFP,通过检查抗体结合强度或者荧光强度就可以得到每种蛋白质的分子数。

P45 给出了内质网估算

2.5 高分子组装体 (P47)(特征长度 10nm)

给出了各种高分子组装体的卡通图和功能简介。

2.6 病毒 (P52 给出了 HIV 病毒的估算)

HIV 有一个显著的特点:不同病毒体并不完全一样!

2.7 细胞的分子构造

给出了氨基酸侧链性质的明确分类。

- **疏水**: 丙氨酸 A(Ala), 缬氨酸 V(Val), 异亮氨酸 I(Ile), 半胱氨酸 C(Cys), 亮氨酸 L(Leu), 蛋氨酸 M(Met), 酪氨酸 T(Tyr), 苯丙氨酸 F(Phe), 色氨酸 W(Trp)
- 2. **极性侧链**: 甘氨酸 G(Gly), 丝氨酸 S(Ser), 苏氨酸 T(Thr), 天冬酰胺 N(Asn), 谷氨酰胺 Q(Gln), 脯氨酸 P(pro)
- 3. **带电侧链**谷氨酸 E(Glu),组氨酸 H(His),天冬氨酸 D(Asp),赖氨酸 K(Lys),精氨酸 R(Arg)

2.8 小结

- 1. 生物结构尺度范围很广
- 2. 从细胞开始,向上或者向下考察了各种生物实体的尺寸和数目
- 3. 有时候生物学对象以大量完全相同的拷贝出现,有时候拷贝数很少
- 4. 细胞内部是拥挤的

3 Chap3

3.1 用来测量生物过程时间的实验

1. 直接测量法

- 2. 固定时间点观测法
- 3. 脉冲追踪法
- 4. 产物积累法

3.2 细胞周期与标准时钟

细菌的细胞周期被定义为从细胞从母细胞分裂出来到这个新细胞自身发生细胞分裂为止的时间段。细胞周期的结构特征是已复制的细菌染色体分离成两组,并且细菌构造出新的细胞壁或者隔膜,将原来的细胞分开成两个子细胞。

P83 给出了大肠杆菌细胞周期的详细估算

4 Chap4 模式系统

P118 给出了血红蛋白相关估算以及全血细胞计数的典型值

协同性: 同一分子的配体在不同位点的结合不是独立的。

协同效应: 血红蛋白与氧结合的过程是一个非常神奇的过程。首先一个氧分子与血红蛋白四个亚基中的一个结合,与氧结合之后的蛋白结构发生变化,造成整个血红蛋白结构的变化,这种变化使得第二个氧分子相比于第一个氧分子更容易寻找血红蛋白的另一个亚基结合,而它的结合会进一步促进第三个氧分子的结合,以此类推直到构成血红蛋白的四个亚基分别与四个氧分子结合。而在组织内释放氧的过程也是这样,一个氧分子的离去会刺激另一个的离去,直到完全释放所有的氧分子,这种有趣的现象称为协同效应。

别构效应: 配体结合在蛋白质的一个位点上能够引起蛋白质上另外一处位点的结合性质改变的效应通常称为别构效应。

5 Chap5 活细胞中的力学和化学平衡 P188

一个重要的关系: $1k_BT = 4.1PNnm = 4.1 \times 10^{-21}J = 25meV$ 蛋白质结构对应于自由能极小态

5.1 疏水性来源于剥夺水分子的部分构型熵

$$\Delta S = k_B \ln 3 - k_B \ln 6 = -k_B \ln 2$$

如果我们不考虑能量效应只考虑熵效应那么吉布斯自由能

$$\Delta G = -T\Delta S$$

那么将一个疏水分子放入水中的疏水成本为

$$\Delta G_{hudrophobic}\langle n \rangle = nk_B T \ln 2$$

其中 n 为靠近非极性分子的水分子数目。也可以换个方式,用 $\gamma_{hydrophobic}$ 来表示水中加入一个疏水分子的自由能成本是

$$\Delta G_{hydrophobic} = \gamma_{hydrophobic} A$$

A 是疏水分子和周围水分子之间的界面面积。书中给出估算: 10 个水分子的覆盖面积大约为 $1nm^2$, 且 $\ln 2 \approx 0.7$,于是可以认为浸在水中的较大的疏水物体的界面自由能为 $7k_BT/nm^2$ 。对于 O_2 而言,界面面积大约 $0.1-0.2nm^2$,因此成本约为 $1k_BT$,很容易溶于水,但是对于大的蛋白质分子或者糖分子则需要足够的自由能才行。

6 Chap6 统计力学应用

 Ω 个格点,L 个配体,总的微观态数目为:

是态数目为:
$$C_{\Omega}^{L} = \frac{\Omega!}{(\Omega-L)!L!}$$
受体被占据的权重 $= e^{-\beta\epsilon_b} \times \sum_{solution} e^{-\beta(L-1)\epsilon_{sol}}$

应用 stirling 近似的一个小结论:

$$\frac{\Omega!}{(\Omega-L)!}\approx \Omega^L$$

课本中给出一个经典计算:

$$p_{bound} = \frac{e^{-\beta\epsilon_b} \frac{\Omega^{L-1}}{(L-1)!} e^{-\beta(L-1)\epsilon_{sol}}}{\frac{\Omega^L}{L!} e^{-\beta L\epsilon_{sol}} + e^{-\beta\epsilon_b} \frac{\Omega^{L-1}}{(L-1)!} e^{-\beta(L-1)\epsilon_{sol}}}$$

定义 $\Delta \epsilon = \epsilon_b - \epsilon_{sol}$ 即结合后的能量减去未结合时候的能量得到能量改变量。可以得到:

$$p_{bound} = \frac{(L/\Omega)e^{-\beta\Delta\epsilon}}{1 + (L/\Omega)e^{-\beta\Delta\epsilon}}$$

再引入参考浓度 $c_0 = \Omega/V_{box}$ 可以再次改写:

$$p_{bound} = \frac{(c/c_0)e^{-\beta\Delta\epsilon}}{1 + (c/c_0)e^{-\beta\Delta\epsilon}}$$

