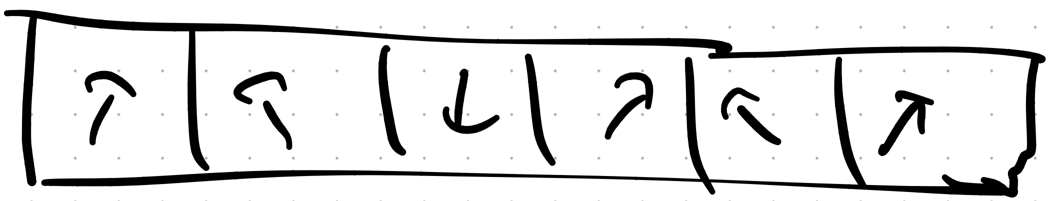
Helix-Coil Model

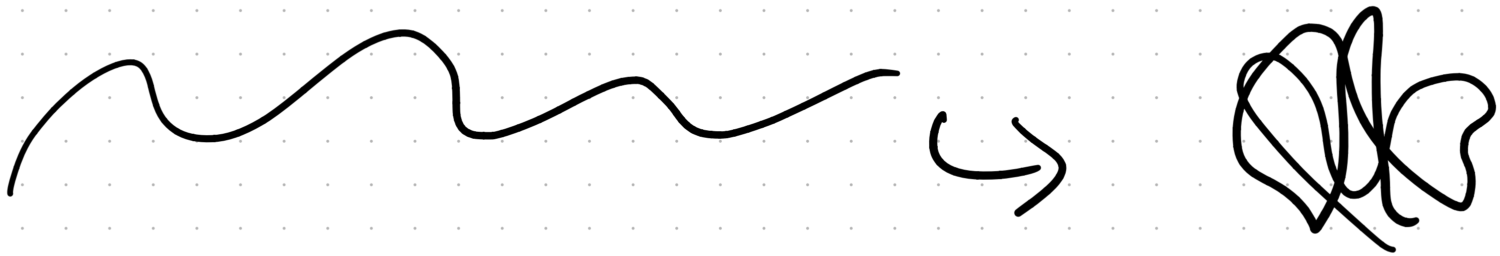
Previously we studied Ising Model



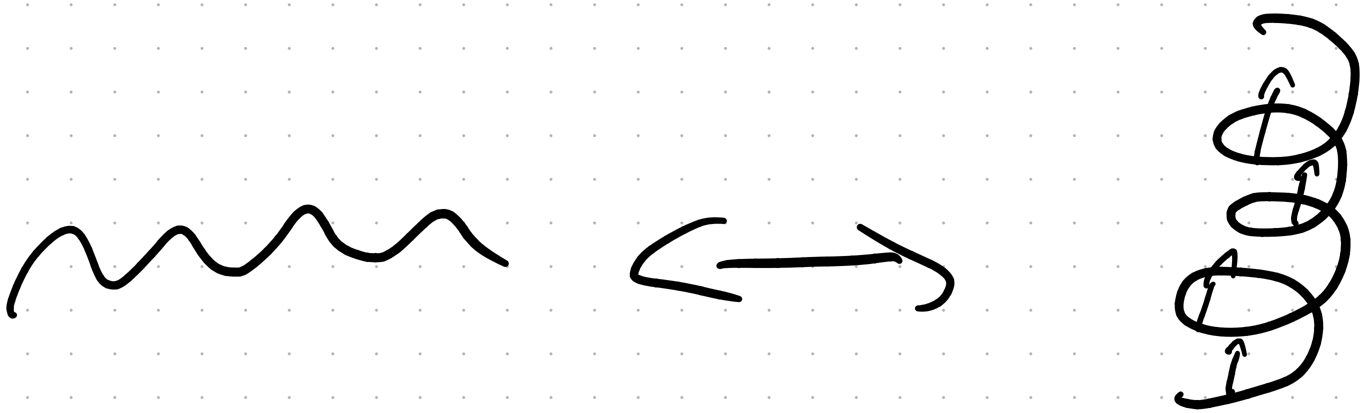
spin $= \pm 1/2$

$$\mathcal{E} = \sum_{i=1}^{N} \underbrace{-h s_i}_{\text{fieldenergy}} \quad \underbrace{- J s_i s_{i+1}}_{\text{neighbor coupling}}$$

proteins – polymers of amino acids (aa)



