Biophysical Chemistry

Applying polymer theory to biomolecules

Jonathan Doye

jonathan.doye@chem.ox.ac.uk

Applying polymer theory to biomolecules

Biology is full of polymers: proteins, DNA, RNA, polysaccharides, lipids.

Q: Can they usefully be modelled using polymer theory?

Outline:

1. Review of freely-jointed chain
2. Applications to genomic DNA
3. Entropic elasticity and pulling DNA
4. Beyond the freely-jointed chain: the Worm-like chain model

Review of freely-jointed chains

Also called random-flight model, Gaussian or ideal chain.

N segments of length l. The direction of a segment is uncorrelated with adjacent segments, i.e. freely jointed. Each segment represented by a vector $\mathbf{l}_i$.

Contour length $L = Nl$.

We can characterise the polymer by its mean square end-to-end distance $\langle R^2 \rangle$:

$$\langle R^2 \rangle = \langle \mathbf{R} \cdot \mathbf{R} \rangle = \left\langle \sum_i^N \mathbf{l}_i \cdot \sum_j^N \mathbf{l}_j \right\rangle = \sum_i^N \sum_j^N \langle \mathbf{l}_i \cdot \mathbf{l}_j \rangle \quad (1)$$

$$= \sum_i^N \langle \mathbf{l}_i \cdot \mathbf{l}_i \rangle + \sum_i^N \sum_{j \neq i}^N \langle \mathbf{l}_i \cdot \mathbf{l}_j \rangle = Nl^2 \quad (2)$$

i.e. $\langle R^2 \rangle^{1/2} = lN^{1/2}$

Can also characterize the polymer by its radius of gyration R_g, which is the mean separation betwen monomers.

For a freely-jointed chain $R_g^2 = \langle R^2 \rangle / 6 = N l^2 / 6$

Although the "chemical" monomers of a real polymer chain will not be freely-jointed, on longer length scales of s monomers a polymer can still effectively behave like a freely-jointed chain.

i.e. $\langle R^2 \rangle = N_K l_K^2$, where $l_K = s l_c$ is the Kuhn length and $N_K = N_c / s$. (c stands for chemical unit).

Deviations in this scaling can occur because of

(i) self-avoidance $\Rightarrow R_g \sim N^{3/5}$

(ii) effective attractions between the polymer (i.e. poor solvent)
$\Rightarrow R_g \sim N^{1/3}$.

Genome size

organism type	example	genome size (bases/base pairs)	contour length	R_g (FJC)
ssDNA virus	STMV	1063	$0.7\,\mu$m	15nm
dsDNA virus	bacteriophage T2	$150\,000$	$50\,\mu$m	$0.9\,\mu$m
prokaryote	E. coli	4.6×10^6	15 mm	$5\,\mu$m
eukaryote	human	$2 \times 3.2 \times 10^9$	2m	0.2 mm

R_g was calculated for double-stranded DNA (dsDNA) using a Kuhn length of 300 base pairs where the rise per base pair is 3.4 Å, and for single-stranded RNA (ssRNA) a Kuhn length of 20 Å and backbone separation between bases of 6.4 Å.

However, these genome dimensions are significantly larger than their ''containers'':

size of STMV capsid: 70 Å
size of T2 capsid: 100 nm
size of E. coli cell: 2 μm long, 1 μm wide
diameter of human nucleus: $6\,\mu$m

Q. Are the R_g reasonable?

Freed DNA!

