



DNA-RNA Hybrid Chain

DNA-RNA 杂合双链 (DRH Chain) 的物理性质与行为学习笔记

By 165830
(2025.10-2025.12)

Sun Yat-sen University
School of Physics

huangyd59@mail2.sysu.edu.cn

目 录

| | |
|--|-----------|
| 1 DRH 在溶液体系中受力拉伸的物理过程 | 2 |
| 1.1 全过程概述 | 2 |
| 1.2 如何从实验看出 DRH 的相变以及 DRH 相变的物理图景 | 2 |
| 1.3 双螺旋相 | 3 |
| 1.4 拉伸过渡相 | 3 |
| 1.5 过拉伸相以及单链相 | 4 |
| 1.5.1 过拉伸相存在的实验证据 | 4 |
| 1.5.2 先经历 S-DRH, 再进行剥离 | 4 |
| 1.5.3 直接进行剥离过程 | 4 |
| 1.5.4 单链相的两种分类 | 4 |
| 2 DRH 链的力学唯象模型 | 5 |
| 2.1 近自由态 dhDRH 的 WLC 模型 | 5 |
| 2.2 拉伸过渡相与过拉伸相的自由能函数近似 | 6 |
| 2.3 单链 DNA 自由能的唯象模型 | 7 |
| 3 碱基相关链自由能 (Sequence Dependence Free Energy) 模型 | 7 |
| 3.1 脱氧核糖碱基对解离能的最近邻 (Nearest Neighbor) 模型 | 7 |
| 3.1.1 碱基解配对过程的热力学函数关系 | 7 |
| 3.1.2 最近邻碱基组的对称性 | 9 |
| 3.1.3 单个最近邻碱基组结合能的数值拟合以及修正项考虑 | 9 |
| 3.1.4 Santa-Lucia 模型对 DRH 链与 DNA 链的不同 | 10 |
| 3.2 最近邻模型与 DRH 链自由能模型刻画平行单链态 | 11 |
| 3.3 基于碱基相关自由能的剥离态演化动力学 | 12 |
| 3.3.1 理论模型 | 12 |
| 3.3.2 实验结果分析 | 14 |
| 4 复合 (Reannealing) 过程与滞回现象 | 14 |
| 4.1 滞回现象的物理图景 | 14 |
| 4.2 滞回现象的数学模型 | 15 |
| 4.2.1 简单滞回模型 (SHM) | 15 |
| 4.2.2 SDH 模型的微分代数分析 | 17 |
| 4.2.3 SDH 模型的数值模拟结果 | 19 |
| 4.3 DRH 复合过程中的滞回现象 | 21 |
| 4.3.1 影响 DRH 复合产生滞回的因素 | 21 |
| 4.3.2 复合过程中力平台的性质 | 21 |
| A SDH 数值模拟代码 | 22 |

若无特别说明，则统一取 $k_B T = 1$ ，未标注符号统一取正

1 DRH 在溶液体系中受力拉伸的物理过程

1.1 全过程概述

对于一条链是 DNA、一条链是 RNA、相互杂合的一对双链，称之为 **DNA-RNA 杂合双链 (DNA-RNA Hybrid Chain, DRH)**。

在受到拉力时，DRH 将表现出一系列相变，以响应拉力变化。主要包括以下过程：

- 双螺旋相 (**Double Helix**) 受力在 65pN 以下，DRH 整体处在近自由态，总长度几乎无变化
- 拉伸过渡相 (**Transition**) 受力一旦超过 65pN，此时 DNA 将不再能维持原有长度保持刚性，发生拉伸并且在某个力阈值处向过拉伸相转变
- 过拉伸相 (**Overstretched**) 经过拉伸过渡相后达到一个较为平稳的位置，可以保持 10min 以上的稳定态 [1]
- 单链相 (**Single Stranded**) 当受力过大时，DRH 难以维持稳定的双链形态，氢键被打破，DNA 和 RNA 以单链形式存在于溶液体系中。单链相分为剥离态 (**Unpeeled**) 和平行单链态 (**Parallel-separated Stranded**)

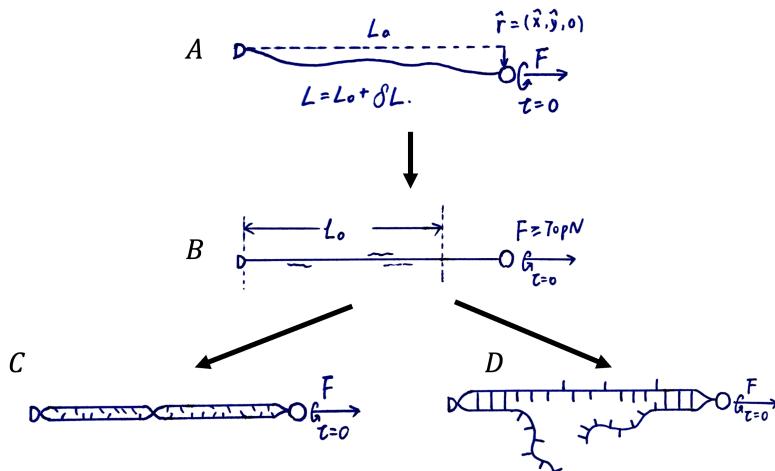


图 1: DRH 拉伸全过程：A. 双螺旋相；B. 过拉伸相；C. 平行单链态；D. 剥离态

1.2 如何从实验看出 DRH 的相变以及 DRH 相变的物理图景

众所周知，在相变过程中往往伴随潜热 Λ 。传统热力学体现在吸放热，但对于力拉伸体系，这个能量主要由拉力来提供。

在实验中观察到， $x - f$ 图像会出现 x 几乎不变而 f 增加的情况，称为**长度平台 (Length Plateau)**；同时也会出现 f 几乎不变而 x 增大的情况，称为**力平台 (Force Plateau)**[4]，如图 (2)。这两个过程分别对应 DRH 的可逆弹性拉伸和 DRH 螺旋结构的不可逆改性。通过这两个实验现象的出现我们就可以判断出现相变。

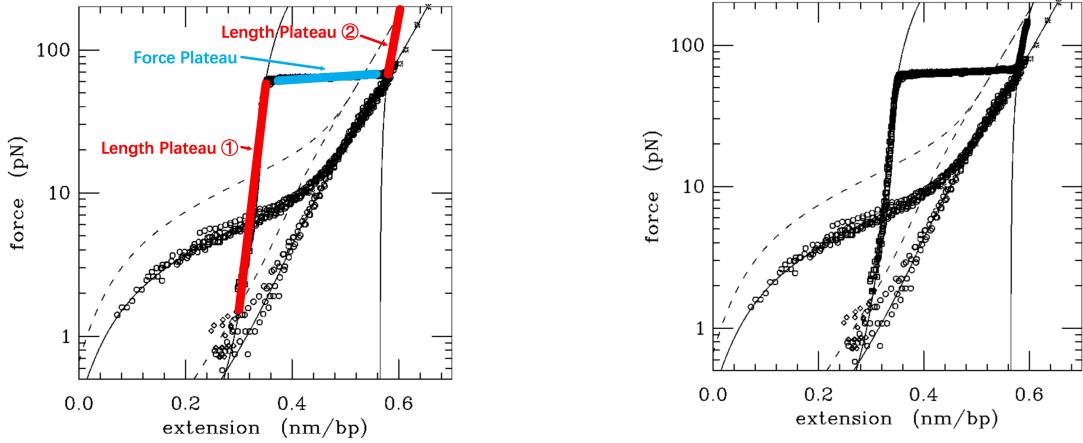


图 2: 力平台和长度平台的实验现象, 有修改 (figure from [3])

对于一个相变过程, 首先需要跨越相变能垒。随后自发吸热进行相变。借用 Van de Waals 方程中的概念, 也即先从失稳分解 (Decomposition) 区跨越到成核长大 (Nucleation) 区, 再自发进行成核长大。长度平台区对应的就是积累弹性势能、逼近成核长大的能垒; 而力平台则对应的是相变潜热 Λ 。

长度平台区对应的是能垒:

$$\Delta E_{\text{barrier}} = \int x(f) df \approx X_{\text{plateau}}(f_1 - f_0) \quad (1)$$

而力平台区对应的是相变潜热:

$$\Lambda = \int f dx \approx F_{\text{plateau}}(x_1 - x_0) \quad (2)$$

全过程的等效 Clapeyron 方程:

$$\frac{dF}{dT} = \frac{\Delta S}{\Delta x} \quad (3)$$

其中 ΔS 为熵变, 可以通过链熵给出近似。从而给出体系关于温度 T 的响应关系。这里的相变潜热是相变非可逆的重要条件, 从潜热的耗散出发, 可以很好地理解 DRH 在松弛弛豫过程中的滞回 (Hysteresis) 现象。

1.3 双螺旋相

这个状态下 DRH 整体比较松弛, 存在垂直于拉力方向 \hat{x} 的振动, 其方向对应为 \hat{y}, \hat{z} 。同时 dhDRH 存在不可忽略的弯曲 (Bend), 由此造成弯曲势能。与此同时由于一定的受力, 在拉伸方向也会存在形变, 因此需要考虑 DRH 的杨氏模量。

1.4 拉伸过渡相

在拉伸试验中, 受力达到 65pN 之后, DRH 的横向振动基本消失, DRH 双螺旋结构作为主体抵抗拉力。经过一个长度平台后, 一旦拉力超过某个阈值 (60pN-80pN)[2], DRH 将迅速进入力平台, 伸长到原长的约 1.7 倍, 并且维持稳定。这个力阈值与溶液离子浓度有关。长度趋近稳定是达到过拉伸相的标志 [1]。

1.5 过拉伸相以及单链相

1.5.1 过拉伸相存在的实验证据

从直观看，DRH 在经受巨大拉力后最终将进入双链解离的单链相，在实际的拉伸过程出现了两次力平台，对应两个相变点。因此可以确定在过拉伸相——即第一个相变区——对应的并不是完全解链的单链相，而是确实存在一个中间的拉伸相。

但是不同的实验给出了对这个状态不同的描述，即 **S-DRH 转变和剥离 (Unpeeling)** 转变。说明这个过程比较复杂，涉及到多种不同的变化与复杂的对应条件 [6]。

1.5.2 先经历 S-DRH，再进行剥离

C.Storm 与 P.C.Nelson 经过实验说明了在拉伸力范围 65pN-120pN 时会出现一个弯曲特征长度约为 10nm 的 DRH 形态，其较 ssDNA 形态仅为 0.7nm 的弯曲特征长度而言具有显著更强的刚性 [7]。无独有偶，J.F.Léger 通过分析力拉伸曲线也分析出在 65pN-140pN 拉力范围内 DRH 转变为和 ssDNA 不同的形态 [1]。除了力的范围，这两个实验都是在生理盐浓度 ($\approx 150 \text{ mM}$) 溶液环境下进行的；且 DRH 链都是完整连接的。进一步的观察发现这个结构具有完整的周期性二级结构，这与杂乱的单链相不符合。具体表现为稳定时中 DRH 螺旋螺距为 $\ell \approx 35 \text{ bp}$ ，对比正常状态下 DRH 的螺距为 $\ell_0 \approx 10 \text{ bp}$ 。这样的螺旋结构暗示了 **S-DRH** 的形成并没有经历大范围的碱基解配对。