这个式子同时也是 n=1 时的希尔函数。

6.1 稀溶液的自由能

 N_{H_2O} 个水分子, N_S 个溶质分子。总的自由能:

$$G_{tot} = N_{H_2O}\mu_{H_2O} + N_S\epsilon_S - TS_{mix}$$

即没有混合时的两部分能量减去混合带来的熵变。

$$W(N_{H_2O}, N_S) = \frac{(N_{H_2O} + H_S)!}{N_{H_2O}! N_S!}$$

代入热力学熵关系:

$$S_{mix} = -k_B \left[N_{H_2O} \ln \frac{N_{H_2O}}{N_{H_2O+N_S}} + N_S \ln \frac{N_S}{N_{H_2O+N_S}} \right]$$

考虑到稀溶液进一步近似最终得到:

$$S_{mix} = -k_B \left(N_S \ln \frac{N_S}{N_{H_2O}} - N_S \right)$$

代入自由能表达式并且利用

$$\mu_S = \left(\frac{\partial G}{\partial N_S}\right) \ c = \frac{N_s}{V_{box}} \ c_0 = \frac{N_{H_2O}}{V_{box}}$$

可以得到:

$$\mu_s = \epsilon_s + k_B T \ln \frac{c}{c_0}$$

也可以写成:

$$\mu_i = \mu_{i0} + k_B T \ln \frac{c_i}{c_{i0}}$$

分母的参考态通常约定为 1mol/L

6.2 渗透压是熵弹性的一种表现形式

热力学第二定律告诉我们:

$$(\mu_1 - \mu_2)dN_1 \le 0$$

之前我们得到了总的吉布斯自由能的表达式并且对溶质分子求导的出了溶质分子的化学势,但是如果我们将体系自由能对 N_{H_2O} 求导就可以得到水的化学势为:

$$\mu_{H_2O}(T,p) = \mu_{H_2O}^0(T,p) - \frac{N_s}{N_{H_2O}} k_B T$$

从而得到容器两端的平衡条件:

$$\mu_{H_2O}^0(T, p_1) = \mu_{H_2O}^0(T, p_2) - \frac{N_s}{N_{H_2O}} k_B T$$

对右边的化学势在 p_1 附近做泰勒展开

$$\mu^0_{H_2O}(T,p_2) = \mu^0_{H_2O}(T,p_1) + \Big(\frac{\partial \mu^0_{H_2O}}{\partial p}\Big)(p_2 - p_1)$$

根据热力学关系 (v 为每个分子所占体积)

$$\frac{\partial \mu}{\partial p} = v$$

我们可以得到范特霍夫公式:

$$(p_2 - p_1) = rac{N_s}{V} k_B T \ \ (V = N_{H_2O} v)$$

P227的估算里并没有考虑另一边的渗透压大小,只是估算了一个量级。

6.3 质量作用定律——平衡常数由最大熵决定

$$A + B \rightleftharpoons AB$$

$$dG = \left(\frac{\partial G}{\partial N_A}\right)_{T,p} + \left(\frac{\partial G}{\partial N_B}\right)_{T,p} + \left(\frac{\partial G}{\partial N_{AB}}\right)_{T,p} = 0 \Rightarrow \sum_i \mu_i dN_i = 0$$

$$dN_i = \nu_i dN_r \Rightarrow \sum_i \nu_i \mu_i = 0$$

我们已经得到了:

$$\mu_i = \mu_{i0} + k_B T \ln \frac{c_i}{c_{i0}}$$

代入上上式就有:

$$\sum_{i} \nu_{i} \mu_{i0} = -k_{B} T \sum_{i} \ln \left(\frac{c_{i}}{c_{i0}}\right)^{\nu_{i}}$$
$$-\frac{1}{k_{B} T} \sum_{i} \nu_{i} \mu_{i0} = \ln \left[\prod_{i} \left(\frac{c_{i}}{c_{i0}}\right)^{\nu_{i}}\right]$$
$$\prod_{i} c_{i}^{\nu_{i}} = \prod_{i} c_{i0}^{\nu_{i}} e^{-\beta \sum_{i} \nu_{i} \mu_{i0}} \equiv K_{eq} = \frac{1}{K_{d}}$$

6.4 配体受体结合再探

$$L + R \rightleftharpoons LR$$

依据解离常数我们可以写出受体被配体占据的概率为:

$$K_d = \frac{[L][R]}{[LR]} \Rightarrow p_{bound} = \frac{[LR]}{[LR] + [R]} = \frac{[L]/K_d}{1 + ([L]/K_d)}$$

从这个式子我们可以看出 K_d 可以很自然地被解释为受体被占据的概率为 $\frac{1}{2}$ 时的配体浓度。并且比对之前得到的关系有:

$$K_d = \frac{1}{v}e^{\beta\Delta\epsilon}$$

要注意一下这里 $[L] = V/\Omega v$

6.5 希尔函数

分析协同效应:

$$L + L + R \rightleftharpoons L_2 R$$

$$(K_d)^2 = \frac{[L]^2 [R]}{[L_2 R]}$$

$$p_{bound} = \frac{[L_2 R]}{[R] + [L_2 R]} = \frac{[L]^2 [R] / K_d^2}{[R] + [L]^2 [R] / K_d^2} = \frac{([L] / K_d)^2}{1 + ([L / K_d])^2}$$

希尔函数的一般形式为:

$$p_{bound} = \frac{([L]/K_d)^n}{1 + ([L/K_d])^n}$$

这里容易产生误会,而维基百科做了一点解释:

$$\theta = \frac{[L]^n}{K_d + [L]^n} = \frac{[L]^n}{(K_A)^n + [L]^n}$$

这里定义的 K_d 是表观解离常数,即我们计算出来的 K_d^2 ,而 K_A 是我们的表达式中的 K_d 。而它给出的 K_A 我们的 K_d 的含义就是产生半数占用时候的配体浓度。

6.6 ATP 水解自由能

$$dG = \sum_{i} \mu_{i} dN_{i} = \sum_{i} \mu_{i} \nu_{i} dN_{r} = \sum_{i} [\mu_{i0} + k_{B}T \ln \frac{c_{i}}{c_{i0}}] \nu_{i}$$

两次使用上式:

$$\Delta G = \sum_{i} [\mu_{i0} + k_B T \ln \frac{c_i}{c_{i0}}] \nu_i$$

$$\Delta G_{ref} = \sum_{i} [\mu_{i0} + k_B T \ln \frac{c_{ref}}{c_{i0}}] \nu_i$$