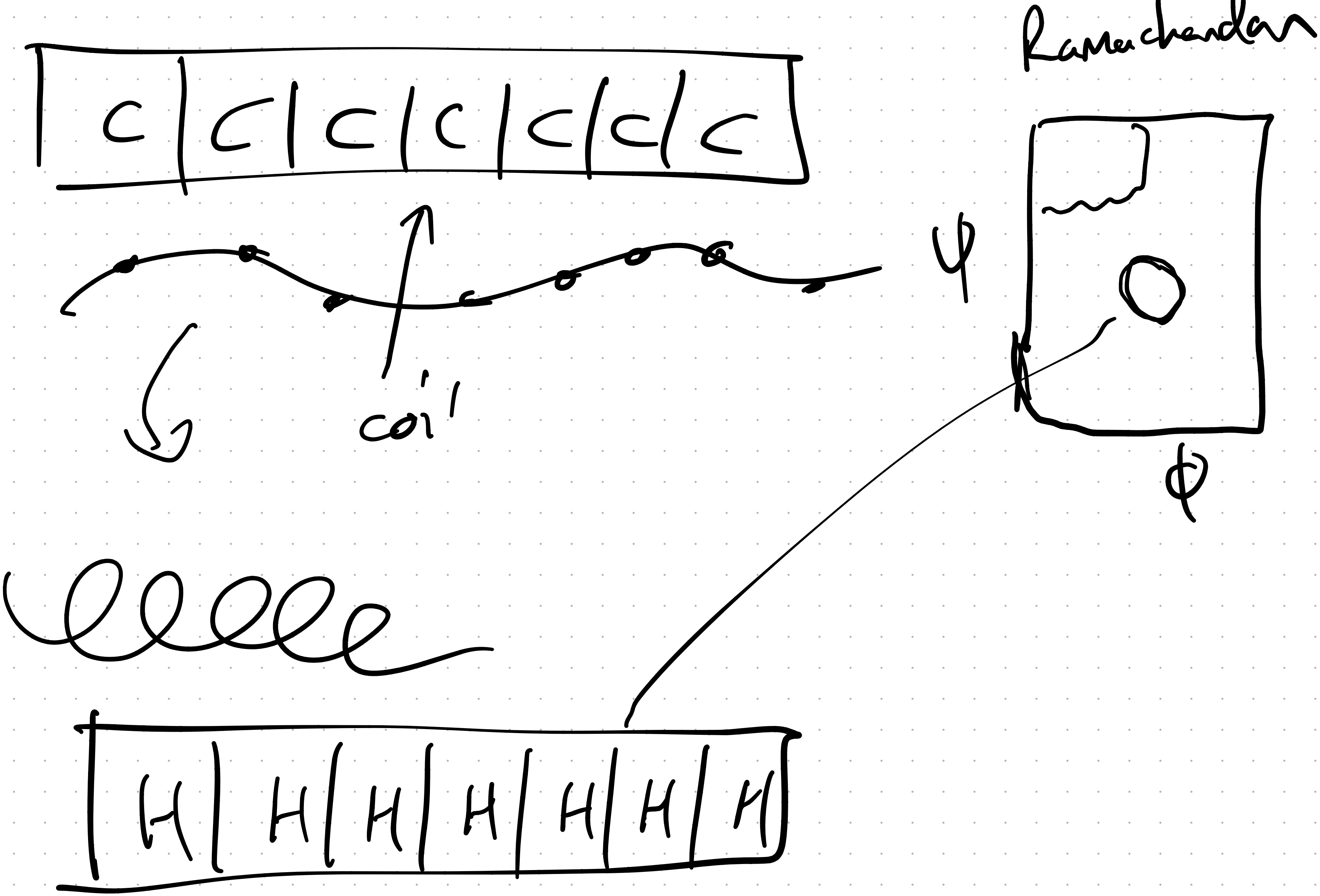
protein 3D structures – made up of Secondery structures