1.5.3 直接进行剥离过程

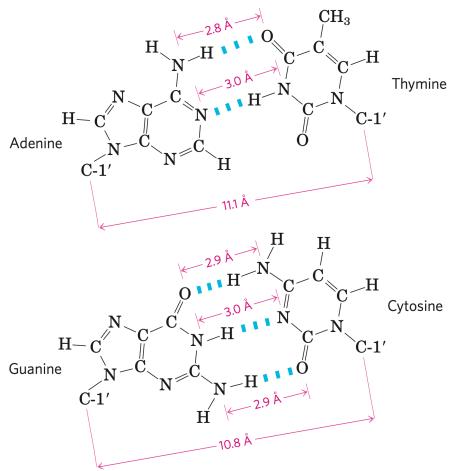


图 3: AT 配对与 GC 配对 (reproduced from [11])

S.Cocco 在低盐浓度 ($\leq 25 \text{ mM}$) 环境下发现对于 **AT 碱基对丰富序列 (AT-Rich Sequence)** 进行拉伸，DRH 没有在 S-DRH 阶段停留，而是直接进入了单链相 [12]。同时 S.Cocco 注意到观察到显著而稳定的 S-DRH 的实验所用的 DRH 链时 GC 碱基对丰富序列。众所周知 GC 配对包含了三个氢键而 AT 只有两个 (见图 (3))，从而 AT 碱基对结合能较弱，更容易发生解配对，引发张力熔解。

1.5.4 单链相的两种分类

剥离过程最终将导致 DRH 转化为单链相，但是单链相有两种不同的状态，剥离态和平行单链态。

J.R.Wenner 实验中发现一种两条平行单链同时承担拉力的状态，也即平行单链态。这种情况往往对应 DRH 结构完整，能够均等地承受拉力。

S.B.Smith 则通过另外一种情况说明了直接进入剥离过程的情况，即当 DRH 中一个链出现缺口 (Nick) 时，剥离过程可以轻易地在缺口处发生，并向 3' 端与 5' 端同时延伸。此时仅有一条链承受拉力。

2 DRH 链的力学唯象模型

对于 DRH，实验室研究其链的力学性质的方式是通过固定其中一端，使用**磁镊 (Magenet Tweezer, MT)** 拉伸其另外一端，从而得到其特征物理量和**拉伸力 (Stretching Force)** 或者**扭矩 (Torque)** 的关系。因此热力学模型中最重要的单位长度自由能 w 也会写为关于 f 的函数：

$$w(f) = \int x(f) df \quad (4)$$

由于拉伸量 x 是可测量量，因此这样一个可积体系可以容易的拟合出自由能的数值形式。

2.1 近自由态 dhDRH 的 WLC 模型

此态的 DRH 处于双螺旋相，其在拉伸中可能具有的能量包括自身弯曲的弯曲势能、受磁镊拉伸导致的拉伸势能以及拉力注入的能量，一般称这种模型为**蠕虫状链模型 (Worm-Like-Chain, WLC)**。至于 DNA 超螺旋导致的扭转势能在低力矩状态下近似不变，不在考虑范围之内。这里仍然正常运算 $k_B T$ 以保证清晰。

采用空间矢量 $\mathbf{r}(s)$ 来描述链的空间形态， s 为随体长度。则曲率写为 $\kappa = \left| \frac{\partial^2 \mathbf{r}}{\partial s^2} \right| = |\partial_s^2 \mathbf{r}|$ ，设单位长度刚度 $k_B T A$ ，进而总弯曲势能表为二次型 [4]：

$$\mathcal{E}_{\text{bend}} = \int \frac{k_B T A}{2} \kappa^2 ds = \int \frac{k_B T A}{2} |\partial_s^2 \mathbf{r}|^2 ds \quad (5)$$

考虑外拉伸力输入的能量，同时先不考虑只跟 x 有关、不存在随 s 变化的拉伸势能。设切向单位向量 $\mathbf{t} = \partial_s \mathbf{r}$ 。因此能量 E 写为：

$$\frac{E}{k_B T} = -\frac{fx}{k_B T} + \int \frac{1}{2} \left(A |\partial_s^2 \mathbf{r}|^2 \right) ds = -\frac{fx}{k_B T} + \int \frac{A |\partial_s \mathbf{t}|^2}{2} ds \quad (6)$$

对切向向量正交分解，设 DNA 链主要延伸方向为 x ，得到 $\mathbf{t} = (t_x, t_y, t_z) = \mathbf{t}_\perp + t_x \hat{x}$ 。由于 $|\mathbf{t}| = 1$ ，因此 $t_x = \sqrt{1 - \mathbf{t}_\perp^2} \approx 1 - \frac{\mathbf{t}_\perp^2}{2}$ 。考虑对 L 的分解，并且考虑一个修正项 δL 用于体现 L 本身的受力伸长，有：

$$\frac{x}{L} = \frac{x}{\sqrt{(x + \delta L)^2 + y^2 + z^2}} \approx \frac{x}{\sqrt{x^2 + 2x\delta L + y^2 + z^2}} = t_x + \frac{\delta L}{L} \approx 1 - \frac{\mathbf{t}_\perp^2}{2} + \frac{\delta L}{L} \quad (7)$$

这里忽略了 δL 与 xy 分量的耦合，这是三阶以上的小量。

将式 (7) 和上式代入式 (6)，设 $\frac{1}{k_B T} \equiv \beta$ 并近似 $t_z = \text{const}$ ：

$$\frac{E}{k_B T} = -f\beta L + \int \frac{ds}{2} \left(f\beta \mathbf{t}_\perp^2 + A |\partial_s \mathbf{t}_\perp|^2 \right) \quad (8)$$

为了将这个关于 s 能量泛函解耦成独立自由度运动，采用 Fourier 变换 $\tilde{\mathbf{t}}_\perp(s) \equiv \int e^{-iqs} \tilde{\mathbf{t}}_\perp(q) \frac{dq}{\sqrt{2\pi}}$ ，进而 $\partial_s \tilde{\mathbf{t}}_\perp(s) = -iq\tilde{\mathbf{t}}_\perp(s)$ ，注意 $\mathbf{t}_\perp \cdot \hat{z} = 0$ ：

$$\int \frac{ds}{2} \left(f\beta \mathbf{t}_\perp^2 + A |\partial_s \mathbf{t}_\perp|^2 \right) = \frac{1}{2} \int \frac{dq}{2\pi} (f\beta + Aq^2) \left| \tilde{\mathbf{t}}_\perp(q) \right|^2 \quad (9)$$

在形式上，这等价于一系列谐振子 $\frac{1}{2}k(q)x^2$ 。等效劲度系数为 $k(q) = (f\beta + Aq^2)$ ，自由坐标 x 是

$\tilde{\mathbf{t}}_{\perp}(q)$ 。从统计平均可以得到，对于单个模式 q 的谐振子平均能量有能均分定理：

$$\left\langle \frac{1}{2}(f\beta + Aq^2) |\tilde{\mathbf{t}}_{\perp}(q)|^2 \right\rangle = 2 \cdot \frac{1}{2}k_B T \quad (10)$$

这里的因子 2 是因为 t_{\perp} 方向有 x, y 两个自由度。为了求 $\langle |\tilde{\mathbf{t}}_{\perp}(q)| \rangle$ ，反过来有：

$$\langle |\tilde{\mathbf{t}}_{\perp}(q)| \rangle = \int \frac{dq}{2\pi} \frac{2k_B T}{f\beta + Aq^2} = \frac{1}{\pi\sqrt{Af\beta}} \arctan(\sqrt{\frac{A}{f\beta}}q) \rightarrow \frac{1}{\sqrt{Af\beta}} \quad (11)$$

极限对应的是积分上下限 $(-\infty, +\infty)$ ，这是因为 q 是在 $(-\infty, +\infty)$ 上的模式。

最后加入拉伸修正 $\frac{\delta L}{L} = \frac{f}{Y}$ ， Y 称为力常数 (Force Constant)，也即弹性力学中的杨氏模量。可以得出包含拉伸和横向偏移修正的 $z - f$ 关系：

$$\frac{x}{L} = 1 - \frac{1}{2} \sqrt{\frac{k_B T}{f A}} + \frac{f}{Y} = \frac{x_{\text{dhDRH}}(f)}{L} \quad (12)$$

这便是近自由态的 dhDRH 的力拉伸函数，是最开始、拉力最小时对应的形态。

2.2 拉伸过渡相与过拉伸相的自由能函数近似

实际实验中并不存在式 (1) 的理想关系。通过实验测出恰好达到过渡态阶段的伸长 x_0 以及对应的临界拉力 f_0 ，拉伸过渡相的拉伸函数 $x_{\text{transition}}(f)$ 可以被近似地写成 [10]：

$$x_{\text{transition}} = \frac{f - f_0}{S_1} + x_0 \quad (13)$$

从实验可以看出理想状态中描述的 $L = \text{const}$ 其实只是在说相对过拉伸相而言 L 增长缓慢。相似地，过拉伸相也可以写成：

$$x_S = \frac{f - f_1}{S_2} + x_1 \quad (14)$$

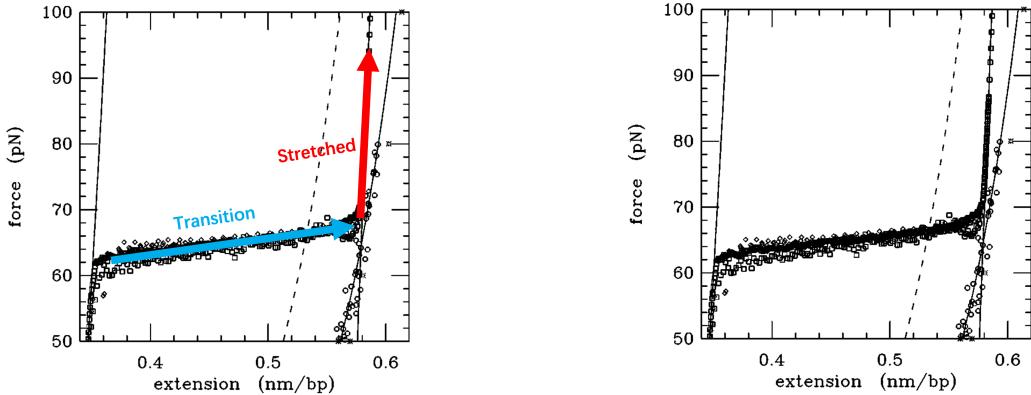


图 4：过渡态和过拉伸相的 $f - x$ 曲线，有修改 (figure from [3])

进而自由能容易被分段积分出来：

$$w_{\text{transition}}(f) = w_{\text{dhDRH}}(f_0) + \frac{1}{2} \left[2x_0(f - f_0) + \frac{(f - f_0)^2}{S_1} \right] \quad (15)$$

$$w_{\text{s-DRH}}(f) = w_{\text{dhDRH}}(f_0) + \frac{1}{2} \left[(x_0 + x_1)(f_1 - f_0) + 2x_1(f - f_1) + \frac{(f - f_1)^2}{S_2} \right] \quad (16)$$

2.3 单链 DNA 自由能的唯象模型

研究表明对于单链 DNA，其 $x - f$ 函数可以写为 [5]：

$$x_{\text{ssDNA}}(f) = L \left(\frac{a_1 \ln(f/A)}{1 + a_3 e^{(-f/B)}} - a_2 - \frac{f}{C} \right) \quad (17)$$

经过对比，这个表达式可以理解为将弹性项 $\frac{f}{C}$ 与固有长度项 a_2 保留，但是剩下一项忽略掉了 \hat{r} 方向的振动，转而考虑了离子静电作用项 $a_1 \ln(f/A)$ 以及本身链的折叠项 $\frac{1}{1 + a_3 e^{(-f/B)}}$ 。