于是有:

$$\Delta G = \Delta G_{ref} + k_B T \ln \left(\frac{\prod_{i=1}^{N} c_i^{\nu_i}}{\prod_{i=1}^{N} c_{i,ref}^{\nu_i}} \right)$$

不妨就取平衡态为参考,那么 $\Delta G_{ref} = 0, \prod_{i=1}^{N} c_{i,ref}^{\nu_i} = K_{eq}$ 于是:

$$\Delta G = k_B T \ln \left(\frac{\prod_{i=1}^{N} c_i^{\nu_i}}{K_{eq}} \right) = -k_B T \ln K_{eq} + k_B T \ln \left(\prod_{i=1}^{N} \frac{c_i^{\nu_i}}{1 mol/L} \right) \equiv \Delta G_0 + k_B T \ln \left(\prod_{i=1}^{N} \frac{c_i^{\nu_i}}{1 mol/L} \right)$$

再拿出来写一遍(注意无论什么函数里的宗量都应该是无量纲的数字, 所以量纲一定要处理好):

$$\Delta G = \Delta G_0 + k_B T \ln \left(\prod_{i=1}^{N} \frac{c_i^{\nu_i}}{1 mol/L} \right)$$

将这个式子应用到 ATP 水解反应

$$ATP \rightleftharpoons ADP + P_i$$

就有:

$$\Delta G = \Delta G_0 + k_B T \ln \frac{[ADP][P_i]}{[ATP]}$$

这里所有浓度都是相对于 1mol/L 的标准状态测量的,采用 $\Delta G_0 = -12.5k_BT/$ 分子 可以大致得到 $\Delta G = -20k_BT/$ 分子

在这里考虑一个习题,也就是作业题的第二题,我们会更加清晰地看到结合能的含义。我们得到有如下关系:

$$p_{bound} = \frac{1}{1 + \frac{N_{NS}}{P} e^{\beta \Delta \epsilon_{pd}}}$$

其中

$$\Delta \epsilon_{pd} = \epsilon_{pd}^S - \epsilon_{pd}^{NS}$$

也就是特异性结合的能量减去非特异性结合的能量。但是我们知道还有一个关系就是将解离常数和结合能联 系起来的一个关系:

$$K_d^S = \frac{e^{\beta \Delta \epsilon^S}}{v} \qquad \quad K_d^{NS} = \frac{\epsilon^{\beta \Delta \epsilon^{NS}}}{v}$$

将两个值做比就有:

$$\frac{K_d^S}{K_d^{NS}} = \exp\{\beta \Delta \epsilon^S - \beta \Delta \epsilon^{NS}\} = \exp\{\beta \epsilon_{pd}^S - \beta \epsilon_{sol} - \beta \epsilon_{pd}^{NS} + \beta \epsilon_{sol}\} = e^{\beta \epsilon_{pd}^S - \beta \epsilon_{pd}^{NS}} \equiv e^{\beta \Delta \epsilon_{pd}^S}$$

而 $\Delta \epsilon$ 就是特异性与非特异性结合的能量差! 也自然可以看作所需**启动能**。

7 Chap7 二态系统

最开始给出了离子通道的一些具体计算 (P244 左右)。

7.1 吉布斯分布 (巨正则系综)

$$\mathcal{Z} = \sum_{i} e^{-\beta (E_s^{(i)} - \mu N_s^{(i)})}$$
$$\langle N \rangle = \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln \mathcal{Z} = \frac{1}{\mathcal{Z}} \sum_{i} N_i e^{-\beta (E_s^{(i)} - \mu N_s^{(i)})}$$

7.2 磷酸化

磷酸化能改变活性态和失活态之间的能量均衡。

用 σ_s 来标记活性 $(0/1),\sigma_p$ 来标记磷酸化状态 (0/1)。蛋白质的自由能可以写为:

$$G(\sigma_p, \sigma_s) = (1 - \sigma_p)[(1 - \sigma_s) \times 0 + \sigma_s \epsilon] + \sigma_p[(1 - \sigma_s)(-I_2) + \sigma_s(\epsilon - I_1)]$$

$$G(\sigma_p, \sigma_s) = \sigma_s \epsilon - I_2 \sigma_p + (I_2 - I_1) \sigma_s \sigma_p$$

那么可以计算得到蛋白质在未磷酸化时处于活性态的概率为:

$$p_{active} = \frac{e^{-\beta \epsilon}}{e^{-\beta \epsilon} + 1}$$

而在磷酸化的蛋白质中概率为:

$$p_{active}^* = \frac{e^{-\beta(\epsilon - I_1)}}{e^{-\beta(\epsilon - I_1) + e^{\beta I_2}}}$$

两者作比来看出活性改变:

$$\frac{p_{active}^*}{p_{active}} = \frac{1 + e^{\beta \epsilon}}{1 + e^{\beta (\epsilon + I_2 - I_1)}}$$

假定 $\epsilon = 5k_BT$, $I_2 - I_1 = -10k_BT$ 那么有:

$$\frac{p_{active}^*}{p_{active}} = 150$$

磷酸化带来的活性增加比通常在 2~1000 左右。

7.3 协同作用

协同性是指对于一个给定的配体,与受体的结合能同该受体上已经结合的配体数目相关。比如血红蛋白的例子,直觉告诉我们,协同性源自配体和蛋白质结合导致的构象变化。从统计力学的角度,协同作用解释为相互作用能,即多个不同配体的结合作用是不能简单相加的。

仍然考虑血红蛋白,但是稍作简化,只考虑一个二聚体,系统的能量可以写为:

$$E = \epsilon(\sigma_1 + \sigma_2) + J\sigma_1\sigma_2$$

血红蛋白二聚体的两个 slot 显然是可以区分的,因此我们不考虑全同性问题,写出巨配分函数:

$$\mathcal{Z} = 1 + 2e^{-\beta(\epsilon - \mu)} + e^{-\beta(2\epsilon + J - 2\mu)}$$

根据

$$\langle N \rangle = \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln \mathcal{Z}$$

