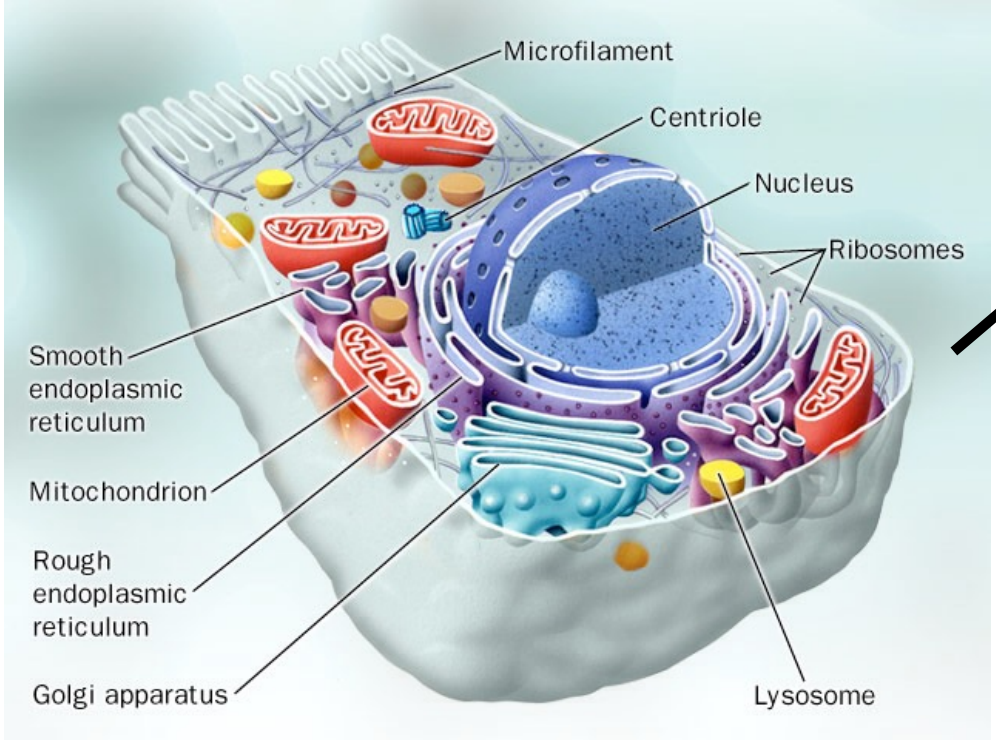
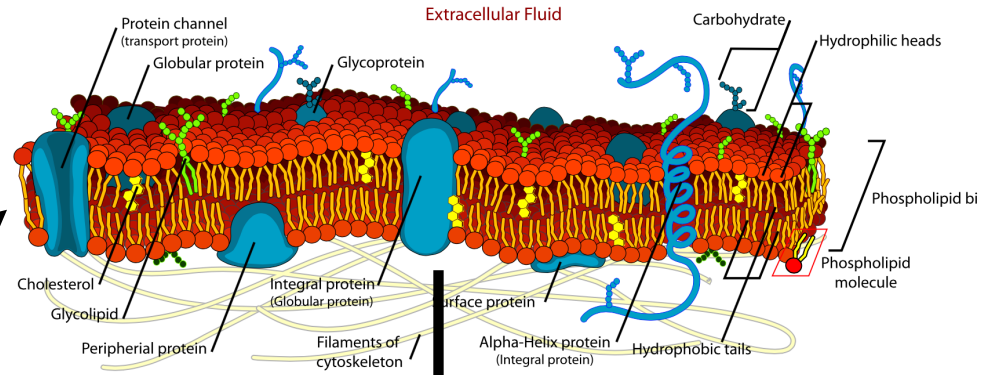
Core Concepts

- Biological polymers are interesting because they carry information at many scales - monomer/polymer/aggregates
- Entropy dominates polymer behaviour
- A simple random walk model of a polymer is surprisingly useful

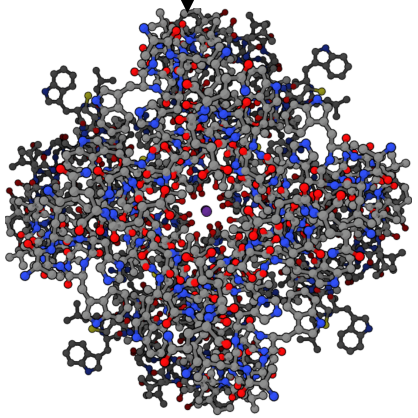
Macromolecules in the cell



www.daviddarling.info



U. Ghent



wikipedia.org

Apart from water, ions, small molecules like ATP, the cell contains many supramolecular structures - membranes, organelles, protein complexes, DNA/chromatin, etc.

Each structure is composed of smaller structures, which are composed of smaller…. down to molecules and atoms in a hierarchy of structural units on different length scales

Polymers and Macromolecules

Industrial polymers have been in use since ~ 1945, and consist of simple repeated units:

Polyethylene $(C_2H_4)n$
Polyethyleneglycol (PEG) $(CH_2\text{-}CH_2\text{-}O)_n$

Industrial polymers are very important, but dull: plastic bags, door knobs, ….

Biological Polymers

DNA/RNA - linear chain of nucleotides

Lipids - approx. 2 - 4 nm - membranes

Sugars - 2 dim. networks

Proteins - linear chains of amino acids

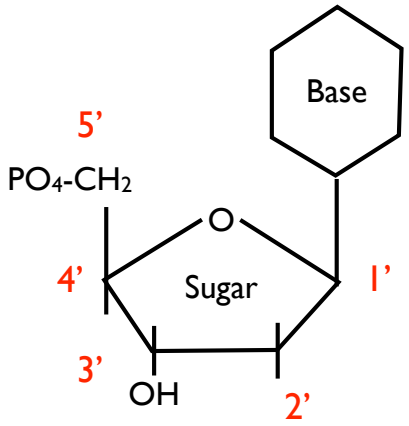
Filament	Monomer	Location	Diameter	Mass
DNA	A, C, G, T Base	Nucleus	2 nm	0.65 kDa
Actin	Actin monomer	cytoplasm	7 nm	43 kDa
Intermediate filaments	α-helical rod	cytoplasm	10 nm	70 - 200 kDa
Microtubules	Tubulin monomer	cytoplasm	25 nm	50 kDa

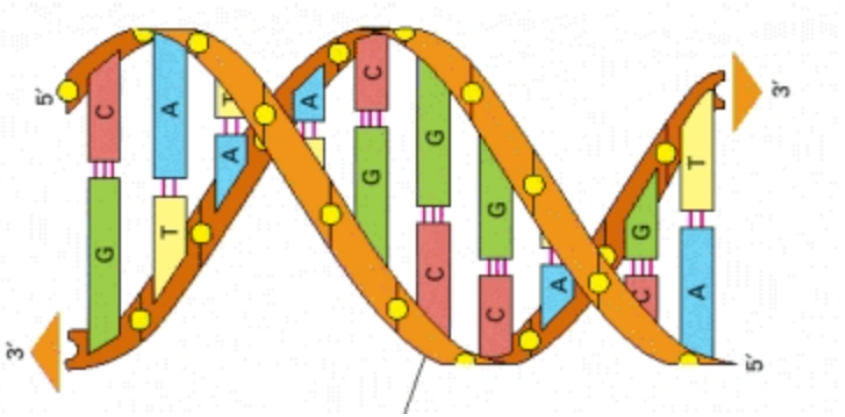
Small, folded proteins can also self-assemble into polymers, e.g., actin filaments, microtubules

Why are biological polymers so much more interesting than industrial polymers?

DNA

Double-stranded helix of nucleotides connected in linear chains





Alberts and Bray et al.

Nucleotide = 5-carbon sugar with an attached base (A, C, G, T) on the 1' carbon, and a phosphate group on the 5' carbon.

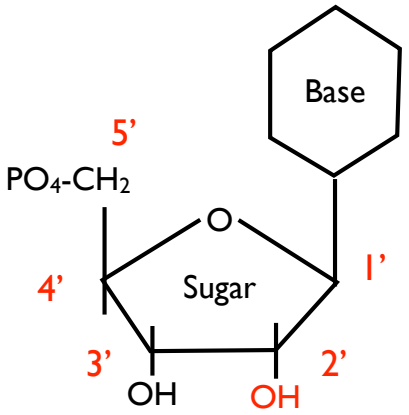
Base = N-containing ring compound, either purine or pyrimidine

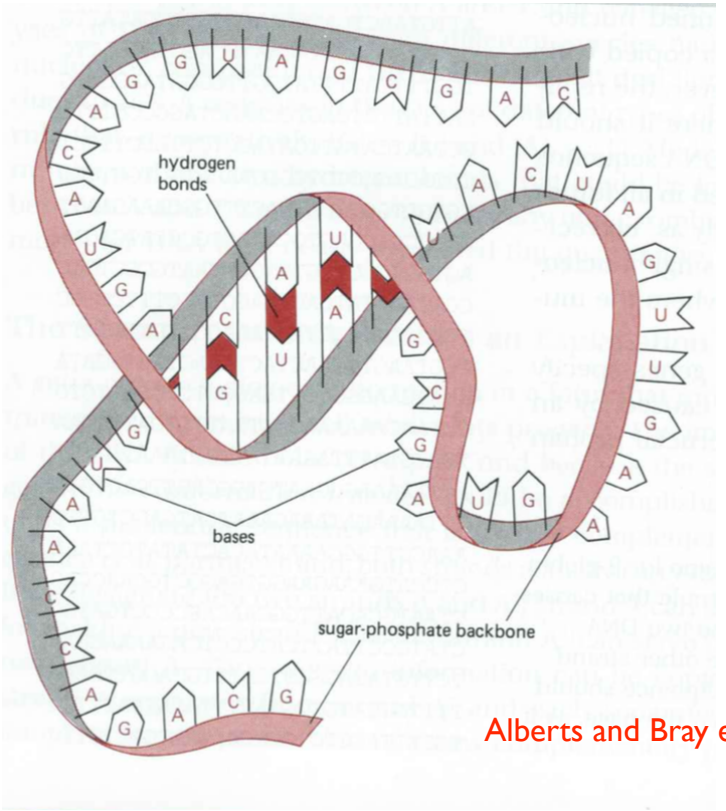
Chain is formed by phosphodiester bonds at 3' and 5' carbons, so the polymer has a direction or *polarity*

First reason biopolymers are interesting: the monomers have structure that carries information (industrial polymers are usually one monomer, e.g., CH_2, repeated)

RNA

Single-stranded linear chain of nucleotides - only difference in the nucleotide to DNA is the OH group on 2' carbon - and T is replaced by U





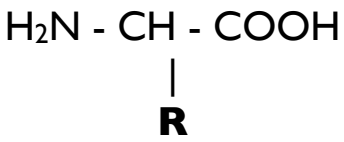
Alberts and Bray et al.

