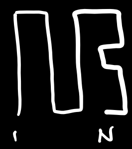
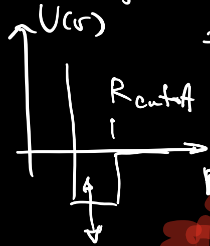
- What makes proteins "foldable"? // Random Energy Model



$\boxed{\vec{a} = \{a_1 \dots a_N\}}$

a.acids $\quad 1 \le a_i \le 20$

$$E = \sum_{i<j} U(r_{ij} < R_{cutoff}, a_i, a_j)$$



$$= \sum_{i<j} \Delta_{ij}\, U(a_i, a_j)$$

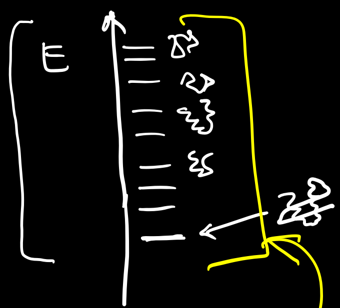
$$\Delta_{ij} = \begin{cases} 1 & \text{if } r_{ij} < R_{cutoff} \\ 0 & \text{otherwise} \end{cases}$$

structure

$\mathcal{M} = g^N$ possible structures

$$= \sum_{i<j} \Delta_{ij} B_{ij}$$

B_{ij}: N×N matrix



$B_{ij} = U(a_i, a_j)$

$0 < i < N$ $\quad \underbrace{\boxed{U(x,y)}}_{20\times 20}$ "known"

- What is the ensemble of structures for a random sequence?

Random matrix B_{ij}

$\Downarrow$

Random energies E

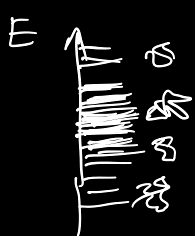
Prob of having a structure with energy E

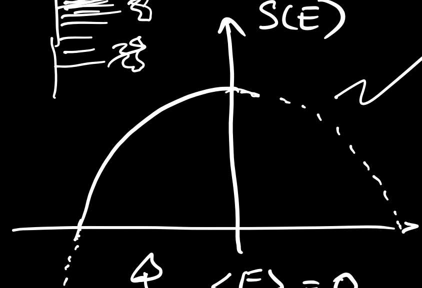
$$p(E) = \frac{1}{\sqrt{2\pi\Sigma^2}}\, e^{-\frac{(E-\langle E\rangle)^2}{2\Sigma^2}}$$

$$S(E) = \log \Omega(E) = \log[\mathcal{M}\cdot p(E)]$$

$$= \underbrace{\log \mathcal{M}}_{\sim N} - \underbrace{\frac{(E-\langle E\rangle)^2}{2\Sigma^2}}_{\sim N^0} - \underbrace{\frac{1}{2}\log(2\pi\Sigma^2)}_{\sim \log N}$$

$\sim N^2/N$





$\langle E\rangle = 0$

$\Sigma^2 = \mathrm{var}(E) \sim N$

$\mathcal{M} \sim g^N$

$\Sigma \sim N^{1/2}$ $\qquad E \sim N^?$

γ - # of neighbors

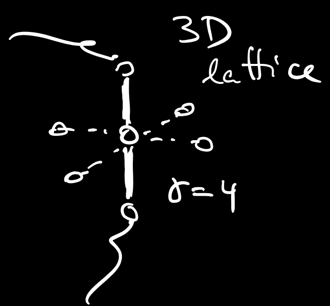
$E \sim N\cdot\gamma$

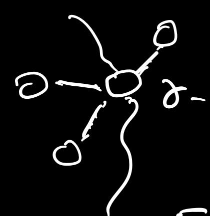
$$\boxed{S(E) = \underbrace{\log \mathcal{M}}_{\sim N} - \underbrace{\frac{E^2}{2\Sigma^2}}_{\sim N}}$$

$\Sigma \sim N$

$E = \sum_{i<j} \Delta_{ij} B_{ij}$

contact map of a protein structure





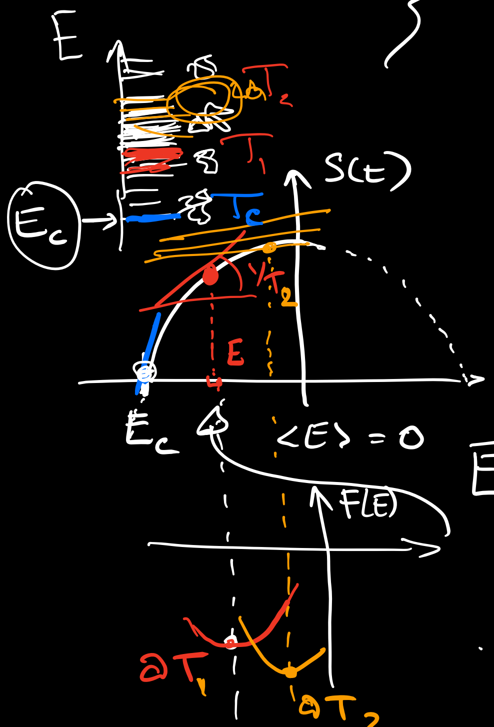
γ - coordination number

$E \sim \gamma \cdot N$

Σ sum of $\gamma \cdot N$ random numbers

$\Sigma^2 \sim N^2$

$\Sigma \sim N$



$F = E - TS(E)$

$\bar{E}: \quad \frac{dF}{dE} = 0 = 1 - T\frac{dS}{dE} \Leftrightarrow \frac{1}{T} = \frac{dS}{dE}$

T_c at which E_c is the minimum of $F(E)$?

$\frac{1}{T_c} = \frac{dS}{dE}\Big|_{E_c} = -\frac{2E_c}{2\Sigma^2} = -\frac{E_c}{\Sigma^2}$

$T_c = -\frac{\Sigma^2}{E_c}$

$$\boxed{S(E) = \underbrace{\log \mathcal{M}}_{N} - \underbrace{\frac{E^2}{2\Sigma^2}}_{N}}$$

$S(E_c) = 0 \Rightarrow \boxed{E_c = -\underbrace{\Sigma}_{\sim N}\underbrace{\sqrt{2\log\mathcal{M}}}_{\sqrt{N}}} \quad \sqrt{N}$

$$\boxed{T_c = \frac{\Sigma}{\sqrt{2\log\mathcal{M}}}}$$

@ T_c the structure @ E_c is the Free energy minimum

$\mathcal{M} \sim g^N$