可以得到:

$$\langle N \rangle = \frac{2e^{-\beta(\epsilon-\mu) + 2e^{-\beta(2\epsilon + J - 2\mu)}}}{1 + 2e^{-\beta(\epsilon-\mu)} + e^{-\beta(2\epsilon + J - 2\mu)}}$$

再利用 $\mu = \mu_0 + k_B T \ln \frac{c}{c_0}$ 得

$$\langle N_{bound}\rangle = \frac{2(c/c_0)e^{-\beta\Delta\epsilon} + 2(c/c_0)^2e^{-\beta(2\Delta\epsilon+J)}}{1 + 2(c/c_0)e^{-\beta\Delta\epsilon} + (c/c_0)^2e^{-\beta(2\Delta\epsilon+J)}}$$

 $\Delta \epsilon = \epsilon - \mu_0$, 我们令 J = 0 可以得到:

$$\langle N_{bound} \rangle = 2 \frac{(c/c_0)e^{-\beta\Delta\epsilon}}{1 + (c/c_0)e^{-\beta\Delta\epsilon}}$$

7.4 MWC 模型 (以及之后的模型简单了解一下)

8 **Chap8 随机行走 **

Note: 结构的统计描述强调的是平均尺寸和形状, 而不是原子坐标。

8.1 无规行走高分子的平均尺寸和链节数目的平方根成正比

给出一个简单但是重要的证明:

$$\langle R \rangle = \langle \sum_{i} x_{i} \rangle = 0$$

$$\langle R^{2} \rangle = \langle \sum_{i} i j x_{i} x_{j} \rangle$$

$$\langle R^{2} \rangle = \sum_{i} \langle x_{i}^{2} \rangle + \sum_{ij}' \langle x_{i} x_{j} \rangle = \sum_{i} \langle x_{i}^{2} \rangle + \sum_{i} \langle x_{i} \rangle \sum_{j} \langle x_{j} \rangle = Na^{2}$$

$$\Rightarrow \sqrt{\langle R^{2} \rangle} = \sqrt{Na}$$

8.2 聚合物链的熵弹性

我们很轻易地可以计算得到:

$$p(n_r; N) = \frac{N!}{n_r!(N - n_r)!} (\frac{1}{2})^N$$

利用它来给出首末端距 $R=(n_r-n_l)a$ 的概率分布(a 为链节长度也称为库恩长度),利用条件 $n_r+n_l=N$ 得到 $n_r=\frac{N+\frac{R}{a}}{2}$ 于是之前得到的 $p(n_r;N)$ 可以改写为:

$$P(R; N) = \frac{N!}{(\frac{N}{2} + \frac{N}{2a})!(\frac{N}{2} - \frac{R}{2a})!}$$

经过一系列的斯特林近似我们就可以得到首末端距的概率分布函数:

$$P(R;N) = rac{1}{\sqrt{2\pi N a^2} e^{-R^2/2Na^2}}$$

事实上如果拓展到三维情况就有:

$$P(R;N) = \left(rac{3}{2\pi N a^2}
ight)^{3/2} e^{-3R^2/2Na^2}$$

8.3 驻留长度是聚合物保持大致为直构型的长度尺度

$$\langle t(s) \cdot t(u) \rangle = e^{-|s-u|/\xi_p}$$

我们可以认为聚合物骨架切向量在大于驻留长度尺度上无关联,可以把聚合物看成由 $N=L/\xi_p$ 个链节组成,链节之间的指向完全随机。为了搞清楚**驻留长度\xi_p** 和**库恩长度** a 的关系,我们进行推导:

$$R = \int_0^L ds \boldsymbol{t}(s)$$

$$\langle R^2 \rangle = \langle \int_0^L ds \boldsymbol{t}(s) \cdot \int_0^L du \boldsymbol{t}(u) \rangle$$

$$RHS = 2 \int_0^L ds \int_s^L du e^{-(u-s)/\xi_p} = 2 \int_0^L ds \int_0^\infty dx e^{-x/\xi_p} = 2L\xi_p$$

而之前我们根据无规行走推导得到的结果是:

$$\langle r^2 \rangle = Na^2 = aL$$

对比可以看出

$$a=2\xi_p$$

8.4 基因组尺寸

为了更加精确地衡量聚合物尺寸,我们引入回转半径的概念:

$$R_G^2 = \frac{1}{N} \sum_{i=1}^{N} \langle (\boldsymbol{R_i} - \boldsymbol{R_{CM}})^2 \rangle$$

根据繁杂的数学我们可以推导出:

$$R_G = \sqrt{\frac{L\xi_p}{3}}$$

进一步我们可以得出:

$$L = 0.34nm \times N_{bp} \Rightarrow R_G \approx \frac{1}{3} \sqrt{N_{bp} \xi_p} nm$$

8.5 染色体被系连在细胞核的不同位置上

我们首先要准备两组 DNA 结合位点,它们的基因距离是分开而且已知的,我们可以通过测量这些结合位点在空间中的物理距离来绘制出基因组上不同位点的空间分布情况。如果没有系连或者只有一个系连点时,染色质的无规行走模型预言两个荧光标记之间的距离 r 是高斯分布:

$$P(\mathbf{r}) = \left(\frac{3}{2\pi N a^2}\right)^{3/2} e^{-3\mathbf{r}^2/2Na^2}$$

若存在两个系连点,一个系连点和其中一个荧光标记重合,另一个系连点在两个标记之间并且与第一个系连点之间的距离为 \mathbf{R} ,那么两个系连点之间的距离是一个有偏高斯分布:

$$P(\mathbf{r}) = \left(\frac{3}{2\pi N' a^2}\right)^{3/2} e^{-3(\mathbf{r} - \mathbf{R})^2 / 2N' a^2}$$