RNA is an intermediate between DNA and proteins (Central dogma: DNA —> RNA —> Protein)

But RNA is also a structural part of cellular machines like Ribosomes, it can catalyse reactions, and has some complementary binding.

Proteins

Linear polymers of amino acids directionally connected by peptide bonds from the amine end (N terminus) to the carboxylic acid end (C terminus):



R can be: Acidic, Basic, Uncharged polar, Non-polar

Proteins fold intro 3D shapes that minimise their energy by arranging charged polar/acidic/basic side-chains to cover the surface and shield non-polar side-chains from water.

Folded state is not covalently bonded, there must be many weak non-covalent interactions to stabilise the folded state agains thermal fluctuations.

Proteins can de-nature if solvent conditions change (pH, temperature, other proteins, …)

Name	Symbol	Type
Alanine	A	Non-polar
Cysteine	C	Non-polar
Aspartic acid	D	Acidic
Glutamic acid	E	Acidic
Phenylalanine	F	Non-polar
Glycine	G	Non-polar
Histidine	H	Basic
Isoleucine	I	Non-polar
Lysine	K	Basic
Leucine	L	Non-polar
Methionine	M	Non-polar
Asparagine	N	Uncharged polar
Proline	P	Non-polar
Glutamine	Q	Uncharged polar
Arginine	R	Basic
Serine	S	Uncharged polar
Threonine	T	Uncharged polar
Valine	V	Non-polar
Tryptophan	W	Non-polar
Tyrosine	Y	Uncharged polar

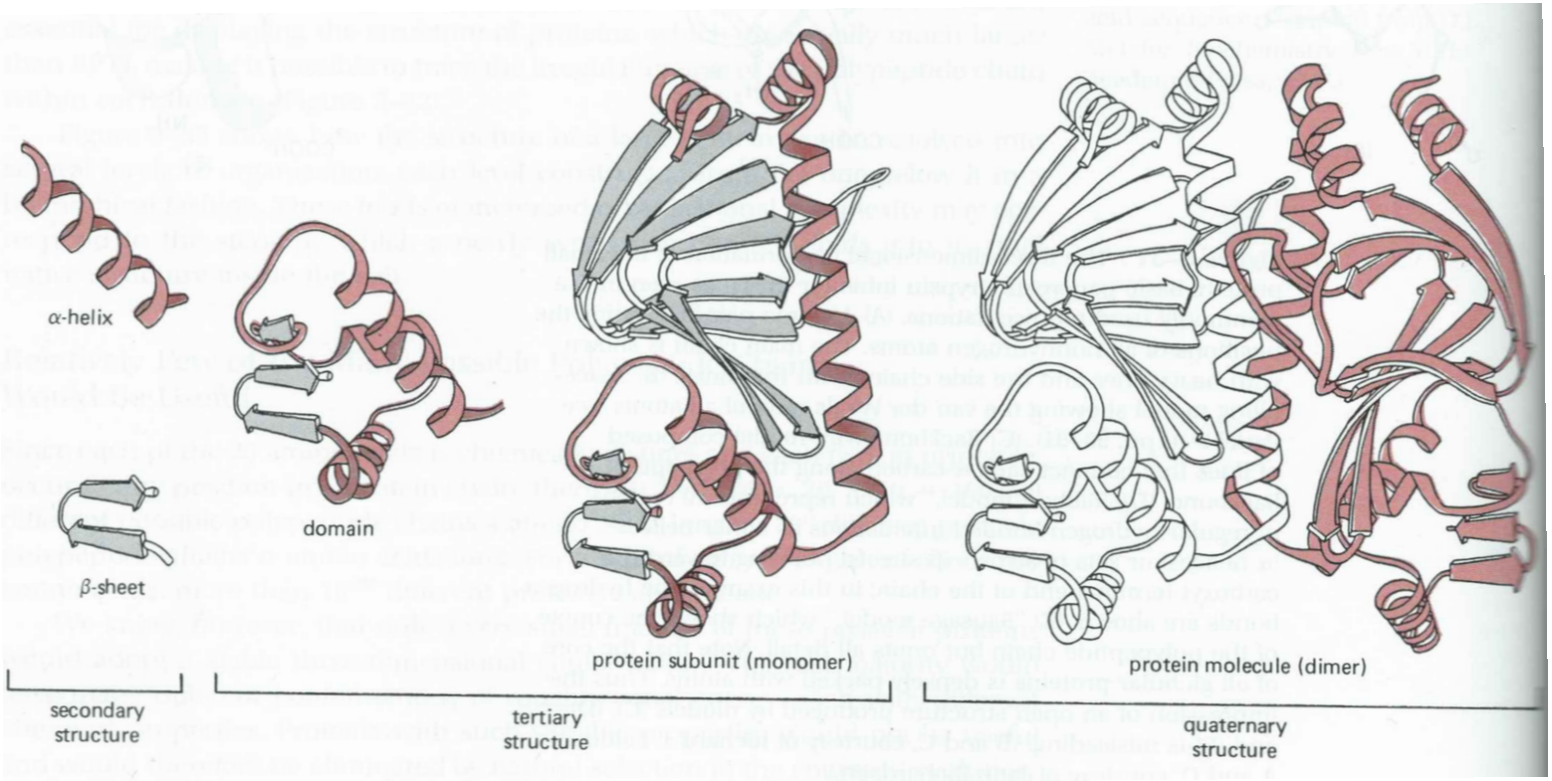
3D structure of proteins

Primary structure = sequence of amino acids

Secondary structure = H-bonding of contiguous aa's into α helices and β sheets

Tertiary structure = domains of globular units in single protein chain

Protein complexes = assembly of several proteins by non-covalent bonds



Second reason: biological polymers fold into shapes to perform functions (not just random flapping about)

Alberts and Bray et al.

Lipids

Lipids are fatty acids that are composed of:

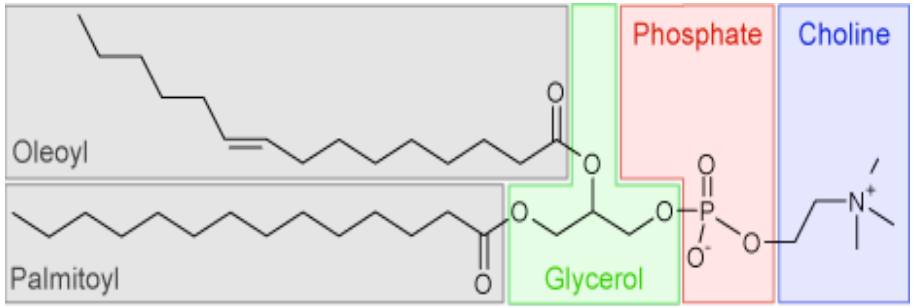
Headgroup - soluble in water, may be charged, uncharged, polar, small or bulky

Oily tails - C-C chains with typical length 8 - 24, saturated or unsaturated, only soluble in organic solvents

Lipids are called:

Amphiphiles = love water and oil - frustrated in either bulk solution

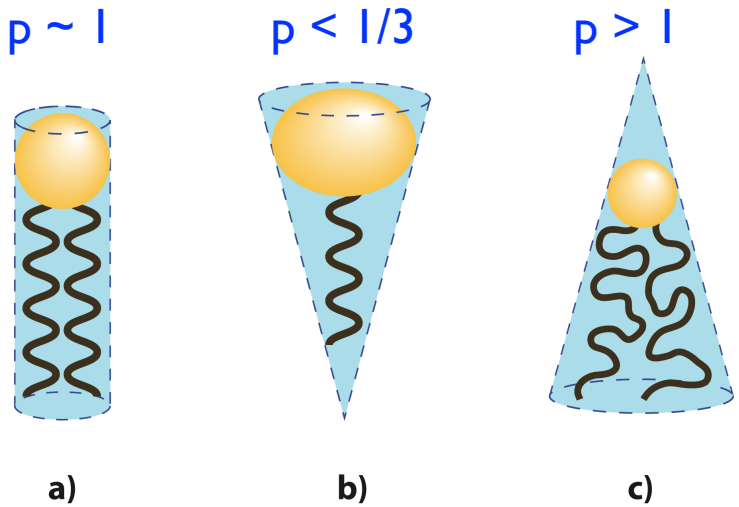
Typical phospholipids look like POPC below: palmitoyl-oleoyl phosphatidylcholine