forms H-bonds
reduce energy

folding decreases
entropy

Consider each residue as a state on a lattice



Non-cooperative Model

What if each residue adopts H/C regardless of neighbors

$\varepsilon_{coil} = 0$

$\varepsilon_H = -\epsilon$ (favorable)

2-state model

For one reside

$$q = \sum_{n=0}^{1} e^{-\beta \varepsilon_n} = \underbrace{1}_{C} + \underbrace{e^{+\beta\epsilon}}_{H}$$

coil/helix

$$q = 1 + e^{\beta\epsilon} \qquad C \rightleftharpoons H$$

$$P_H = \frac{e^{\beta\epsilon}}{q} = \frac{e^{\beta\epsilon}}{1+e^{\beta\epsilon}} \qquad P_C = \frac{1}{1+e^{\beta\epsilon}}$$

$$K_{eq} = \frac{[H]}{[C]} = \frac{P_H}{P_C} = e^{+\beta\epsilon}$$

think of ϵ as a free energy

$$\Delta G = -k_B T \ln K_{eq}$$

$$P_H = \frac{K}{1+K}$$

K is Boltzmann factor for a Helix (1 for C)

H->C transition of a chain

N residues, currently independent
distinguishable

$$q = 1 + k$$

$$Q = q^N = (1+k)^N$$

Eq: 4 sites

$$Q = (1+k)^4 = 1 + 4k + 6k^2 + 4k^3 + k^4$$

coefs are $\binom{N}{n_H}$

HCCC
CHCC
CCHC
CCCH

$$Q = (1+k)^N = \sum_{n_H=0}^{N} \binom{N}{n_H} k^{n_H}$$

from an energy perspective

$$Q = \sum_{\text{states } i=0}^{2^N} e^{-\beta \varepsilon_n}$$

$$= \sum_{\varepsilon} \omega(\varepsilon) e^{-\beta \epsilon} \qquad \epsilon_{\text{lattice}} = -n_H \epsilon$$

$$= \sum_{\epsilon} \omega(\varepsilon) \Big(\underbrace{e^{+\beta \epsilon}}_{K}\Big)^{n_H}$$

$$Q = (1+K)^N = (1+e^{\beta\epsilon})^N$$

how many residues are in a H configuration on average $\langle n_H \rangle$

$$f_H = \frac{n_H}{N}$$

$$Q = \sum_{n_H=0}^{N} \binom{N}{n_H} K^{n_H}$$

want $f_H = \left\langle \frac{n_H}{N} \right\rangle = \sum_{n_H=0}^{N} \left(\frac{n_H}{N} \right) P(n_H)$

$$f_H = \sum_{n_H=0}^{N} \left(\frac{n_H}{N}\right) \underbrace{P(n_H)}_{\binom{N}{n_H} k^{n_H} / Q} = \frac{1}{N} \sum_{n_H=0}^{N} n_H \underbrace{\binom{N}{n_H}}_{\text{binomial}} \frac{k^{n_H}}{Q}$$

$$Q = \sum_{n_H=0}^{N} \binom{N}{n_H} k^{n_H}$$

$$\frac{\partial \ln Q}{\partial k} = \frac{1}{Q} \cdot \sum_{n_H=0}^{N} \binom{N}{n_H} \cdot (n_H) k^{n_H - 1}$$

