Take-Home Message

Cells often use multiple weak forces instead of one strong one

Indirect forces control a lot of cellular behaviour

Reversibility is a sign of health; irreversibility a sign of disease

Constructing Mathematical Models: Biomolecular condensates / membraneless organelles

- Forces - direct, indirect, hydrophobic, entropic, membrane-mediated
- Compartmentalisation - lipid-bounded, condensates
- Phase separation of Intrinsically Disordered Proteins
- Experimental manipulation of condensates
- Model of membraneless organelles

Forces and compartmentalisation

How to classify forces relevant in cells? Direct versus Indirect

Direct: Gravity, Coulomb, Dipole, VdW, H-bonds, Covalent and ionic bonds

- Operate between two molecules "independently" of others ⇒ pairwise additive
- Typically short-ranged (not gravity! or Coulomb force in a vacuum)

Indirect: Hydrophobic effect, Depletion, Capillary forces, Membrane-mediated, Fluctuation-induced forces

- Arise between many molecules simultaneously, strongly affected by other objects ⇒ NOT pairwise additive
- Can be long ranged (~ 1/R)
- Proportional to temperature (so pressure is *entropic*… PV = RT)

Direct forces

Gravitational force between two point masses:

$F = G M_1 M_2 / R^2$ $\quad G = 6.67\ 10^{-11}\ N.m^2/Kg^2$

Nearly negligible for cells, but they can sink under gravity

Coulomb force between charges in vacuum

$F = k Q_1 Q_2 / R^2$ $\quad k = 1 / 4\pi\varepsilon_0 = 9\ 10^9\ N.m^2 / C$

in a material ε_0 is replaced by $\varepsilon_0\,\varepsilon$, where ε is the relative permittivity of the material, and makes a big difference between water ($\varepsilon \sim 80$) and oil ($\varepsilon \sim 1$).

Bare force is usually screened out by ions under physiological conditions

Screened Coulomb force

$F = e^{-\kappa R} / R$ $\quad \kappa^{-1}$ = screening length ~ 1 nm in physiol. conditions

$F_G / F_C \sim G M_1 M_2 / k Q_1 Q_2 \sim 10^{-10}\ (10^{-27})^2 / 10^{10}\ (10^{-19})^2 \sim 10^{-36}$

so gravitational force between bare ions is utterly negligible.

Chemical and physical bonds

Bond Type	Strength (kJ/Mol)	Strength (k_BT/bond)	Length (nm)	Description
Covalent	500	200	0.154 (C-C)	Shared outer e^-
Ionic	~ 880	~ 355	0.276 (NaCl nn)	e^- donated/summed
Hydrogen	10 - 40	4 - 16	~ 0.176 (O…H)	small H - electroneg. atom
"Van der Waals"	~ 1	~ 0.4	$1/R^6$	fluctuating induced dipoles

NB 1 kJ/Mol ~ 0.4 kBT per particle

J. Israelachvili, Intermolecular and Surface Forces, Academic Press, 2nd ed. London 1992.

If Strength ~ k_BT, the force has no effect as thermal noise overwhelms it:

- Covalent/ionic bonds cannot be broken by (room) temperature (in vacuum)
- H-bonds may be broken by a large fluctuation
- VDW must occur as many bonds to have any effect

Indirect forces

Indirect forces originate in Coulomb's law (electrostatics) but depend on the temperature, and usually involve many weak interactions supporting each other rather than one dominant strong interaction.

This makes them hard to calculate, and almost impossible to guess their functional form or even their sign.

Importantly: indirect forces can be "free", i.e., they arise from the thermodynamics of the environment at room temperature and not on the atomic structure of the interacting molecules (cp. polymer not depending on monomer identity).

Entropic forces

hydrophobic effect of oily chains in water

depletion (molecular crowding)

fluctuation-induced forces

Membrane-mediated forces

curvature

composition

thickness mismatch

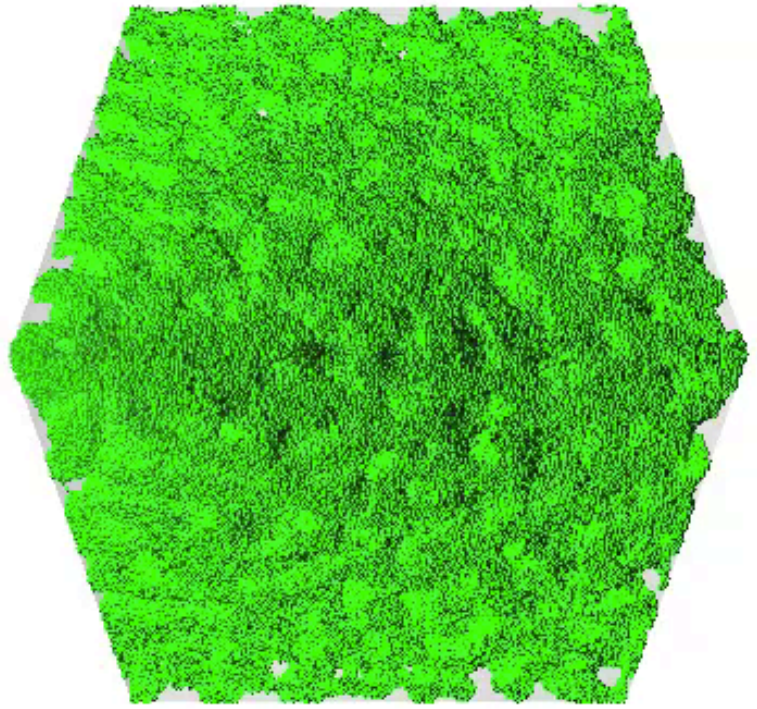
Hydrophobic effect

Water - an homogenous liquid with no bulk structure beyond ~ 1nm, that possesses the unique property of a hydrogen bonded network between the H_2O molecules.

This network has a huge entropy as the molecules continually make and break H-bonds.

The hydrophobic effect results from the prevention of H-bonding between water by the presence of non-polar molecules that drives them to aggregate in one place.

C. Tanford, The Hydrophobic Effect, Wiley, New York 1980



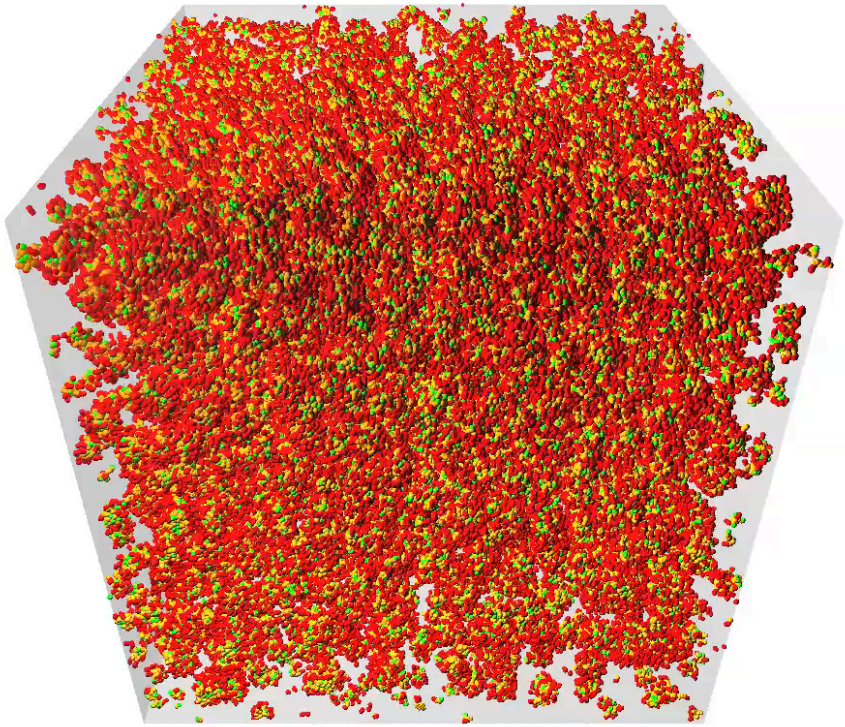
NB. Movies produced from DPD simulations and visualized with Povray and Quicktime. Water in the simulation box is invisible for clarity.

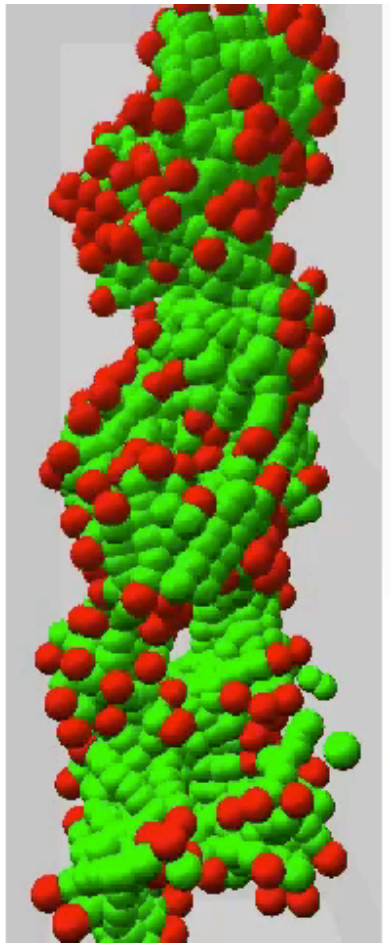
Mix oil and water ⇒ droplets appear by phase separation and create a (thin) bounding surface between the phases.

Amphiphiles in water form complex aggregates

Mix lipids and water ⇒ interfaces and *compartments* appear with interfacial width ~ lipid length

Compartments support gradients, and gradients can be used to do work, e.g., ion concentrations can differ across the neuronal plasma membrane.





Aggregate type is encoded in the molecular shape: no external control is needed: we expect that simulations will be useful in predicting lipid phases if we can capture their amphiphilic nature.

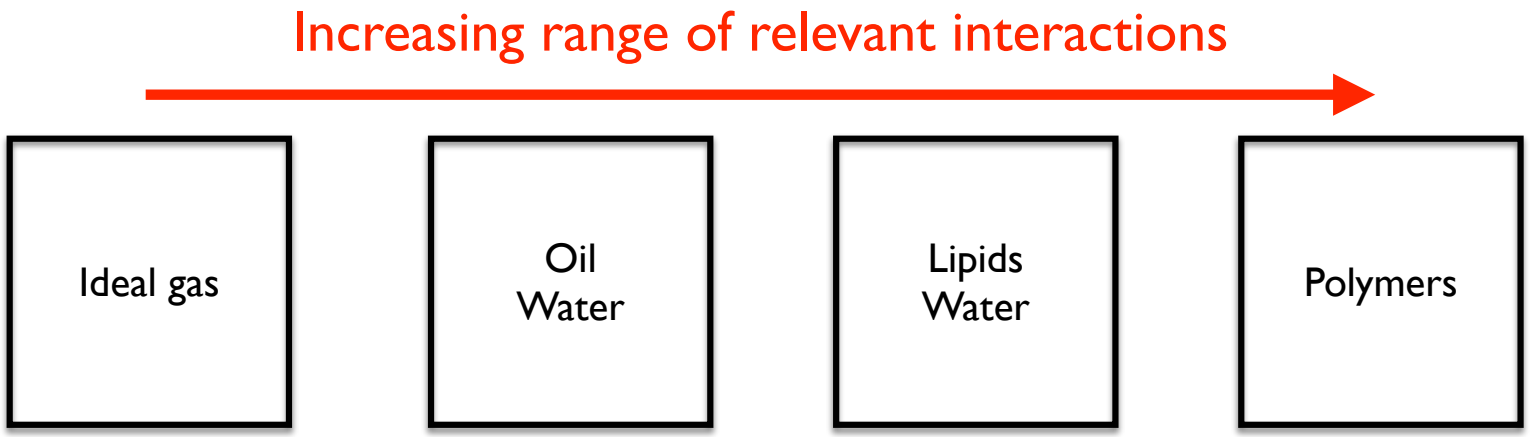
Proteins self-assemble into complex aggregates

- Oil in water can only form spheres with a smooth interface ~ size of oil molecule
- Lipids form aggregates with one dimension comparable to their size ~ 2 nm (micelles, planar bilayers, vesicles, cubic structures of bilayers, etc.)

What structures could emerge from the self-assembly of *large* molecules?

Folded proteins are one type of self-assembled aggregate; are there others?

Ask yourself: what is the range of the interactions driving the self-assembly?

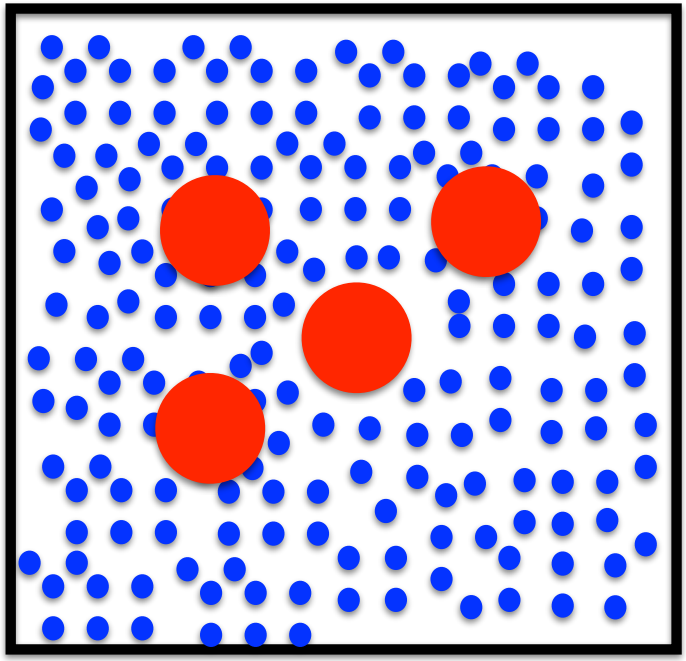


Depletion forces

Depletion forces arise when some entities reduce the freedom to move (or fluctuate) of others.