静电项对数的来源是等离子体中的 Debye 屏蔽效

应 $\phi_e = \phi_0 e^{-x/\lambda_D}$ ，其中 $\lambda_D = \sqrt{\frac{\varepsilon_0 \varepsilon_r k_B T}{e^2 N_A I}}$, $I = \sum c_i Z_i$, 为

Debye 屏蔽长度。其中 I 定义式中 Z_i 为离子价态, c_i 为离子浓度。

e 指数分母项来源于 ssDNA 的折叠 (Fold) 效应。设实际长度为 ℓ_{eff} , 原长度 L 的 ssDNA 形成折叠结构, 形成能量为 ϵ , 打开时长度增加 Δ 。从而拉力注入能量表示为 Δf , 这样考虑 Boltzmann 分布有：

$$\ell_{\text{eff}} = \frac{L}{1 + e^{\beta(\epsilon - f\Delta)}} \equiv \frac{1}{1 + a_3 e^{(-f/B)}} \quad (18)$$

上面的效果叠加在一起的到式 (17)。而对于单链 RNA, 也即 DRH 的另外一链, 可以采用式 (12) 的 WLC 模型, 只需要更改系数为 A' 与 Y' [10], 也即：

$$x_{\text{ssRNA}}(f) = L \left(1 - \frac{1}{2} \sqrt{\frac{k_B T}{f A'}} + \frac{f}{Y'} \right) \quad (19)$$

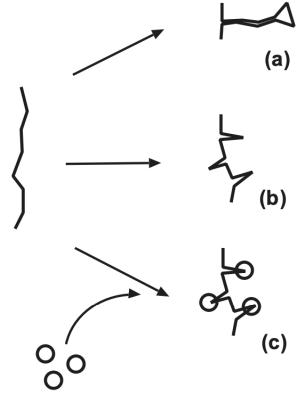


图 5: ssDNA 在溶液体系中的折叠与发卡 (Zipped) 结构 (reproduced from [5])

3 碱基相关链自由能 (Sequence Dependence Free Energy) 模型

上面的研究都集中在 DRH 链的力学特性，但是从双链态的 S-DRH 或者拉伸过渡相的 DRH 转变为单链态必然涉及碱基解配对，S.Cocco 也发现剥离现象的出现与碱基对有关 [12]。因此需要研究碱基对的独立结合能。

3.1 脱氧核糖碱基对解离能的最近邻 (Nearest Neighbor) 模型

3.1.1 碱基解配对过程的热力学函数关系

SantaLucia 通过研究碱基配对与解配对过程的热力学函数关系，得出碱基所具有的结合能 [13]。虽然这个研究最早是针对 DNA 的，但是对于 RNA-DNA 杂合，碱基互补配对过程几乎是完全相同的，因此完全可以迁移。

从化学反应的角度出发，考虑到 DNA 的反义对称性导致的自互补 (Self-complementary)(也就是常说的回文序列) 同时忽略自互补序列自己的对折，DNA 的随机螺旋单链 (Random Coil) 和双链

(Duplex), 之间的转化本质上就是两种可逆反应:



化学平衡原理指出, Gibbs 自由能 ΔG 与标准状态 Gibbs 自由能 ΔG° 、浓度商 Q 的关系为:

$$\Delta G = \Delta G^\circ + RT \ln Q \quad (22)$$

平衡态下 $\Delta G = 0$, $Q = K$, 进而可以得出:

$$\Delta G^\circ = -RT \ln K \quad (23)$$

带入 Gibbs 自由能关系 $\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$, 移向得出平衡常数 K 的关系, 这也是 van't Hoff 方程的积分形式:

$$\ln K = -\frac{\Delta H^\circ}{RT} + \frac{\Delta S^\circ}{R} \quad (24)$$

如果能求出 ΔS° 、 ΔH° , 对于可以轻易测出的 T , 就可以算出 ΔG° 以作为 DNA 脱离反应的结合能。规定转化率 $\alpha = 0.5$ 为平衡状态, 认定两个反应加入的总 DNA 浓度为 C_T 并通过紫外分光光度仪进行测定, 且假定第一个反应 $C_A = C_B$, 进而得到:

$$K_{AB} = \frac{\alpha(C_T/2)}{(C_A/2) \cdot (C_B/2)} = \frac{C_T/4}{(C_T/4) \cdot (C_T/4)} = \frac{4}{C_T} \quad (25)$$

$$K_{AA} = \frac{\alpha(C_T/2)}{((1-\alpha)C_T)^2} = \frac{1}{C_T} \quad (26)$$

带入后将 C_T 作为自变量, T_M 为 $\alpha = 0.5$ 时的温度作为 DNA 链熔解的特征温度。得到关系:

$$\frac{1}{T_M} = \frac{1}{\Delta H^\circ} (\Delta S^\circ - R \ln K) \quad (27)$$

$$= \frac{1}{\Delta H^\circ} (\Delta S^\circ + R \ln C_T - (1 - \delta_{0,C_B}) R \ln 4) \quad (28)$$

其中 $1 - \delta_{0,C_B}$ 为指征回文序列存在性的 Dirac delta 函数。这对应了式 (25) 中多的系数 4。

如此一来, 对 $\frac{1}{T_M}$ 和 $\ln C_T$ (或者 $\ln C_T/4$) 进行拟合, 如图 (6)。斜率 k 和截距 b 分别对应:

$$\Delta H^\circ = \frac{R}{k}, \Delta S^\circ = \frac{bR}{k} \quad (29)$$

但是这个是针对某条 DNA 链的解离过程, 因此算出的 ΔG° 也是针对在整条链所有碱基自身以及其近邻作用之下所表现出的总结和能。可将总结能写为不同碱基的结合能的线性组合, 也即:

$$\Delta^i G_{\text{tot}} = \sum_j n_{ij} \Delta G_j + C_i + \Delta^i G_{\text{sym}} \quad (30)$$

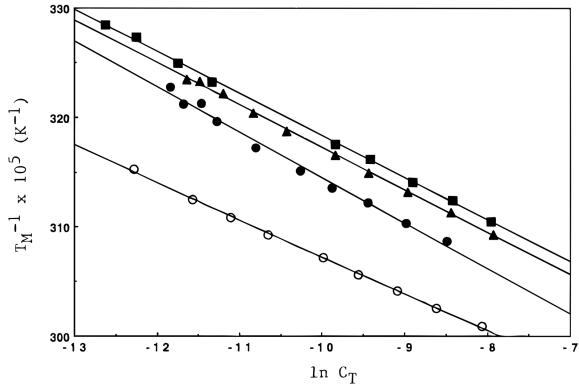


图 6: 序列 GTTGCAAC(\bullet), GTACGTAC(\blacktriangle), GGACGTCC(\circ) 与 CGATATCG(\blacksquare) 的 T_M^{-1} - $\ln C_T$ 拟合 (reproduced from [13])

其中指标 i 是指一条整体 DNA 链的序号 (这些序列长度相同), 指标 j 代表整条链中第 j 种最近邻碱基组 (Nearest pair) (如图(7)), n_{ij} 代表第 i 个链中第 j 种最近邻碱基组的个数。 $C_i \approx 1.8 - 2.8 \text{ cal} \cdot \text{mol}^{-1}$, 代表启动解链过程的起始自由能 [12], 类似于静摩擦力向滑动摩擦力转变中需要的额外边界能量。G-C 起始时较大, A-T 起始时最小。 $\Delta^i G_{\text{sym}} \approx 0.4 \text{ cal} \cdot \text{mol}^{-1}$, 为自互补链的修正自由能项 [13]。引入这个修正的原因是自互补序列同时也存在单体互补配对的情况 (如图 (8.B)), 造成自互补序列更容易解链。这个情况在式 (28) 中被忽略了。

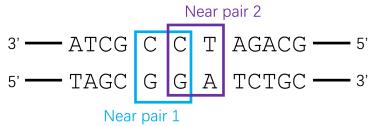


图 7: 近邻碱基组的取法

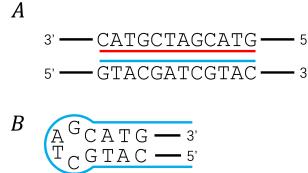


图 8: A. 自互补链两体配对, B. 自互补链
单体配对

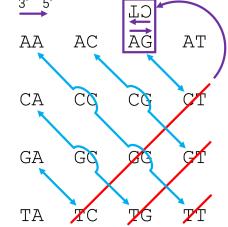


图 9: 对称的近邻碱基对

3.1.2 最近邻碱基组的对称性

因为最近邻碱基组的取法是有限的, 且因为双链 DNA 具有对称性, 任取一对相邻组等价于取与其互补的相邻组。线性代数的角度来看, 将碱基指标化为: $A = 1, C = 2, G = 3, T = 4$, 进而以 $a_{ij} = (i, j)$ 代表近邻碱基对, 则:

$$a_{ij} = (i, j) = (5 - j, 5 - i) = a_{5-j, 5-i} \quad (31)$$

注意 DNA 单链是具有方向性的, 所以需要在指标相减后交换。上文的互补也指的是考虑 DNA 链方向的反向互补而不是只看碱基顺向互补。最后得出最近邻总数为 $4^2 - 6 = 10$, 如图 (9)。当然如果考虑的是 DRH, 那么反过来的链即使也是互补的, 也是与模板链不同的分子, 自然就没有这个对称性了。

3.1.3 单个最近邻碱基组合能的数值拟合以及修正项考虑

从统计学来看, 对于 n 个碱基的 DNA 链, 其共有 4^n 种。但是待求的最近邻碱基组仅有 10 个, 因此求解会变成超定方程组。考虑取 $m < 4^n$ 种长度为 n 的链, 方程组形式为:

$$\begin{bmatrix} \Delta^1 G_{\text{tot}} - \Delta^1 G_{\text{sym}} - C_1 \\ \Delta^2 G_{\text{tot}} - \Delta^2 G_{\text{sym}} - C_2 \\ \vdots \\ \Delta^m G_{\text{tot}} - \Delta^m G_{\text{sym}} - C_m \end{bmatrix} = \begin{bmatrix} n_{1,1} & n_{1,2} & \cdots & n_{1,10} \\ n_{2,1} & n_{2,2} & \cdots & n_{2,10} \\ \vdots & \vdots & \ddots & \vdots \\ n_{m,1} & n_{m,2} & \cdots & n_{m,10} \end{bmatrix} \begin{bmatrix} \Delta G_1 \\ \Delta G_2 \\ \vdots \\ \Delta G_{10} \end{bmatrix} \quad (32)$$

令 $\Delta \mathbf{G}_{\text{eff}} \equiv \Delta \mathbf{G}_{\text{tot}} - \Delta \mathbf{G}_{\text{sym}} - \mathbf{C}$, 化为向量-张量积形式, 便于书写最小二乘解:

$$\Delta \mathbf{G}_{\text{eff}} = \mathbf{N}_{m \times 10} \cdot \Delta \mathbf{G} \quad (33)$$

从而残差为 [13]:

$$\chi^2 = \sum_{i,j} \left| \frac{\Delta^i G_{\text{tot}} - N_{ij} \Delta^j G}{\sigma_i} \right|^2 \quad (34)$$

χ^2 取最小，对应最小二乘解，正是对应正规方程的解。这个方程是满秩的：

$$\mathbf{N}^T \Delta \mathbf{G}_{\text{eff}} = \mathbf{N}^T \mathbf{N} \cdot \Delta \mathbf{G} \quad (35)$$