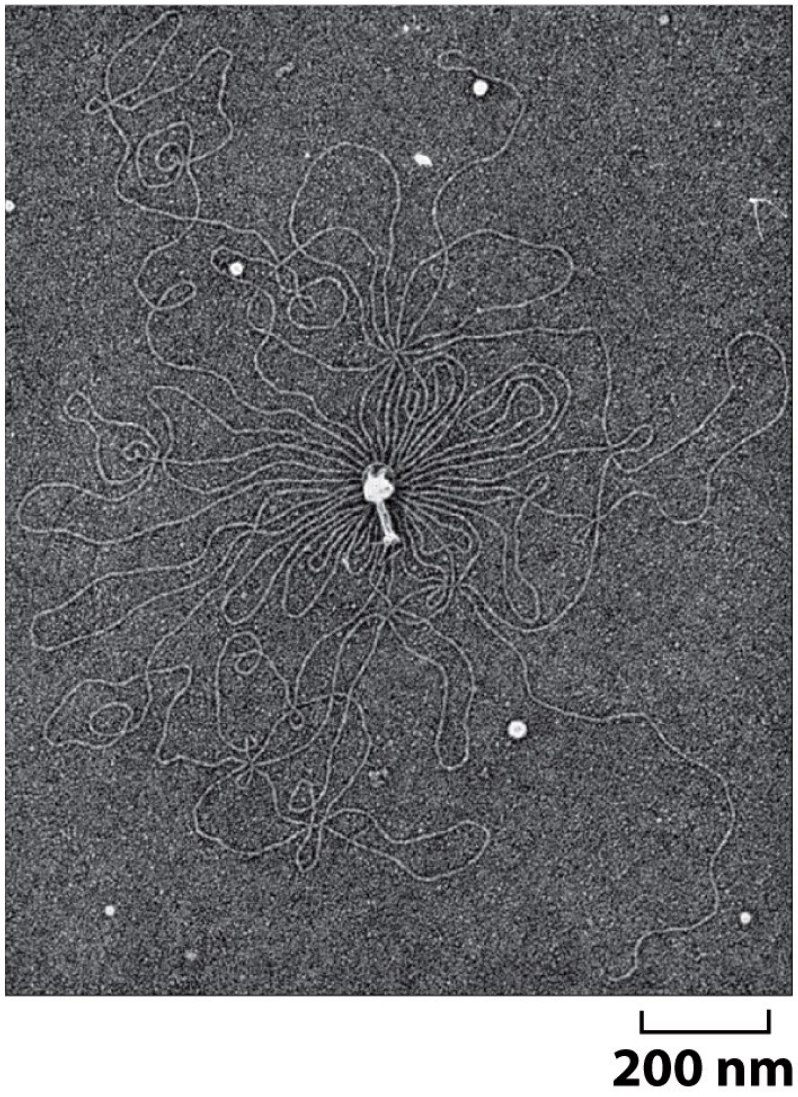


Figure 1.13 Physical Biology of the Cell (© Garland Science 2009)