Large solute particles (red) exclude the smaller solvent molecules (blue) from around them.

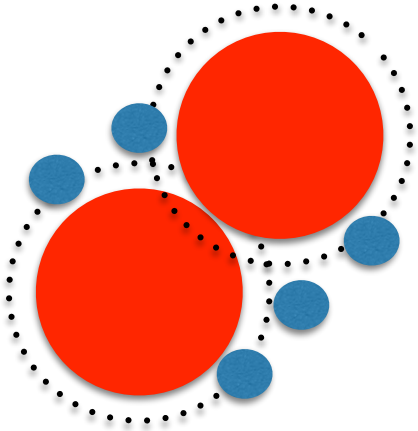
If the red ones cluster, some volume is freed up, and there is more space for the solvent to move around, hence the entropy increases.



Not to scale

Depletion forces don't only arise in bulk solvent, they also arise within membranes (see later).

They are independent of the type of entity (atom, molecule, nanoparticle, membrane) and only require the molecules to be *mobile* and *repulsive* at short range (steric or hard-core repulsion).

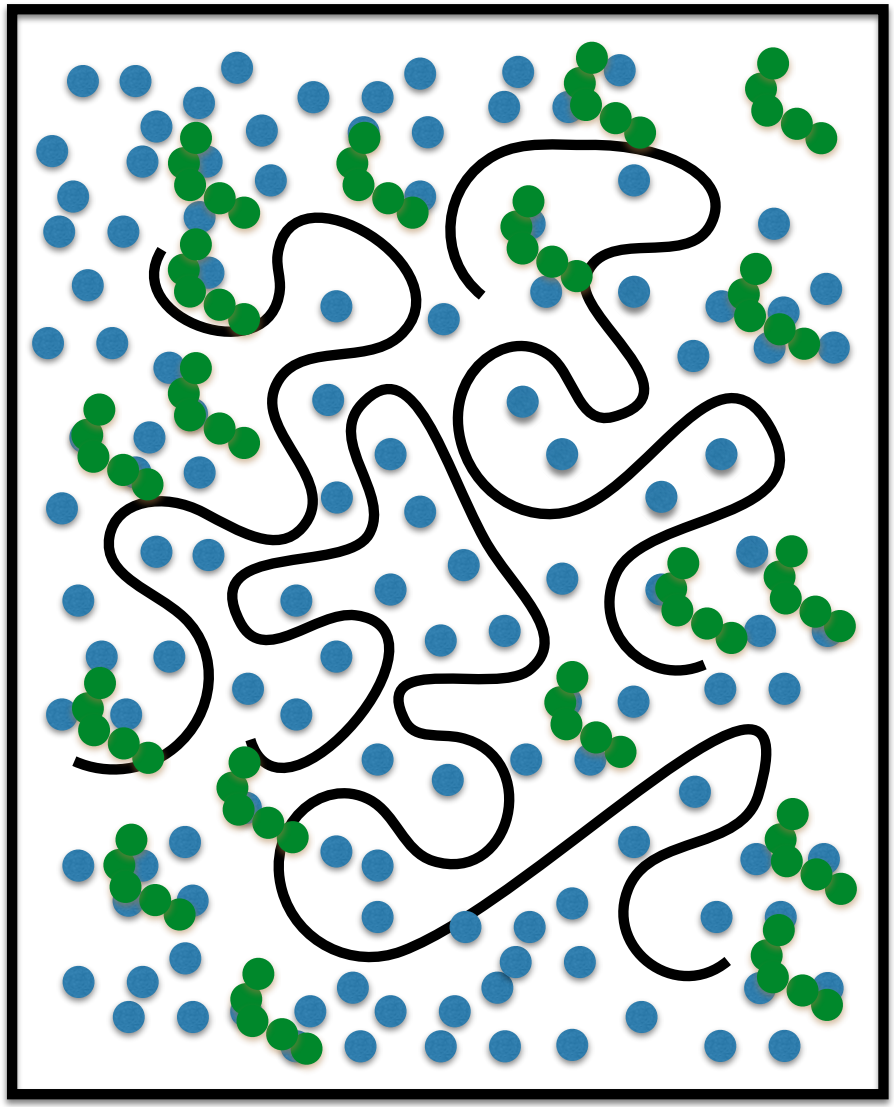


Molecular crowding

Consider a dilute solution of long polymers in water in the presence of concentrated short polymers.

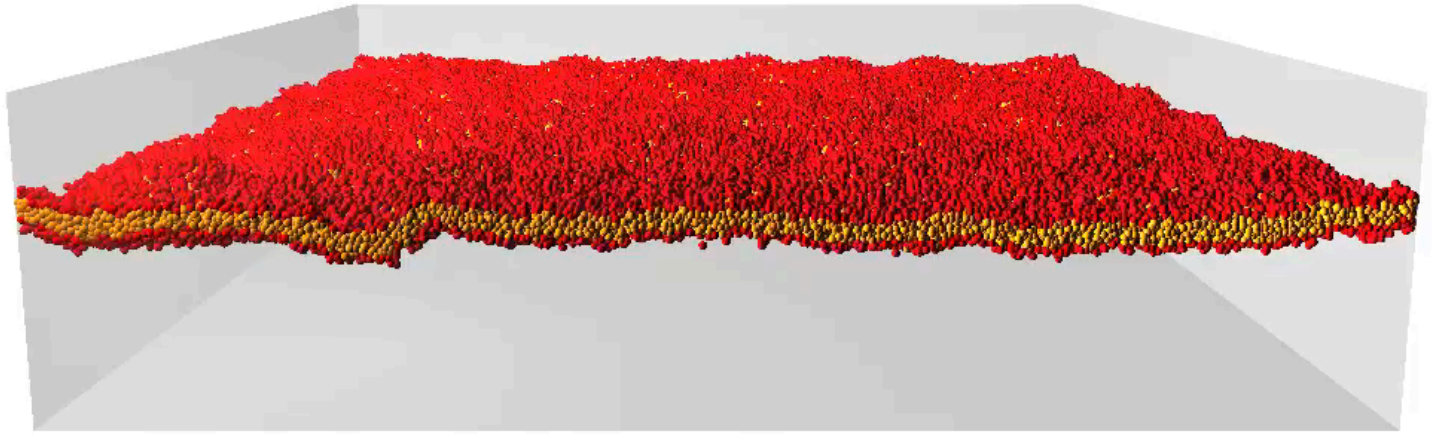
If the two polymer types are sterically repulsive, the long ones restrict the motion and shape fluctuations of the short ones.

This leads to a force on the long ones that squeezes them into an aggregate to maximise the entropy of the short ones.



Membrane fluctuation force

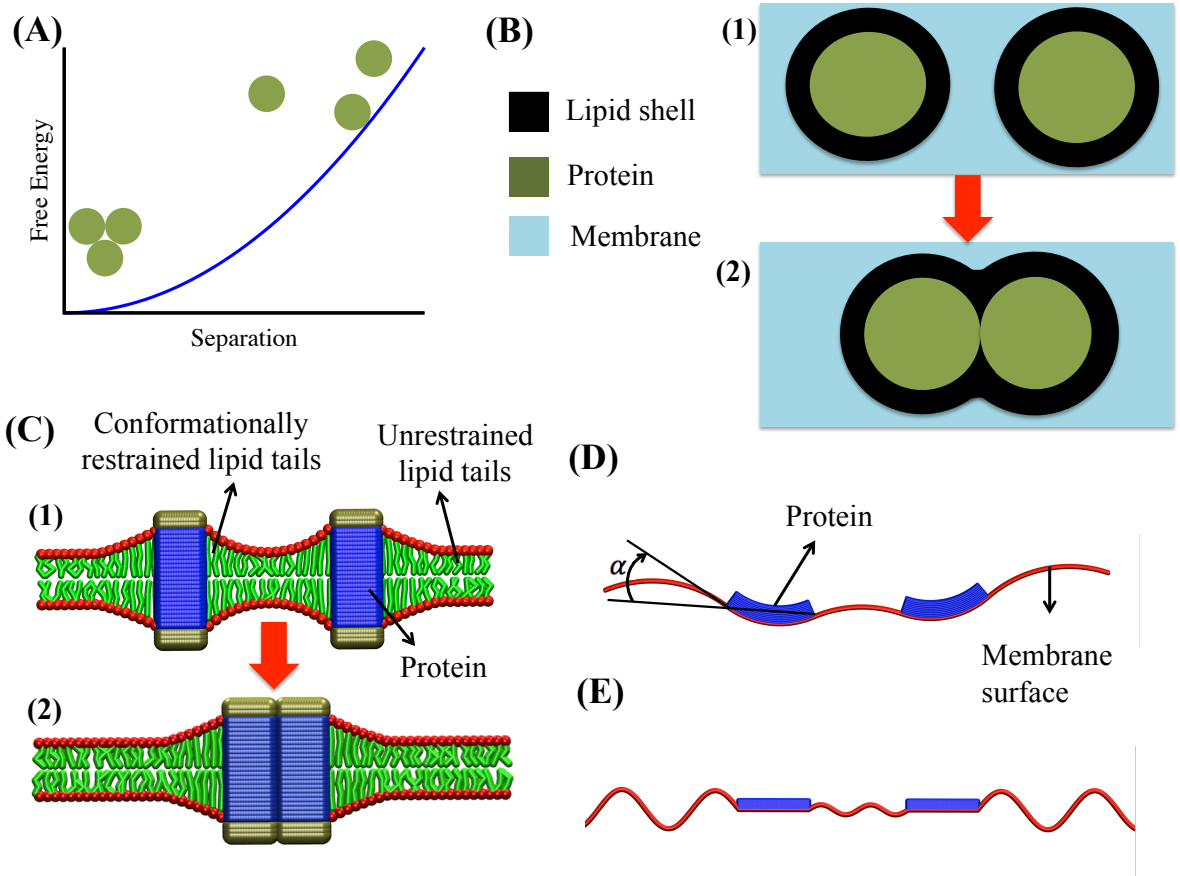
The thermal fluctuations of a lipid membrane give rise to a repulsive force on an object or material that approaches and suppresses its fluctuations: this would lower its entropy and therefore is opposed.



Can a membrane exert forces on nearby or adsorbed objects?

Indirect forces: membrane-mediated

Unlike bare forces, membrane-mediated forces arise when two (or more) proteins/nanoparticles adsorb to/embed in a membrane and perturb its state.



Johannes et al. TCB 2018

A) All operate by lowering the total free energy of membrane+proteins

B) Capillary force/line tension

C) Depletion force

D) Curvature force

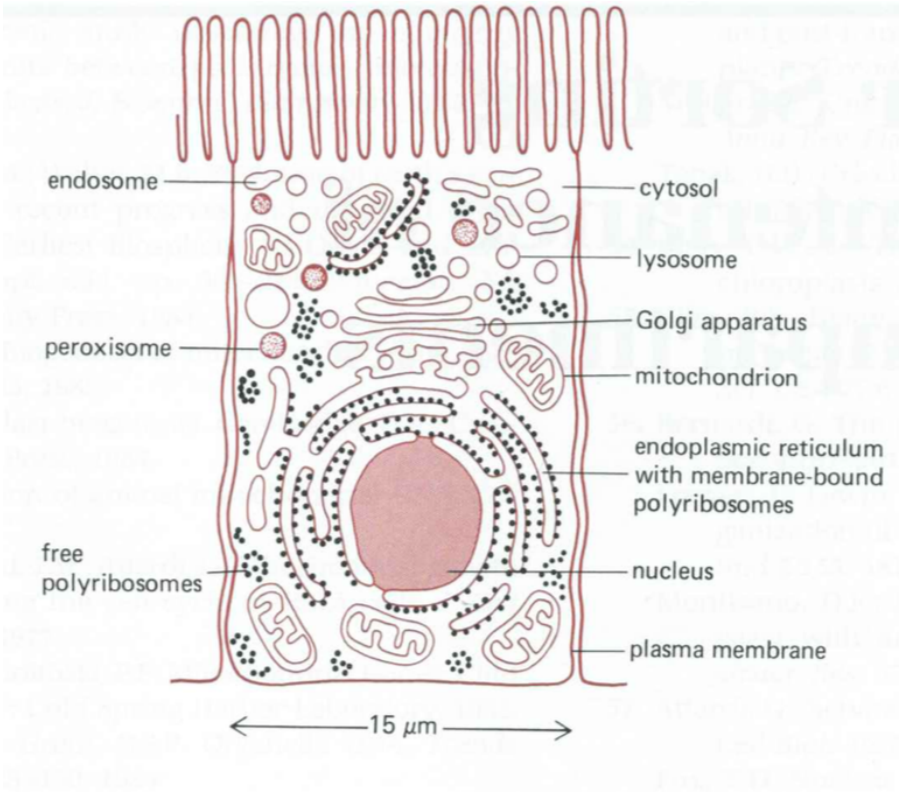
E) Fluctuation-induced force (Lecture 12)

Indirect forces create compartments

Compartments allow spatial localisation of function: reactions can be kept apart, acidic environment in one place, neutral in another, gradients can form:

- Nucleus - separates transcription from translation
- Rough ER - ribosomes synthesize proteins
- Smooth ER - lipid synthesis
- Golgi - sorting proteins for destination
- Mitochondria - e^- transport, H^+ gradient
- Endosome - imports material into cell

Alberts and Bray, Fig. 8.1



and compartments come for free if you make the right molecules (lipids, proteins) and rely on the equipartition theorem.