最后得到的结果如图：

| propagation sequence | ΔH° (kcal/mol) | ΔS° (eu) | ΔG°_{37} (kcal/mol) |
|----------------------------------|--------------------------------|--------------------------|-------------------------------------|
| AA/TT | -8.4 ± 0.7 | -23.6 ± 1.8 | -1.02 ± 0.04 |
| AT/TA | -6.5 ± 0.8 | -18.8 ± 2.3 | -0.73 ± 0.05 |
| TA/AT | -6.3 ± 1.0 | -18.5 ± 2.6 | -0.60 ± 0.05 |
| CA/GT | -7.4 ± 1.1 | -19.3 ± 2.9 | -1.38 ± 0.06 |
| GT/CA | -8.6 ± 0.7 | -23.0 ± 2.0 | -1.43 ± 0.05 |
| CT/GA | -6.1 ± 1.2 | -16.1 ± 3.3 | -1.16 ± 0.07 |
| GA/CT | -7.7 ± 0.7 | -20.3 ± 1.9 | -1.46 ± 0.05 |
| CG/GC | -10.1 ± 0.9 | -25.5 ± 2.3 | -2.09 ± 0.07 |
| GC/C | -11.1 ± 1.0 | -28.4 ± 2.6 | -2.28 ± 0.08 |
| GG/CC | -6.7 ± 0.6 | -15.6 ± 1.5 | -1.77 ± 0.06 |
| initiation at G•C ^b | (0) | (-5.9 ± 0.8) | $+1.82 \pm 0.24$ |
| initiation at A•T ^c | (0) | (-9.0 ± 3.2) | ($+2.8 \pm 1$) |
| symmetry correction ^d | 0 | -1.4 | +0.4 |
| 5'-terminal T•A bp ^e | +0.4 | 0 | +0.4 |

图 10: 1 M NaCl 溶液中不同最近邻碱基对的解离能 (reproduced from [13])

在此方程中考虑到的所有修正如下：

- **首位激发能 (Initiate Energy) C_i** : 当 i 号链以 GC 起始, $C_i \approx 2.8 \text{ cal} \cdot \text{mol}^{-1}$; 当 i 号链以 AT 起始, $C_i \approx 1.8 \text{ cal} \cdot \text{mol}^{-1}$
- 自互补对称能: $\Delta^i G_{\text{sym}} \approx 0.4 \text{ cal} \cdot \text{mol}^{-1}$, 仅当 i 号链为自互补序列时存在
- 自互补反应熵: 因为自互补序列进行两体反应时化学过程本质与非自互补序列不同, 表现在计算 $\ln K$ 时多出一项 $\Delta^i S = -R \ln 4$, 记在 i 链的反应熵上
- 末端自解配 (Fraying) 修正: 对于氢键不够牢靠的 T⁵-A³(即 5' 端为 T, 3' 端为 A) 序列, 末端碱基容易打开, 因此其自由能要补充一个 $\Delta i_{\text{fray}}^G \approx 0.4 \text{ cal} \cdot \text{mol}^{-1}$

3.1.4 Santa-Lucia 模型对 DRH 链与 DNA 链的不同

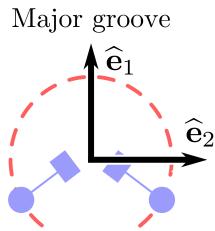


图 11: DRH 中 DNA-RNA 的非对称结合结构

另外, 首位激发能 $C \approx 3.1 \text{ cal} \cdot \text{mol}^{-1}$, 其能量有所提高。同时增加了首位激发焓 (Initiate Enthalpy), 表示为 $\Delta i^H \approx 1.9 \text{ cal} \cdot \text{mol}^{-1}$ 。这是因为 DRH 空间结构上不是完全对称的, 而是存在一定的偏角 (如图 (11))。在解配对时首先要将偏转的某条链拉到对称的位置, 再进行解配对, 从而产生了焓变。

因为两条链化学本质的不同, DRH 链中所有的结合反应都是 $A + B \rightleftharpoons AB$, 因此可以统一取 $\ln C_T / 4$ 作为自变量拟合。同样的, 也不存在自互补序列修正 ΔG_{sym} 。末端自解配能也因为 RNA 的加入而不复存在。同时这也导致所有的等效最近邻碱基对从原来的 10 个变成了完整的 $4^2 = 16$ 个。但是因为互补对称性仍然存在, 因此对于一种 DNA 最近邻碱基对, 对应的互补 RNA 最近邻碱基对与之等价, 故只需要考虑 DNA 的全部 16 种情况。

| Base i and $i+1$ ($5' \rightarrow 3'$) | ΔH (kcal/mol) | ΔS (cal/mol·K) | J_i (unit in $k_B T$) |
|--|-----------------------|------------------------|--------------------------|
| rAA | -7.8 | -21.9 | 1.62 |
| dTT | | | |
| rAC | -5.9 | -12.3 | 3.41 |
| dTG | | | |
| rAG | -9.1 | -23.5 | 2.92 |
| dTC | | | |
| rAU | -8.3 | -23.9 | 1.46 |
| dTA | | | |
| rCA | -9 | -26.1 | 1.46 |
| dGT | | | |
| rCC | -9.3 | -23.2 | 3.41 |
| dGG | | | |
| rCG | -16.3 | -47.1 | 2.76 |
| dGC | | | |
| rCU | -7 | -19.7 | 1.46 |
| dGA | | | |
| rGA | -5.5 | -13.5 | 2.11 |
| dCT | | | |
| rGC | -8 | -17.1 | 4.38 |
| dCG | | | |
| rGG | -12.8 | -31.9 | 4.71 |
| dCC | | | |
| rGU | -7.8 | -21.6 | 1.79 |
| dCA | | | |
| rUA | -7.8 | -23.2 | 0.97 |
| dAT | | | |
| rUC | -8.6 | -22.9 | 2.44 |
| dAG | | | |
| rUG | -10.4 | -28.4 | 2.60 |
| dAC | | | |
| rUU | -11.5 | -36.4 | 0.32 |
| dAA | | | |

图 12: 1 M NaCl、25°C 的最近邻碱基对结合能数据

最终使用类似的方法可以算出 16 种最近邻碱基对能量 [14]，换算到 $k_B T$ 单位，得到如图 (12)

3.2 最近邻模型与 DRH 链自由能模型刻画平行单链态

由于平行单链有两条承担拉力的链，因此分别设对应拉力为 f_{RNA} 、 f_{DNA} 以及总拉力 f ，显然有：

$$f_{\text{RNA}} + f_{\text{DNA}} = f \quad (36)$$

以 i 作为碱基的序号，以 b_i 为指标， $b_i = 1$ 对应双螺旋相， $b_i = 2$ 对应单链相， $b_i = 3$ 对应过拉伸相。按碱基位数的链自由能写为：

$$\mathcal{F}_i(f) = \begin{cases} w_{\text{dhDRH}}(f) & b_i = 1 \\ w_{\text{ssDNA}}(f_{\text{DNA}}) + w_{\text{ssRNA}}(f_{\text{RNA}}) & b_i = 2 \\ w_{\text{S-DNA}}(f) & b_i = 3 \end{cases} \quad (37)$$

如果考虑近邻碱基效应，第 i 位碱基导致的能量可以利用 Kronecker δ 写为 [12]：

$${}^i G_{bp} = C(1 - \delta_{b_i, b_{i+2}}) + J_i \delta_{b_i, 2} \delta_{b_{i+1}, 2} \quad (38)$$

C 为首位激发能， $J_i = \Delta {}^i G^\circ$ ，为近邻碱基的结合能。

$\delta_{b_i, 2} \delta_{b_{i+1}, 2}$ 指标的含义为当且仅当 i 、 $i+1$ 位碱基都是 ssDNA 时才代表这一对近邻碱基被完全打开，于是考虑结合能； $(1 - \delta_{b_i, b_{i+1}})$ 指标的含义为当 i 、 $i+1$ 位碱基状态不同时，说明这部分存在解链过程，因此需要付出一个激发能。加上 DRH 的链自由能，则总能量 G 写为：

$${}^i G = w_{\text{dhDRH}}(f) \delta_{b_i, 1} + (w_{\text{ssDNA}}(f_{\text{DNA}}) + w_{\text{ssRNA}}(f_{\text{RNA}})) \delta_{b_i, 2} + w_{\text{S-DNA}}(f) \delta_{b_i, 3} + {}^i G_{bp} \quad (39)$$

从单个碱基到整条链需要两步计算。首先需要对一种链的从 1 到 $N - 1$ 的所有碱基求和一遍（不求到 N 是因为最近邻模型会在 $N - 1$ 的位置考虑 N ），然后需要对每种 b_i 的取值进行讨论。这一步需要用 Boltzmann 分布将 G 引入配分函数，然后进行分类加法，求和总共 3^N 中情况的配分函数。表示为：

$$\Xi = \sum_{b_1} \sum_{b_2} \cdots \sum_{b_N} \left(\prod_{i=1}^{N-1} \exp^i G \right) \quad (40)$$

考虑到这个求和式描述的是 i 到 $i+1$ 位的关联能量累加，因此容易想到将 i 与 $i+1$ 的 b 指标当作行、列指标，将关联能量的形式写成一个 3×3 矩阵，也即：

$${}^i G = \begin{bmatrix} 1 & e^{-C} & e^{-C} \\ e^{-C} & e^{-J_i} & e^{-C} \\ e^{-C} & e^{-C} & 1 \end{bmatrix} \quad (41)$$

这样一来，考虑将列矢 $[w_{\text{dhDRH}}, w_{\text{ssDNA}} + w_{\text{ssRNA}}, w_{\text{S-DNA}}]^T$ 左乘在这个矩阵，乘在矩阵元 G_{jk} 上的分量就可以理解为对于第 $i+1$ 位的三种情况指标 k ，分别匹配三种不同的自由能函数。而列指标作为第

i 位的情况，用于衔接前一个关联矩阵 ${}^{i-1}G$ 的列指标，从而满足了两个近邻组首尾相接的关系。每一次运算都缩并了中间的指标，本质就是矩阵乘法。于是配分函数可以被写为非常优美的形式：

$$\Xi = [1, 1, 1] \left(\prod_{i=1}^{N-1} \begin{bmatrix} 1 & e^{-C} & e^{-C} \\ e^{-C} & e^{-J_i} & e^{-C} \\ e^{-C} & e^{-C} & 1 \end{bmatrix} \right) \begin{bmatrix} w_{\text{dhDRH}} \\ w_{\text{ssDNA}} + w_{\text{ssRNA}} \\ w_{\text{S-DNA}} \end{bmatrix} \quad (42)$$

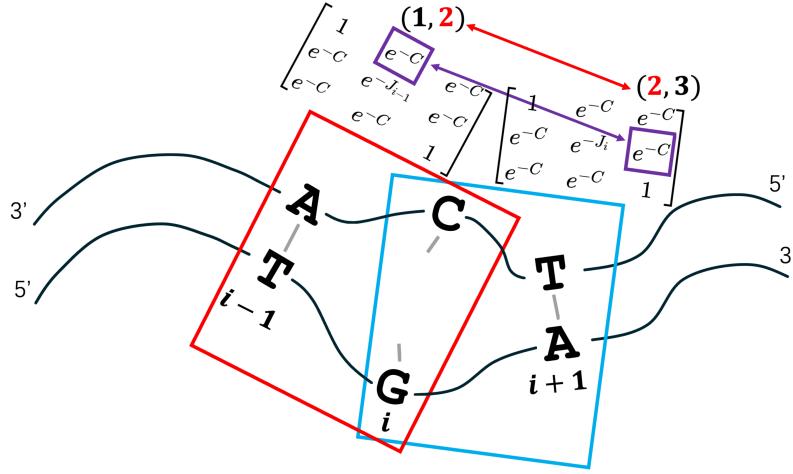


图 13: 传态矩阵的表现 (以 Bubble 产生的情况为例)