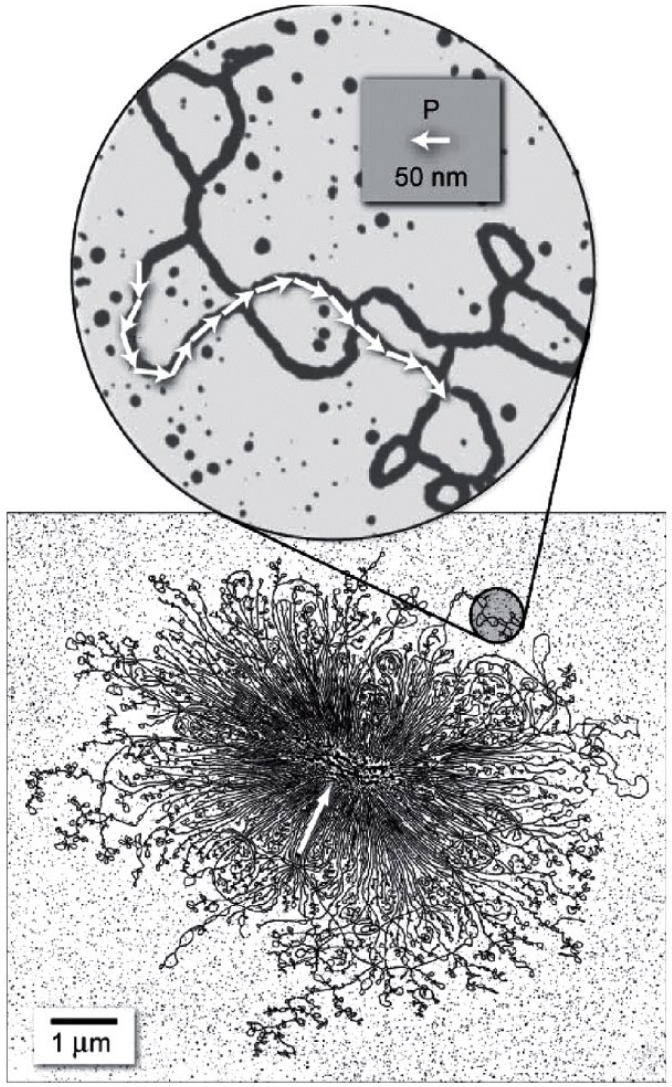
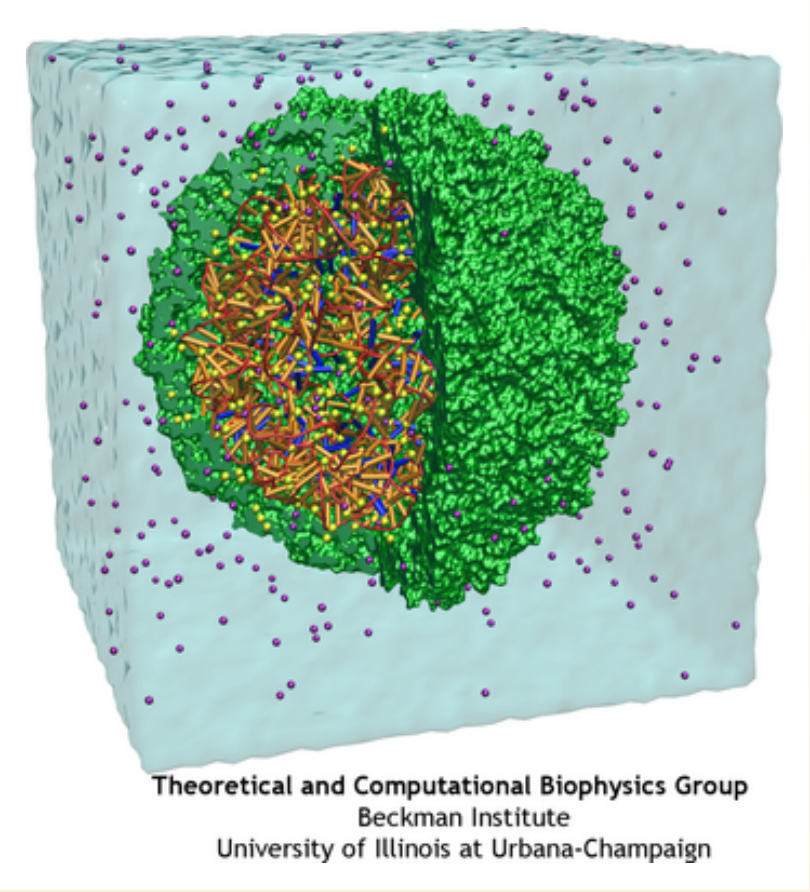


Figure 8.6 Physical Biology of the Cell (© Garland Science 2009)

Packaging DNA: ssRNA virus



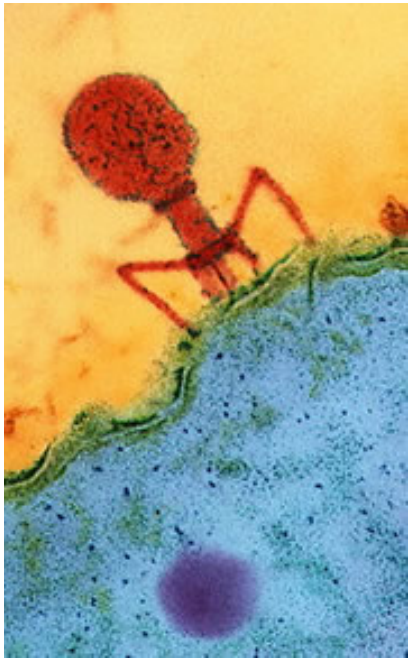
The active virus particle can spontaneously self-assemble from a solution containing the genome and the virus capsid protein.