Self-assembly is a consequence of Equipartition Thm.

- Self-assembled structures form because they minimise a system's free energy by **phase separating** rather than staying mixed
- The cell uses self-assembly because it's **free**
- It's free because of the Equipartition Theorem

The total energy of a molecular system is continually redistributed among all its atoms and molecules by random thermal motion; this allows the system to eventually discover its state of lowest free energy.

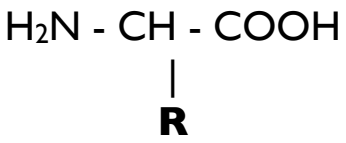
We will cover lipid membrane self-assembly in a later lecture

Membrane-bound compartments self-assemble from dispersed lipids because of the hydrophobic effect, but maybe there are other forces between other molecules that can stabilise large aggregates or compartments.

Are all cellular compartments bounded by self-assembled lipid membranes?

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

e.g., R=H for Glycine

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

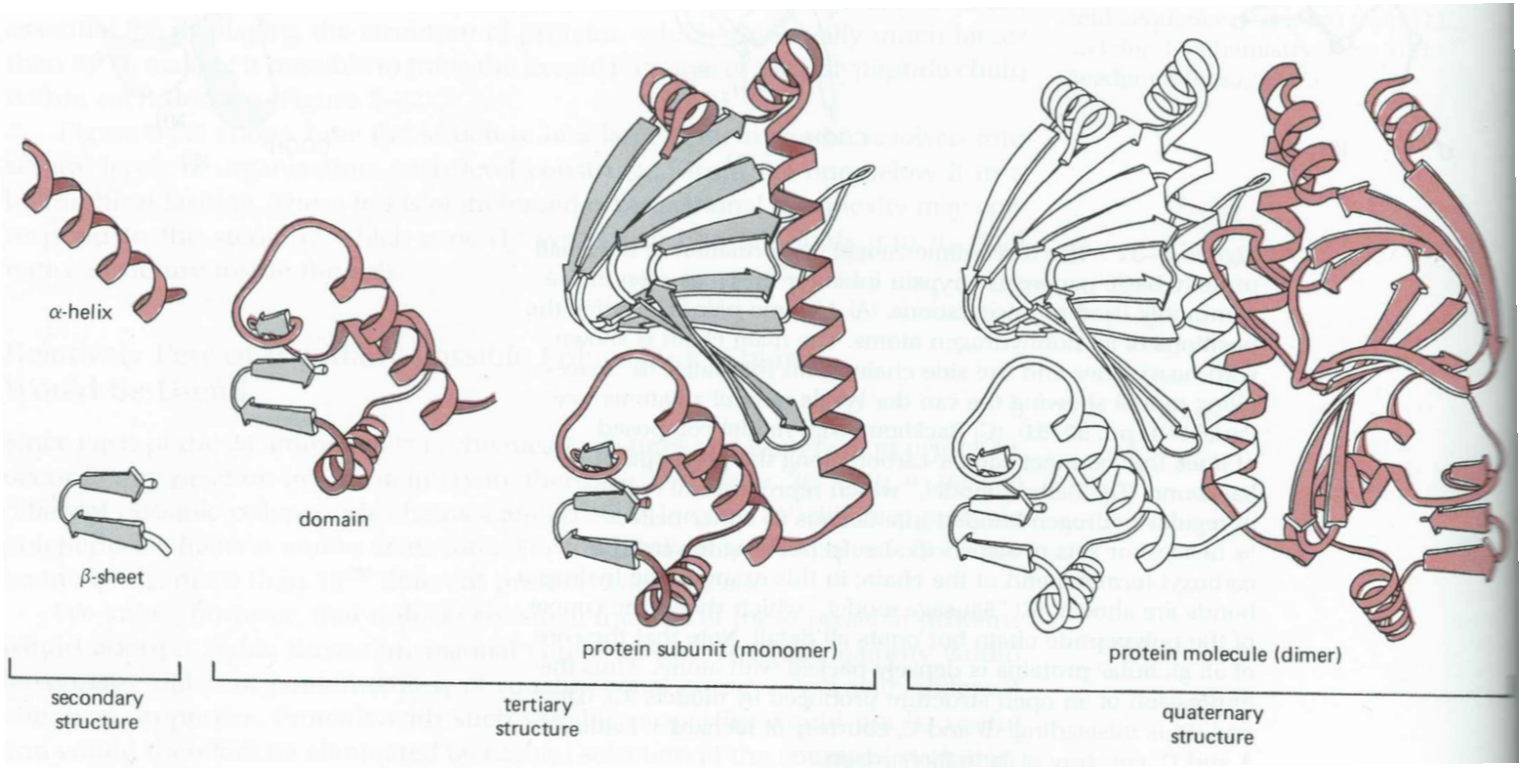
3D structure of proteins

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

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

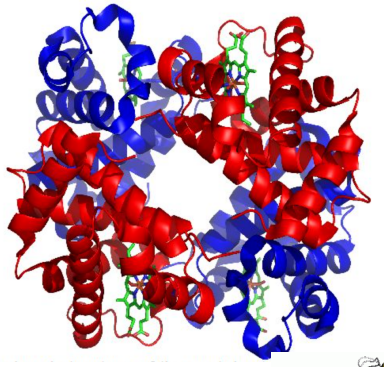


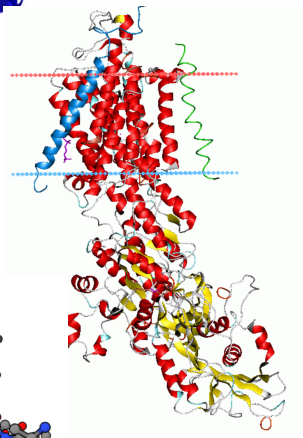
Alberts and Bray et al.

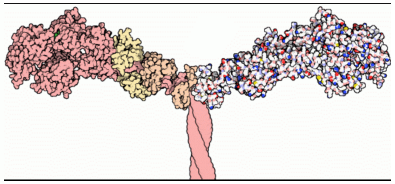
Typical proteins are molecular machines

What are some typical proteins?

- Haemoglobin - binds oxygen
- Actin - forms filaments, cytoskeleton
- Myosin motor - pulls cargo along a filament
- Na-K ATPase - pump ions against gradient
- Ion channels - allow ions to flow down gradient
- Respiratory chain proteins - makes ATP using energy of e- transport

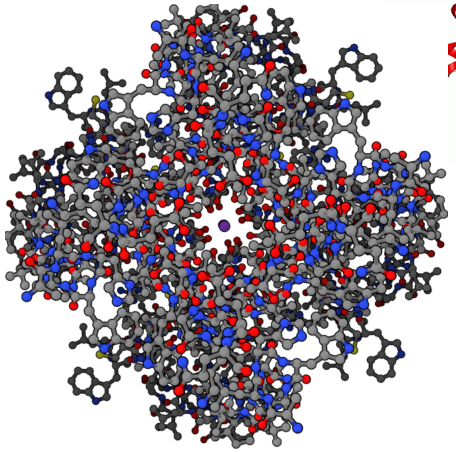






wikipedia.org

Machines because they are precisely arranged (mutate one aa and it can destroy the protein) and they operate relatively independently of other proteins



Proteins fold into precise, unique, lowest-energy states that are known from crystal structures: but is this always true of all proteins?

Are all proteins folded?

Do all proteins fold into lowest-energy states? ~25 years ago answer would have been yes.

Common picture of proteins was precisely-folded molecular *machines* that made lock-and-key binding with ligands or other proteins. Specific sequences of amino acids determined who bonded to whom, and that was it.

Very mechanical - very Newtonian - not thermodynamically accurate

There are other types of protein - *intrinsically-disordered proteins (IDP)*- that have no stable structure because of their amino acid sequence, but do exhibit biological activity, and play a fundamental role in organising cellular dynamics.

Examples

- Casein in milk, disordered structure makes digestion easier
- Calcineurin activates immune T cells on Ca binding, needs accessible binding sites, and surrounds its binding partner
- PSD-95, scaffold protein in neuronal spines, binds membrane receptors

Intrinsically-Disordered Proteins

IDPs are common and behave very differently to folded proteins - human proteome has
30 - 50% disordered residues (mainly E, K, P, Q, S) *

- IDPs are long, flexible chains of amino acids (just like proteins that fold)
- No unique folded state/no average shape/no mean atom coordinates
- Contain regions or domains of low complexity (repeated amino acids)
- Mechanically floppy, fluctuate due to random thermal motion
- Multiple weak, non-specific binding sites
- Soluble at low concentration (contain few hydrophobic residues)
- Coalesce above a critical concentration into spherical, fluid droplets known as:

Biomolecular Condensates

Intracellular Condensates

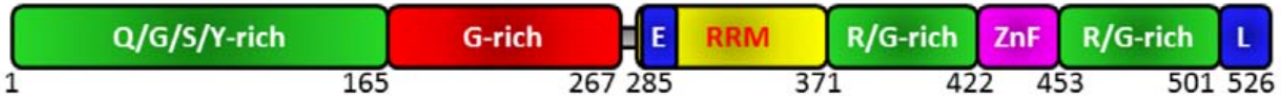
Membraneless Organelles

Protein droplets

* C. J. Oldfield and A. K. Dunker, Ann. Rev. Biochem. 83:553 (2014)

Observations from the literature

- Eukaryotes have more IDRs than prokaryotes; 30-50% of human proteome are disordered regions [1]
- Membraneless organelles compartmentalise biochemistry [5]
- Disease-associated proteins are rich in disorder [1] and loss of fluidity accompanies disease [2]
- FUS, 526 residues: N-term. LC seq (QGSY) - RNA recog, motif; C-term LCD; phase separated FUS is disordered [3]; RRM interacts with membranes [6]

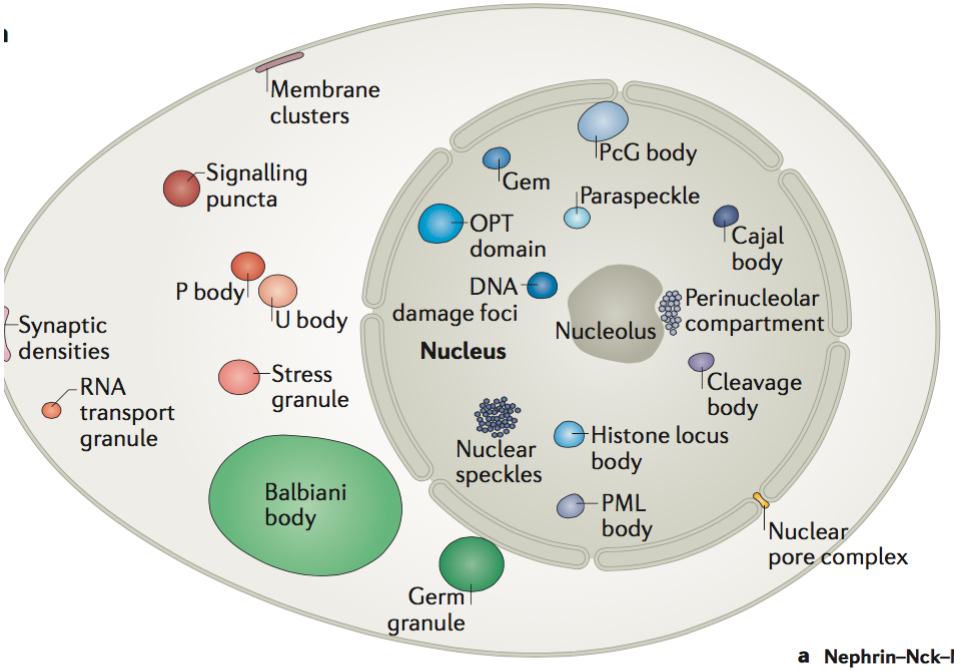


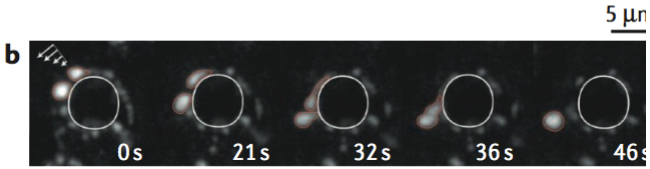
- p53 is disordered at the N, C termini while the central DNA-binding region is ordered; it has tens of binding sites at each terminus (> 40 at N term.) [1]
- BRCA1 has ~40 binding sites in its central IDR, and ~ 400 binding partners [1]
- IDRs are rich in Ser, Arg, Tyr; PTM sites are often located in IDRs [1]
- Replication Protein A has an IDR linker with a *variable* sequence but *conserved rigidity and disorder* [1]
- PSD-95 and SynGap spontaneously phase separate into fluid droplets in vitro [4]

IDR = intrinsically-disordered region
LC = low complexity, repeated amino acids
PTM = post-translational modification

[1] Oldfield and Dunker, Ann. Rev. Biochem. **83**:553 (2014)
[2] Alberti and Hyman, Bioessays **38**:959 (2016)
[3] Burke et al. Molecular Cell **60**:231 (2015)
[4] Zeng et al. Cell **166**:1163 (2016)
[5] Holehouse and Pappu, Biochemistry DOI: 10.1021/acs.biochem.7b01136 (2018)
[6] Lu et al. Bioarxiv **doi:** https://doi.org/10.1101/122671

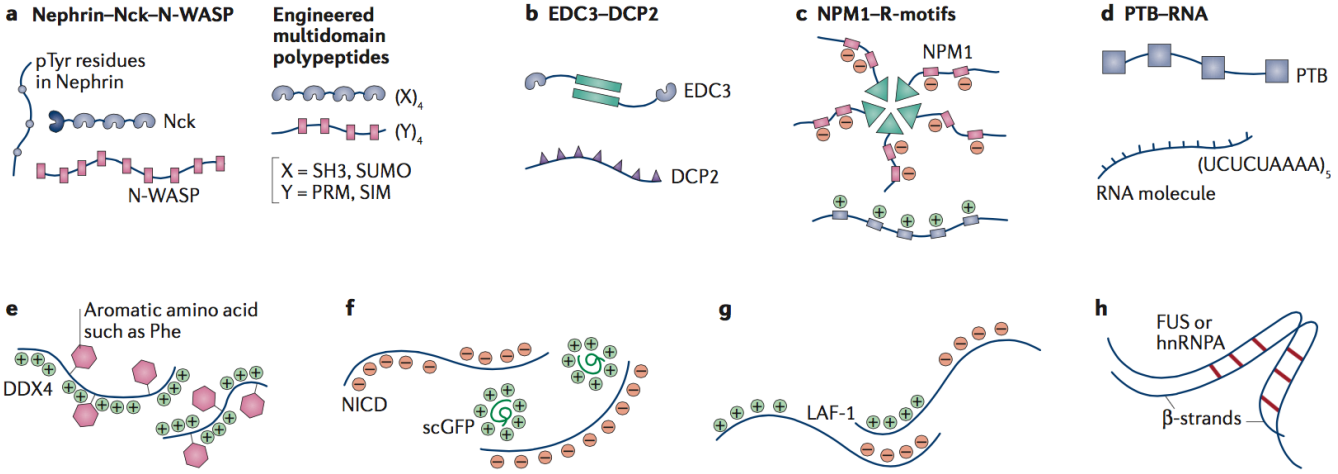
Biomolecular condensates: a new phase of cellular matter?