这样一来可以给出总解离碱基数 $n = \frac{\partial \ln \Xi}{\partial (w_{\text{ssDNA}} + w_{\text{ssRNA}})}$ 、拉伸长度 $l = k_B T \cdot \frac{\partial \Xi}{\partial f}$

3.3 基于碱基相关自由能的剥离态演化动力学

3.3.1 理论模型

描述碱基的剥离动力学所需要的是自由能的时空演化。而碱基能量的空间分布和碱基位数是完全相关的，时间演化则由不含涨落项的 Langevin 方程决定：

$$\frac{dn}{dt} = \dot{n} = -r \frac{dG}{dn} \quad (43)$$

这个方程可以理解为 $J = kf = k\nabla G$ ，即广义输运方程。 r 代表输运系数， $f = \nabla G$ 代表广义驱动力， J 作为“碱基打开”这个状态的广义流密度。

对于 $G(n)$ 的形式，可以根据碱基相关自由能定义分支函数 $g(n_0, n)$ ，以描述从 n_0 位碱基起始，相对位置为 n (以延伸向 3' 为正，向 5' 为负号， n_0 位记为 0) 的碱基配对能量：

$$g(n_0, n) = \begin{cases} C & n = \pm 1 \\ J_{n_0+n-1} & n \geq 2 \\ J_{n_0+n+1} & n \leq -2 \end{cases} \quad (44)$$

同时要补充的还有链自由能之差。需要注意对于 DRH，剥离的链可能是 DNA 或者 RNA，在选择链自由能函数时需要区分。这里统一记为 $w_{\text{ss}}(f)$ 。与这个能量作差的则是过拉伸相的自由能 $w_{\text{S-DNA}}(f)$ 或

者双螺旋相的自由能 $w_{\text{dhDRH}}(f)$, 取决于拉伸过程是直接从 dhDRH 跃变到 ssDNA 与 ssRNA, 还是要经过 S-DNA。这里统一记为 $w_{\text{ds}}(f)$ 。于是得到从 n_0 开始沿正方向的半侧能量:

$$\Delta G_+ = n_+[w_{\text{ss}}(f) - w_{\text{S-DNA}}(f)] + \sum_{i=1}^{n_+} g(n_0, i) \quad (45)$$

沿负方向的半侧能量:

$$\Delta G_- = n_- [w_{\text{ss}}(f) - w_{\text{S-DNA}}(f)] + \sum_{i=1}^{n_-} g(n_0, -i) \quad (46)$$

加上基准能 $Nw_{\text{S-DNA}}$, 其中 N 为总碱基数。同时定义总解链数 $n = n_+ + n_-$, 从而得到自由能:

$$G_f(n_+, n_-) = \Delta G_- + \Delta G_+ - Nw_{\text{S-DNA}}(f) = \sum_{i=1}^{n_-} g(n_0, -i) + \sum_{i=1}^{n_+} g(n_0, i) - nw_{\text{ss}}(f) - (N - n)w_{\text{ds}}(f) \quad (47)$$

但是这还没有完全描述完。假定拉伸力 f , 从而得到总自由能:

$$G_x = Nx f - nw_{\text{ss}}(f) - (N - n)w_{\text{ds}}(f) + \sum_{i=1}^{n_-} g(n_0, -i) + \sum_{i=1}^{n_+} g(n_0, i) \quad (48)$$

这里总拉伸 x 需要满足几何关系:

$$N \cdot x = nx_{\text{ss}}(f) + (N - n)Nx_{\text{ds}}(f) \quad (49)$$

将其对 n_+ 求导, 同时注意在整个拉伸过程中, 还有拉伸力的能量注入。这个能量的注入过程中正逢相变的力平台, 因此 $x = \text{const}$:

$$\begin{aligned} \frac{dG_x}{dn_+} &= Nx \frac{\partial f}{\partial n_+} - w_{\text{ss}}(f) + w_{\text{ds}}(f) - n \frac{\partial w_{\text{ss}}(f)}{\partial f} \frac{\partial f}{\partial n_+} - (N - n) \frac{\partial w_{\text{ds}}(f)}{\partial f} \frac{\partial f}{\partial n_+} + \\ &\quad \frac{d}{dn_+} \left(\sum_{i=1}^{n_-} g(n_0, -i) + \sum_{i=1}^{n_+} g(n_0, i) \right) \\ &= g(n_0, n_+) - w_{\text{ss}}(f) + w_{\text{ds}}(f) \end{aligned} \quad (50)$$

这里前一部分运用了关系 $\frac{\partial w_{\text{ss}}(f)}{\partial f} = x_{\text{ss}}$ 、 $\frac{\partial w_{\text{ds}}(f)}{\partial f} = x_{\text{ds}}$, 后一部分运用了积分-求和等价性以及等价的变上限积分求导法则:

$$\frac{d}{dn_+} \left(\sum_{i=1}^{n_-} g(n_0, -i) \right) = \frac{d}{dn_+} \left(\int_1^{n_-} g(n_0, x) dx \right) = 0 \quad (51)$$

$$\frac{d}{dn_+} \left(\sum_{i=1}^{n_+} g(n_0, i) \right) = \frac{d}{dn_+} \left(\int_1^{n_+} g(n_0, x) dx \right) = g(n_0, n_+) \quad (52)$$

如果对 n_- 求导, 过程是完全一致的, 只需要将最后一项求导消去的求和项修改即可。

再考虑定常拉伸速度 v , 则有 $dt = \frac{dx}{v}$, 用此消掉式 (43) 中的 dt , 最终得到:

$$\frac{dn_+}{dx} = \frac{r}{v} (g(n_0, n_+) - w_{\text{ss}}(f) + w_{\text{ds}}(f)) \quad (53)$$

$$\frac{dn_-}{dx} = \frac{r}{v} (g(n_0, n_-) - w_{\text{ss}}(f) + w_{\text{ds}}(f)) \quad (54)$$

如果要求数值模拟, 可以将不可积的 $g(n_0, n)$ 跃变函数替换为 Gaussian 分布的近似函数:

$$g_i = \frac{J_i}{\sqrt{2\pi\delta^2}} \exp \frac{-n^2}{2\sigma^2} \quad (55)$$

参数取为: $\delta = 15\text{bp}$

3.3.2 实验结果分析

利用计算模拟，可以得到恒 f 拉伸过程中和恒 x 拉伸过程中不同的自由能情况。如图 (14)、图 (15)。

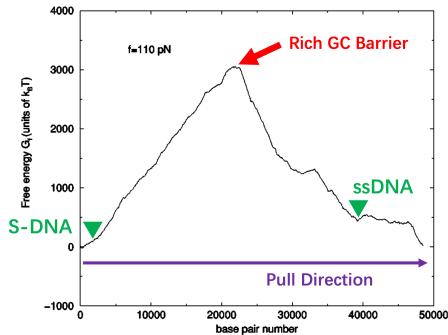


图 14: f 恒定的拉伸 (reproduced from [12])

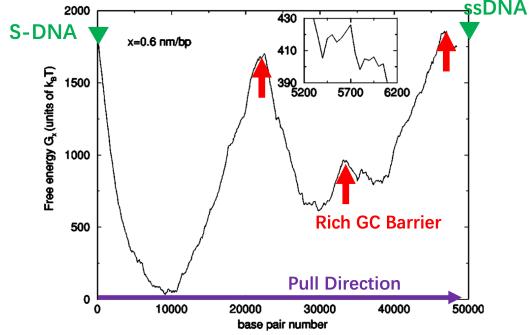


图 15: x 恒定的拉伸 (reproduced from [12])

其中图 (14) 表现了在拉力恒定过程中，因为 GC 碱基对堆叠造成的大能垒，使得从过拉伸态跨越到单链态几乎不可能。但是当固定地调制 $x = 0.6\text{nm}/\text{bp}$ ，这一过程则改变了原有的自由能分布，造成了多处峰值和平衡势阱。在实际的拉伸过程中，这些势垒与势阱造成全过程拉伸难度分布极不均匀，进而产生了粘-滑运动 (“stick-slip” motion)。在粘-滑运动中，产生大量的耗散能量，从而造成了体系的滞回现象。

通过控制序列，不同态的自由能可以通过势能分布来调控自然状态下链的解链过程，进而调控生命活动。

4 复合 (Reannealing) 过程与滞回现象

4.1 滞回现象的物理图景

保守力做功与路径无关，这保证了在固定起终点的过程中可以任意选取路径而不会造成能量变化。因此单纯保守力做功完全不会出现滞回现象。如果是理想的耗散力，例如干摩擦模型 $f = \mu N$ ，则运动参数 x, v, a 对于驱动参量 f, N 的响应是瞬时的，也无法造成滞回效果。但是当过程中出现响应参量和驱动参量的延迟响应，例如响应物理量突然不随驱动物理量变化的死区 (Dead Zone)。

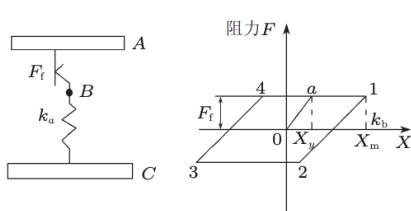


图 16: 延时响应 (reproduced from [16])

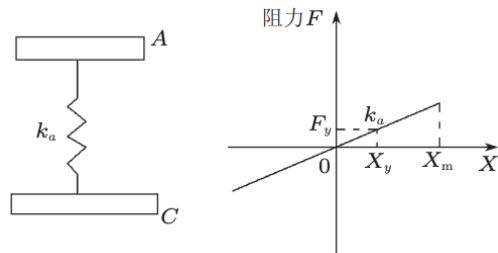


图 17: 即时响应 (reproduced from [16])

延迟响应后，体系响应变量最终在驱动变量达到最大值时未达到对应的极值，从而驱动力转为反向驱动后，响应变量不能在驱动量回归初态的时候同时回归，造成了回复曲线整体的右移 (如图 (16))。到达反向对称点后，则会对正向过程一对称的方式发展 (如左移曲线 (如图 (16)))，从而形成一个闭环，称为

滞回环 (Hysteresis Loop)。在此过程中，能量在死区被耗散。全过程中耗散的能量正是整个环的面积，可以类比热力学循环中的 $p - V$ 图来证明。

4.2 滞回现象的数学模型

为了描述一个滞回现象，光有一个响应变量 Φ 与一个主动变量 $\sigma(t)$ 是可以的。但是一旦涉及到含时间的反馈效果，就必须引入一个附带反馈函数 $\eta(\sigma)$ 。与此同时，测得的主动变量和响应变量都是它们的含时演化结果，因此时间变量是串起数个变量之间的最重要的参数。

4.2.1 简单滞回模型 (SHM)

如果不考虑反馈条件，那么最简单的滞回模型可以写为：

$$\Phi_S(\sigma) = \begin{cases} 1, & |\sigma| < \delta, \dot{\sigma} \leq 0 \\ -1, & |\sigma| < \delta, \dot{\sigma} > 0 \\ sgn(\sigma), & |\sigma| \geq \delta \end{cases} \quad (56)$$

这个模型没有反馈过程，只是描述了一种唯象的、取决于 σ 演化方向的分段函数。方程中不存在含时的反馈项，因此也没有办法讨论其演化。于是称这种滞回模型为简单静态滞回 (Simple Static Hysteresis, SSH)[17]。