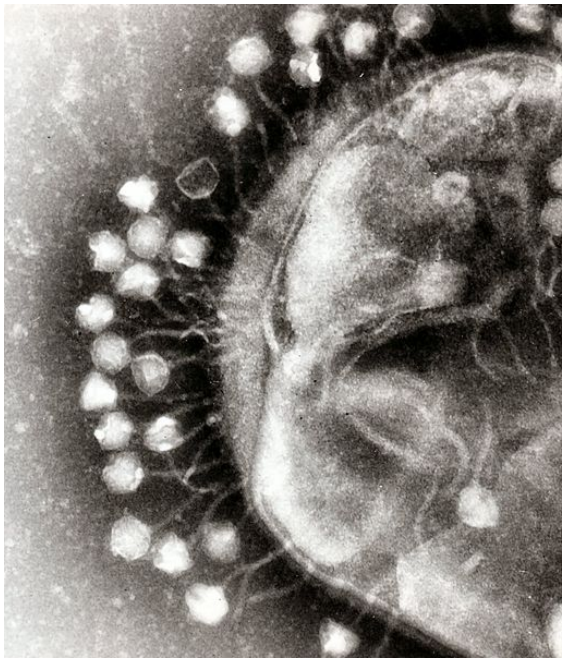
The compression of the RNA must be compensated by attractive interactions with the inside of the capsid.

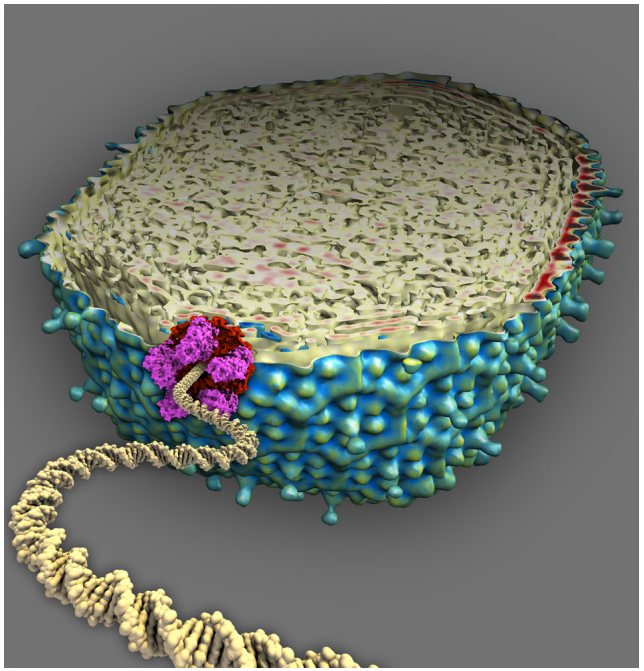
Packaging DNA: dsDNA virus

Double-stranded DNA is much stiffer than single-stranded RNA (or DNA) and so a different approach is needed.

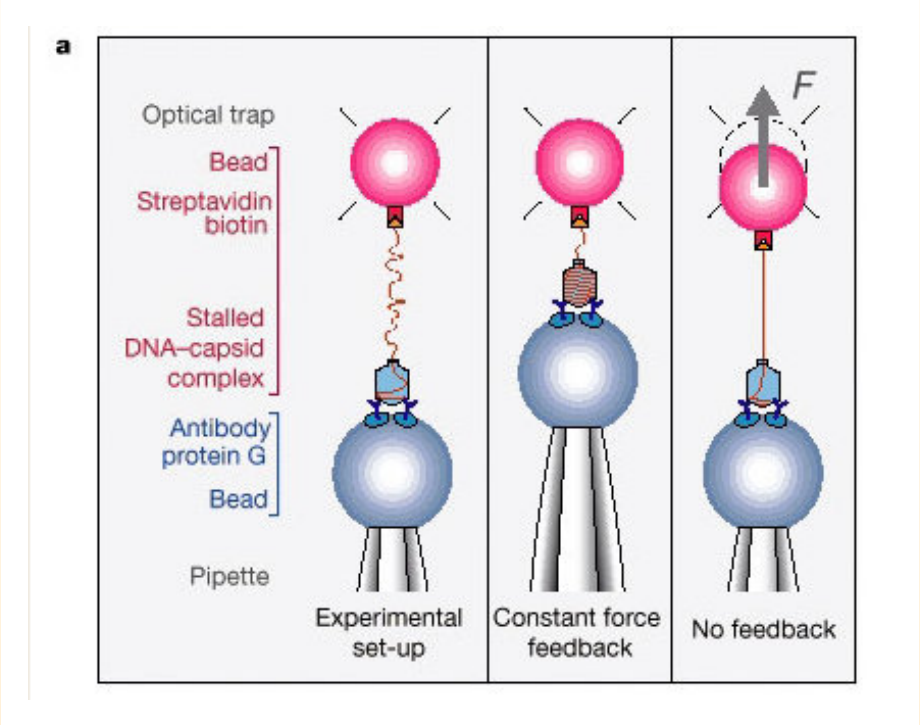
Bacteriophages have a packaging motor located at one vertex of the capsid that is able to push DNA into the empty capsid. It uses the cell's ATP, as the source of energy for the work that it does.