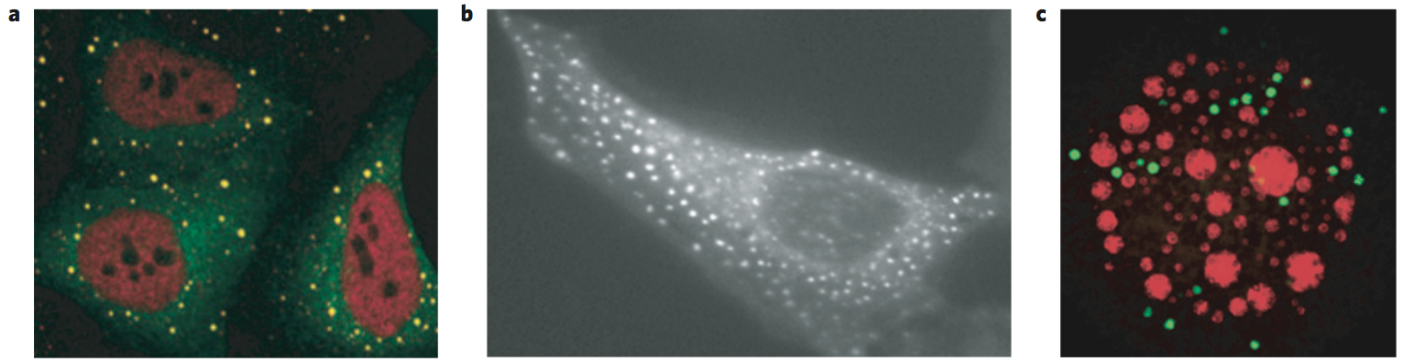
There's a lot!

All formed from long, flexible proteins (Intrinsically-Disordered Proteins)



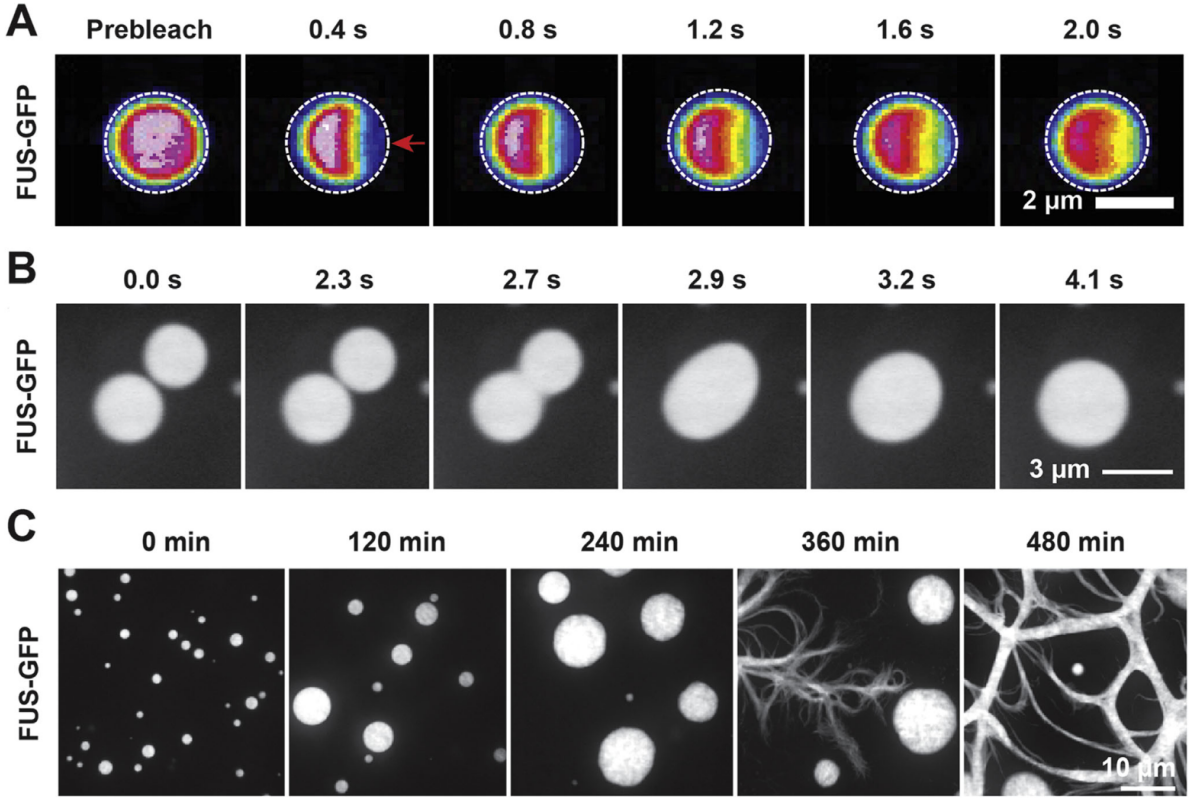
S. F. Banani et al. Biomolecular condensates: organizers of cellular biochemistry, Nat Rev. Mol. Cell Biol. 18:285 (2017)

Physical properties of condensates



Brangwynne et al.
Nature Phys. 11:899 (2015)

Figure 1 | Examples of membrane-less bodies in cells. a, P bodies (yellow) in tissue culture cells (adapted from ref. 63, NPG). **b**, Purinosomes (adapted from ref. 3, AAAS). **c**, Nucleoli (red) and histone locus bodies (green) in the nucleus of a large *X. laevis* oocyte (adapted from ref. 14, NPG).



A) FRAP of FUS-GFP after bleaching at red arrow (in vitro)

B) Merging of FUS-GFP droplets under flow (in vitro)

C) Aging of droplets into fibrous aggregates (in vitro)

Alberti and Hyman,
Bioessays 38:959 (2016)

Opposite extremes of protein models

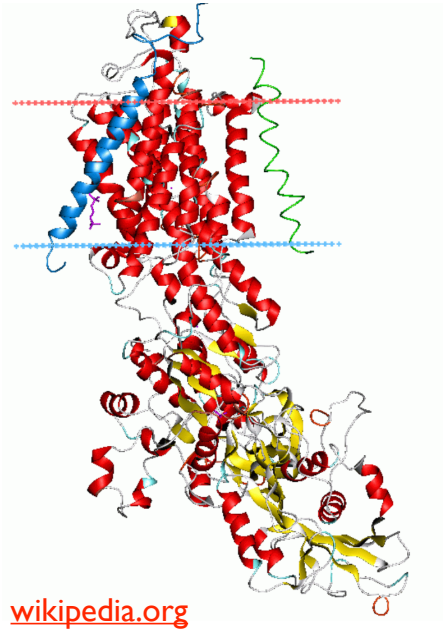
EPFL

Folded protein

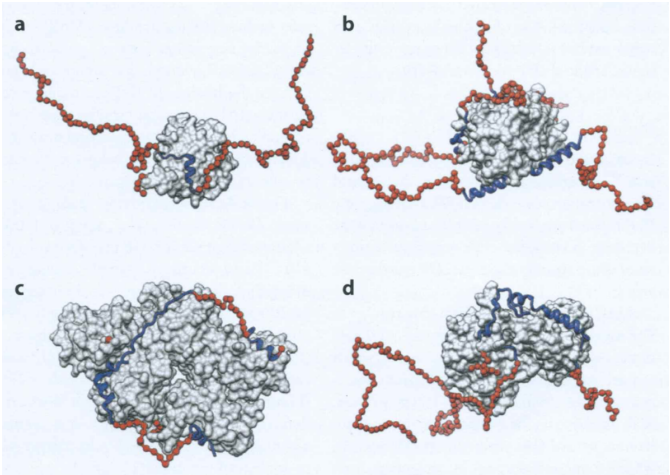
- Unique folded state
- Lowest energy (energy dominated)
- Precise shape
- Precise functions
- Disrupted by single aa mutation
- No model, need the actual protein

Intrinsically Disordered protein

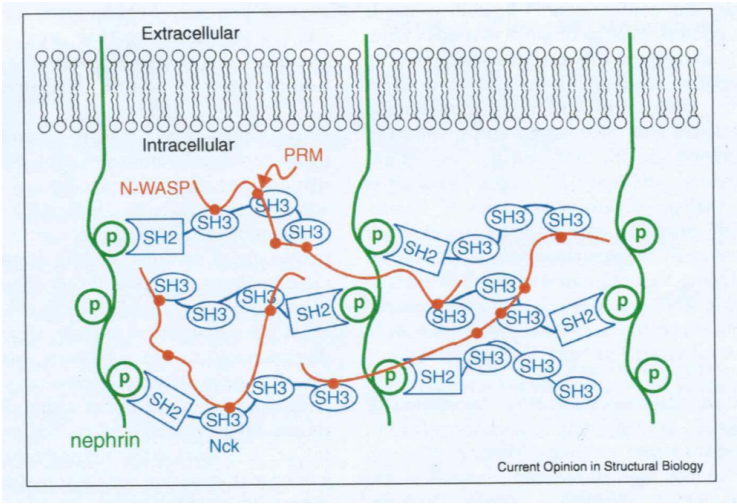
- No unique folded state
- Many conformations of similar energy (entropy dominated)
- Generic binding via multiple, weak sites
- Sequence not conserved but properties are
- Can fold/unfold on binding
- Model it as a phantom chain?



wikipedia.org

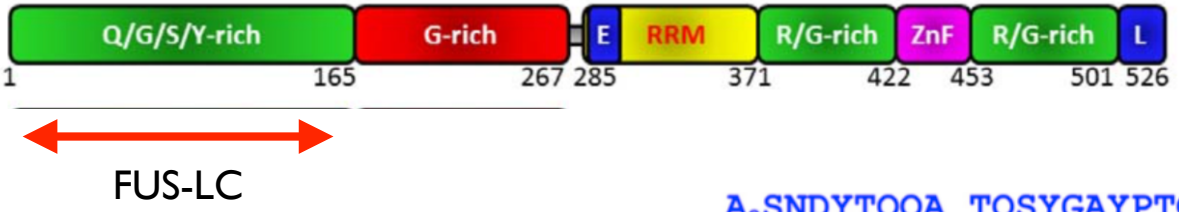


Oldfield and Dunker
Ann. Rev. Biochem.l. 83:553 (2014)



Chong and Forman-Kay,
Curr. Op. Struct. Biol. 41:180 (2016)

Fused in Sarcoma - a "typical" IDP



A_2SNDYTQQA TQSYGAYPTQ PGQGYSQQSS QPYGQQSYSG YSQSTDTSGY$_{50}$
GQSSYSSYGQ SQNTGYGTQS TPQGYGSTGG YGSSQSSQSS YGQQSSYPGY$_{100}$
GQQPAPSSTS GSYGSSSQSS SYGQPQSGSY SQQPSYGGQQ QSYGQQQSYN$_{150}$
PPQGYGQQNQ YNSSSGGGGG GGGGGNYGQD QSSMSSGGGS GGGYGNQDQS$_{200}$
GGGGSGGYGQ QDRG$_{214}$

163 residues M.Wt ~ 16.4 kDa

Why FUS-LC? only 2 charged residues (Asp/D), many Tyr/Y "stickers", little or no secondary structure in solution ⇒ good for coarse-grained simulations

(Wang et al. Cell 174:1 2018)

Intramolecular interactions for FUS-LC in dense phase are transient and weak - "flickering"

(Burke et al. Mol. Cell. 60:231 2015)

Droplets of FUS are ~ 65% water by volume;
Dense phase concentration is ~ 27.8 mM

(Murthy et al. Nature. Struc. Mol. Biol. 26:637 2019)

Residue frequency in FUS 1-163:

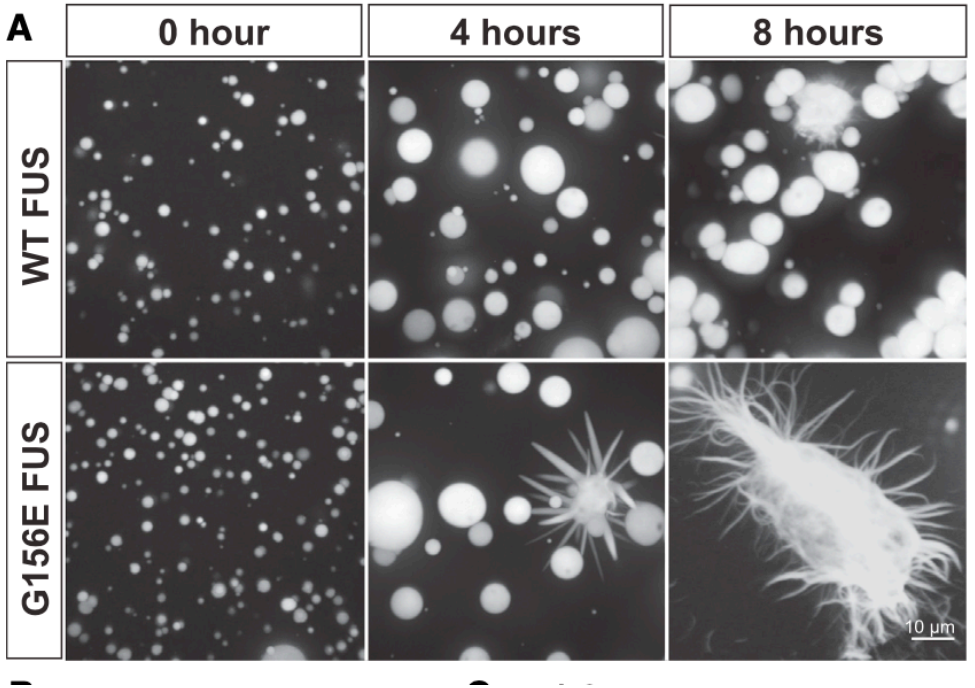
Q - 37
G - 27
S - 43
Y - 24

(if random, we expect ~ 8)

How do we construct a simulation model of IDPs?