如果考虑一个反馈项，可以先定性地分析一下反馈的过程，以便于选取反馈函数。首先从初态开始，驱动力正向，从而 σ, Φ 都增大。此时因为反馈的滞后性，反馈函数为 0。当驱动力反向且已经存在了，此时反馈函数 $\eta(\sigma)$ 产生效果，将 Φ 向反向拉，并且随着 σ 愈大，反馈效果越强。反过来则同理。

根据这些要求，引入简单动态滞回 (Simple Dynamic Hysteresis, SDH)：

$$\frac{d\eta}{dt} = \frac{k}{\tau} [\sigma(t) - \Delta(\eta)] \quad (57)$$

其中 $\Delta(\eta)$ 是 $\eta(\sigma)$ 的反馈函数，形式为：

$$\Delta(\eta) = \begin{cases} \frac{1}{k} (\eta - f_0), & \eta > f_0 \\ 0, & |\eta| \leq f_0 \\ \frac{1}{k} (\eta + f_0), & \eta < -f_0 \end{cases} \quad (58)$$

输出函数 $\Phi_D(\sigma)$ 则表示为：

$$\Phi_D = \eta - k\sigma \quad (59)$$

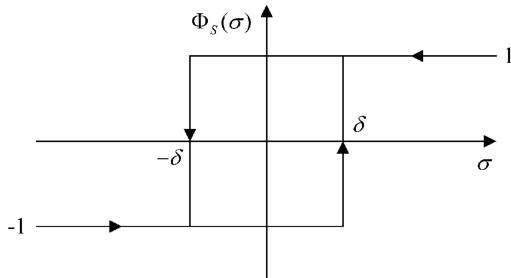


图 18: $\Phi_S - \sigma$ 图 (reproduce from [17])

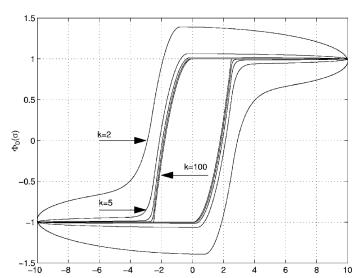


图 19: $\Phi_D - \sigma$ 图 (reproduce from [17])

式(57)有显著的非线性和自反馈性。整个反馈过程可以被形象地化成一个分块图 (Block Diagram)。图中表现的实际上是积分关系，但是取变上限积分之后是完全等价的：

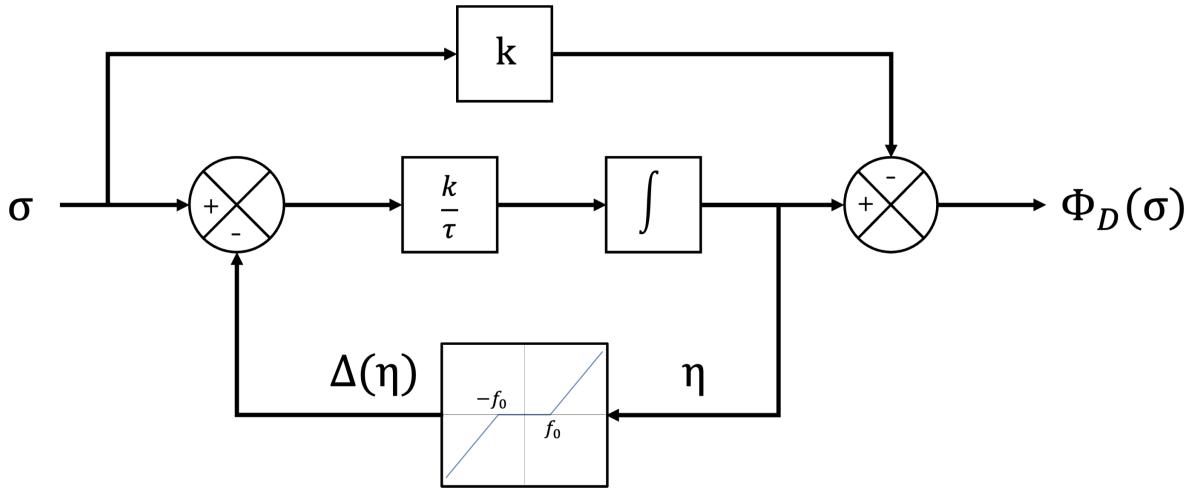


图 20: 式 (57) 的等效控制图

为了分析这个过程，令积分过程取线性近似 $\int f dt = f \Delta t$ 。约定右上角标代表迭代数：

- $N = 0$:

输入 σ^0 ，没有 η 返回，从而直接输出 $\Phi_D^0 = k\sigma^0 \left(\frac{\Delta t}{\tau} - 1 \right)$ 、返回 $\eta^0 = k\sigma^0 \frac{\Delta t}{\tau}$ 。

同时假设 η^0 没有超出 f_0 的截断，返回 $\Delta(\eta^0) = 0$

- $N = 1$:

输入 σ^1 ，总输入值为 $\sigma^1 - \Delta(\eta^0) = \sigma^1$ 。从而直接输出 $\Phi_D^1 = k\sigma^1 \left(\frac{\Delta t}{\tau} - 1 \right)$ ，返回 $\eta^1 = k\sigma^1 \frac{\Delta t}{\tau}$ 。

不妨设 $\sigma^1 > \sigma^0$ ，此时 $\eta^1 > \eta^0$ ，假设 η^1 过了阈值，因此返回 $\Delta(\eta^1) = \frac{1}{k} \left(k\sigma^1 \frac{\Delta t}{\tau} - f_0 \right)$

- $N = 2$:

输入 σ^2 ，总输入值为 $\sigma^2 - \Delta(\eta^1) < \eta^2$ 。不妨设 $\sigma^0 < \sigma^1$ ，从而

直接输出 $\Phi_D^2 = k\sigma^2 \left(\frac{\Delta t}{\tau} - 1 \right) - \frac{\Delta t}{\tau} \left(k\sigma^1 \frac{\Delta t}{\tau} - f_0 \right)$ ，返回 $\eta^2 = \frac{\Delta t}{\tau} \left(k \left(\sigma^2 - \frac{\Delta t}{\tau} \sigma^1 \right) + f_0 \right)$ 。

此时仍假设 η^2 未过阈值，因此返回 $\Delta(\eta^2) = 0$

- ...

从上面的积分代数分析，可以得出几条规律，佐证了前面的纯定性分析：

- 无反馈状态下输出 σ 和输入 Φ_D 是正相关的
- Φ_D 输出的本质是衡量对时积分过程中的实际增量。
- $\Delta(\eta)$ 函数的目的是在 σ 过大的时候增加一个削弱反馈。这样就可以减弱输出，达成滞回效果；这个反馈是作用于下一个 σ 的，因此也包含延时效果

4.2.2 SDH 模型的微分代数分析

考虑一个实际的 DRH 拉伸模型。 σ 等价于伸长 x , 输出量 Φ_D 等价于阻滞力 F , 下面的计算以实际物理量为代数符号。

首先考虑匀速拉伸模型, 此时 x 等价于一个三角波:

$$x(t) = \begin{cases} \frac{x_0}{T_0}t, & 0 < t \leq T_0 \\ \frac{x_0}{T_0}(2T_0 - t), & T_0 < t \leq 3T_0 \\ \frac{x_0}{T_0}(t - 4T_0), & 3T_0 < t \leq 5T_0 \end{cases} \quad (60)$$

阈值设为 $f_0 = \alpha kx_0$ 。较低的 α 将会在很早就产生阻滞效果, 进而产生明显的滞回环; 较高的 α 将对负反馈不敏感。同时也设 $\tau = \beta T_0$ 。

- $t \in (0, T_0]$, $x(t) = \frac{x_0}{T_0}t$

- $\eta < f_0$

此时 $\Delta(\eta) = 0$, 可以容易写出:

$$\eta(t) = \frac{kx_0}{2T_0\tau}t^2 \quad (61)$$

$$F(t) = \frac{kx_0}{T_0} \left(\frac{t}{2\tau} - 1 \right) \quad (62)$$

- $\eta = f_0$

此时有临界时间:

$$t_1 = \sqrt{2\alpha T_0 \tau} \quad (63)$$

- $\eta > f_0$

此时 $\Delta(\eta) = \frac{1}{k}(\eta(x(t)) - f_0)$, 从而微分方程写为:

$$\frac{d\eta}{dt} + \frac{\eta - f_0}{\tau} = \frac{kx_0}{T_0\tau}t \quad (64)$$

带入边界条件 $\eta(t_1) = f_0$ 解得:

$$\eta(t) = f_0 + \frac{kx_0}{T_0} \left(t - \tau + (\tau - t_1)e^{-(t-t_1)/\tau} \right) \quad (65)$$

$$\begin{aligned} F(t) &= \frac{kx_0}{T_0} \left(\alpha T_0 - \tau + (\tau - t_1)e^{-(t-t_1)/\tau} \right) \\ &= f_0 \left(1 - \frac{\beta}{\alpha} + \left(\frac{\beta}{\alpha} - \sqrt{2\alpha\beta} \right) e^{-(t-t_1)/\tau} \right) \end{aligned} \quad (66)$$

- $t \in (T_0, 3T_0]$, $x(t) = \frac{x_0}{T_0}(2T_0 - t)$, 记 $\eta(T_0) = \eta_1$

- $\eta > f_0$ 此时方程变为:

$$\frac{d\eta}{dt} + \frac{\eta - f_0}{\tau} = \frac{kx_0}{T_0}(2T_0 - t) \quad (67)$$

解为:

$$\eta(t) = \frac{kx_0}{T_0}(2T_0 - t - \tau) + f_0 + C e^{-t/\tau} \quad (68)$$

带入边界条件有:

$$C = \left(\eta_1 - kx_0 - f_0 + \frac{kx_0}{T_0} \tau \right) e^{T_0/\tau} \quad (69)$$

从而也有 $F(t)$:

$$F(t) = \frac{f_0}{\alpha} \left[(\beta - \alpha) \left(1 - e^{-\frac{(t-T_0)}{\tau}} \right) + \left(1 - \beta + (\beta - \sqrt{2\alpha\beta}) e^{-(1-\sqrt{2\alpha\beta})/\beta} \right) \right] \quad (70)$$

很显然这个区段下 $\eta(t)$ 是单调递减的, 因此一定会回落到死区。

- $\eta = f_0$

设定回落到死区的时刻 t_2 , 其满足超越方程:

$$\begin{aligned} \eta(t_2) &= \frac{kx_0}{T_0} (2T_0 - \textcolor{red}{t}) + f_0 - \frac{kx_0}{T_0} \tau + C e^{-\frac{t}{\tau}} \\ &= \frac{f_0}{\alpha} \left[\left(2 - \frac{\textcolor{red}{t}}{T_0} \right) - (\alpha - \beta) \left(1 - e^{-\frac{(t-T_0)}{\tau}} \right) + \left(1 - \beta + (\beta - \sqrt{2\alpha\beta}) e^{-(1-\sqrt{2\alpha\beta})/\beta} \right) \right] \\ &= f_0 \end{aligned} \quad (71)$$

不妨令 $t_2 = \gamma T_0$ 以便讨论。

- $-f_0 < \eta < f_0$

容易解出:

$$\eta(t) = f_0 + \frac{kx_0}{2T_0\tau} (4T_0 - t_2 - \textcolor{red}{t}) (t - t_2) \quad (72)$$