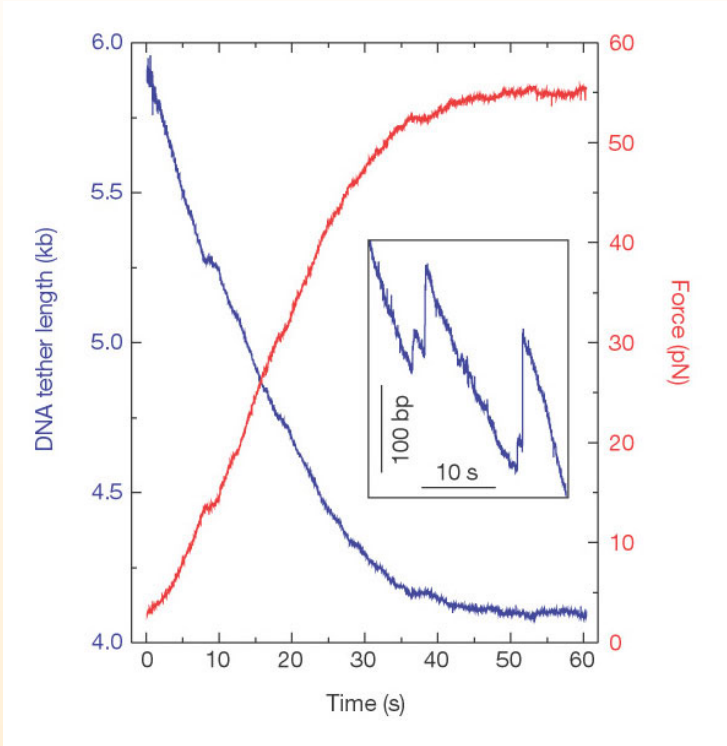






Single-molecule experiments have studied the forces that the motor is able to exert on DNA as the genome is packaged.