N'是第二个系连点和第二个标记之间的库恩链节数目。这个是容易理解的,因为当中间有一个链节点的时候,

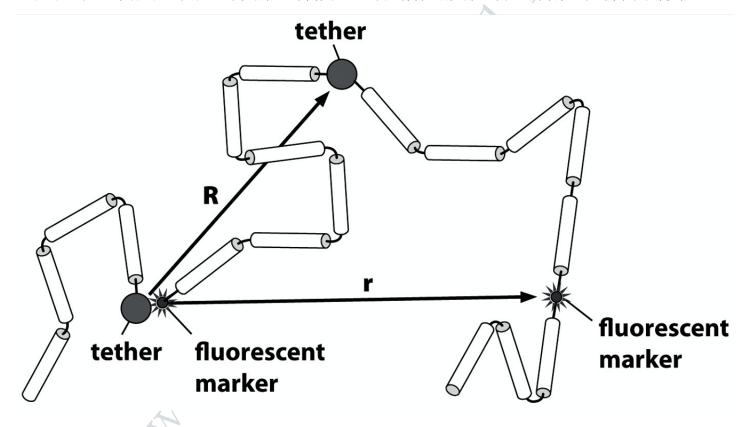


图 1: 两个系连点

 $r = R + r_{22}$ 。但是之前的结论都是矢量等式不直观,我们希望得到标量式:

$$P(r') = \left(\frac{3}{2\pi N a^2}\right)^{3/2} 4\pi (r')^2 e^{-3(r')^2/2Na^2}$$

若存在 2 个系连点:

$$P(r') = \left(\frac{3}{2\pi N' a^2}\right)^{1/2} \frac{r}{R} \left[e^{-3(\mathbf{r} - \mathbf{R})^2 / 2N' a^2} - e^{-3(\mathbf{r} + \mathbf{R})^2 / 2N' a^2} \right]$$

染色体上两个标记的基因座之间的轮廓长度 Na 可以用两个标记之间的基因距离表示成 $Na=N_{bp}/\nu$, ν 是染色质 DNA 的包装线密度。

$$\langle R^2 \rangle = Na^2 = N_{bp}a/\nu$$

。为了加深印象,这里给出一个直观的例子,假设在 30nm 粗的染色质纤维上,两个基因距离为 $N_{bp}=100kb$ 的基因座之间的轮廓长度大约是 $100kb/(100bp/nm)=1\mu m$ 【迷糊了,这怎么出来的?】

8.6 DNA 成环

我们考虑一维无规行走返回原点的概率:

$$p_0 = rac{$$
成环构型数目 $}{$ 总构型数目 $} = rac{rac{N!}{(N/2)!(N/2)!}}{2^N}$

N 是库恩链节的数目,使用斯特林近似就有:

$$p_0 \approx \sqrt{\frac{2}{\pi N}}$$

考虑三维情况,我们也不一定要让他完全返回原点,只需要首末端距保持在 $\delta \ll \sqrt{Na^2}$ 内即可:

$$p_0 = \int_0^{\delta} 4\pi R^2 P(R; N) dR = \left(\frac{6}{\pi N^3}\right)^{1/2} \left(\frac{\delta}{a}\right)^3 \propto N^{-\frac{3}{2}}$$

8.7 PCR,DNA 解旋和 DNA 泡

解旋是熵和能量竞争的结果,熵增原理喜欢解旋,而碱基配对形成氢键则符合能量最低原理的口味。我们只考虑单泡模型,因为生成多个泡引起的熵增相比于要产生他们而增加的能量实在是微乎其微。泡长为 n 个碱基对的平衡态概率:

$$p_{1}(n) = \frac{e^{-\Delta G_{1}(n)/k_{B}T}}{Z}$$

$$Z = \sum_{n=1}^{N} e^{-\Delta G_{1}(n)/k_{B}T}$$

$$\Delta G_{1}(n) = E_{in} + nE_{el} - k_{B}T \ln[\Omega_{0}(n)(N - n + 1)]$$

$$\Omega_{0}(n) = 2^{2n}p_{0}(2n) = \frac{2^{2n}}{\sqrt{\pi n}}$$

代入可以得到一个方程:

$$\Delta \epsilon + \frac{1}{2n^*} + \frac{1}{N - n^* + 1} = 0$$
 $\Delta \epsilon = E_{el}/k_B T - 2 \ln 2$

8.8 HP 模型

课本第 300 页给出了氨基酸的 HP 分类.

9 **Chap13**

9.1 浓度场与扩散动力学

菲克定律:

$$j = -D \frac{\partial c}{\partial x}$$

j 是单位时间的流密度,可视作单位时间通过单位面积的粒子数;D 是扩散系数。下面开始导出扩散方程:

$$N_{box}(x, y, z, t) = c(x, y, z, t)\Delta x \Delta y \Delta z$$

粒子数守恒,引起盒子内粒子数变化的唯一因素就是穿过盒子表面的流:

$$\frac{\partial N_{box}}{\partial t} = \frac{\partial c}{\partial t} \Delta x \Delta y \Delta z = j(x, y, z) \Delta y \Delta z - j(x + \Delta x, y, z) \Delta y \Delta z$$

泰勒展开:

$$\frac{\partial c}{\partial t} \Delta x \Delta y \Delta z \approx j(x, y, z) \Delta y \Delta z - \left[j(x, y, z) + \frac{\partial j}{\partial x} \Delta x \right] \Delta y \Delta z$$

整理可以得到:

$$\frac{\partial c}{\partial t} = -\frac{\partial j}{\partial x}$$

结合菲克定律就可以得到一维扩散方程:

$$\boxed{\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2}}$$

这是一种比较宏观的推导方式,我们也可以选择一种微观的视角:即对微观轨迹求和导出扩散方程。具体来 说就是用给定时间在指定位置发现粒子的概率密度 $p(\boldsymbol{x},t)$ 来描述随机轨迹,则 t 时刻中心处于位置 x 宽度为 Δx 的盒子内发现粒子的概率为: $p(x,t)\Delta x$, 为了简化情况, 我们只考虑一维的情形。那盒子里的粒子只有向 左,不动和向右三种选项。当然我们认为向左和向右的概率相同为 $k\Delta t$,那么原地不动的概率就为 $1-2k\Delta t$, 于是可以轻易地计算出:

$$\langle \Delta x \rangle = 0, \langle \Delta x^2 \rangle = 2a^2k\Delta t, \langle x_{tot}^2 \rangle = 2(a^2k)t$$