从而 $F(t)$:

$$\begin{aligned} F(t) &= f_0 + \frac{kx_0}{2T_0\tau} (4T_0 - t_2 - t) (\textcolor{red}{t} - t_2) - \frac{kx_0}{T_0} (2T_0 - \textcolor{red}{t}) \\ &= f_0 \left[1 + \frac{1}{\alpha\beta} \left(-\frac{t^2}{T_0^2} + \frac{4+2\beta}{T_0} t + (4\gamma + \gamma^2 - 4\beta) \right) \right] \end{aligned} \quad (73)$$

- $\eta \leq -f_0$

由于 η 取在二次函数下降段, 因此会快速反向离开死区, 从而重新具有式 (68) 的类似形式, 只是将 f_0 换为 $-f_0$:

$$\eta(t) = \frac{kx_0}{T_0} (2T_0 - \textcolor{red}{t} - \tau) - f_0 + C e^{-\frac{t}{\tau}} \quad (74)$$

C 由边界条件 $\eta(t_3) = -f_0$ 决定。此后由于 x 始终减小, 因此 $\Delta(\eta)$ 会反向地产生作用, 也即让 η 上升。

- $t \in (3T_0, 5T_0]$, $x(t) = \frac{x_0}{T_0} (\textcolor{red}{t} - 4T_0)$, 记 $\eta(3T_0) = \eta_3$

- $\eta \leq -f_0$

此时 η 具有形式:

$$\eta(t) = \frac{kx_0}{T_0} (\textcolor{red}{t} - 4T_0 - \tau) - f_0 + C e^{-\frac{t}{\tau}} \quad (75)$$

带入边界条件有:

$$C = \left(\eta_3 + kx_0 + f_0 + \frac{kx_0}{T_0} \tau \right) e^{3T_0/\tau} \quad (76)$$

随 η 增长, 在 $t = t_4$ 回到死区, 对应 t_4 有关系 $\eta(t_4) = -f_0$ 。

- $-f_0 < \eta < f_0$ 此区间可以解出:

$$\eta(t) = -f_0 + \frac{kx_0}{2T_0\tau}(\textcolor{red}{t} - t_4)(\textcolor{red}{t} + t_4 - 8T_0) \quad (77)$$

如果取到 $t_4 > 4T_0$, 则这显然是一个单调递增函数, 因此会增加到 $\eta = f_0$, 从而再次脱离死区。

对应临界 $\eta(t_5) = f_0$

- $\eta > f_0$

这个过程解的形式仍然是类似 (75):

$$\eta(t) = \frac{kx_0}{T_0}(\textcolor{red}{t} - t_5 - T_0 - \tau) + f_0 + Ce^{-\textcolor{red}{t}/\tau} \quad (78)$$

带入边界条件有:

$$C = \left(kx_0 + \frac{kx_0}{T_0}\tau \right) e^{t_5/\tau} \quad (79)$$

- 随着三角波跨过 T_5 , x 又将减小, 从而 η 也随之减小, 再次进入死区, 如此往复。

上面的讨论实际上是非常不完善的, 根据参数的选择, 实际的 SDH 微分方程的解具有极多分枝, 这里取的系数性质较好, 从而保证了这个循环充分展现 SDH 方程的特性。

- α 是描述阈值和最大反馈力之间关系的无量纲系数。较高的 α 将减弱反馈的效果。
- k 是反馈函数斜率的倒数, 也是正向积分的系数。 k 越高, 滞回的反应会更不显著。这是因为更快进入阈值, 阈值中的反馈更小, 从而很快就脱出阈值了。反之同理。
- β 是描述特征时间 τ 和周期 T_0 的比例关系。 β 越高, 整个积分会被压缩。进而造成更难进入死区, 同时进入死区的时间更长, 能量耗散更多。

4.2.3 SDH 模型的数值模拟结果

取系列参数 (量纲取 SI 单位制):

$$\tau = 0.1, \quad k = 10, \quad x_0 = 5, \quad T_0 = 5, \quad f_0 = 1 \quad (80)$$

数值计算得到结果, 代码见A, 得到了佐证上面半定量分析的结论。此处仍然以 σ 为输入参量。

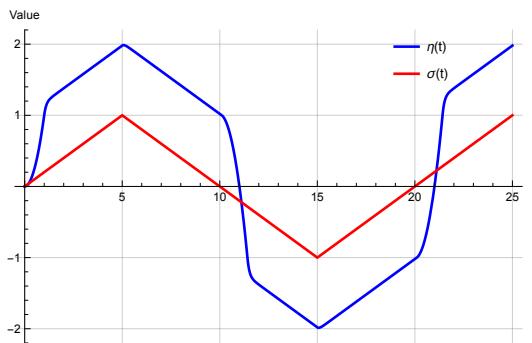


图 21: $\eta - t$ 与 $\sigma - t$ 图

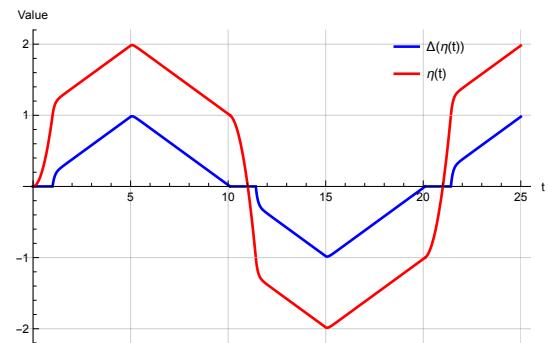


图 22: $\Delta(\eta) - t$ 与 $\eta - t$ 图

可以看出这样的参数选取保证了在每个区段中存在两次从死区的进出 ($(0, T_0]$ 只有一次)。同时也保证了不同周期 $\Delta(\eta)$ 函数取值的连续性。

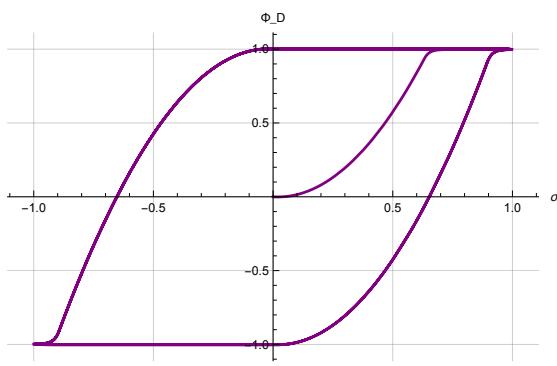


图 23: 滞回环结构

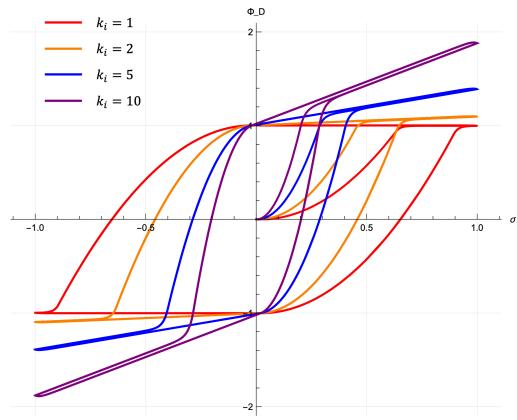


图 24: 滞回环和 k 之关系

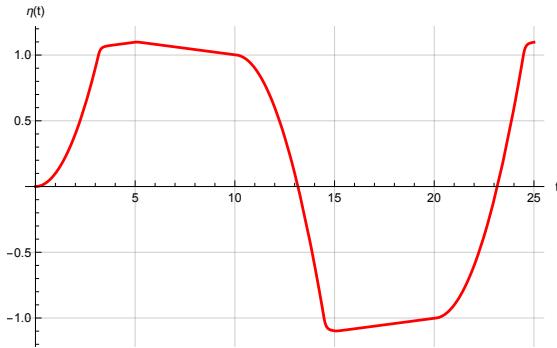


图 25: $\eta - t$ 图

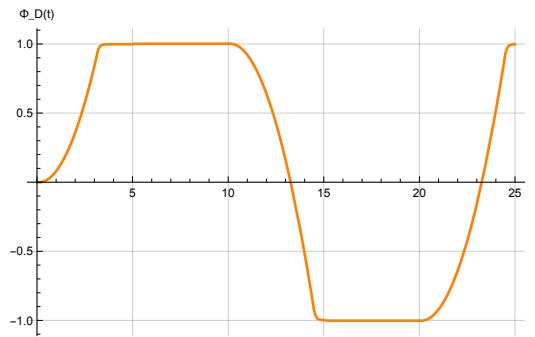


图 26: $\Phi_D - t$ 图

图 (23) 展示了滞回环的基本形态。图 (24) 则表现了不同的 $k_i = \frac{1}{k\tau}$ 情况下滞回环的不同。可以看出 k_i 大对应的是较好地统一了 k 与 τ 对滞回不同的影响。 k_i 值越大，滞回环面积越小，滞回效果越不显著； k_i 值越小，滞回环面积越大，滞回效果越显著。

图 (25) 较好地表现了指数函数和二次函数之间的转化。图 (26) 则显示了对于在负反馈区 $|\eta| > f_0$ 时，输出的结果是近似不变的，这是因为取 τ 较小，在式子 (66) (70) 中含时项较小，从而没有明显的体现。在图 (27)、(28)、(29) 可以看出这个结论。

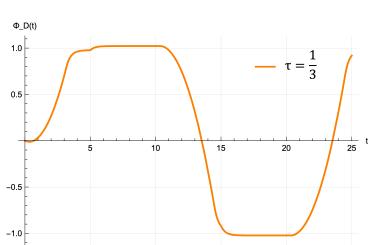


图 27: $\Phi_D - t$ 图 $\left(\tau = \frac{1}{3}\right)$

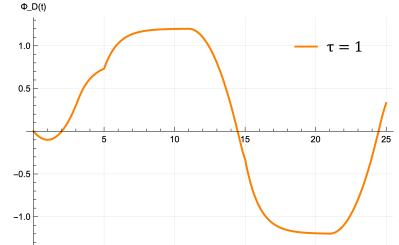


图 28: $\Phi_D - t$ 图 $(\tau = 1)$

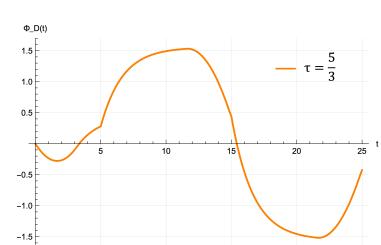


图 29: $\Phi_D - t$ 图 $\left(\tau = \frac{5}{3}\right)$

4.3 DRH 复合过程中的滞回现象

4.3.1 影响 DRH 复合产生滞回的因素

DRH 拉伸-复合过程中产生滞回的原因是多样的。Wei-Chang Lo 发现在周期电场的作用下反复 ss-DNA 会产生一个沿着 DNA 主链的管状区域，ssDNA 在其中收到外部流体粘滞阻力的作用从而产生耗散滞回 [19]。单纯的碱基自由能作用和堆叠效应也会有对应的结果，如3.3所展现的“Stick-Slip”过程。这说明复合过程和碱基序列也是有关的。

同时在链溶液体系之下，需要考虑复合过程链之间的竞争关系，以及伴随而生的链侵入（Invasion）动力学。A.Walbrun 在研究中发现通过力对 RNA/DNA 链的拉伸可以相对链在自由状态复合消耗更少的能量 [18]，如图（30）：