Lipids do not covalently bond to each other, but self-assemble into distinct 3d structures driven by their frustration: **tails want to hide from water, headgroups want to be solvated.**

Lipids have a shape

Lipids have a *shape* controlled by relative size/interactions of their headgroup and tails



J. C. Shillcock, Fig. 3, Ch. 26 in Biomolecular Simulations,
ed. L. Monticelli and E. Salonen, Methods in Mol. Biol. 924, Humana Press 2012

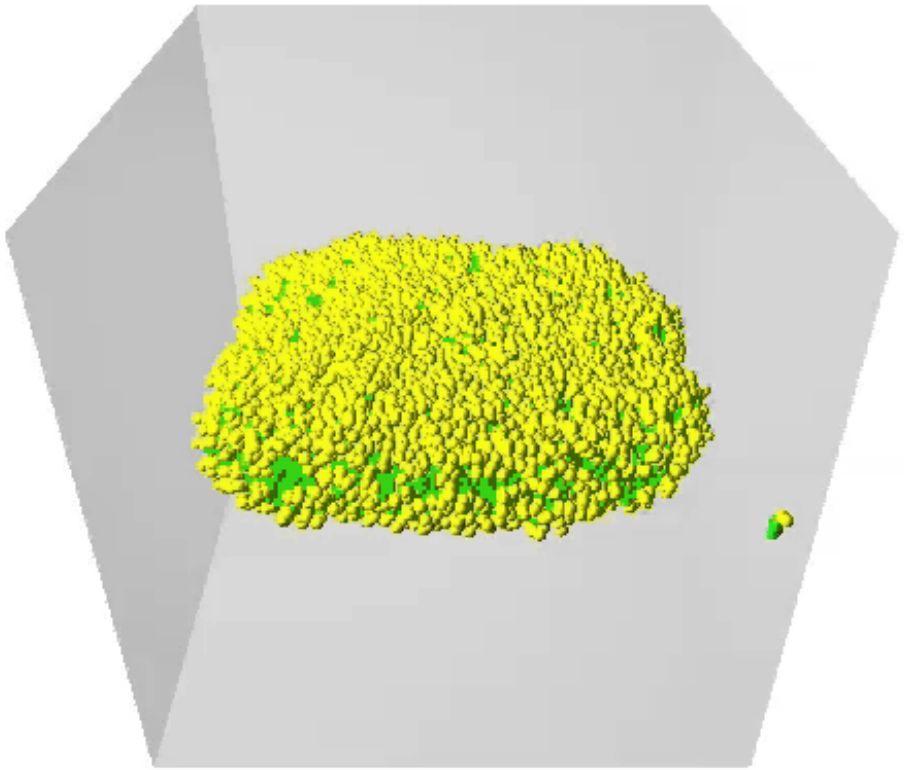
If the headgroup has the same "cross-sectional area" as the tails the molecule is like a cylinder; otherwise the molecule is like a cone.

Israelachvili devised a *packing parameter* to quantify the "shape" of lipids in the fluid phase:

$$p = v / a_0.l_c$$

v = equilibrium volume of the molecule (depends on environment, temperature, etc!)
a_0 = equilibrium cross-sectional area of the molecule (" ")
l_c = maximum extension of the hydrocarbon chains

Lipids have many phases



DPD simulation of 3428 lipids in a planar disk: the lipids on the edge have tails exposed to water; the patch bends so that it can close off its edge and forms a vesicle (water is invisible).

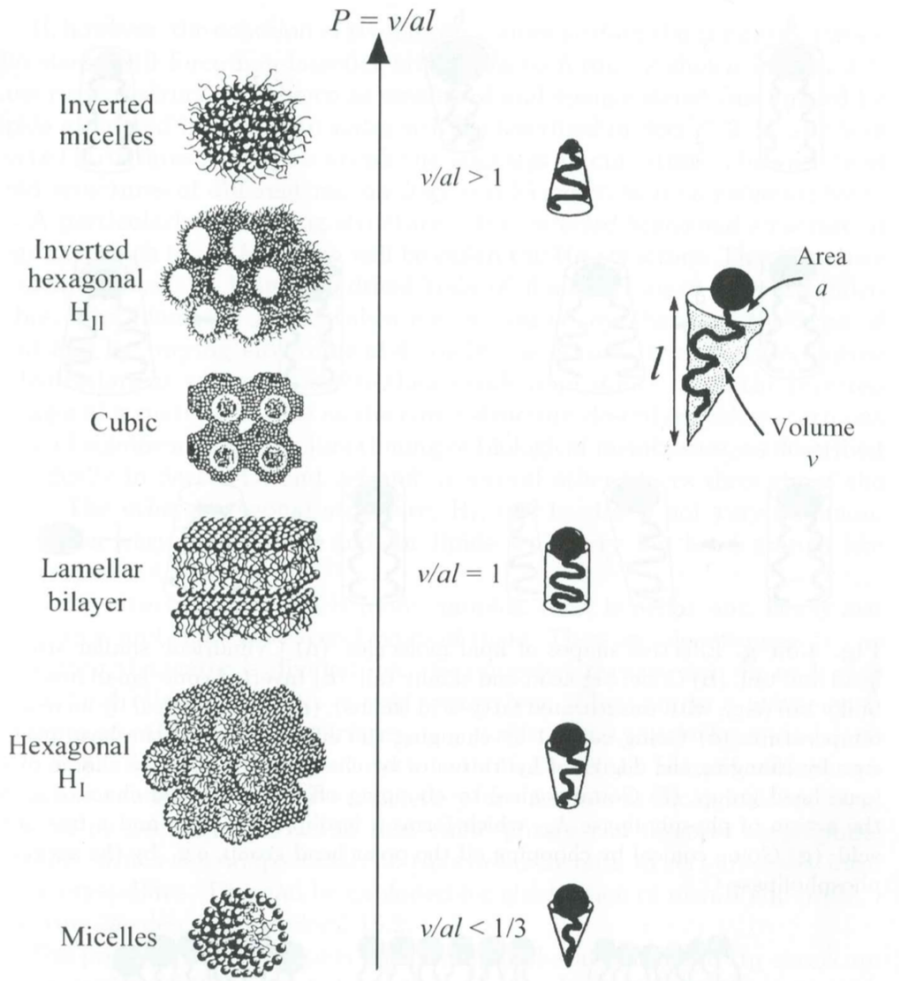


Fig. 4.4. Schematic illustration of lamellar and non-lamellar lipid aggregates formed in water. The different structures have different curvature and are arranged in accordance with the value of packing parameter $P = v/al$

Third reason: biopolymers non-covalently form aggregates that can respond to their environment by changing shape or structure.

Life as a Matter of Fat,
O. G. Mouritsen, Springer, Berlin 2005

What's special about lipids?

Lipids are often viewed as just providing a bounding membrane around cells that separates interior from exterior and is a solvent for proteins.

But they have important active roles too:

Signalling - endocannabinoids and steroid hormones

Disease - important lipid-binding proteins: PLAs (inflammation), ApoE (Alzheimer's), alpha synuclein, PUFAs (Parkinson's)

Synaptic dynamics - neural membranes are 1/2 PUFAs; modify ion channel dynamics

B. Davletov and C. Montecucco, Lipid function at synapses. Curr. Op. Neurobiol. 20:543 (2010)

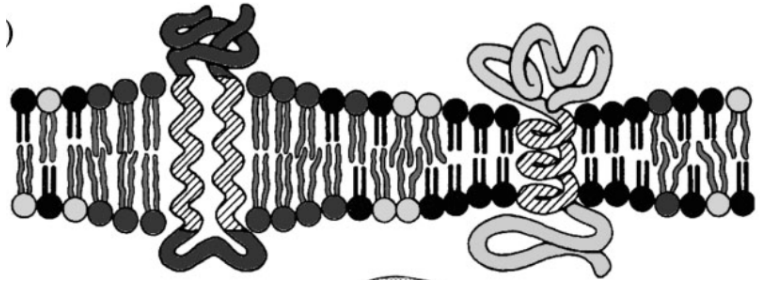
Why are there so many different types of lipid in the body?