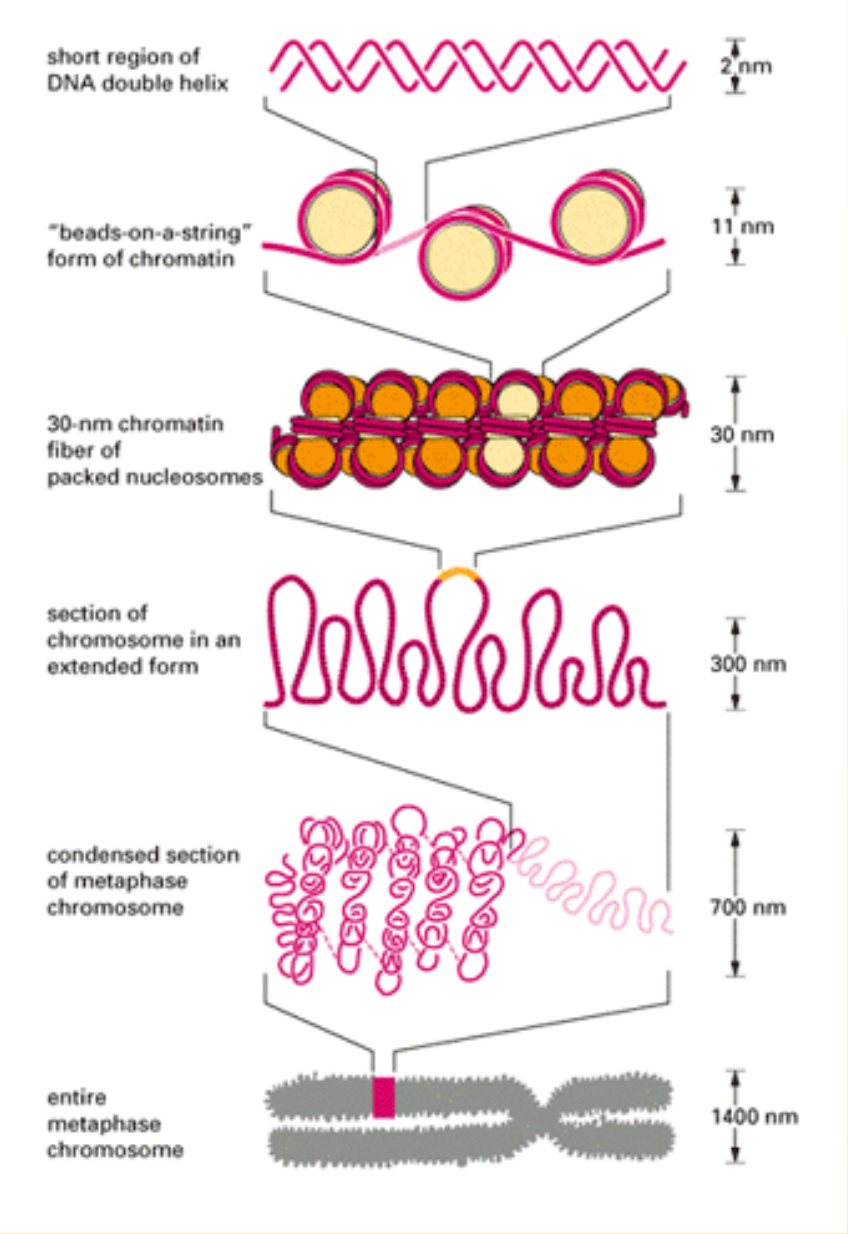




It has been estimated that the motor has to work against an internal force of $\sim 50pN$ when packaging the final part of the genome, and that the internal pressure is 6 MPa.

The pressure provides the driving force for the injection of the DNA into a cell during infection.

Packaging DNA: eukaryotic cells



Free energy cost for bending DNA around histones is compensated by binding free energy.

Elasticity in the FJC model: Hooke

A freely-jointed chain can show elastic behaviour that is purely entropic.

When subject to a force there is a free energy cost to pay associated with forcing the chain to adopt an entropically less likely conformation.

It can be shown that probability distribution for the end-to-end vector $\mathbf{R}$ for a freely-jointed chain is a Gaussian:

$$p\left(\mathbf{R}\right) \propto \exp\left(\frac{-3R^2}{2Nl^2}\right) \tag{3}$$

Hence, the entropy is $S = k\log(p(\mathbf{R})) + c$.

As all conformations of a freely-jointed chain have the same energy

$$F = -\frac{dA}{dz} = kT\frac{d}{dz}\log\left(p(\mathbf{R})\right) = kT\frac{d}{dz}\left(-\frac{3z^2}{2Nl^2}\right) = -\frac{3kTz}{Nl^2} \tag{4}$$

A freely-jointed chain obeys Hooke's law with a force constant that increases with temperature.

Elasticity in the FJC model: Beyond Hooke

The previous derivation assumed that the chain continues to have a Gaussian probability distribution even when subject to a force.

