

8. Polymer collapse

After their synthesis, proteins collapse into a very specific 3D shape, the **native** state, where they can perform their functions. A thorough analysis of the arrangement of amino-acids reveals that hydrophobic ones are prevalently hidden from water in the protein core, with polar, hydrophilic ones on the surface.

This has led to the conclusion that protein folding proceeds in a hierarchical way: hydrophobicity provides most of the "force" to drive the collapse to a compact state, whereas the other interactions provide the forces for the local rearrangements that specify the particular shape.

Of course, this is an oversimplistic view, but, as a first step, it is a reasonable one.

So, let's assume that all monomers are identical (homopolymer) and have a mutual attraction.

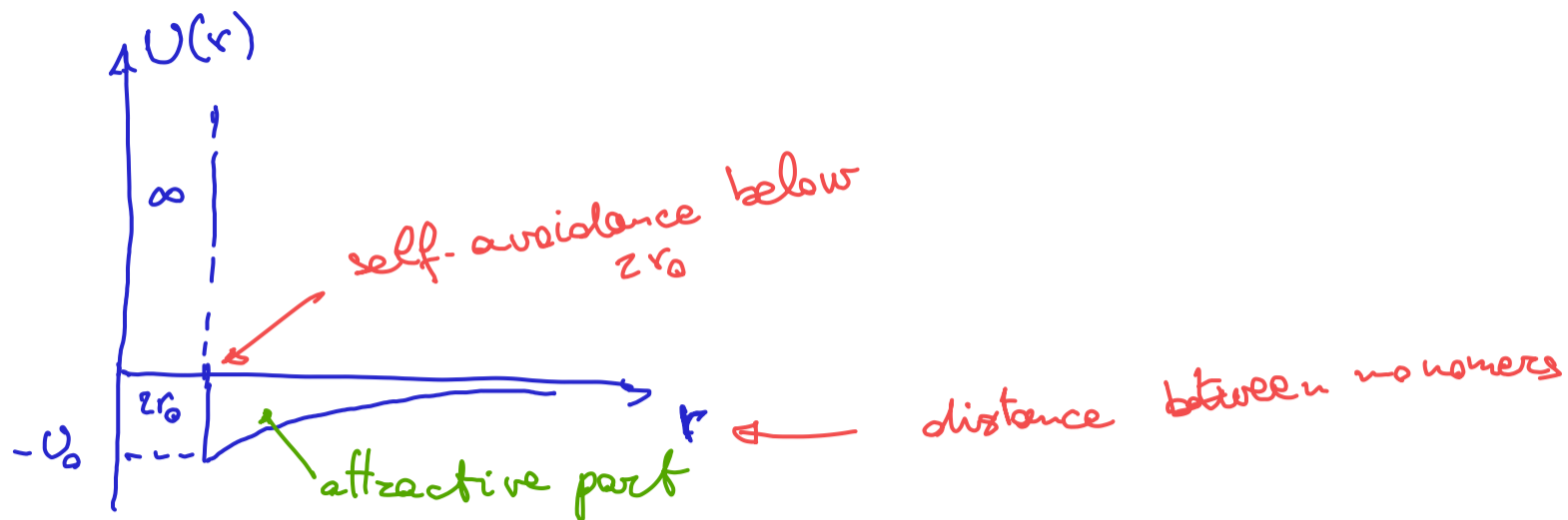
Can we understand polymer collapse with a Flory-like approach?

We must revisit the meaning of the B_2 factor

$$P(R_{ee}) \propto \exp \left[- \frac{d R_{ee}^2}{2b^2 N} - B_2 \frac{N^2}{R_{ee}^3} \right]$$

It was coming from the necessity of monomers to avoid each other, in the dilute gas approximation.