- If all nature needed was a barrier, a few lipid types would be sufficient - but there are 1000s of different types of lipid in cell membranes (cp. 4 bases in DNA and 20 amino acids)
- Membrane is a fluid so all components could diffuse freely - but different proteins need distinct environments due to thickness, tension, and their function depends on local lipid composition: membrane is not an ideal mixture

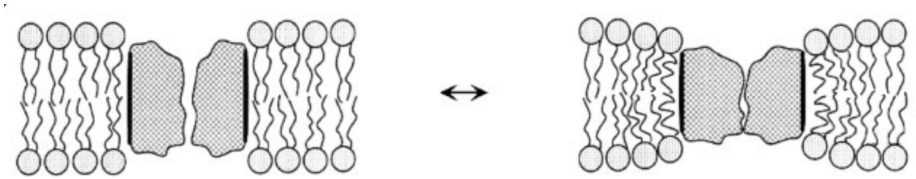
Lipids influence membrane material properties

Thickness, bending modulus, stretch modulus, fluidity

Protein hydrophobic region must match the membrane **thickness** **(or pay high energy cost)**, hence they prefer to be surrounded by lipids with matching thickness



Channel proteins can have their open/closed equilibrium changed by **local membrane constituents**



Membrane **bending stiffness** (or degree of saturation of lipids) controls their shape; flexible lipids can form curved membranes, vesicles, tubes, rigid lipids form flat bilayers;

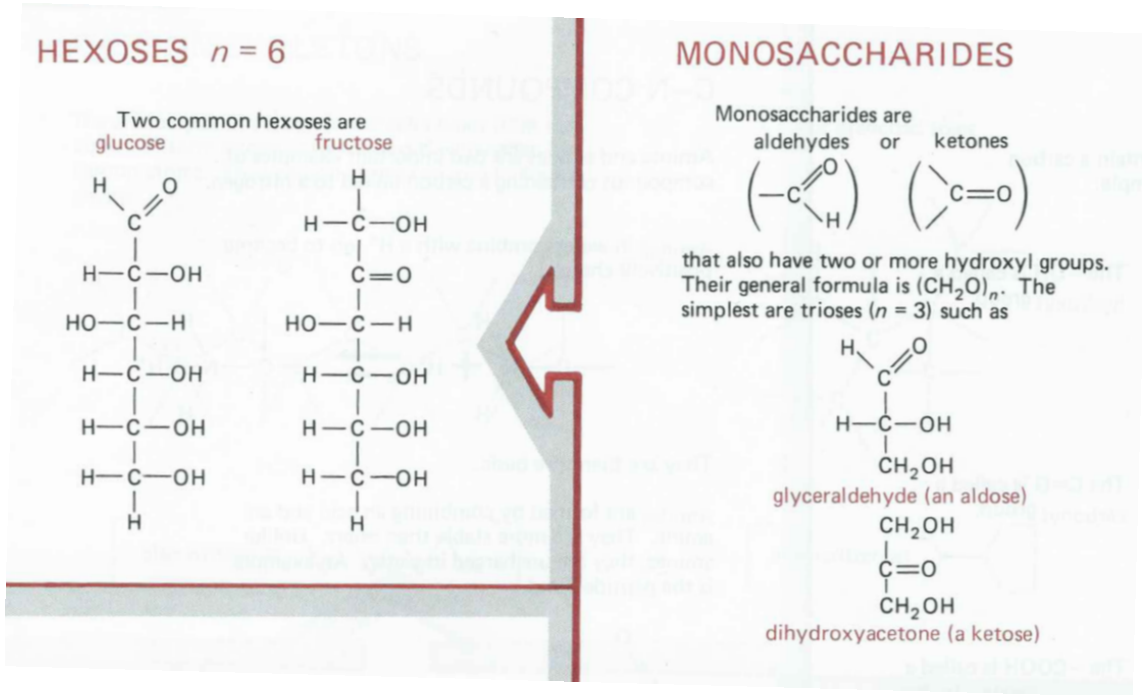
Stretch modulus is high which maintains surface area constant

Fluidity of membrane components allows cell to create/remove lipids that associate by diffusion; transient domains can form to aid signalling, tension can be relieved by lipid flow

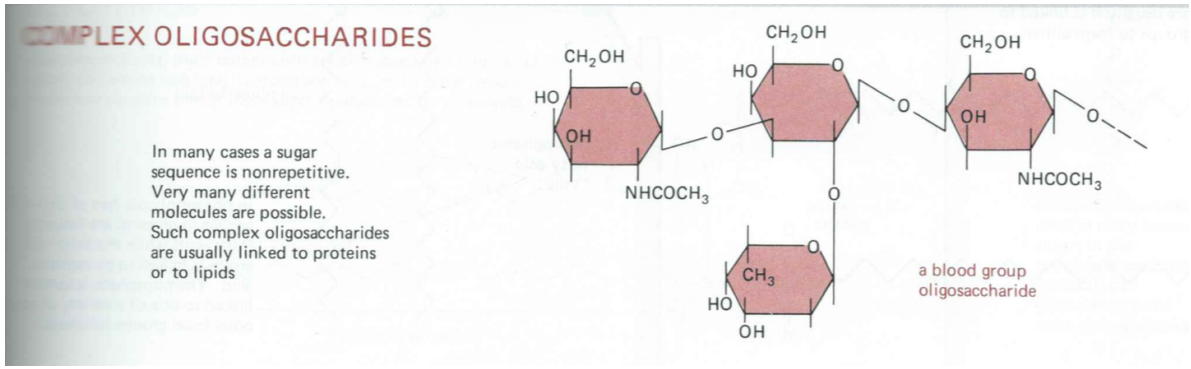
Sugars

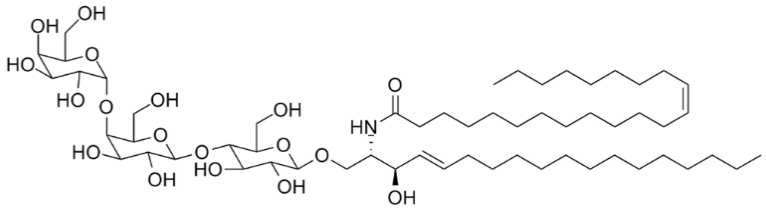
Monosaccharides have formula C_n-H_{2n}-O_n

Glucose is a hexose - C_6-H_{12}-O_6
Ribose is a pentose - C_5-H_{10}-O_5



Complex oligosaccharides often bind to proteins or lipids (glycoprotein, glycolipid) to encode information in the immune system, signalling, pathogen entry to cells.



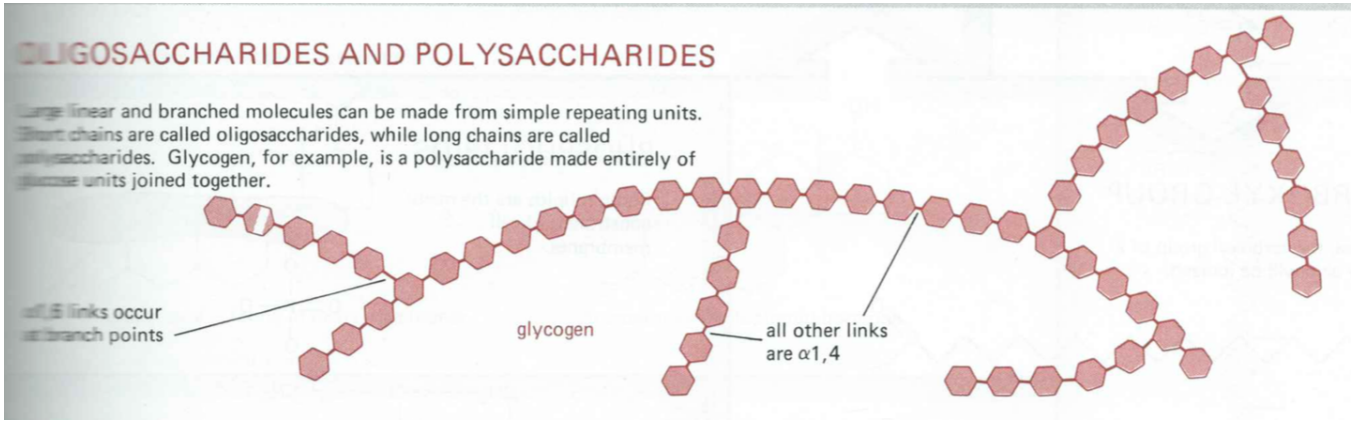


Shiga toxin binds to Gb3 to enter cell

Alberts and Bray

Sugars can be structurally complex

Sugars can be linear or ring-like, and can link to form long linear chains (e.g., glycogen), or branched networks (via OH groups).



(residues). Because each monosaccharide has several free hydroxyl groups that can form a link to another monosaccharide (or to some other compound), the number of possible polysaccharide structures is extremely large. Even a simple disaccharide consisting of two glucose residues can exist in 11 different varieties (Figure 2–4), while three different hexoses ($C_6H_{12}O_6$) can join together to make several thousand different trisaccharides. For this reason it is very difficult to determine the structure of any particular polysaccharide; with present methods it takes longer to determine the arrangement of half a dozen linked sugars (for example, those in a glycoprotein) than to determine the nucleotide sequence of a DNA molecule containing many thousands of nucleotides.

43 The Chemical Components of a Cell

Alberts and Bray

Macromolecules form complex aggregates

DNA/RNA - helical chains, DNA (2 nm) has hierarchical folding of histones (11 nm), packed nucleosomes (30 nm), chromosome (300 nm);

RNA makes supramolecular machines with proteins (Ribosomes) and phase separated droplets (stress granules, stalled in translation)

Proteins - rigid rods, floppy polymers, globular structures, non-covalently bonded aggregates, RNA/protein complexes

Lipids - micelles, membranes, vesicles, lipid-protein aggregates (e.g., LDL, HDL)

Sugars - flexible chains, 2 dim. networks, bonded to proteins/lipids to form glycoproteins/glycolipids

Biological polymers are interesting because …

1) Monomers have structure that carries information, e.g., nucleotides, amino acids (industrial polymers are usually one monomer, e.g., CH_2, repeated or random)

2) Biopolymers can fold into well-defined shapes to perform functions (not just random flapping about)

3) Biopolymers can non-covalently assemble into aggregates that respond to their environment by changing shape or structure, e.g., actin monomers fold into rigid shapes that subsequently self-assemble into filaments.