$\times \frac{k}{N}$ $\qquad$ $\times \frac{k}{N}$

$$\boxed{f_H = \frac{k}{N} \frac{\partial \ln Q}{\partial k}} \qquad \underbrace{\qquad\qquad}_{f_H}$$

$$f_H = \frac{k}{N}\frac{\partial \ln Q}{\partial k}$$

$$Q = (1+k)^N$$

independent res model

$$f_H = \frac{k}{N} \cdot \frac{\partial}{\partial k} \ln(1+k)^N$$

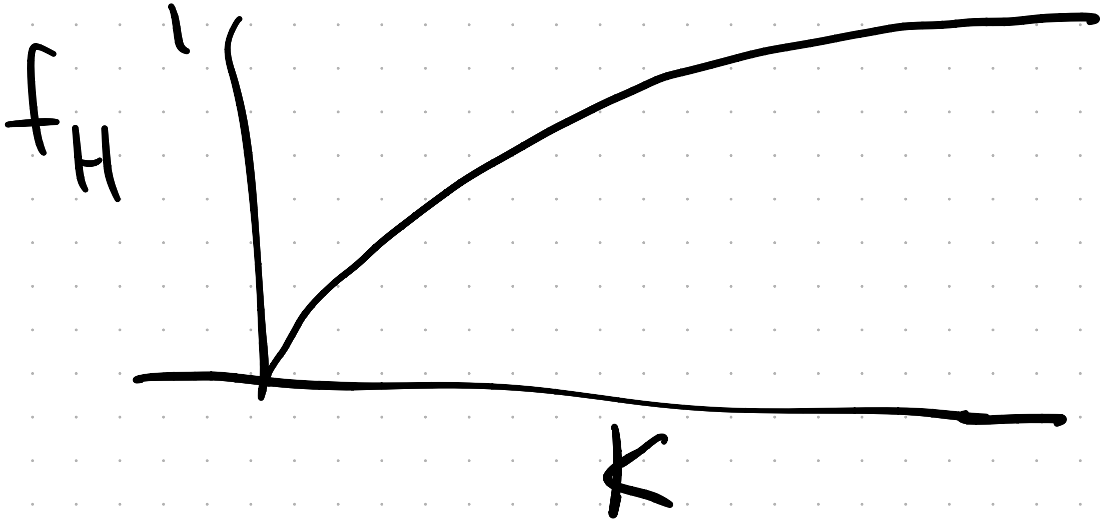
$$\ln(a^x) = x \ln(a)$$

$$\frac{\partial \ln x}{dx} = \frac{1}{x}$$

$$= k \frac{\partial}{\partial k} \ln(1+k)$$

$$= k/(1+k) \cdot \frac{\partial(1+k)}{\partial k} = \underbrace{\frac{k}{1+k}}_{\text{for } N \text{ resides}}$$

doesn't depend on N



$$f_H = \left(\frac{\partial \ln Q}{\partial K}\right)\frac{K}{N}$$

How does this connect to the energy description

$$\langle E_{total} \rangle = -\epsilon \cdot \langle n_H \rangle$$

$$\langle E \rangle = -\frac{\partial \ln Q}{\partial \beta} = -\frac{\partial \ln Q}{\partial K}\frac{\partial K}{\partial \beta}$$

$$K = e^{\beta\epsilon} \qquad \frac{\partial K}{\partial \beta} = \epsilon e^{\beta\epsilon} = \epsilon K$$