Now we have interacting monomers, with a potential that we can, schematically, represent as



The total energy is

$$E = \sum_{i=1}^{N-1} \sum_{j=i+1}^N U(r_{ij})$$

$r_{ij} = |\vec{R}_i - \vec{R}_j|$

$\underbrace{U(r_{ij})}_{U_{ij}}$

and the free energy is (for a gas of molecules):

$$F = -k_B T \ln Z$$

and

$$Z = \frac{1}{N! \lambda^{3N}} \int d\vec{r}_1 \dots d\vec{r}_N e^{-\beta E} = \int d\vec{r}_1 \dots d\vec{r}_N \prod_{i=1}^{N-1} \prod_{j=i+1}^N e^{-\beta U_{ij}} =$$

$$= \frac{1}{N! \lambda^{3N}} \int d\vec{r}_1 \dots d\vec{r}_N \prod_{i=1}^{N-1} \prod_{j=i+1}^N \left(1 + \underbrace{\psi(r_{ij})}_{\text{Mayer function}} \right) =$$

Mayer function

$$\psi(r_{ij}) = e^{-\beta U(r_{ij})} - 1$$

develop $\rightarrow = \frac{1}{N! \lambda^{3N}} \int d\vec{r}_1 \dots d\vec{r}_N \left\{ 1 + \sum_{\substack{(i,j) \\ \text{pairs}}} \varphi(r_{ij}) + \right.$

$$+ \sum_{\substack{(i,j,k) \\ \text{triplets}}} \left[\varphi(r_{ij})\varphi(r_{ik}) + \varphi(r_{ij})\varphi(r_{jk}) + \varphi(r_{ik})\varphi(r_{jk}) + \varphi(r_{ij})\varphi(r_{ik})\varphi(r_{jk}) \right] + \dots \left. \right\} =$$

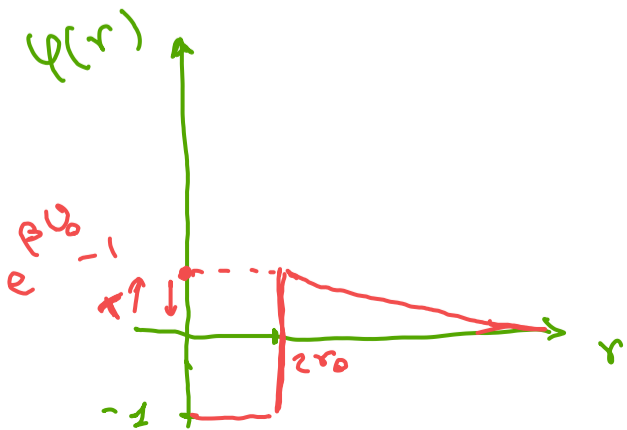
$$= \frac{1}{N! \lambda^{3N}} \left[V^N + \binom{N}{2} V^{N-2} \int d\vec{r}_i d\vec{r}_j \varphi(|\vec{r}_i - \vec{r}_j|) + \right.$$

$$+ \binom{N}{3} V^{N-3} \int d\vec{r}_i d\vec{r}_j d\vec{r}_k \left[\dots \right] + \dots \left. \right] =$$

$$= \frac{1}{N! \lambda^{3N}} \left[V^N + \binom{N}{2} V^{N-2} \cdot \underbrace{V \int_V d\vec{r} \left(e^{-\beta U(\vec{r})} - 1 \right)}_{\varphi(r)} + \right.$$

$$+ \binom{N}{3} V^{N-2} \int_V \int_V d\vec{r}_i d\vec{r}_j \left[\dots \right] + \dots \left. \right]$$

$$\int_V d\vec{r} \underbrace{\left(e^{-\beta U(\vec{r})} - 1 \right)}_{\varphi(r)} = B_2$$



B_2 is made of two parts:

a negative contribution in the repulsive part, equal to the volume of a sphere of radius $z r_0$

and a positive part where the potential is attractive.

The negative part is mostly temperature independent, whereas the positive part decreases with temperature.

Thus, at high temperature $B_2 < 0$

at low temperature $B_2 > 0$

There is an intermediate temperature T_0 where $B_2 = 0$

The free-energy, if we stop at the second virial coefficient (B_2) is

$$F = -k_B T \ln Z = -k_B T \ln \left[\frac{V^N + \binom{N}{2} V^{N-1} B_2}{\lambda^{3N} N!} \right] =$$

$$= -k_B T \ln \frac{V^N}{\lambda^{3N} N!} - k_B T \ln \left[1 + \binom{N}{2} \frac{B_2}{V} \right] \stackrel{\text{dilute gas}}{\approx}$$

$$= F_{\text{perfect gas}} - k_B T \binom{N}{2} \frac{B_2}{V} \rightarrow Z = e^{-\beta F} = e^{-\beta F_{\text{PG}} + k_B T \binom{N}{2} \frac{B_2}{V}}$$