Industry is interested in making *smart* polymers that are responsive to their environment and can perform functions:

Smart polymeric materials: emerging biochemical applications
Roy and Gupta, Chemistry and Biology 10:1161-1171 (2003)

Advancing nanogel engineering for drug delivery
Raemdonck et al. Soft Matter 5:701-715 (2009)

Natural polyphenol surfactants: solvent-mediated spherical nanocontainers and their stimuli-responsive release of molecular payloads, Payra et al. Chemistry of Materials 30:8025-8033 (2018)

Advances in understanding the stimuli-responsivle phase behaviour of IDP polymers
Ruff et al. J. Mol. Biol. 430:4619-4635 (2018)

Self-assembly of filaments

Unlike DNA/RNA, which are always in polymer form, other filaments in a cell only form under certain conditions: they self-assemble.

Actin filaments - Actin monomers, a globular protein, 5-7 nm diameter, forms double stranded polar helical filaments

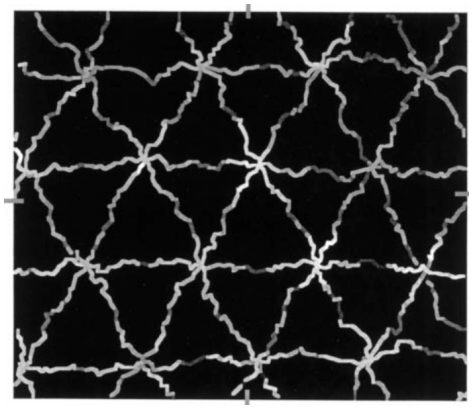
Microtubules - Tubulin monomers, a globular protein, forms a polar, hollow tube of 13 protofilaments, 25 nm diameter

Intermediate filament - Vimentin, Keratinin, etc, made of filamentous proteins, 10 nm

Cortical cytoskeleton - Actin/Spectrin/ankyrin network attached to the plasma membrane, a mix of hexagonally-connected spectrin with actin at the vertices and attached to the PM by ankyin, band 3 protein, etc.

Unstretched length of spectrin ~ 200 nm
Edge length in the *fish-net* ~ 75 nm
Distance from PM ~ 15 nm

So, the network is pre-stressed, but can also shear and rearrange under low stress



D.H. Boal, Computer simulation of a model network for the erythrocyte cytoskeleton, Biophys. J. 67:521-529 (1994)

Cytoskeletal function

What does a cell do with its polymers?

Actin filaments

Support - forms linear filaments in cytoplasm and an hexagonal (fishing-net like) network associated with the plasma membrane for structural support

Motility, filopodia, lamellipodia - treadmilling of actin filaments drives cell motility

Microvili and cilia - oscillatory beating or rotations drive bacterial motion

Transport - myosin motor proteins move along actin filament pulling e.g., vesicles

Microtubules

Support - linear filaments

Transport - Dynein and Kinesin move along microtubules

Mitotic spindle for chromosome separation

Why does the cell make filaments that self-assemble?

Cell has to minimise energy consumption, and monomers diffuse around for free (*Equipartition theorem again*)

50% of G-actin in a cell is monomeric and the equilibrium between this pool of free monomers and growing filaments is controlled by many actin-binding proteins. These proteins prohibit or stimulate the filament growth, and may cause disassembly as well.

Spatially-varying concentrations of these proteins allows the cell to control the direction in which filaments grow - this is the basis of cell crawling and neuron growth.

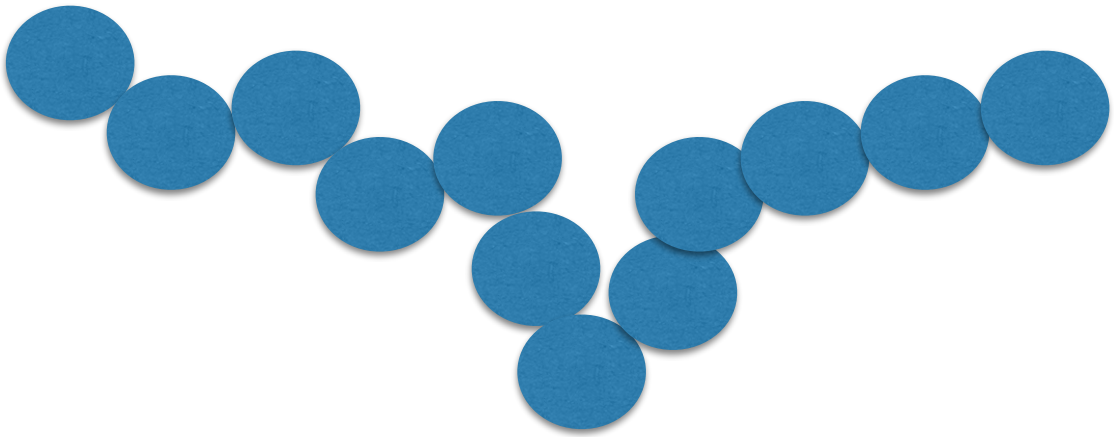
The cell doesn't *make* polymers - it makes monomers that *want* to self-assembly into polymers

How can we describe these growing/shrinking filaments and, more generally, any polymer?

Start with a random walk model …

Mathematical models of polymers

A polymer is a chain of N (here identical) monomers bonded together in a (usually linear) chain:



e.g., Polystyrene, Polyethylene glycol (PEG), polyethyl-ethylene (PEE), etc.

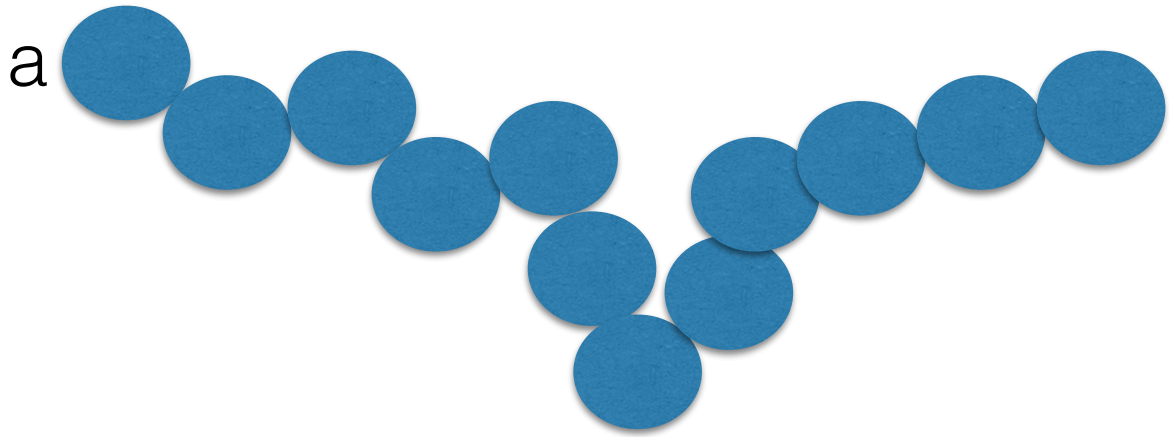
What determines the polymer's *average* length, other static properties, and its dynamics?

If the polymer is "floppy", i.e., its monomers can move around relatively freely (subject to remaining connected) there are an enormous number of configurations available to the molecules. This is very different to small molecules like methane, water, short-chain alcohols, etc. Even lipids have only 10-24 methyl groups in their chains, but polymers can have 1000s of monomers.

This huge conformation space means that entropy dominates these polymers' behaviour, and the identity of the monomers is relatively unimportant beyond a few nm (i.e., monomer diameters).

Polymer Model

Let a polymer be a chain of N+1 identical monomers, of size a, bonded together in a linear chain



We can represent this as a random walk in 3D of **N** steps each of a fixed size **a** (this is a BAD cartoon of a RW!)

There are many models for linear polymers:

A) Completely flexible = freely-jointed chain = *phantom* polymer
B) Self-avoiding chain
C) Gaussian chain
D) Freely-rotating chain
E) Worm-like chain
F) Infinitely rigid rod

The Theory of Polymer Dynamics, M. Doi and S. F. Edwards, Oxford Science Pub., Clarendon Press, Oxford, UK.

A) Freely-Jointed chain

Polymer is a chain of N+1 identical monomers connected in a linear chain by bonds of fixed length **a** that are free to point in any direction independently of all others (aka *ideal* chain, *phantom* chain):

The conformation of the chain is given by N+1 position vectors: $\{\mathbf{R}\} = (\mathbf{R}_0, \mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3, \ldots \mathbf{R}_N)$ or, alternatively, by the N bond vectors $\mathbf{r}_i = \mathbf{R}_i - \mathbf{R}_{i-1}$:

$$\{\mathbf{r}_i\} = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \ldots \mathbf{r}_N)$$

The size of the polymer can be represented by the end-to-end length:

$$\mathbf{R}_{ee} = \mathbf{R}_N - \mathbf{R}_0 = \sum \mathbf{r}_i \qquad \text{where the sum is over } i = 1 \ldots N$$

By construction, $< \mathbf{r}_i > = 0$, but $< \mathbf{R}_{ee}^{\,2} >$ is not zero. (Optional derivation)

$$< R_{ee}^2 > = N.a^2 + < \sum r_i \,.\, r_j > \qquad \text{where the sum is over } i \neq j = 1 \ldots N.$$

which for the freely-jointed chain reduces to $< R_{ee}^2 > = N.a^2$ as $< r_i \,.\, r_j > = 0$.

Quantifying the size of a polymer

$$< R_{ee}^2 > = N.a^2 + < \sum r_i \, . \, r_j >$$

where the sum is over $i \neq j = 1 \ldots N$. For many models of polymers the dependence on N is the same. The behaviour of real polymers often does NOT depend on the chemical details of the monomers, but only on their length. Other polymer models: bond lengths have a distribution, bond angles are fixed, bond angles have a distribution etc. These models differ in the $\sum r_i \, . \, r_j$ term, i.e., on the correlations between successive monomers in the chain.