下面开始推导 p(x,t) 的方程: 从 t 时刻出发, 我们考虑会在 $t + \Delta t$ 时刻到达 x 处的所有轨迹的求和, 那显然 只有三种可能, 一是原本处于 x 的粒子这一次选择不动, 原本在 x-a 的粒子选择向右, 原本在 x+a 的粒子选 择向左。我们可以轻易地写出概率:

$$p(x,t+\Delta t) = \underbrace{(1-2k\Delta t)\times p(x,t)}_{\text{\vec{r}-$zd}} + \underbrace{k\Delta t \times p(x-a,t)}_{\text{\vec{p}-$fz}} + \underbrace{k\Delta t \times p(x+a,t)}_{\text{\vec{p}-$zd}}$$

这显然是一个马尔可夫过程,此刻的概率和它经历过的历史并无关联。我们再次使用泰勒展开:

$$p(x, t + \Delta t) \approx p(x, t) + \Delta t \frac{\partial p(x, t)}{\partial t}$$
$$p(x \pm a, t + \Delta t) \approx p(x, t) \pm a \frac{\partial p(x, t)}{\partial x} + \frac{a^2}{2} \frac{\partial^2 p(x, t)}{\partial x^2}$$

$$p(x \pm a, t + \Delta t) \approx p(x, t) \pm a \frac{\partial p(x, t)}{\partial x} + \frac{a^2}{2} \frac{\partial^2 p(x, t)}{\partial x^2}$$

把这些泰勒展开时带入之前得到的概率式中就可以得到:

$$\boxed{\frac{\partial p(x,t)}{\partial t} = (a^2k)\frac{\partial^2 p(x,t)}{\partial x^2}}$$

对比之前由菲克定律得出的扩散方程就可以得到之前未给出的系数 $D=a^2k$ 。不过这样的推导仅仅给出了 Δt 时间,我们现在开始考虑任意场的时间,我们将时间划分成长度为 Δt 的间隔。这也就是爱因斯坦推导方 法:

$$\underbrace{c(x,t+\Delta t)}_{\text{位于 x 的现时浓度}} = \int_{-\infty}^{+\infty} \underbrace{c(x,+\Delta,t)}_{\text{位于 } x+\Delta \text{ 的早先浓度 跳跃步长为 } \Delta \text{ 的概率}} d\Delta$$

我们假设跳跃的概率形式是对称的,而且满足归一化条件,再次将上式泰勒展开,令

$$D \equiv \frac{1}{2\Delta t} \int_{-\infty}^{+\infty} \Delta^2 \phi(\Delta) d\Delta$$

就又得到了扩散方程。

9.2 扩散方程的解与性质

对于 t=0 时刻 N 个分子集中于 x=0 的无限小邻域内,浓度的演化为:

$$c(x,t) = \frac{N}{\sqrt{4\pi Dt}}e^{-x^2/4Dt}$$

这个解被称为扩散方程的格林函数,高斯展宽为 $\sqrt{4Dt}$ 根据这个式子我们还可以直接计算出展宽 $\langle x^2 \rangle$ 随着时间的变化:

$$\langle x^2 \rangle = \frac{\int_{-\infty}^{+\infty} x^2 \frac{N}{\sqrt{4\pi Dt}} e^{-x^2/4Dtdx}}{N} = 2Dt$$

这里的处理书中写的是 t 时刻在 x 处发现粒子的概率为 c/N,可以有两个方式理解,一个是由于体系中有 N 个粒子,因此这个浓度并不是归一化的,因此需要做个归一化处理,或者直接从数学上理解去归一化为一个概率分布函数再求期望。这个结果表明了扩散时间与位移成正比。

$$\langle x^2 \rangle = 2Dt$$

9.3 FRAP和FCS

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2}$$

$$c(x,0) = \begin{cases} c_0 & -L < x < -a \\ 0 & -a < x < a \\ c_0 & a < x < L \end{cases}$$

边界条件: $x = \pm L$ 时 $\partial c/\partial x = 0$, 即边界上无物质流动,细胞内的东西不能流出去。我们首先对浓度进行傅立叶展开:

$$c(x,t) = A_0(t) + \sum_{n=1}^{\infty} A_n(t) cos\left(\frac{x}{L}n\pi\right)$$

代入方程对比系数可以得到:

$$c(x,t) = c_0 \left[1 - \frac{a}{L} - \sum_{n=1}^{\infty} \frac{2\sin(n\pi a/L)}{n\pi} e^{-(Dn^2\pi^2/L^2)t} \cos(\frac{x}{L}n\pi) \right]$$
$$c_{\infty} = c_0 \left[1 - \frac{a}{L} \right]$$

很显然,如果时间很长会导致浓度在 2L 区间内均匀分布。我们可以根据 $N_f(t) = \int_{-a}^a c(x,t) dx$ 来计算出恢复曲线:

$$N_f(t) = 2c_0a(1 - \frac{a}{L}) \times \left[1 - \frac{1}{(a/L)(1 - a/L)} \sum_{n=1}^{\infty} \frac{2}{n^2\pi^2 sin^2(n\pi a/L)e^{-(Dn^2\pi^2/L^2)t}}\right]$$

9.4 爱因斯坦关系

考虑有外力存在时候的扩散,比如电场或者粘滞的存在,我们就会得到有漂移的扩散。

$$F = \gamma v$$

当半径为 a 的球形粒子在粘滞系数 η 的液体中移动时, $\gamma = 6\pi\eta a$ 。由于这个外力引起的净流可以写为:

$$J_F = \frac{v\Delta tc}{\Delta t} = \frac{F}{\gamma}c$$

并且需要改写总的概率流写为随机跳动部分和外力作用部分之和:

$$J(x) = -D\frac{\partial c}{\partial x} + \frac{F}{\gamma}c$$

当平衡时,净流量为0就有:

$$D\frac{\partial c}{\partial x} = \frac{F}{\gamma}c$$

采用分离变量法求解

$$\gamma D \frac{dc}{c} = F dx$$

为了方便起见,我们只考虑由于势能导致的力为 F = -dU/dx, 对该式积分就可以得到:

$$\frac{c(x)}{c(0)} = \frac{e^{-U(x)/\gamma D}}{e^{-U(0)/\gamma D}}$$

这个式子实际上描述的是经典玻尔兹曼分布,只需要设定 $D = k_B T/\gamma$ 就是爱因斯坦关系。

9.5 扩散理论的简单生物学应用

我们希望计算单位时间内信号分子结合到受体的数目。假设细胞半径为 a, 球面上共有 M 个受体,再假设信号分子浓度为球对称分布 $[c(\mathbf{r})=c(r)]$,而且 $c_0=c(\infty)$ 。考虑三维的扩散方程的稳态解实际上就是一个拉普拉斯方程:

$$D\nabla^2 c = 0 \Rightarrow \nabla^2 c = 0$$

化简:

$$\frac{1}{r^2}\frac{d}{dr}(r^2\frac{dc}{dr}) = 0 \Rightarrow r^2\frac{dc}{dr} = A$$