Droplets of FUS LC:

- contain many flexible FUS polymers
- usually spherical
- weakly interacting
- highly porous (> 65% water by volume)
- no electrostatics

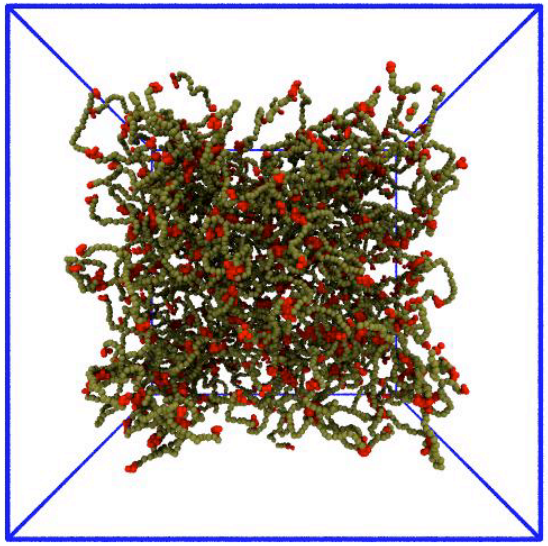


FUS droplets
Patel et al. Cell 2015

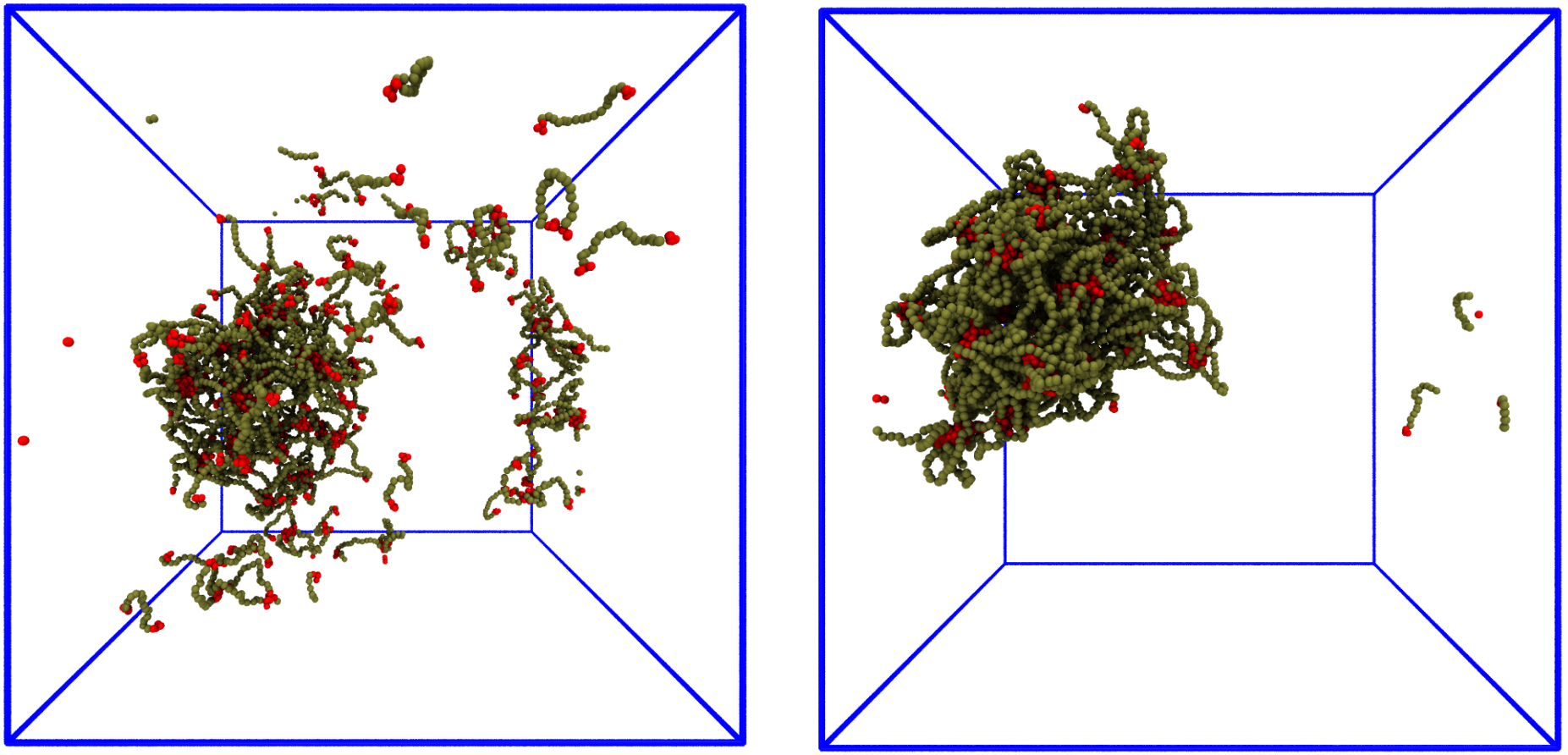
Represent FUS-LC as a linear polymer with sticky sites:

Explore its behaviour as the concentration/length/# and affinity of binding sites are varied.

Measure spatial structure, diffusion in dense phase, rheology, …



Increasing affinity drives phase separation



260 polymers E-B_{16}-E in 48^3 with: (left) a_{EE} = 6 and (right) affinity a_{EE} = 4
(recall that lower aEE means less repulsion, so more attractive)

Methods for characterizing the material properties of biomolecular condensates

Ibraheem Alshareedah, Taranpreet Kaur, and Priya R. Banerjee*
Department of Physics, University at Buffalo, Buffalo, NY, United States
*Corresponding author: e-mail address: prbanerj@buffalo.edu

Abstract

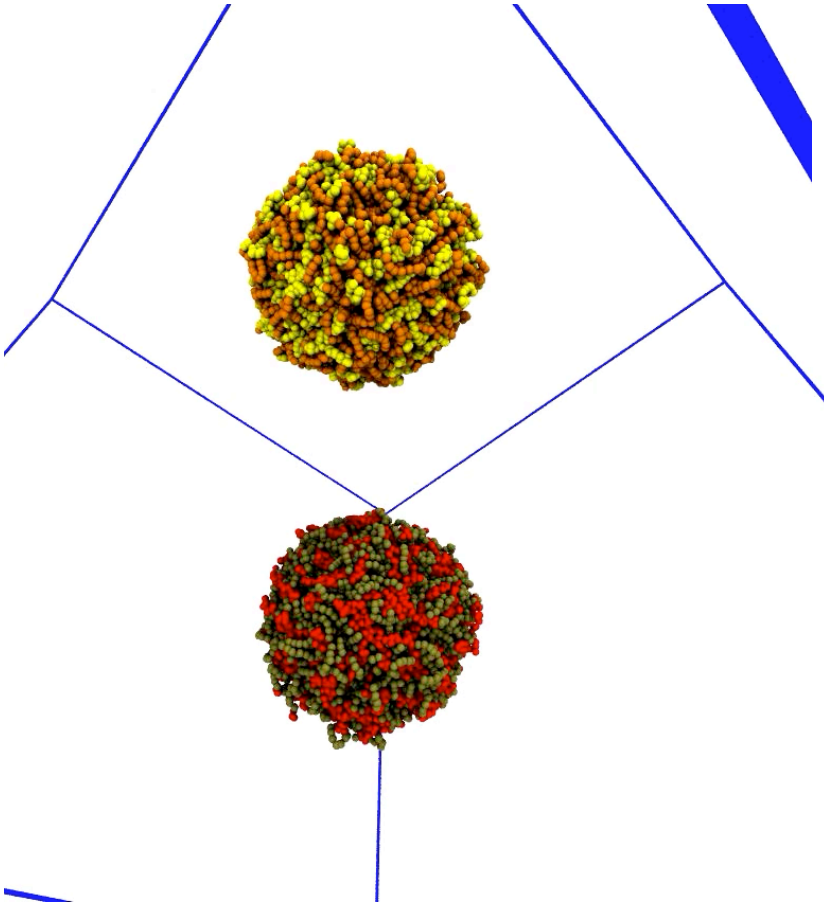
Biomolecular condensates are membrane-less sub-cellular compartments that perform a plethora of important functions in signaling and storage. The material properties of biomolecular condensates such as viscosity, surface tension, viscoelasticity, and macromolecular diffusion play important roles in regulating their biological functions. Aberrations in these properties have been implicated in various neurodegenerative disorders and certain types of cancer. Unraveling the molecular driving forces that control the fluid structure and dynamics of biomolecular condensates across different length- and time-scales necessitates the application of innovative biophysical methodologies. In this chapter, we discuss major experimental techniques that are widely used to study the material states and dynamics of biomolecular condensates as well as their practical and conceptual limitations. We end this chapter with a discussion on more advanced tools that are currently emerging to address the complex fluid dynamics of these condensates.

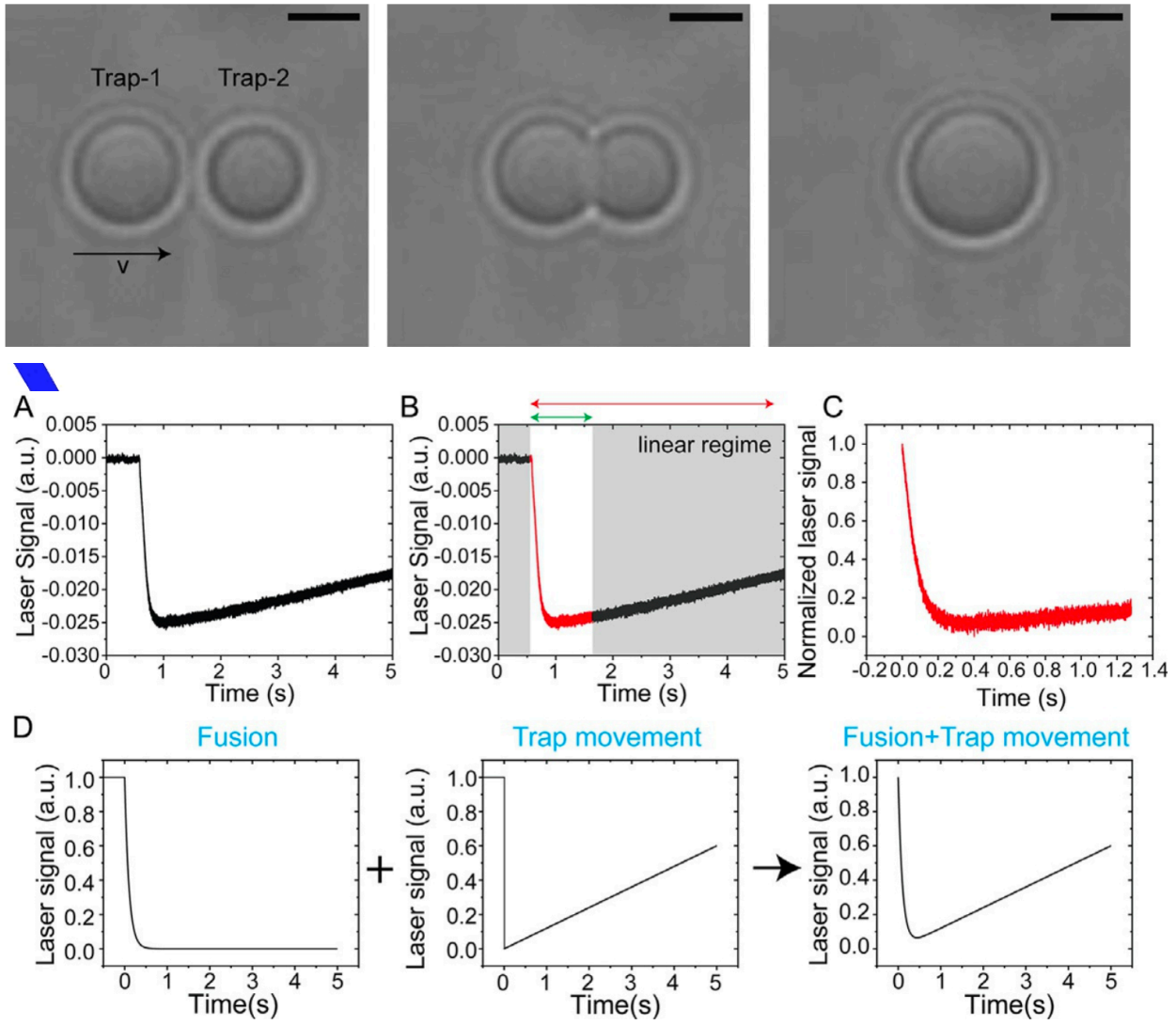
Project ideas in here …

Chapter: https://www.sciencedirect.com/science/article/pii/S007668792030269X

Whole volume:
https://www.sciencedirect.com/bookseries/methods-in-enzymology/vol/646/suppl/C