The *Characteristic Ratio* of a polymer is defined as:

$$C_N = < R_{ee}^2 > / N.a^2$$

Centre of mass of a polymer (with N monomers):

$$\mathbf{R_{cm}} = 1/N \sum \mathbf{R}_i$$

Radius of gyration of polymer:

$$<\mathbf{R_g}^2> = 1/N \sum (\mathbf{R}_i - \mathbf{R}_{cm})^2 = 1/2N^2 \sum_{i \neq j} \mathbf{R}_{ij}^2 \qquad \mathbf{R}_{ij} = \mathbf{R}_i - \mathbf{R}_j$$

(Optional derivation: show that the radius of gyration is given by this result)

The ideal chain, for which, $C_N = 1$, $< R_{ee}^2 > = N.a^2$, and $<R_g^2> = 1/6. N.a^2$, has the same end-to-end length scaling as a random walk.

Relating a real polymer to a random walk

How do we relate the scaling law $< R_{ee}^2 > = N.a^2$ to a real polymer?

Is **N** the same as the number of monomers?
Is the monomer size equal to **a**? (No!)

Define a new length (the *Kuhn length*) $\mathbf{l_k}$ as the ratio of the real polymer's $<R_{ee}^2>$ to its contour length, and define N* as the ratio of the contour length to this new length:

$$l_k = <R_{ee}^2> / L$$

$$N^* = L / l_k$$

Therefore

$$< R_{ee}^2> = N^* . l_k^2$$

So, a real polymer is equivalent to a RW with step size l_k and number of steps equal to the number of *Kuhn lengths* in its contour length.

The Kuhn length is the *step size* in the equivalent RW model because the real monomers are highly correlated in space (backbones have some stiffness) while succeeding steps in an RW must be uncorrelated. The Kuhn length is roughly the distance required for a segment of the polymer to be uncorrelated in direction with the previous segment.

Visualising the Kuhn length

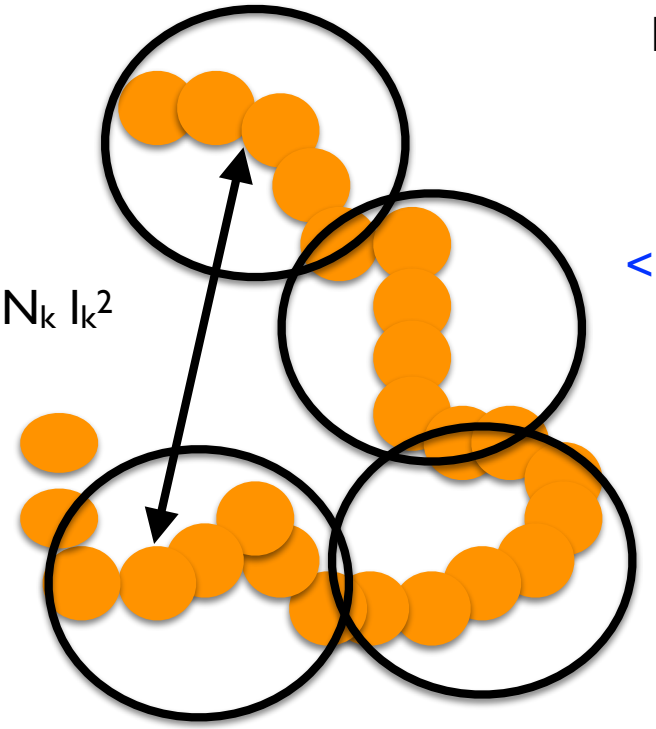
What is the relation between a real polymer and the equivalent phantom one?

A real polymer has *correlations* between adjacent monomers- some *stiffness* along the chain

A phantom polymer has *no correlations* (aka freely-jointed chain)

But if the real polymer is sufficiently long, we can represent it - to some degree of accuracy - as a phantom chain with a smaller number of monomers each of which is larger: this is the Kuhn length.

$< R_{ee}^2 > = N_k \, l_k^2$

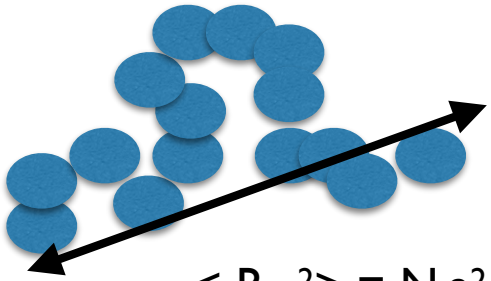


$l_k = <R_{ee}^2> / L$

$N_k = L / l_k$

$< R_{ee}^2> = N_k \, . \, l_k^2$

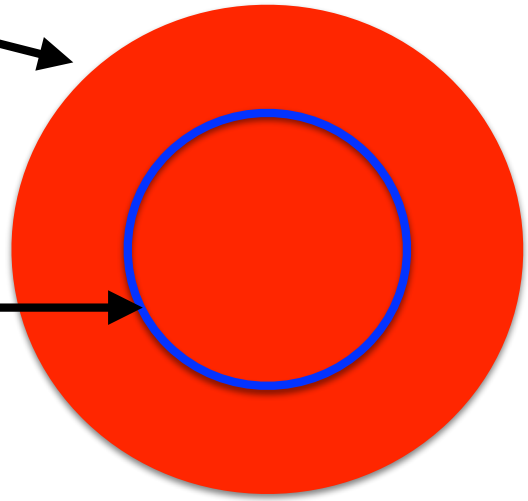
Random walk



$< R_{ee}^2> = N.a^2$

How big is a hard sphere?

R_h = actual sphere radius



R_g ~ distribution of mass in space

$$M R_g^2 = \int r(m)^2 \, dm$$

e.g.,

Spherical shell: $R_g^2 = R^2$

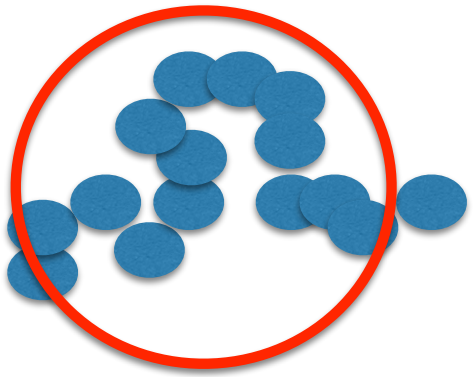
Solid sphere: $R_g^2 = 3/5 \, R^2$ which is smaller than for the shell because the interior mass pulls Rg to smaller values

How big is a polymer?

For a polymer we have the following "sizes":

$< R_{ee}^2 > = N.a^2 + < \sum r_i \, . \, r_j >$ end to end length

$<R_g^2> = 1/N \sum (R_i - R_{cm})^2 = 1/2N^2 \sum R_{ij}^2$ radius of gyration



hydrodynamic radius

all are useful in some circumstances, but can be misleading

We have

$$< R_{ee}^2 > = N \cdot l_k^2$$

and for polyethylene glycol (PEG) with a M.Wt of 20,000 Da, we have:

M.Wt ~ 20,000 Da
Monomer size (-CH_2CH_2O-) ~ 0.44 nm and mass 44 Da
Kuhn length = 1.8 nm

So, PEG has 20,000/44 = 455 real monomers, and contour length L = 455*0.44 = 200 nm.

There are $N = L / l_k = 200/1.8 = 111$ Kuhn lengths in the polymer.

So, we predict the polymer's mean size to be:

$$\sqrt{<R_{ee}^2>} = \sqrt{111} * 1.8 \text{ nm} \sim 19 \text{ nm}$$

And this is the size we would expect to measure in a light scattering experiment, for example.

We have

$$\langle R_{ee}^2 \rangle = N \cdot l_k^2$$

For the E. Coli chromosome:

Size = 4.64 Mbp
Monomer size (1 base pair) ~ 0.34 nm
Kuhn length = 100 nm

What are N* and the $\sqrt{\langle R_{ee}^2 \rangle}$ for this strand of DNA?

Calculate these values

What does this result tell you about the packing of DNA in the bacterium?

B) Self-avoiding random walk

We know for a simple random walk that:

$$< R_{ee}^2 > = N.a^2$$

But real polymers cannot pass through each other, so this is only an approximation.

If we impose the condition that monomers cannot intersect each other in space, we can no longer calculate the static properties of the polymer by summing over all conformations as the self-avoidance introduces non-local correlations into monomer positions.

But we can generate SAWs on a computer and count the number of configurations of N monomers. We then find that self-avoiding random walks satisfy the relation:

$$< R_{ee}^2 > = N^{2\nu}.a^2$$

The exponent ν is called the *Flory Exponent*, and is 1/2 for simple RWs, and has the following value in d dimensions for SAWs:

$$\nu = 3 / (d + 2)$$

Soft Condensed Matter Physics in Molecular and Cell Biology,
WCK Poon and D Andelman, Taylor and Francis, USA, 2006.

C) Gaussian chain model

The phantom polymer (aka RW) with N monomers has mean and maximum (stretched) lengths:

$$< R_{ee}^2 > = N.a^2 \qquad \text{Max } R_{ee}^2 = (N.a)^2$$

Typically, there will be many more conformations clustered around a length Na^2 than much shorter or longer lengths - it is highly unlikely that thermal fluctuations will stretch the polymer near to its maximum size or compress it to a volume the size of a few monomers.