再代入边界条件可以得到: $c(r) = c_0(1 - \frac{a}{r})$ 。再利用菲克定律求出流密度:

$$j(a) = -D\frac{dc}{dr}\Big|_a = -\frac{Dc_0}{a}$$

最后乘上球面的面积即可得到总的粒子数为:

$$\left|rac{dn}{dt}=-j(a)4\pi a^2=4\pi Dac_0
ight|$$

但是我们没有考虑到如果受体不够,来不及和大量的配体结合呢?我们假设细胞对信号分子的结合速率 k_{on} 有限。

$$\frac{dn}{dt} = Mk_{on}c(a)$$

为了保证稳态,必然有扩散流等于结合流:

$$-j(r)4\pi r^2 = D\frac{dc}{dr}4\pi e^2 = Mk_{on}c(a)$$

积分并且利用边界条件得到:

$$c(a) = \frac{c_0}{1 + (Mk_{on}/4\pi Da)}$$

考虑极限条件,如果受体俘获信号分子速率远远大于扩散过程带来的信号分子速率那表面上浓度基本为0,如果完全赶不上那就为 c_0 ,十分直观。这个例子告诉我们太多的受体修饰细胞表面并不能提高表面对配体的结合能力。我们考虑如果结合速率是扩散速率的极值的一半:

$$\frac{dn}{dt} = 2\pi Dac_0 = Mk_{on}c_a$$

可以得到:

$$M = \frac{4\pi Da}{k_{on}}$$

9.6 扩散限速化学反应的普适速度

$$k_{\text{diffusive}} = \frac{8k_BT}{3\eta}$$

10 Chap14 无序与拥挤的环境

10.1 拥挤与结合

我们将反应域分割成 Ω 个体积为 ν 的单元,格点可以背溶剂分子、配体或者拥挤试剂占据。设其中有 L 个配体 C 个拥挤试剂。考虑溶液中只有配体和拥挤试剂存在的情况

$$Z_{sol}(L,C) = C_{\Omega}^{L} C_{\Omega-L}^{C} e^{-\beta L \epsilon_{L}^{sol}} e^{-\beta C \epsilon_{C}^{sol}} = \frac{\Omega!}{L!C!(\Omega - L - C)!} e^{-\beta L \epsilon_{L}^{sol}} e^{-\beta C \epsilon_{C}^{sol}}$$

由于受体只有两种情况,要么与配体结合要么没有结合,因此配体与蛋白质相结合的概率为:

$$p_{bound} = \frac{Z_{sol}(L-1,C)e^{-\beta\epsilon_L^b}}{Z_{sol}(L-1,C)e^{-\beta\epsilon_L^b} + Z_{sol}(L,C)}$$

我们假设 $\Omega - L - C \gg 1$ 因此可以简化得到:

$$p_{bound} = \frac{1}{1 + \frac{\Omega - L - C}{L} e^{\beta \Delta \epsilon_L}}$$

上式中 $\Delta \epsilon_L \equiv \epsilon_L^b - \epsilon_L^{sol}$ 是配体-受体结合与配体在溶液中的能量差。

我们可以画图来观察一下结合概率的曲线,我们会发现其实只有当 $C \sim \Omega$ 时,拥挤分子才会对配体与受体的结合起到显著的影响。因为 C 的增多导致了配体无处可去,只能和受体结合。

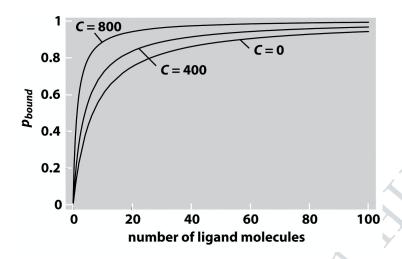


图 2: 不同拥挤分子浓度下,配体分子结合到蛋白质结合位点的概率

10.2 解离常数与拥挤的关系

实验告诉我们随着拥挤分子的增加,解离常数不断减小,结合反应更加容易发生。我们采用大配体的格子模型来分析:一共有 Ω 个大盒子,每个大盒子由r个小盒子构成,一个小盒子只能容纳一个聚乙二醇分子。考虑L个蛋白,C个拥挤分子组成的体系的配分函数

$$Z_{sol}(L,C) = C_{\Omega}^L C_{r\Omega-rL}^C e^{-\beta L \epsilon_L^{sol}} e^{-\beta C \epsilon_C^{sol}} = \frac{\Omega!}{L!(\Omega-L)!} \frac{(r\Omega-rL)!}{C!(r\Omega-rL-C!)} e^{-\beta L \epsilon_L^{sol}} e^{-\beta C \epsilon_C^{sol}}$$

同样有结合的概率为:

$$p_{bound} = \frac{1}{1 + \frac{Z_{sol}(L,C)}{Z_{sol}(L-1,C)}} e^{\beta \Delta \epsilon_L} = \frac{1}{1 + \frac{\Omega}{L} (1 - \phi_C)^r e^{\beta \Delta \epsilon}}$$

其中 $\phi_c = C/r\Omega$ 是拥挤试剂在溶剂中的体积分数,我们再次假设 $L \ll \Omega$ 根据质量作用定律给出解离常数:

$$L + R \rightleftharpoons LR$$
$$K_d = \frac{[L][R]}{[LR]}$$

显然受体被配体占据的概率为:

$$p_{bound} = \frac{[LR]}{[R] + [LR]} = \frac{[L]/K_d}{1 + [L]/K_d}$$

可以得到:

$$\frac{K_d(\phi_c)}{K_d(\phi_c = 0)} = (1 - \phi_c)^r$$

但是根据我们的理论配体和拥挤分子质量猜测 r 约 15, 而最佳拟合为 45。理论一定有缺陷。我们没有考虑到排空力! 粒子所占据的有限空间比它们的物理体积大。

10.3 拥挤溶液中的渗透压

对于血红蛋白溶液,填满的格子代表蛋白,空格子代表溶剂分子,这个模型的配分函数为:

$$Z_{sol}(H,\Omega) = \frac{\Omega!}{H!(\Omega-H)!} e^{-\beta H \epsilon_H^{sol}}$$