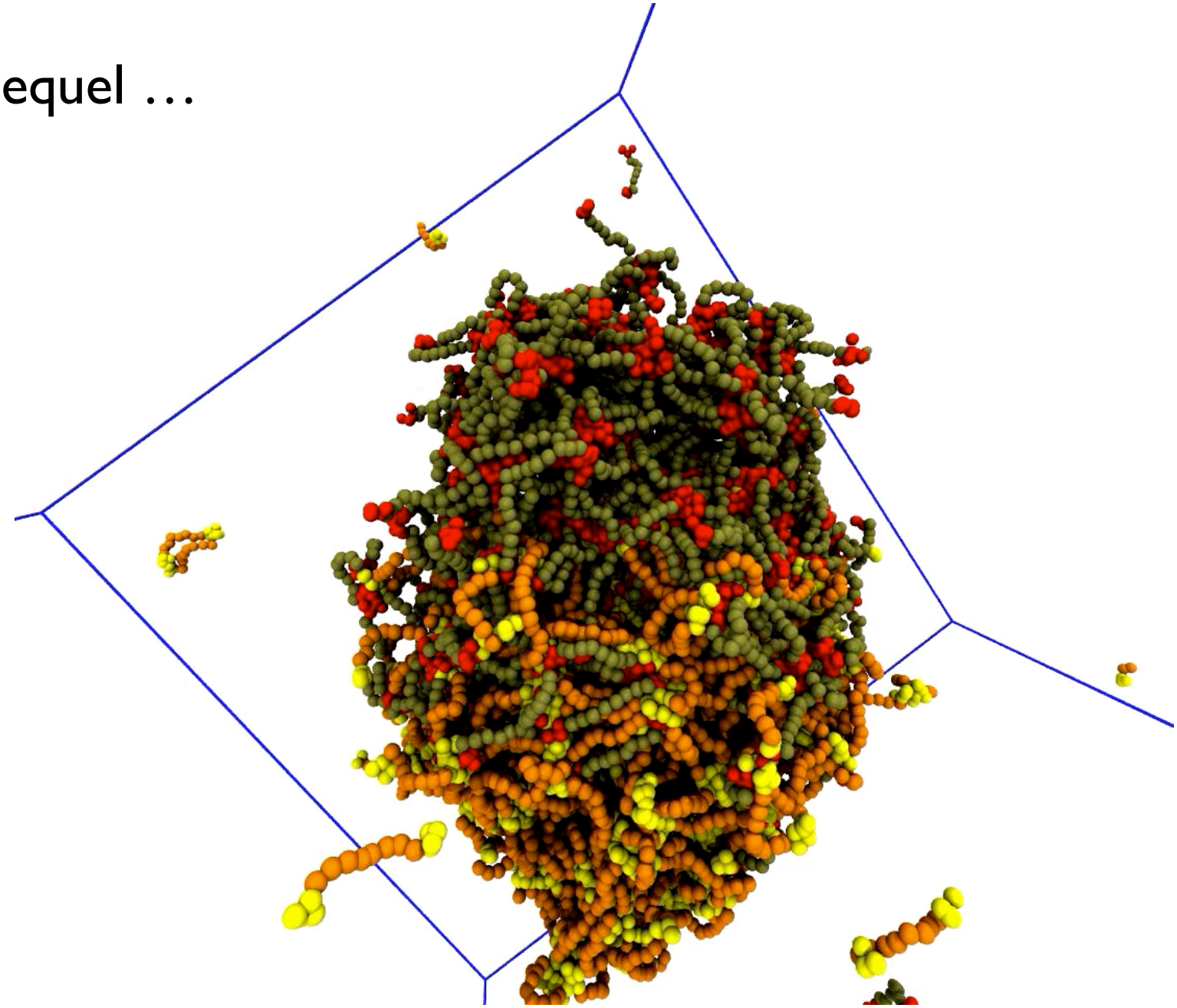
Fusion of two droplets ⇒ fluid

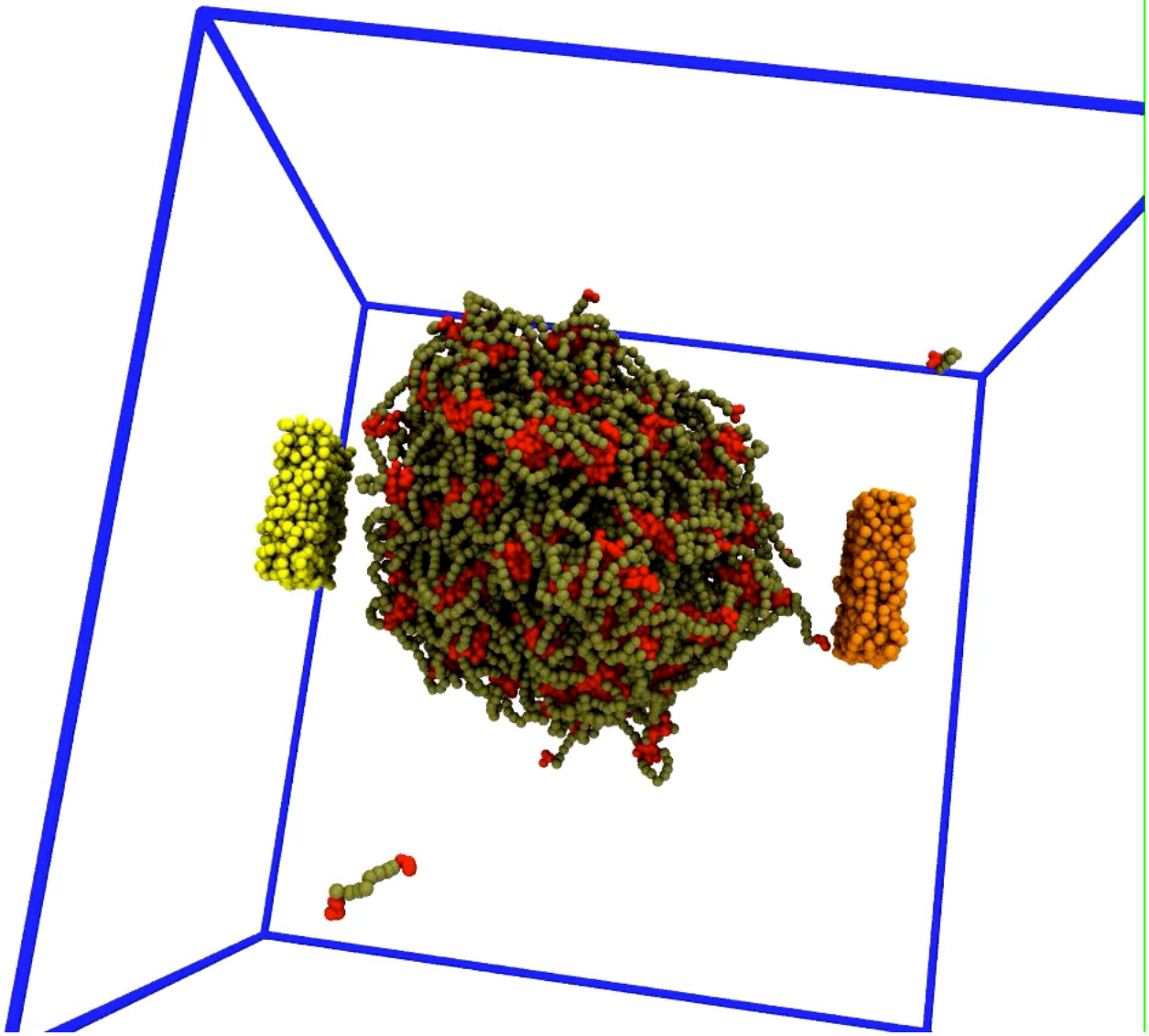




Banerjee et al. Methods Enzymology 2021,
optical trap-driven fusion of two droplets

the sequel …





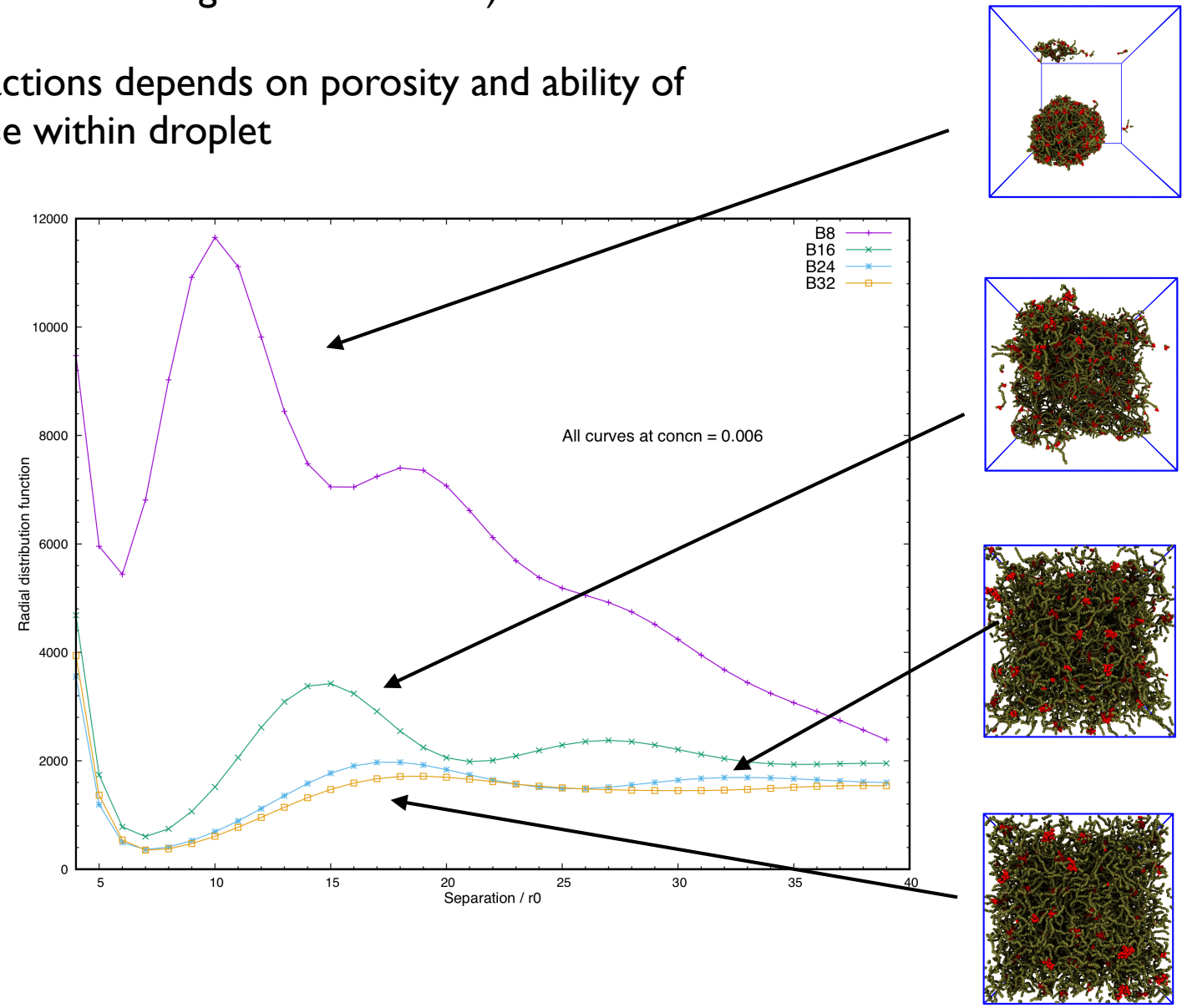
Fast compression,
retains deformed shape
droplet is visco-elastic

Note slab is not “frozen” - we don’t want to break momentum conservation

Structural properties of the dense phase

Radial distribution function reveals the geometric structure of the dense phase (cp., neutron-scattering measurements?)

Speed of enzymatic reactions depends on porosity and ability of client proteins to diffuse within droplet



Making a theoretical model of IDPs?

EPFL

Relevant facts:

- an IDP protein is a flexible polymer with multiple, weak binding sites
- aggregate into spherical droplets with a (small) surface tension
- droplets have low density, not densely packed like oil droplets

Relevant questions:

- How does a polymer's average size in solution scale with its molecular weight?
- At what concentration do proteins in solution "notice" each other?
- What is their shape in solution and in the aggregate (if they form one)?
- What are typical energies of the proteins' interactions?
- What is their entropy?
- Which free energy should we consider?
- What is the shape of the aggregate?

Start with a simple model of an oil droplet

Why do oil molecules aggregate? The hydrophobic effect, but how do we model this?