⇒ Translated in the polymer language:

$F_{\text{perfect gas}}$

→

$P_{\text{FC}}(R_{ee}) = R_{ee}^{d-1} e^{-\frac{d}{2b^2N} R_{ee}^2}$

$$P(R_{ee}) = \frac{1}{Z} R_{ee}^{d-1} e^{-\frac{d}{2b^2N} R_{ee}^2 + k_B T \frac{N(N-1)}{2} \frac{B_2}{R_{ee}^d}}$$

Since at high T we have $B_2 < 0$ (only self-avoidance, which is infinitely strong, matters) we recover the formula of Flory.

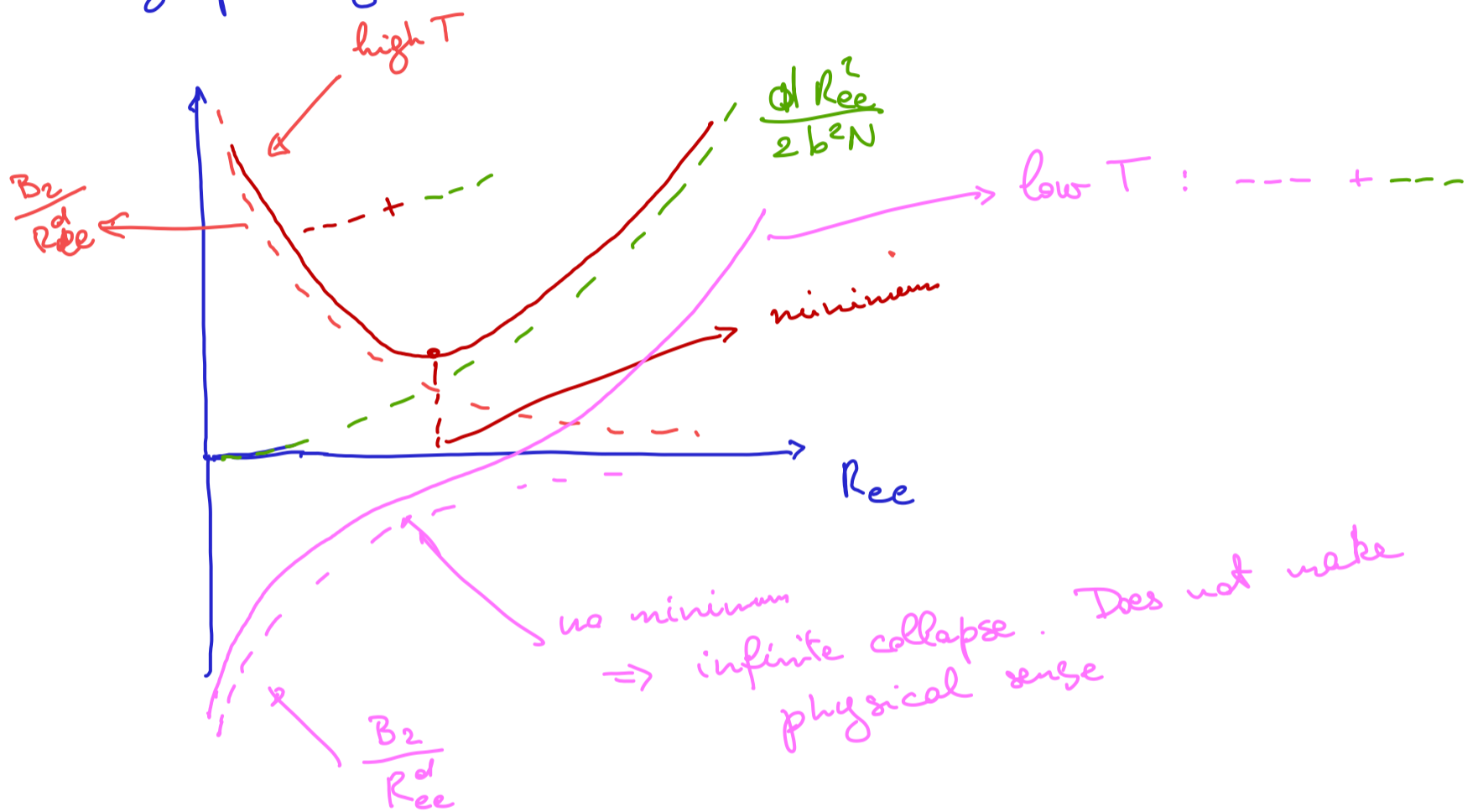
What if temperature is low?