Here we calculate the partition function of a freely-jointed chain subject to a force:

$$Z = \int d\hat{\mathbf{l}}_1 \int d\hat{\mathbf{l}}_2 \ldots \int d\hat{\mathbf{l}}_N \exp\left(-E\left(\{\hat{\mathbf{l}}_i\}\right)/kT\right) \tag{5}$$

We can make progress by noting that the energy is separable

$$E\left(\{\hat{\mathbf{l}}_i\}\right) = -Fz = -F\mathbf{R}\cdot\hat{\mathbf{z}} = -F\sum_i^N \mathbf{l}_i\cdot\hat{\mathbf{z}} = -Fl\sum_i^N \cos\theta_i. \tag{6}$$

As each link is independent $Z = Z_1^N$ where

$$Z_1 = \int_0^{2\pi} d\phi_1 \int_0^{\pi} \sin\theta_1 e^{Fl\cos\theta_1/kT} d\theta_1 = 2\pi\left[-\frac{kT}{Fl}e^{Fl\cos\theta_1/kT}\right]_0^{\pi} \tag{7}$$

$$= 4\pi\frac{kT}{Fl}\sinh\left(\frac{Fl}{kT}\right) \tag{8}$$

This can then be used to obtain a expression for the force-extension curve:

$$\langle z \rangle = -\frac{\partial A}{\partial F} = kT\frac{\partial \log Z}{\partial F} = NkT\frac{\partial}{\partial F}\left(\log\left(\sinh\left(\frac{Fl}{kT}\right)\right) - \log F\right) \quad (9)$$

$$= Nl\left(\coth\left(\frac{Fl}{kT}\right) - \frac{kT}{Fl}\right) \quad (10)$$

At low force ($F \ll kT/l$) this simplifies to Hooke's Law:

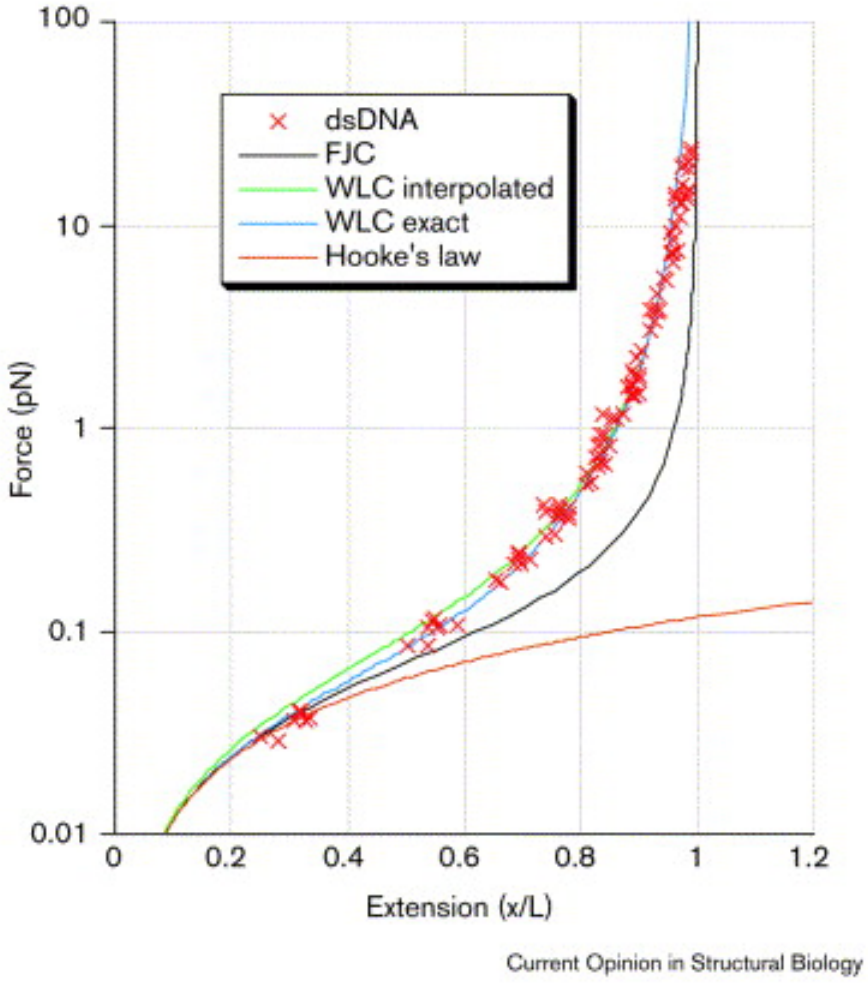
$$F = \frac{3kT}{Nl^2}z \quad (11)$$

and at high force ($F \gg kT/l$) to

$$F = \frac{kT}{l}\frac{1}{1 - z/L} \quad (12)$$

The force diverges as the polymer approaches its contour length.