Each oil molecule is repelled from water with a small interaction energy

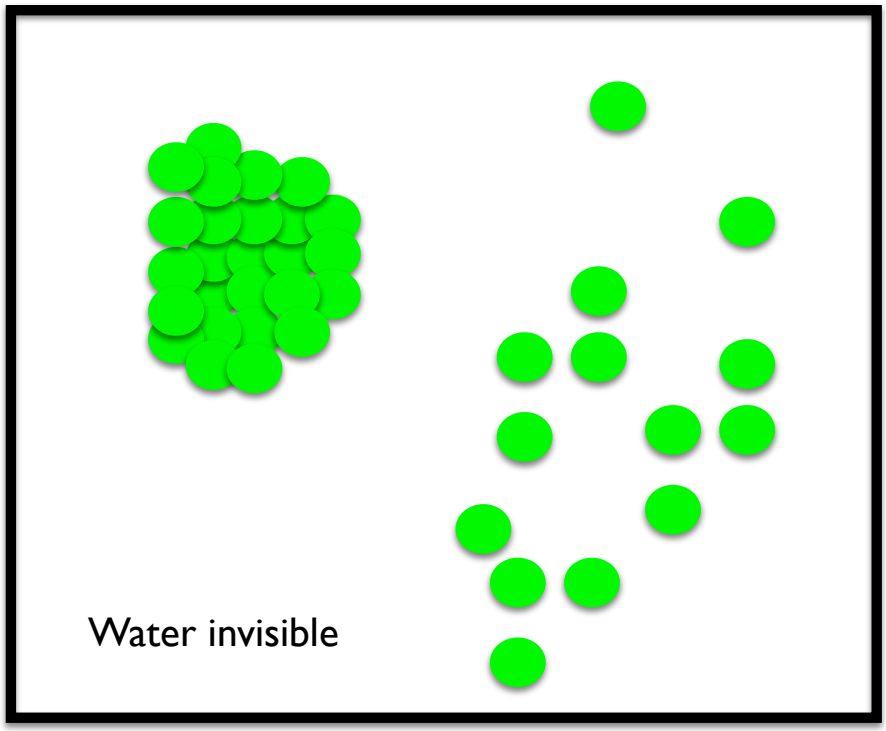
When oil molecules aggregate into a sphere, they reduce their energy by a term proportional to the volume (γ), but still have a repulsive surface energy proportional to the surface area (σ):

$$U(R) \sim 4\pi\sigma R^2 - 4/3\pi\gamma R^3$$

What preferred size droplets form?

Note there is no entropy term here.

(Blackboard plot of U(R), find dU/dR, asymptotes, zeroes, …)

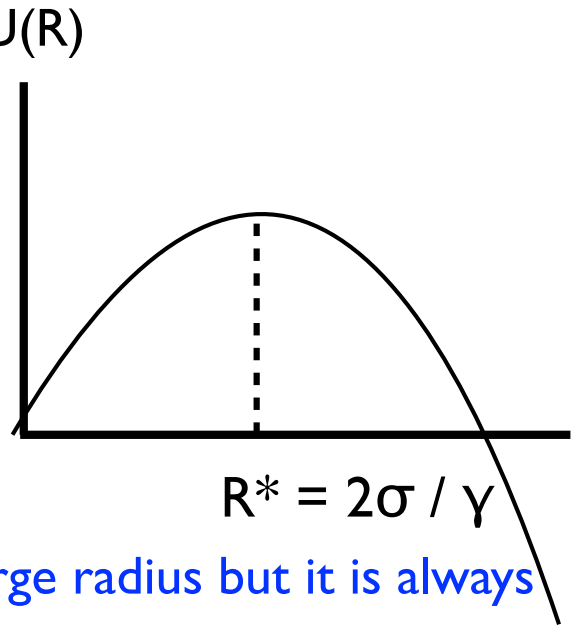


Oil and water phase separate

$U(R) \sim 4\pi\sigma R^2 - 4/3\pi\gamma R^3$

$dU/dR = 8\pi\sigma R - 4\pi\gamma R^2 = 0$

So, the energy has a peak at droplet size R*



If the ratio(σ/γ) is high, the barrier is large and at large radius but it is always energetically favourable to increase the droplet size.

There is no equilibrium droplet size: smaller ones will break up while a larger one would grow without bound ⇒ phase separation.

Is this a good model of IDPs forming membraneless organelles?

Problem is that IDPs are NOT hydrophobic - they are soluble in water. We need a different mechanism to make them aggregate

From Oil to IDPs

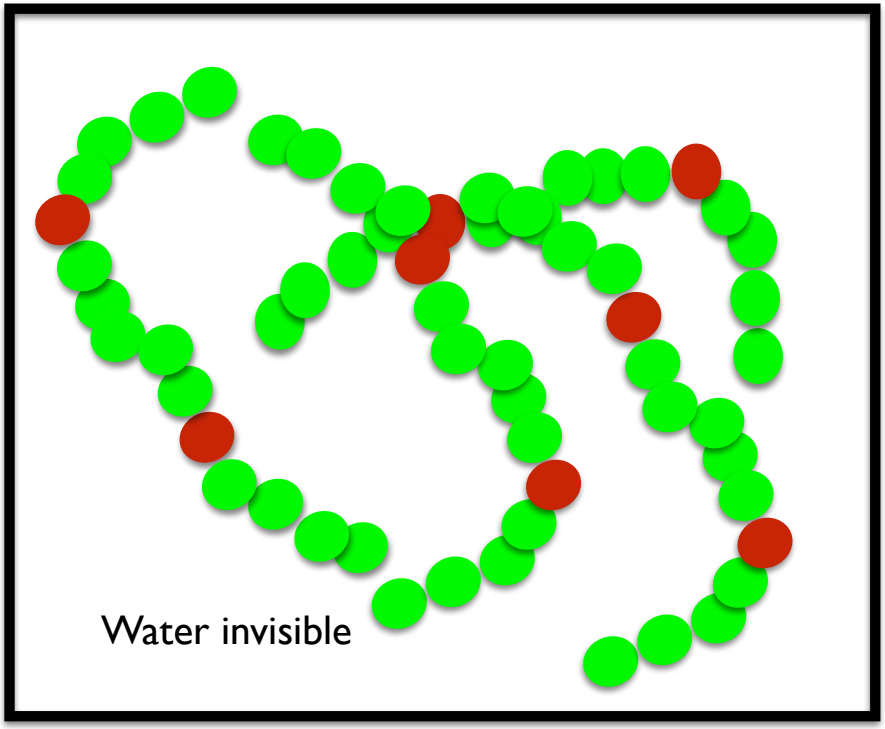
The IDPs are NOT strongly repelled from water (they have few hydrophobic residues), and membraneless organelles have a low surface tension. BUT IDPs have attractive binding sites. So we need to modify the attractive term in the energy U(R).

An important property of IDPs is their conformational **entropy** - they are constantly fluctuating in position.

This suggests we could model them as a phantom chain (from last week)

$\langle R_{ee}^2 \rangle = Na^2$

Question: Oil molecules are long flexible polymers too, why didn't we worry about their fluctuations?



Building a model of a membraneless organelle

EPFL

We assume:

Single disordered protein = a phantom chain with N monomers of size a (/nm): $\langle R_{ee} \rangle = aN^{\nu}$ (ν is a constant that controls how $\langle R_{ee} \rangle$ varies with N: $\nu = 0.5$ for phantom chain but can be different for more complex models)

Aggregated proteins = **spherical** droplet of radius R with a surface tension σ (J/m^2)

Why?

- Experiments show they are spherical
- Experiments show droplets are fluid (FRAP)
- Non-zero surface tension from experiments
- Volume of the system does not change whether polymers are dispersed or aggregated, so pressure is not important (is this accurate?)

We use the Helmholtz free energy: $F = U - TS$ as volume and temperature are constant.

Can we make a model like the oil droplet one?

Model predictions

U = surface energy of spherical aggregate = surface tension x area

P(R) = prob. of polymer with R_{ee} having a size of R ~ exp($-(R - R_{ee})^2 / 2R_{ee}^2$) - why?

The Helmholtz free energy of N polymers in a sphere of radius R is then:

$F = U - TS = 4\pi \sigma R^2 - Nk_BT \log P(R)$

$= 4\pi \sigma R^2 + Nk_BT (R - R_{ee})^2 / 2R_{ee}^2$

Minimising this gives:

$$R = R_{ee} / (1 + 8\pi \sigma R_{ee}^2 / Nk_BT)$$

What does this look like? Is it dimensionally correct? (Blackboard-gnuplot)

Is it correct? Have we made hidden assumptions? We've ignored the binding sites.

Better model - include binding energy

U = surface energy of spherical aggregate = surface tension x area

P(R) = prob. of polymer with R_{ee} having a size of R ~ $\exp(-(R - R_{ee})^2 / 2R_{ee}^2)$

Add the binding energy of M(R) active binding sites per polymer: $-\varepsilon$

The Helmholtz free energy is now:

$F = U - TS = 4\pi \sigma R^2 - N M(R) \varepsilon / 2 - N k_B T \log P(R)$

1) But how many binding sites M do the IDPs make?

2) And how does M depend on the aggregate's radius? Are they all active or only some depending on the conformation of the polymer?

This is where is rapidly gets hard …

Break / Exercise

- Creating a rigid nanoparticle by command in DPD
- Micelle formation and SAXS scattering function
- Creating a rigid fibril and diffusing nanoparticles

Creating a rigid nanoparticle by command in DPD

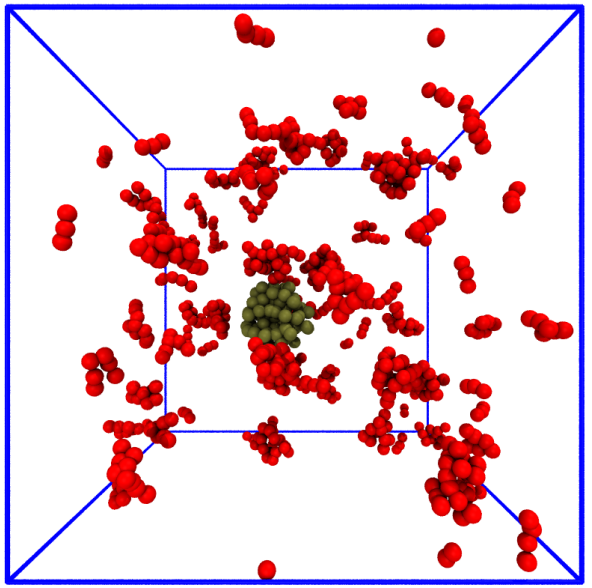
```
Bead  W
      0.5
      25
      4.5

Bead  O
      0.5
      75   25
      4.5  4.5



Bond      O O   128 0.5
BondPair  O O O   15.0  0.0

Polymer Water  0.99  " (W) "
Polymer Oil    0.01  " (O O O) "
```

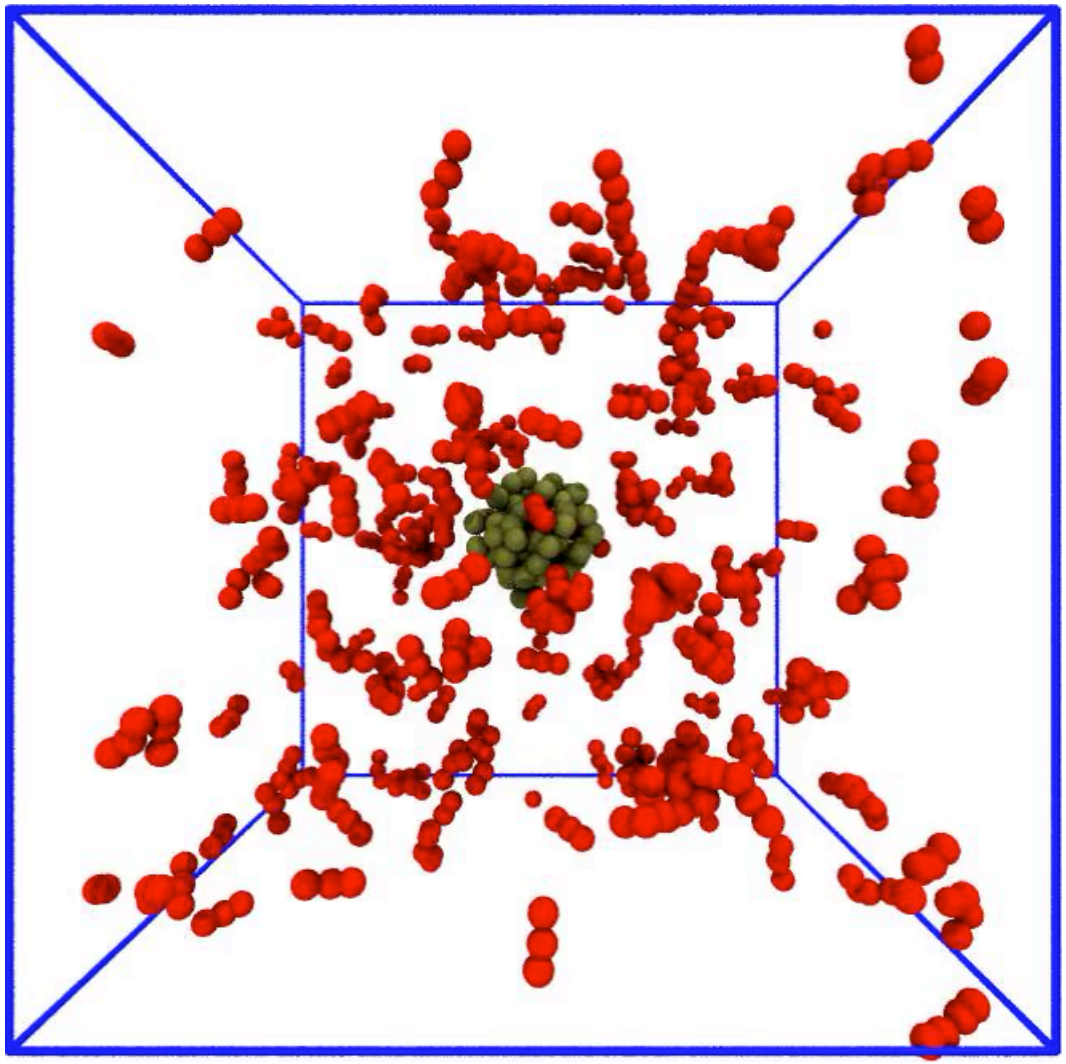


```
Command SelectBeadTypeInSphere          100 Ball1Beads    W     0.5 0.5 0.5 0.0 2.0
Command SelectPolymerTypeHeadInSphere   100 Ball1        Water  0.5 0.5 0.5 0.0 2.0
Command PolymerisePolymersInTarget      100 Ball1 12 1.5 1.0 128.0 0.5
Command SetTargetDisplayId              100 Ball1 1
Command ChangeNamedBeadType             100 Ball1Beads Nanoparticles

Command SetDPDBeadConsInt               100 Nanoparticles W 20
```