渗透压可以通过体系的总体积减少单个盒子的体积 ν 导致体系自由能改变来计算,书本中用 G 代表自由能, 但实际上应该是 F。

$$pv = F(\Omega - 1) - F(\Omega) = -k_B T \ln \frac{Z_{sol}(H, \Omega - 1)}{Z_{sol}(H, \Omega)}$$

于是可以得到:

$$p = \frac{k_B T}{\nu} \ln \frac{\Omega}{\Omega - H} = -\frac{k_B T}{\nu} \ln(1 - [H]\nu)$$

** 排空力 ** 10.4

我在这里采用 F 而不用书本里的 G。由于排斥体积引发的自由能改变为:

$$F_{ex} = -Nk_BT \ln \frac{V_{box} - V_{ex}}{\nu} + Nk_BT \ln \frac{V_{box}}{\nu}$$

假定排斥体积远远小于整个盒子的体积就有:

$$F_{ex} = Nk_B T \frac{V_{ex}}{V_{hox}}$$

我们可以定义 $Nk_BT/Vbox$ 为盒子中小分子理想气体的渗透压强,那么排斥体积导致的自由能改变等于压强 在小分子占据体积减少 V_{ex} 时对小分子气体做的功。这里不去推导排空力的具体表达式了。具体见书本P496, 之后的一些内容也不在这里赘述。

10.5 拥挤动力学

底物转化成产物的反应速率依赖于酶的浓度、底物的浓度及酶的固有的翻转速率。分子伴侣可以帮助蛋白质 进行折叠。

10.5.1 拥挤环境中的扩散

这里有一点不一样,我们假设粒子向左或者向右跳的概率为 $\frac{1}{2}(1-\phi)$,而留在原地的概率为 ϕ 于是可以计算 得到:

$$\langle x^{2}(\tau) \rangle = a^{2} \times p_{right} + a^{2} \times p_{left} + 0 \times p_{stay} = a^{2}(1 - \phi)$$
$$\langle x^{2}(t) \rangle = \frac{t}{\tau} \langle x^{2}(\tau) \rangle = \frac{t}{\tau} a^{2}(1 - \phi)$$
$$D = D_{0}(1 - \phi) \qquad D_{0} = a^{2}/2\tau$$

$$D = D_0(1 - \phi) \qquad D_0 = a^2 / 2\tau$$

速率方程与细胞内的动力学 11

11.1 针对轨迹的统计力学

我们将时间 t 离散化为 N 个时间段, 分子在前 N 段保持基态, N+1 段突然发生转变, 这一过程的概率为

$$p(t)\Delta t = (1 - k\Delta t) \times (1 - k\Delta t) \times ... \times (1 - k\Delta t) \times k\Delta t$$

注意了 p(t) 为概率密度, $p(t)\Delta t$ 才是 $t \sim t + \Delta t$ 内发生的概率。

$$p(t)\Delta t = (1 - k\Delta t)^N k\Delta t = (1 - \frac{kt}{N})^N \approx ke^{-kt}\Delta t = \frac{1}{\tau}e^{-t/\tau}\Delta t$$

11.2 衰变的过程可以用两条轨迹来描述

我们要求的是使得香农熵达到最大值的分子轨迹所对应的概率密度分布函数 p(t)。

$$S = -\int_0^\infty p(t) \ln p(t) dt - \lambda \left(\int_0^\infty p(t) dt - 1 \right) - \mu \left(\int_0^\infty t p(t) dt - \tau \right)$$

令泛函导数为 0, 可以得到:

$$-1 - \ln p(t) - \lambda - \mu t = 0$$

这个式子看起来就像直接求导,但是其实不是,不过我也不想深究。寿命为 t 的轨迹的概率为:

$$p(t)\Delta t = e^{-1-\lambda-\mu t}\Delta t = \frac{1}{\tau}e^{-t/\tau}\Delta t$$

11.3 一种物质的衰变对应另一种物质的生成

略

11.4 快速平衡

$$A \rightleftharpoons B \rightarrow C$$

物质转化的速率方程为:

$$\begin{cases} \frac{dA}{dt} = -k_{+}A + k_{-}B \\ \frac{dB}{dt} = k_{+}A - (k_{-} + r)B \\ \frac{dC}{dt} = rB \end{cases}$$

我们做一下变换记 $k=k_+/k_-, \tau=k_-t, \epsilon=r/k_-$ 那么就有

$$\begin{cases} \frac{dA}{dt} = -kA + B \\ \frac{dB}{dt} = kA - (1 + \epsilon)B \\ \frac{dC}{dt} = \epsilon B \end{cases}$$

前两个方程与第三个没有关系,可以不管第三个先求解前两个我们假设最初只有 A, 也就是 (A(0), B(0)) = (1,0)。用矩阵的形式表示方程为:

$$\frac{d}{d\tau} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} -k & 1 \\ k & -(1+\epsilon) \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix}$$

涌解为:

$$\begin{bmatrix} A \\ B \end{bmatrix}(\tau) = a_1 \begin{bmatrix} A_1 \\ B_1 \end{bmatrix} e^{\omega_1 \tau} + a_2 \begin{bmatrix} A_2 \\ B_2 \end{bmatrix} e^{\omega_2 \tau}$$

其中的参数根据求解本征值方程计算本征矢得到,具体过程见书本 P527。具体的计算很繁杂。最后得到的结论是 A 和 B 虽然都缓慢衰减,但是它们的相对浓度保持不变,B(t)/A(t)=k

11.5 米-曼酶动力学

分析生化动力学最有效的工具之一是米-曼动力学模型。任何一种酶可以近似用两个参数来刻画:最大速率和米氏常数。最大速率:底物浓度饱和时酶能达到的最大翻转率(turn over rate),米氏常数指反应速率为最大速率一半时的底物浓度。

具体内容见书本 P531。过程和之前的分析十分类似,只是稍微复杂一些。