Heteropolymers

20 natural amino acids

a different K for every AA

for independent residue model

$$q_i = (1 + k_i)$$

$k_{alanine}$ k_{valine}

$$Q_N = (1+k_1)(1+k_2)\cdots(1+k_N)$$

$$= \prod_{i=1}^{N} (1+k_i)$$

Eg 2 amino acid types

$$Q_N = (1+K_A)^{n_A}(1+K_B)^{n_B}$$

$$n_A + n_B = N$$

Interacting Residues

H/C configuration should really depend on neighbors

$\underline{C}\,\underline{H}\,\underline{C}\,HHC \in$ configuration

$B.F. = (1)(k)(1)(k)(k)(1) = k^3$

$CHCHCH$

$BF = (1)(k)(1)(k)(1)(k) = k^3$

want to prefer this

add a term (τ) for neighboring H's

C H C H H C

$$BF = (1)(K)(1)(K\tau)(K)(1) = K^3\tau$$

If $\tau > 1$ get stretches of Helices

$\tau > 1$ is like a $J > 1$ in ising model

Zipper Model

$\tau >> 1$

only H's will be in a string of consecutive H's

HCHHCCCH $\quad K^4 \tau$

HHHHCCCC

CHHHHCCC

etc

$K^4 \tau^3 >>$

$\# = N - n_H + 1 \quad n_H \geq 1$

$$Q = \sum_{n_H=0}^{N} BF_{n_H}$$

$\# \; 1 \quad n_H = 0$

$$Q = 1 + \sum_{n_H=1}^{N} (N - n_H + 1) K^{n_H} \tau^{n_H - 1}$$

N C's

n_H's $\quad \underbrace{HHHH\cdots}_{n_H - 1}$

$$= 1 + \frac{1}{\tau} \sum_{n_H=1}^{N} (N - n_H + 1)(K\tau)^{n_H}$$

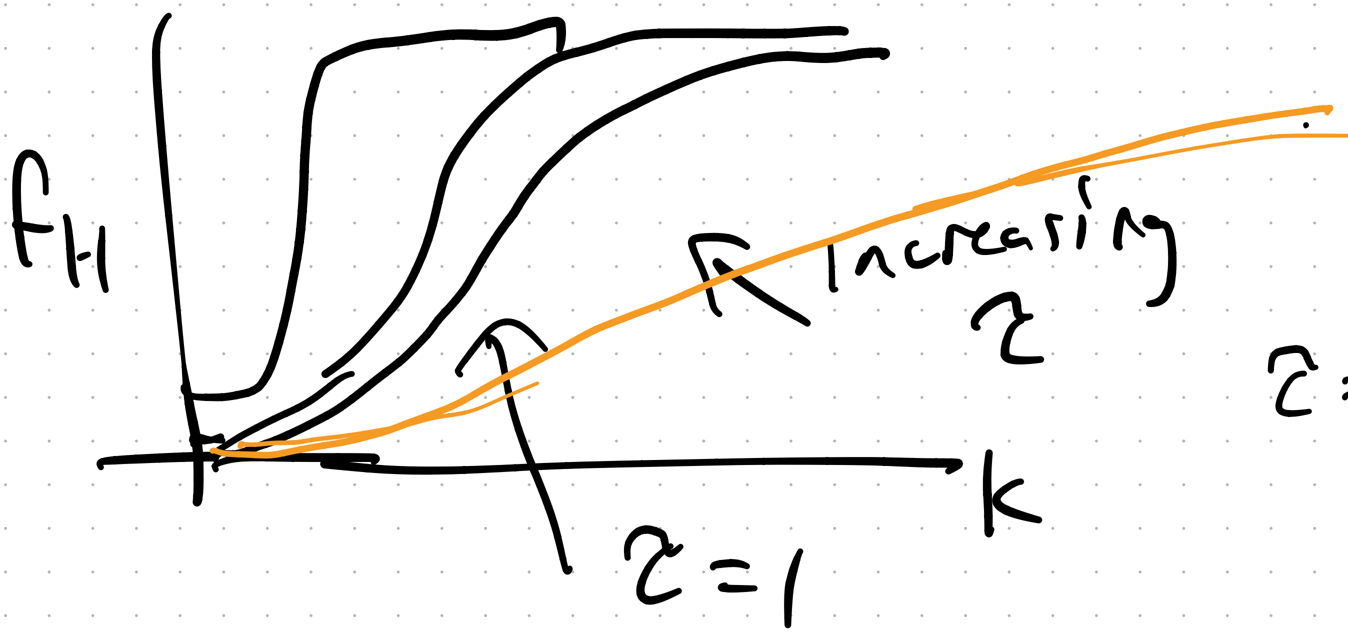
$$(N+1)(K\tau)^{n_H} \quad - n_H (K\tau)^{n_H}$$