$$B_2 > 0$$

and the free energy to minimize is

$$F(R_{ee}) = \frac{d}{2b^2N} R_{ee}^2 - k_B T \frac{N^2}{2} \frac{B_2}{R_{ee}}$$

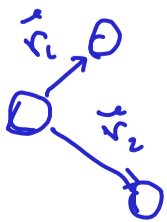
Graphically



Where does the analysis break down? As the polymer collapses, the density increases and we cannot stop at B_2 anymore.

We must also consider at least B_3 .

$$B_3 = \int d\vec{r}_1 d\vec{r}_2 \left[3 \varphi(\vec{r}_1) \varphi(\vec{r}_2) + \varphi(\vec{r}_1) \varphi(\vec{r}_2) \varphi(\vec{r}_2 - \vec{r}_1) \right] =$$



$$= 3 \left[\int \varphi(\vec{r}_i) d\vec{r}_i \right]^2 + \int d\vec{r}_1 d\vec{r}_2 \varphi(\vec{r}_1) \varphi(\vec{r}_2) \varphi(\vec{r}_2 - \vec{r}_1) =$$

$$= 3B_2^2 + \int d\vec{r}_1 d\vec{r}_2 \underline{\varphi(\vec{r}_1)\varphi(\vec{r}_2)\varphi(\vec{r}_2 - \vec{r}_1)}$$

If interactions are weak, the attractive part barely dominates over the repulsive one, but taking the products like in the underlined expression does not reduce the contribution of the repulsive part (always equal, in absolute value, to 1), but reduces the others.

When the volume collapses enough, the repulsive parts are relevant for the three bodies



giving a negative contribution (-1) before the integral, with the result that, for small enough volumes, the integral of the triple product is negative $\Rightarrow B_3 < 0$

We can thus write again the expression for the non-ideal gas:

$$F = -k_B T \ln Z \approx -k_B T \ln \left[\frac{V^N + \binom{N}{2} V^{N-1} B_2 + \binom{N}{3} V^{N-2} B_3}{\lambda^{3N} N!} \right] =$$

$$= -k_B T \left[\frac{V^N}{\lambda^{3N} N!} \right] - k_B T \ln \left[1 + \binom{N}{2} \frac{B_2}{V} + \binom{N}{3} \frac{B_3}{V^2} \right] \approx$$

$$= F_{P.G.} - \binom{N}{2} k_B T \frac{B_2}{V} - \binom{N}{3} k_B T \frac{B_3}{V^2} + \dots$$

$$\Rightarrow Z(V) = e^{-\beta F_{p.g.}} e^{\binom{N}{2} \frac{B_2}{V} + \binom{N}{3} \frac{B_3}{V^2}}$$

$$e^{-\frac{d R_{ee}^2}{2b^2 N}}$$

$$\Downarrow P(R_{ee}) = \frac{1}{2} R_{ee}^{d-1} e^{-\frac{d R_{ee}^2}{2b^2 N} + \frac{B_2}{2} \frac{N^2}{R_{ee}^d} + \frac{B_3}{6} \frac{N^3}{R_{ee}^{2d}}}$$

neglected the singular part: can be absorbed in the prefactor

We have now

$$B_3 < 0$$

$$B_2 \begin{cases} < 0 & \text{at high } T \\ = 0 & \text{at } T_0 \\ > 0 & \text{at low } T \end{cases}$$