Stretching DNA



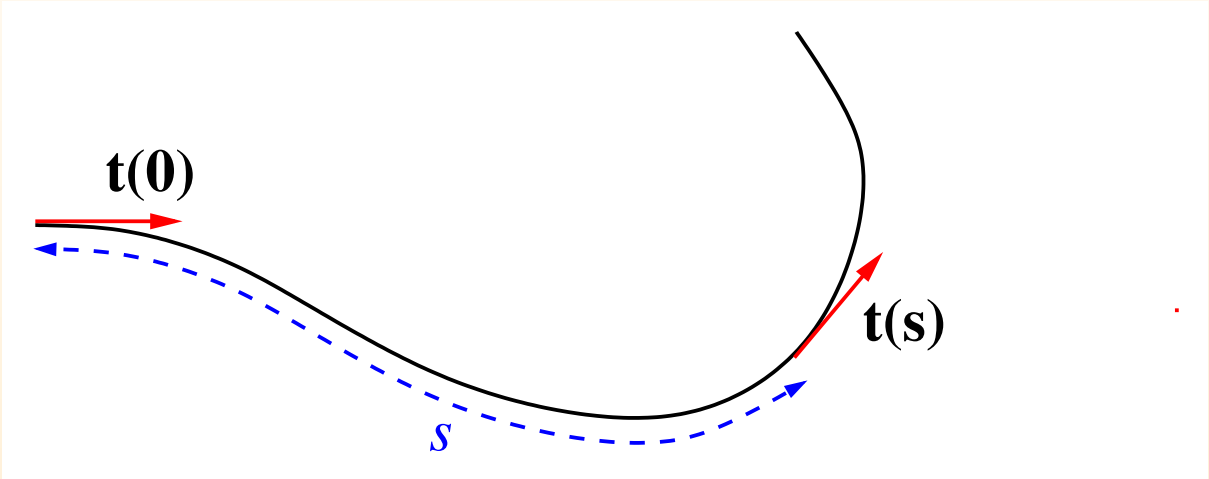
The force-extension curve can be obtained for a single molecule of DNA can be obtained using an optical trap set-up.

The freely-jointed chain provides a good description of the stretching, but there are deviations at higher force.

The worm-like chain model provides an excellent description.

Worm-like chain model

The worm-like chain model usually provides a better representation of relatively stiff polymers, such as double-stranded DNA, actin fibres and microtubules.



The polymer is represented by a continuous flexible rod that can be described in terms of a bending modulus or a persistence length l_p.

The persistence length describes the decay in correlations of the tangent vectors:

$$\langle \mathbf{t}(0) \cdot \mathbf{t}(s) \rangle = e^{-s/l_p} \tag{13}$$

This allows the mean square end-to-end distance to be calculated:

$$\langle R^2 \rangle = \langle \mathbf{R} \cdot \mathbf{R} \rangle = \left\langle \int_0^L \mathbf{t}(s) ds \cdot \int_0^L \mathbf{t}(s') ds' \right\rangle \tag{14}$$

$$= \int_0^L ds \int_0^L \langle \mathbf{t}(s) \cdot \mathbf{t}(s') \rangle ds' = \int_0^L ds \int_0^L e^{-|s-s'|/l_p} ds' \tag{15}$$

$$= 2 l_p L \left(1 + \frac{l_p}{L} \left(e^{-L/l_p} - 1 \right) \right) \tag{16}$$

For $L \gg l_p$ this simplifies to the freely-jointed chain result where $l_K = 2 l_p$.

The force-extension relation can also be derived:

$$F = \frac{kT}{l_p} \left[\frac{1}{4 \left(1 - z/L \right)^2} - \frac{1}{4} + \frac{z}{L} \right] \tag{17}$$