Q has a "simple" formula

$$f_H = \frac{K}{N} \frac{\partial \ln Q}{\partial K}$$ – complicated formula

physically in Zipper model

more coop for big N, big τ



$\tau = 1$ wrong

Exact problem (like 1d ising model)

$$Q = \sum_{x_1=0,1} \sum_{x_2=0,1} \cdots \sum_{x_N=0,1} \prod_{i=1}^{N} \tilde{K}^{x_i} \tau^{x_i x_{i+1}}$$

↖ sum w/ 2^N terms

$x = 0$ being C

$\;\;\, = 1$ being H

HCCHHHCCH one of 2^9 configs

weight $K^1 k^0 k^0 k^1 k^1 k^1 k^0 k^0 k^1 \tau^0 \tau^0 \tau^0 \tau^1 \tau^1 \tau^0 \tau^0 \tau^0 \tau^0 = k^5 \tau^2$

Generate sum "transfer matrices"

multiply matrices, get sums
as combo of terms

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} w & x \\ y & z \end{pmatrix} = \begin{pmatrix} aw+by & ax+bz \\ cw+dy & cx+dz \end{pmatrix}$$

i+1

i

$i-1 \backslash i$	$\underline{h_i}$	c_i
$\underline{h_{i-1}}$	$k\tau$	1
$\underline{c_{i-1}}$	k	1

transfer matrix

$$\omega = \begin{pmatrix} k\tau & 1 \\ k & 1 \end{pmatrix}$$

$$W = \begin{pmatrix} k^2\tau & 1 \\ k & 1 \end{pmatrix}$$

for $N = 1$

$Q = 1 + k$

$$W^2 = \begin{pmatrix} k^2\tau & 1 \\ k & 1 \end{pmatrix}\begin{pmatrix} k^2\tau & 1 \\ k & 1 \end{pmatrix} = \begin{pmatrix} - & - \\ k^2\tau + k & k+1 \end{pmatrix}$$

sum $k^2\tau + k + k + 1$

Chain of 2:	CC	CH	HC	HH
	1	k	k	$k^2\tau$

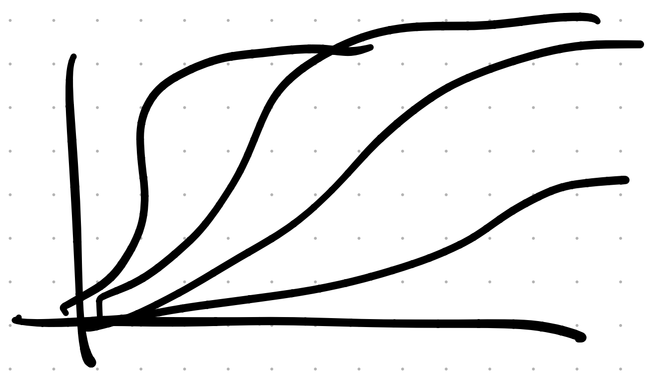
$$\underline{(0\ \ 1)}\begin{pmatrix}a & b\\ c & d\end{pmatrix} = (0+c\ \ \ 0+d) = (c\ \ \ d)$$

$$(c\ \ \ d)\begin{pmatrix}1\\ 1\end{pmatrix} = c+d$$

$$Q = (0\ \ 1)\ \omega^N \begin{pmatrix}1\\ 1\end{pmatrix}$$

general formula

$$f_H = \frac{k}{N}\frac{\partial \ln Q}{\partial k}$$



Other variations Zimm-Bragg

$$CCCCCC \underset{\sigma}{\rightleftharpoons} CHCCCC \underset{s}{\rightleftharpoons} CHHCC$$

big

σ small

$$\sigma \ll 1 < s$$

σ is nucleation

s represents growth