The *Gaussian chain model* assumes a polymer is ideal (no energy cost to bending it) but extends the allowed lengths to infinity to make integrating over polymer conformations simpler. The probability for a Gaussian polymer to have an end-to-end length of R is:

$$p(R) = \text{const. } \exp(-d.R^2/2.N.a^2)$$

where d = space dimension = 1, 2, 3. The constant is determined by normalization of the probability:

$$\int p(R)\, d^3R = 1$$

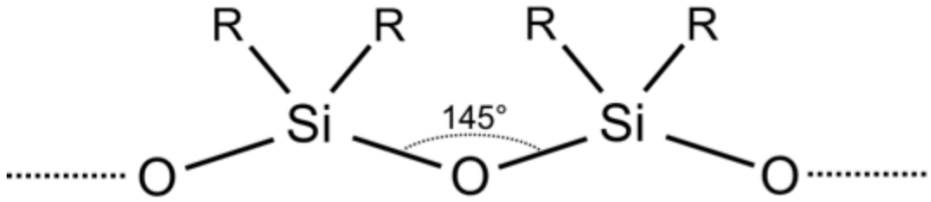
and the mean square end-to-end length is:

$$<R_{ee}^2> = \int R^2.p(R)\, d^3R$$

What are $<R_{ee}>$, $<R_{ee}^2>$ and for the Gaussian chain? (Optional derivation)

D) Freely-Rotating chain model

The polymer is again represented as a chain of N+1 identical monomers connected in a linear chain but each bond (of length a) is fixed at a given angle to its predecessor, the remainder of the chain is free to rotate about any bond, e.g., polysiloxane (NB, this actually has two angles but we assume they are the same for simplicity).



The conformation of the chain is again represented by the bond vectors:

$$\{\mathbf{r}_i\} = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \ldots \mathbf{r}_N)$$

where $\mathbf{r}_i = \mathbf{R}_i - \mathbf{R}_{i-1}$.

Clearly, $\mathbf{r}_1 \cdot \mathbf{r}_2 = -a^2 \cos\theta$, $\mathbf{r}_2 \cdot \mathbf{r}_3 = -a^2 \cos\theta$, etc.

What are $\mathbf{R}_{ee}$ and C_N now?

The mean-square end-to-end length and characteristic ratio are:

$$\langle R_{ee}^2 \rangle = N.a^2 + 2.a^2.\left(N.(\alpha - \alpha^N) / (1 - \alpha) - (\alpha - \alpha^{N+1} - N.\alpha^N + N.\alpha^{N+1}) / (1 - \alpha)^2 \right)$$

$$C_N = (1 + \alpha) / (1 - \alpha) - (2.\alpha / N) .(1- \alpha^N) / (1 - \alpha)^2$$

where $\alpha = \cos \theta$

In the limit of long polymers, $N \rightarrow \infty$

$$\langle R_{ee}^2 \rangle = N.a^2 (1 + \alpha) / (1 - \alpha)$$

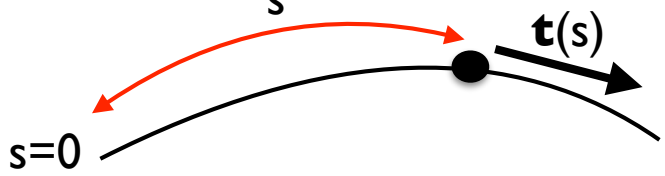
$$C_N = (1 + \alpha) / (1 - \alpha)$$

which clearly diverge for $\cos \theta = 1$, i.e., rigid polymer (but the finite N expressions do not).

E) Worm-like chain model

The WLC model is more complex than the ideal chain as it includes an energy cost to bending the polymer. This makes it a better model for moderately stiff polymers.

Consider a polymer as a curve in 3d space, and define the tangent vector to the curve at point s as $\mathbf{t}(s) = \partial \mathbf{r}/\partial s$, where s is the arc length.



The energy H of a given conformation of the curve $\mathbf{r}(s)$ is determined by the energy cost of bending it into shape:

$$H(\mathbf{r}(s)) = \kappa/2 \int |\partial \mathbf{t}/\partial s|^2 \, ds$$

where κ is the bending energy of the polymer. The probability of this conformation is given by the Boltzmann weight

$$\text{Prob.} \sim e^{-\beta H(\mathbf{r}(s))}$$

The usefulness of the WLC is that one can do many integrals exactly with this form, and calculate results more simply than performing simulations of self-avoiding walks but more accurately than ideal chains.

F) Rigid rod model

This is a trivial model in which all bonds are parallel, there are no fluctuations, and so

$$< R_{ee} > = N.a$$

$$< R_{ee}^2 > = N^2.a^2$$

$$C_N = N$$

But if the polymer is not *infinitely* rigid, there will be shape fluctuations, or wiggles, along the polymer that modify its long-range orientation, and lead to the idea of a *persistence length*.

Let a polymer be represented as a curve $\mathbf{r}(s)$ in space with tangent vector $\mathbf{t}(s)$ where s measures the arc length from a fixed point. The energy of the polymer is:

$$H(\mathbf{r}(s)) = \kappa/2 \int | \partial \mathbf{t}/\partial s |^2 \, ds$$

The parameter κ is the *Bending Rigidity* of the polymer, and has units Energy*Length. This allows us to define a *Persistence Length* L_p as:

$$L_p = \kappa / k_B T$$

The persistence length is the typical distance along a polymer for which the initial and final orientations are uncorrelated: or the distance needed for the polymer to "forget" the original orientation.

Persistence length of polymers

Filament	Monomer	Location	Diameter	Persistence Length	Contour Length
DNA	A, C, G, T Bases	Nucleus	2 nm	~50 nm [1]	~ 1m
Actin	Actin monomer	cytoplasm	7 nm	~10 µm [2]	50 nm - 1 µm
Intermediate filaments	α-helical rod	cytoplasm	10 nm	~1 µm [3]	?
Microtubules	Tubulin monomer	cytoplasm	25 nm	1-10 mm [4]	10s µm

1 - Manning, **Biophys. J. 91**:3607 (2006)
2 - Isambert et al. **J. Biol. Chem. 270**:11437 (1995)
3 - Block et al. **BBA. 1853**:3053 (2015)
4 - Martin, Ch. 2 in **Methods in Cell Biology, 115** (Elsevier, 2013)

Q. How could a cell modify a polymer's persistence length?

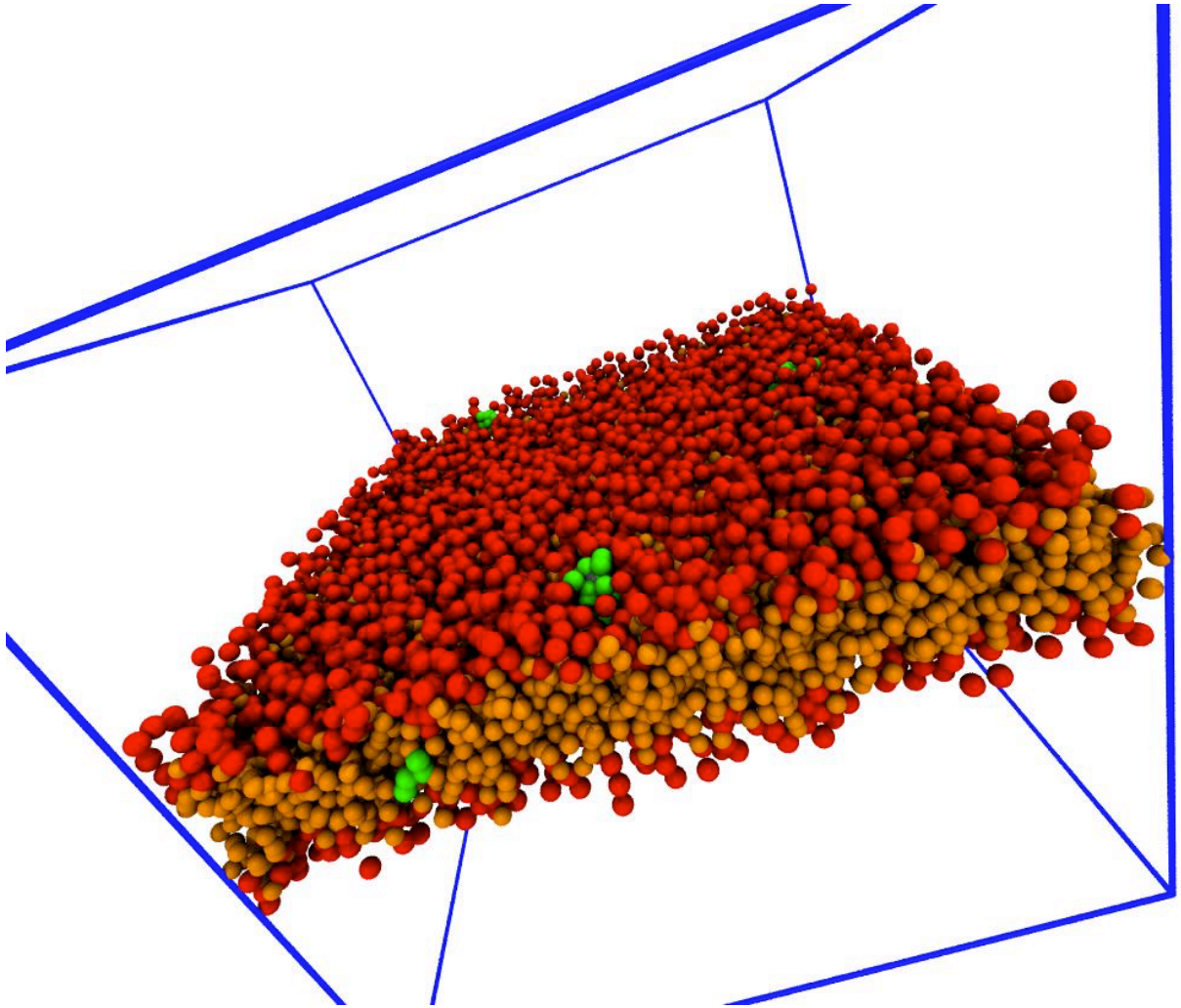
- Break
- Overview of more homework exercise
- Select and work on one of them

Homework exercises/Projects

EPFL

3) What does GFP to do lipid diffusion/fluctuations?