$$\Rightarrow B_2 \approx -b_2 (T - T_0)$$

with $b_2 > 0$

At high T , then, we are exactly in the same situation as before \Rightarrow SAW

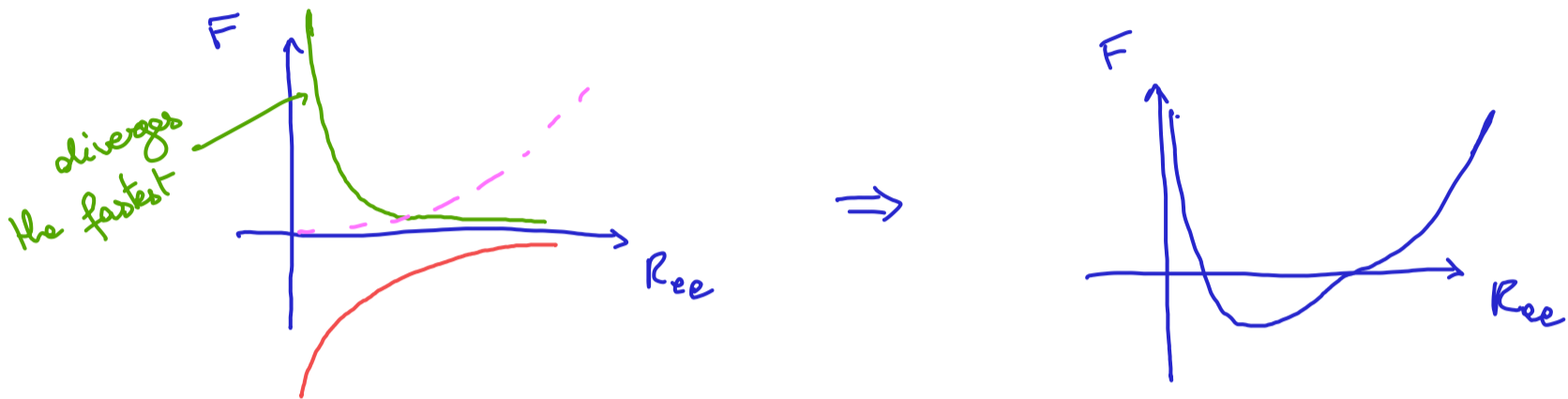
At low T , instead, it gets more complicated:

without B_3 , as we have seen, the polymer would indefinitely collapse. How does B_3 control it?

We can try to drop some term of the free-energy, one at a time, and see what happens:

$$\frac{F}{k_B T} = \frac{d}{2 b^2 N} R_{ee}^2 - \frac{B_2}{2} \frac{N^2}{R_{ee}^d} - \frac{B_3}{6} \frac{N^3}{R_{ee}^{2d}}$$

Graphically:



The third term cannot be neglected.
Which of the others can?

① Drop $\frac{R_{ee}^2}{N}$

$$\Rightarrow F \approx - \frac{B_2}{2} \frac{N^2}{R_{ee}^d} - \frac{B_3}{6} \frac{N^3}{R_{ee}^{2d}}$$

$$\Rightarrow \frac{dF}{dR_{ee}} = d \frac{B_2}{2} \frac{N^2}{R_{ee}^{d+1}} + 2d \frac{B_3}{6} \frac{N^3}{R_{ee}^{2d+1}} = 0$$

$$\Rightarrow R_{ee} \sim \left(\frac{B_3}{B_2} \right)^{\frac{1}{d}} N^{\frac{1}{d}}$$

$$B_3 \sim (\text{volume})^2$$

$$B_2 \sim \text{volume} \Rightarrow \left(\frac{B_3}{B_2} \right)^{\frac{1}{d}} \sim \text{length}$$

↓
OK

This is correct: once we reach max compactness, the volume is the sum of the volumes of the monomers, and the radius is just the $\frac{1}{d}$ power of the volume

Would the first term be dominant?

The second is

$$\frac{N^2}{R_{ee}^d} \Big|_{R_{ee} \sim N^{1/d}} \sim N^{2-1} = N$$

The third is

$$\frac{N^3}{R_{ee}^{2d}} \Big|_{R_{ee} \sim N^{1/d}} \sim N^{3-2} = N$$

The first is

$$\frac{R_{ee}^2}{N} \Big|_{R_{ee} \sim N^{1/d}} \sim N^{\frac{2}{d}-1} = N^{\frac{(2-d)}{d}}$$

$$\frac{2-d}{d} < 1 \quad \text{if } d > 1$$

\Rightarrow negligible

② Drop

$$B_2 \frac{N^2}{R_{ee}^d}$$

$$\Rightarrow F \approx \frac{d R_{ee}^2}{2b^2 N} - \frac{B_3}{6} \frac{N^3}{R_{ee}^{2d}}$$