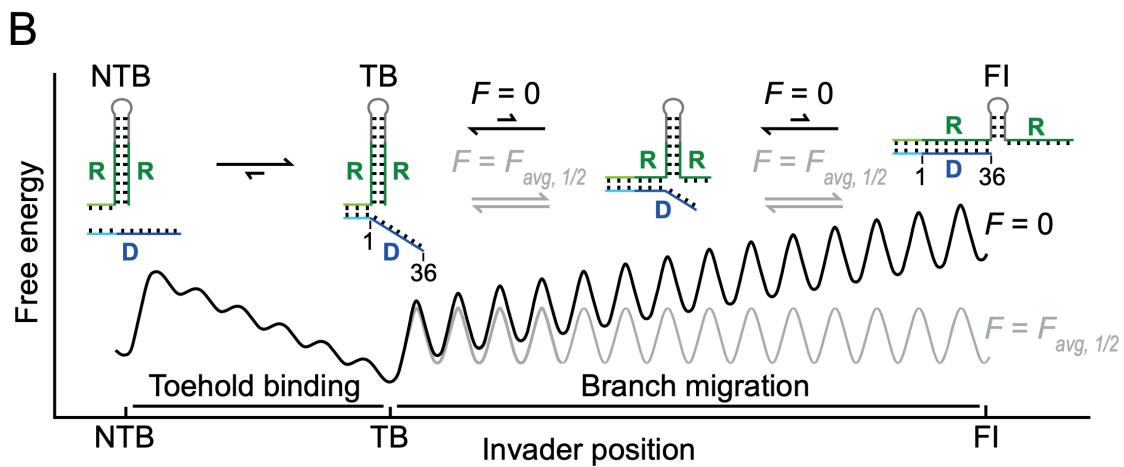


图 30: 侵入过程的自由能比较 (reproduced from [18])

4.3.2 复合过程中力平台的性质

容易想到，在复性过程中出现力平台的原因是引发的复合过程需要一个稳态的力来保证复性成功。在这个过程中，因为从 ssDNA 态转变为 S-DRH，整体螺旋度增加，从而有效长度降低。表现为拉力不变，伸长量减小的反向力平台。

这样来看，较高的力平台说明复合过程难度较高，需要将原 DNA/RNA 链拉的比较长才能进行复合过程。自然进行的势垒过大，需要高外力的拉伸来人为提高初态能量，以便进行复合。

环境因素上，较高的溶液离子浓度会提高力平台的大小，这也印证了前面的论断。因为高离子浓度会促进 ssDNA 的折叠，并且减弱碱基之间氢键的静电相互作用。

A SDH 数值模拟代码

```
1 ClearAll
2  $\tau = 1/10$ ; (*Dead zone斜率*)
3  $ki = 1$ ; (*积分系数*)
4  $k = 1/(\tau \cdot ki)$ ; (* $k=1/(\tau \cdot ki)$ *)
5  $x_0 = 5$ ; (*最大拉伸长度*)
6  $T_0 = 5$ ; (*半周期*)
7  $v_0 = x_0/T_0$ ; (*等效拉动速度 $v_0=x_0/T_0$ *)
8  $F_0 = 1$ ; (*阈值力*)

9 (*三角波输入*)
10 sigma[t_] := TriangleWave[{-x_0/T_0, x_0/T_0}, t/(4*T_0)]
11 (* $v_0 T_0 \sin[\pi/(2T_0)t]$ *)

12 (*死区函数\Delta(\eta)*)
13 delta[eta_] :=
14 Piecewise[{{1/(ki \tau)*(eta - F_0), 
15 eta > F_0}, {0, -F_0 \leq eta \leq F_0}, {1/(ki \tau)*(eta + F_0), 
16 eta < -F_0}}]

17 (*SDH方程*)
18 equation = eta'[t] == ki*(sigma[t] - delta[eta[t]]);

19 (*初始条件*)
20 initialCondition = eta[0] == 0;

21 (*数值求解微分方程*)
22 solution =
23 NDSolve[{equation, initialCondition}, eta, {t, 0, 15*T_0},
24 MaxSteps \rightarrow 10000, PrecisionGoal \rightarrow 8, AccuracyGoal \rightarrow 8];
25 etaFunc[t_] := eta[t] /. solution[[1]];

26 (*SDH输出*)
27 phiD[t_] := etaFunc[t] - (1/k)*sigma[t];
28 Print["模拟参数: \u03c4= ", \u03c4, ", \u03c4, ", ki= ", ki, ", x_0= ", x_0,
29 ", T_0= ", T_0, ", v_0= ", v_0, ", k= ", k];

30 (*输入信号\sigma(t)*)
31 Plot[sigma[t], {t, 0, 15*T_0}, PlotLabel \rightarrow "\u03c3(t)", AxesLabel \rightarrow {"t", "\u03c3(t)"}, 
32 PlotStyle \rightarrow {Blue, Thick}, GridLines \rightarrow Automatic, ImageSize \rightarrow 400]

33 (*状态反馈变量\eta(t)*)
34 Plot[etaFunc[t], {t, 0, 15*T_0}, PlotLabel \rightarrow "\u03b7(t)", 
35 AxesLabel \rightarrow {"t", "\u03b7(t)"}, PlotStyle \rightarrow {Red, Thick}, 
36 GridLines \rightarrow Automatic, ImageSize \rightarrow 400]

37 (*输出变量\Phi_D(t)*)
38 Plot[phiD[t], {t, 0, 15*T_0}, 
39 PlotLabel \rightarrow 
40 "\u03a6_D(t) = \u03b7(t) - (1/k)\u03c3(t)", 
41 AxesLabel \rightarrow {"t", "\u03a6_D(t)"}, 
42 PlotStyle \rightarrow {Orange, Thick}, GridLines \rightarrow Automatic, 
43 ImageSize \rightarrow 400]

44 (*滞回环*)
45 ParametricPlot[{sigma[t], phiD[t]}, {t, 0, 15*T_0},
```

```

56 PlotLabel ->"SDH滞回环",
57 AxesLabel ->{" $\sigma$ ", " $\Phi_D$ "},
58 PlotStyle ->{Purple, Thick}, AspectRatio ->1/GoldenRatio,
59 GridLines ->Automatic, PlotRange ->All, ImageSize ->400]
60
61 (*不同ki值下的滞回环*)
62 plotHysteresisForDifferentKi[kVal_, v0Val_, T0Val_] :=
63 Module[{kiValues = {1, 2, 5, 10}, colors, plots, i},
64 colors = {Red, Orange, Blue, Purple};
65 plots = Table[ki = kiValues[[i]];
66 solution =
67 NDSolve[{eta'[t] == ki*(sigma[t] - delta[eta[t]]), eta[0] == 0},
68 eta, {t, 0, 15*T0Val}];
69 etaFuncLocal[t_] := eta[t] /. solution[[1]];
70 phiDLocal[t_] := etaFuncLocal[t] - (1/kVal)*sigma[t];
71 ParametricPlot[{sigma[t], phiDLocal[t]}, {t, 0, 15*T0Val},
72 PlotStyle ->colors[[i]], {i, Length[kiValues]}];
73 Show[plots, PlotRange ->All, GridLines ->Automatic,
74 ImageSize ->500, AspectRatio ->0.8,
75 PlotLabel ->"ki值对滞回环的影响 (k=" <> ToString[kVal] <> ")",
76 AxesLabel ->{" $\sigma$ ", " $\Phi_D$ "},
77 PlotLegends ->
78 Placed[LineLegend[colors, kiValues, LegendLabel ->"ki值"], Right]]];
79 plotHysteresisForDifferentKi[k, v0, T0]

```

参考文献

- [1] J.F.Léger. *L'adn : une flexibilite structurale adaptee aux interactions avec les autres macromolecules de son environnement*[D]. 1999, l'Université Louis Pasteur Strasbourg I, 170 p.
- [2] S.B.Smith, Y.Cui, C.Bustamante *Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules*[J]. Science, 1996, **271**: p.795.
- [3] John F.Marko. *Torque and dynamics of linking number relaxation in stretched supercoiled DNA*[J]. Phys.Rev.E, 2007, **76**(2): p.21926.
- [4] S.Cocco, J.F.Marko. *Overstretching and force-driven strand separation of double-helix DNA*[J]. Phys.Rev.E, 2004, **70**(1): p.011910.
- [5] S.Cocco, J.F.Marko, R.Monasson, A.Sarkar, J.Yan. *Force-extension behavior of folding polymers*[J]. Eur.Phys.J.E , 2003, **10**: p249-263.
- [6] J.R.Wenner ,M.C.Williams, I.Rouzina, and V.A.Bloomfield *Salt Dependence of the Elasticity and Overstretching Transition of Single DNA Molecules*[J]. Biophys., 2002, **82**(6): p3160-3169.
- [7] C.Storm and P.C.Nelson, *Theory of high-force DNA stretching and overstretching*[J]. Phys.Rev.E, 2003, **67**(5): p.051906.
- [8] I.Rouzina and V.A.Bloomfield, *Force-Induced Melting of the DNA Double Helix: 1. Thermodynamic Analysis*[J]. Biophys.J, 2001, **80**: p.882-893.
- [9] I.Rouzina and V.A.Bloomfield, *Force-Induced Melting of the DNA Double Helix: 2. Effect of Solution Conditions**[J]. Biophys.J, 2001, **80**: p.894 -900.
- [10] Dong-Ni Yang, Zhen-Sheng Zhong, Wen-Zhao Liu, Thitima Rujiralai, Jie Ma. *Theoretical study of overstretching DNA-RNA hybrid duplex*[J]. Chin.Phys.B, 2019, **28**(6): p.068701.
- [11] David L.Nelson, Michael M.Cox, *Lehninger Principles of Biochemistry*[M]. W.H.FREEMAN AND COMPANY. 2014:988-989,991,994
- [12] S.Cocco, J.F.Marko,a, R.Monasson, A.Sarkar, J.Yan *Force-extension behavior of folding polymers*[J]. Eur.Phys.J.E, 2003, **10**:p249-263.
- [13] J.SantaLucia Jr., H.T.Allawi, P.A.Seneviratne *Improved Nearest-Neighbor Parameters for Predicting DNA Duplex Stability*[J]. Biochemistry, 1996, **35**:p3555-2562.
- [14] Naoki Sugimoto, Shu-ichi Nakano, Misa Katoh, Akiko Matsumura, Hiroyuki Nakamura, Tat-suo Ohmichi, Mari Yoneyama, Muneo Sasaki *Thermodynamic Parameters To Predict Stability of RNA/DNA Hybrid Duplexes1*[J]. Biochemistry, 1995, **34**:p11211—11216.
- [15] S.K.Nomidis, E.Skoruppa, E.Carlon, J.F.Marko *Twist-bend coupling and the statistical mechanics of the twistable wormlike-chain model of DNA: Perturbation theory and beyond*[J]. Phys.Rev.E, 2019, **99**:p032414.

- [16] Ding Qian, Zhai Hongmei *The Advance In Researches Of Friction Dynamic In Mechanics System*[J]. ADVANCES IN MECHANICS, 2013, **43**(1):p112-131.
- [17] P.Gurfil, B.Friedland *New simple model of dynamic hysteresis*[J]. Transactions of the Institute of Measurement and Control, 2003,**25**(3):p202-216.
- [18] A.Walbrun, T.Wang, M.Matthies *Single-molecule force spectroscopy of toehold-mediated strand displacement*[J]. Nat Commun, 2024,**15**:p7654.
- [19] Wei-Chang Lo, H.C.Hong, H.J.Chi, Pik-Yin Lai, C.K.Chan *Stretching and migration of DNA by solvent elasticity in an oscillatory flow*[J]. Phys.Rev.E, 2011,**82**:p021802.