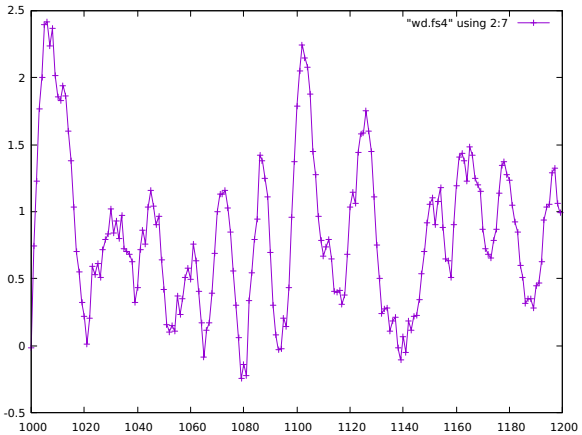
(see **dmpci.ex3** on moodle for today)



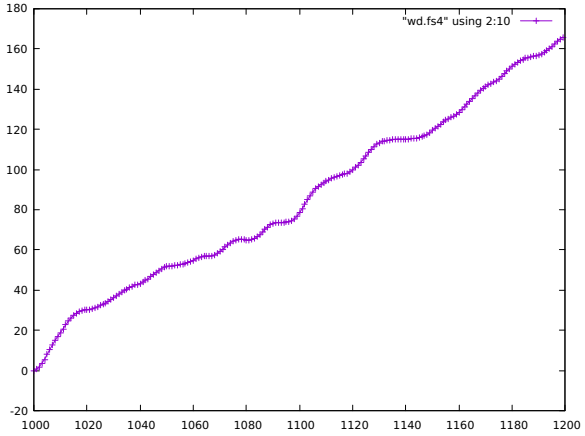
4) WD in pulling a lipid out of the membrane

(see **dmpci.fs4**)

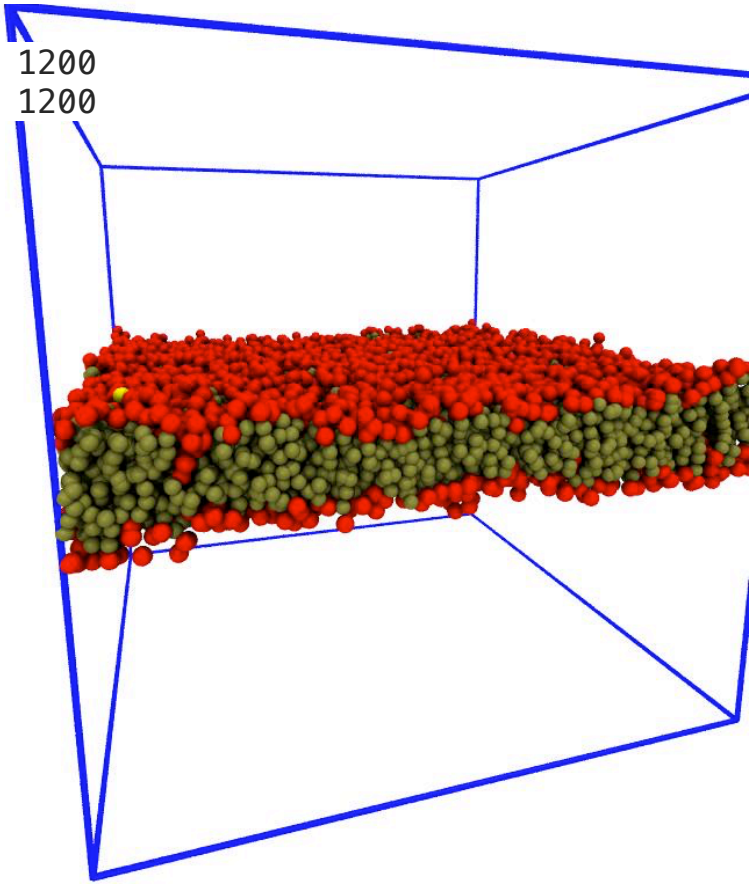
```
Command Comment   1 // Following commands apply a force upwards to the single lipid
with the H1 bead  //

Command SelectBeadTypeInSimBox          1  head  H1
Command SelectPolymerTypeInSimBox       1  labelledLipid  Lipid1
Command CountBeadsInTarget              1  head

Command ConstantForceOnTarget           1000  head  fh   0 0 1   30.0

Command DistanceMovedByTarget           1000  head fh  dm1  1000 1200
Command ExternalWorkOnTarget            1000  head fh  ew1  1000 1200

Command RemoveCommandTargetActivity   1300   fh
```



Work done per time-step



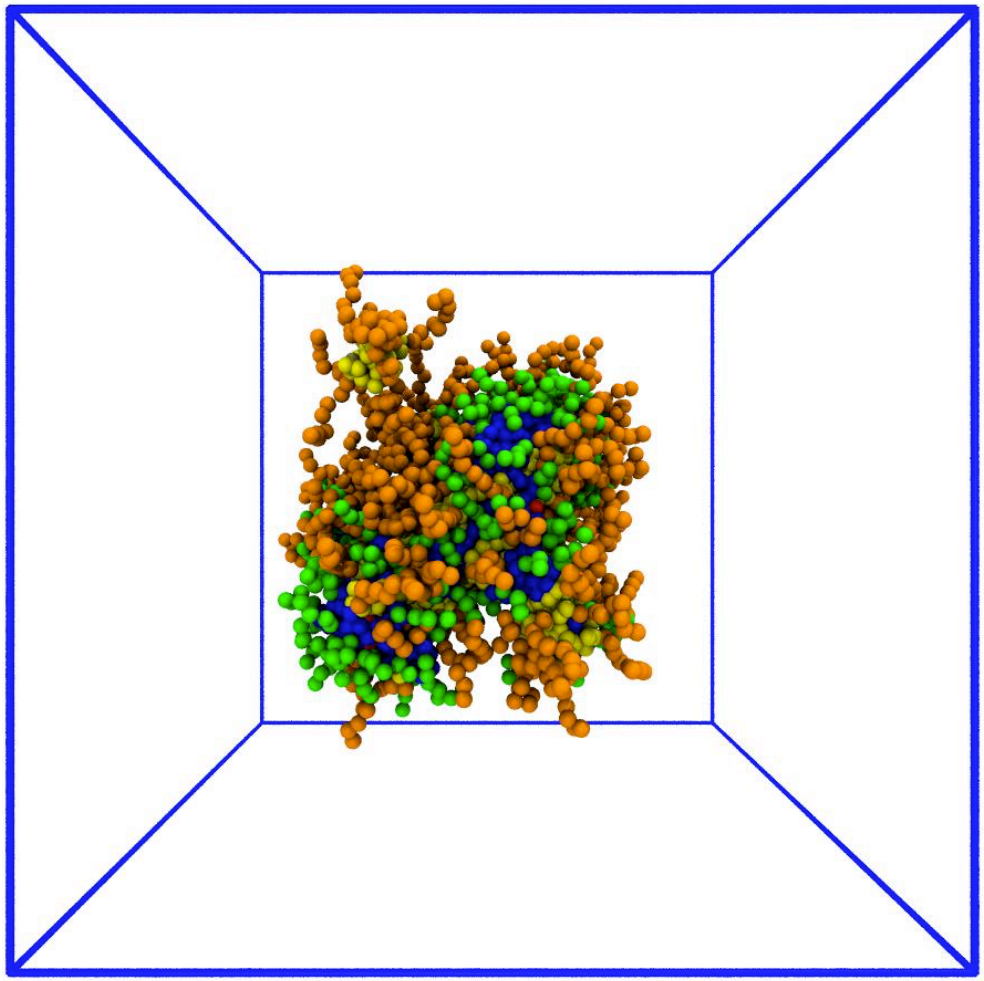
Cumulative work done



Homework exercises/Projects

(see **dmpci.ring-micelle**)

Why is this a ring?



```
" (H1 (8 (B B B (* (S (6 S) S)) B B B)) T1) "
" (H1 (8 (B1 B1 B1 (* (S1 S1 S1 S1)) B1 B1 B1)) T1) “

B, B1 are strongly hydrophobic; S, S1 are hydrophilic
```

Total number of beads in box: Nb = ρ*L**3

Assume you have 1 polymer of interest in water with n_p beads per polymer, and water with a single bead $n_w = 1$. Then

ρ_p = number fraction of the polymer which creates N_p polymers
ρ_w = number fraction of water, N_w

Clearly, $\rho_p + \rho_w = 1$

and by definition of number fractions:

$\rho_p = N_p / (N_p + N_w)$, $\rho_w = N_w / (N_p + N_w)$.

$N_b = Np * n_p + Nw*n_w$

Solve for ρ_p in terms of N_p, n_p, n_w and box size.

Runtime= Nb*Nt / 5 10**10 in cpu-days