$$\frac{dF}{dR_{ee}} = \frac{d}{b^2 N} R_{ee} + 2d \frac{B_3}{6} \frac{N^3}{R_{ee}^{2d+1}} = 0$$

$$\Rightarrow R_{ee}^{2(d+1)} = \left(-\frac{B_3 b^2}{3} \right) N^4$$

$$\Rightarrow R_{ee} = \left(-\frac{B_3 b^2}{3} \right)^{\frac{1}{2(d+1)}} N^{\frac{2}{d+1}}$$

Is the second term negligible?

First term: $\frac{R_{ee}^2}{N} \Big|_{R_{ee} \sim N^{\frac{2}{d+1}}} \sim N^{\frac{3-d}{d+1}}$

Third term: $\frac{N^3}{R_{ee}^d} \Big|_{R_{ee} \sim N^{\frac{2}{d+1}}} \sim N^{3 - \frac{4d}{d+1}} = N^{\frac{3-d}{d+1}}$

Second term: $\frac{N^2}{R_{ee}^d} \Big|_{R_{ee} \sim N^{\frac{2}{d+1}}} \sim N^{2 - \frac{2d}{d+1}} = N^{\frac{2}{d+1}}$

and $\frac{2}{d+1} < \frac{3-d}{d+1}$ if $2 < 3-d \Rightarrow \underline{d < 1}$
never!

The second term is thus non negligible, and we must thus compensate the second and third terms, unless $B_2 = 0 \Rightarrow T = T_0$ in which case

$$\boxed{v_\theta = \frac{2}{d+1}}$$

(Try repeating all the treatment we used for the SAW in this case!)

Interesting:

$$v_\theta = \begin{cases} 1 & d=1 & \text{ok} \\ \frac{2}{3} & d=2 & \Rightarrow \text{exact result} & v_\theta = \frac{2}{3} \\ \frac{1}{2} & d \geq 3 & \Rightarrow \text{correct.} \end{cases}$$

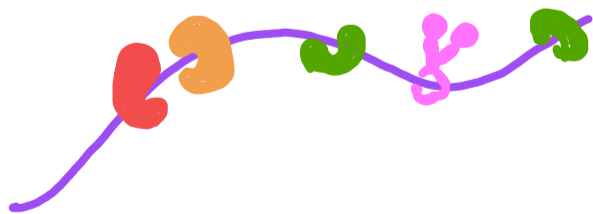
General considerations

In the case of proteins, the attractive interactions leading to the protein collapse (first, but by far not last stage of protein folding) are the hydrophobic ones and electrostatics between positively and negatively charged amino-acids (more on this in the next chapter).

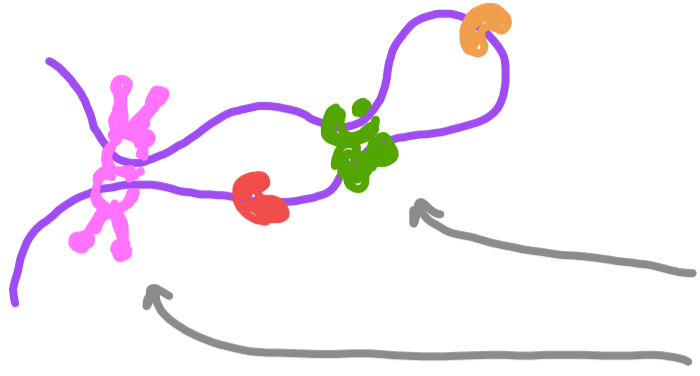
As we saw, DNA must be compactified to fit in the nucleus. But DNA is more or less homogeneously negatively charged! How can be compactified? What are the attractive interactions? That's why physics is useful:

does it really matter what causes the attractive interaction between different parts of DNA? Of course it does at the molecular level (e.g. if we want to intervene with drugs etc.) but as long as they are attractive, the final effect does not depend on the mechanism.

That said, of course the molecular mechanism is important! In the cell, DNA is not "naked", but is covered by a great number of DNA-binding proteins



Some of these proteins can form dimers (or higher order complexes)



they act as attractive potentials.