Other geometric regions that can be specified are:

cylinder, ellipsoid, pentagon, slice, sphere. spherical cap, whole simbox



Note

The oil molecules aggregate into droplets, i.e., non covalent forces

The beads in the nanoparticle are tied together with stiff Hookean springs, i.e., covalent bonds

Micelle formation and SAXS scattering function

```
Bead  H
      0.5
      30
      4.5

Bead  T
      0.5
      35  10
      4.5 4.5

Bead  W
      0.5
      30    75  25
      4.5   4.5 4.5


Bond    H H  128    0.5
Bond    H T  128    0.5
Bond    T T  128    0.5

BondPair    H T T   15.0    0.0
BondPair    T T T   15.0    0.0

Polymer Water  0.99   " (W) "
Polymer Lipid  0.01   " (H H T T T T) "
```

Note

1) Simplest case is single-tailed surfactants H_2T_4 that spontaneously aggregate into micelles.

2) What is the average size of the micelles?

3) Are they equilibrated?

SAXS scattering function command

Scattering function is calculated using the Debye formula (from Paissoni et al. JCTC 2020):

2.2. Hybrid-Resolution SAXS-Driven Metainference Simulations. Given a coarse-grain representation of a molecule of N atoms as a collection of M beads, each comprising a variable number of atoms, if the form factors $F(q)$ of the beads are known, the scattering intensities can be approximated as

$$I(q) = \sum_{i=1}^{M} \sum_{j=1}^{M} F_i(q) F_j(q) \frac{\sin(qR_{ij})}{qR_{ij}}$$

where R_{ij} indicates the distance between the center of mass of beads ij and with the sum running over the number of beads.

Note

This is a VERY expensive function to calculate.

It slows the code by a factor of 10!

Don't sample too often.

JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC

Determination of Protein Structural Ensembles by Hybrid-Resolution SAXS Restrained Molecular Dynamics

Cristina Paissoni,* Alexander Jussupow, and Carlo Camilloni*

Cite This: J. Chem. Theory Comput. 2020, 16, 2825–2834 Read Online

SAXS scattering function command

```
Polymer Water   0.99    " (W) "
Polymer Lipid   0.01    " (H H T T T T) "
```

```
Time                100000
SamplePeriod        1000
AnalysisPeriod      50000
DensityPeriod       100000
DisplayPeriod       1000
RestartPeriod       100000
```

```
Command SaveSAXS                            50000  1  100 0 1 -1
Command SetSAXSProcessBeadElectronNo        50000  1  H  10.0
Command SetSAXSProcessBeadElectronNo        50000  1  T  5.0
```

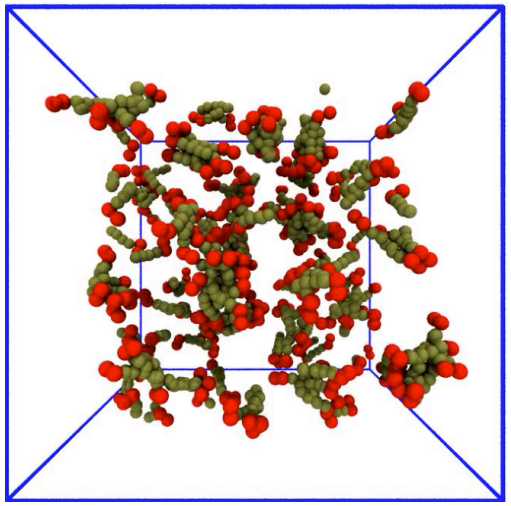
Specify which molecules to include by putting 1 or 0 in the order they are defined. We have to end with -1 because the code doesn't know how many molecule types there are.

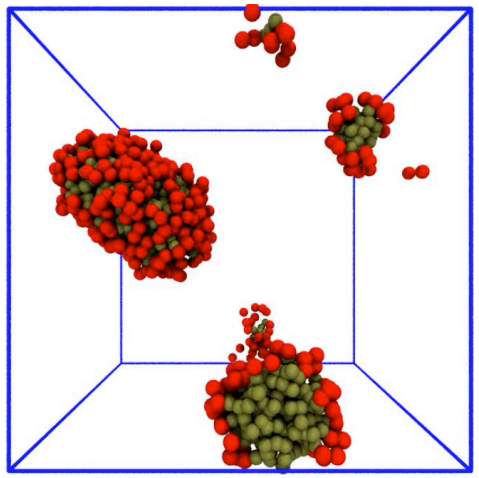
AnalysisPeriods to average over

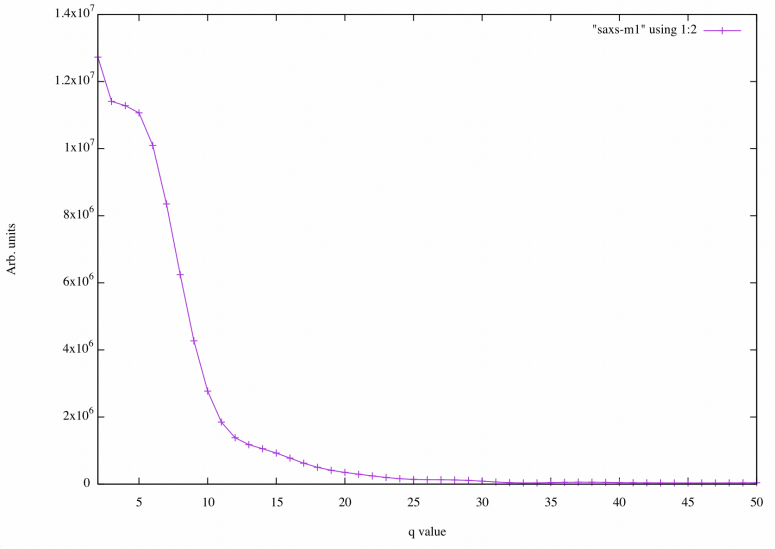
of q values

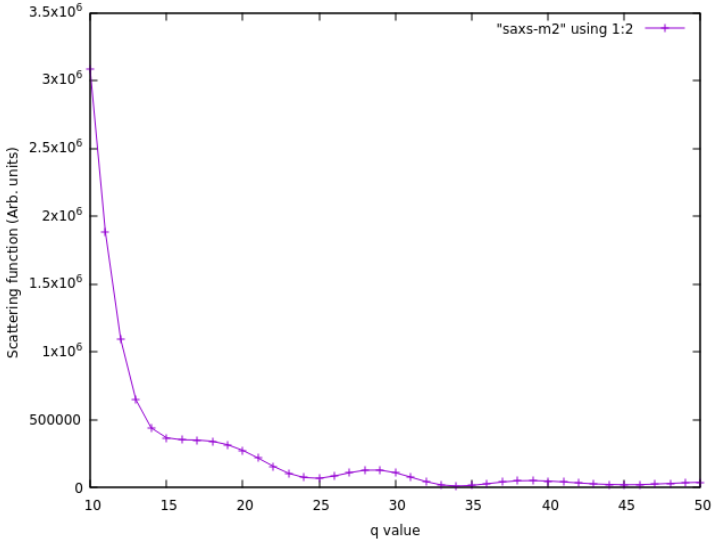
# of electrons per bead of the specified type.

NB A DPD bead can represent many atoms, so you have to multiply by atoms/bead to get this value.









Note

1) The scattering function is calculated over second half of the simulation, but the micelles are still changing size.

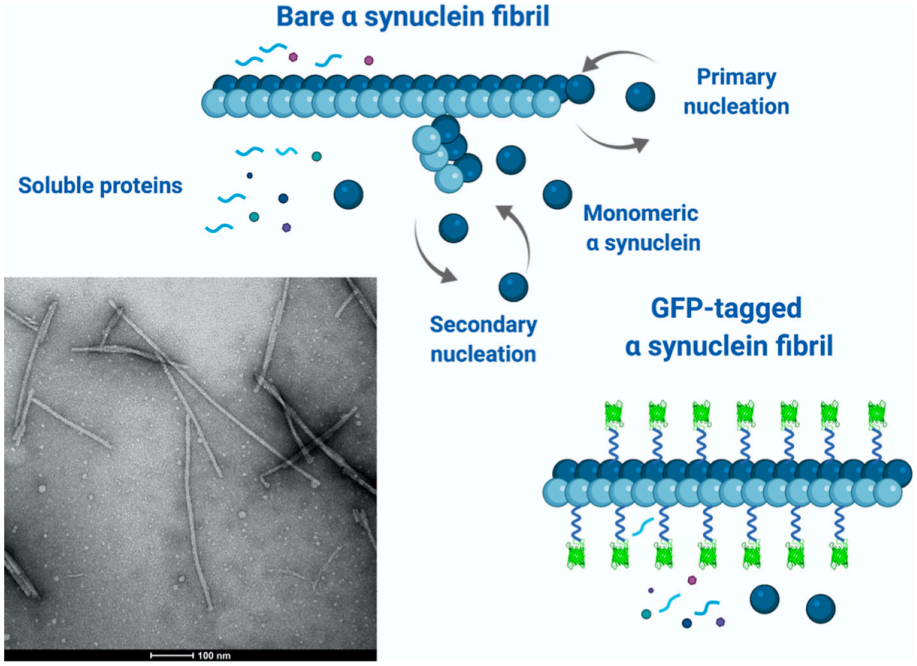
2) The normalisation of the I(q) is not fully defined yet.

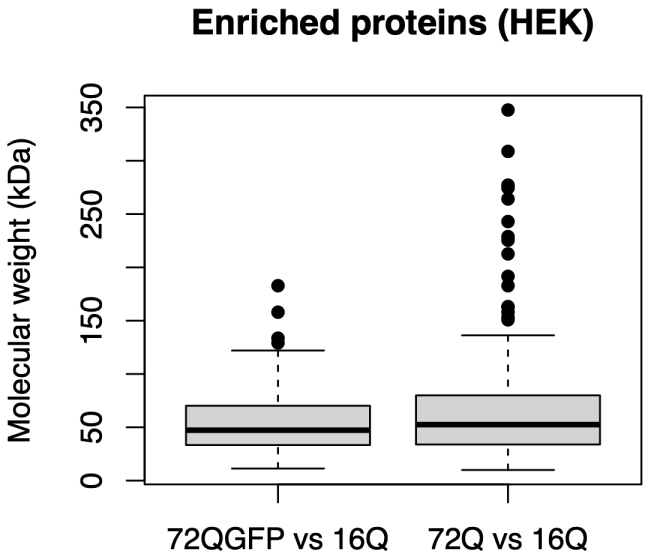
3) The range of q is defined by the number of points the user specifies and the bounds:

$$q_{min} = 2\pi / \text{Box size}$$

$$q_{max} = 2\pi / d_0, \text{ where } d_0 = \text{bead diameter.}$$

GFP tags create a size-dependent molecular sieve around a fibril

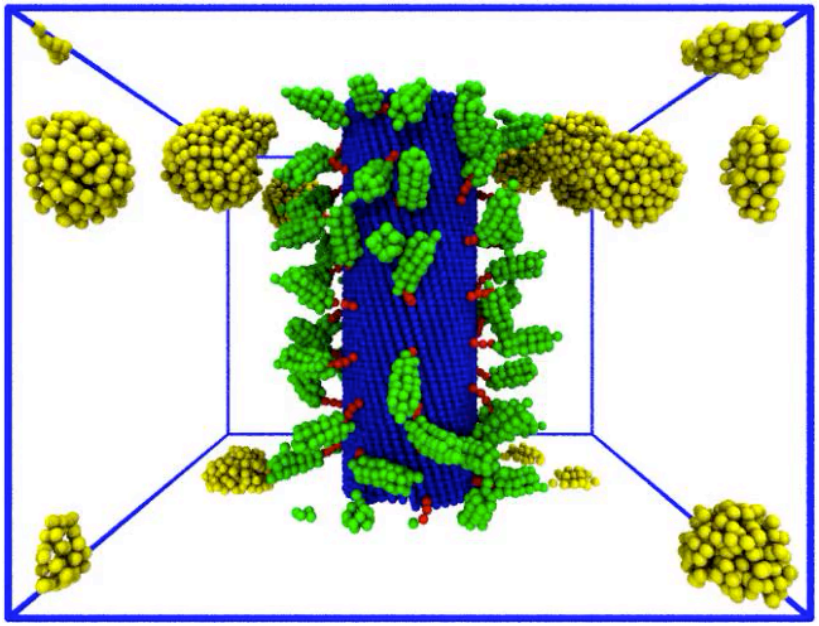


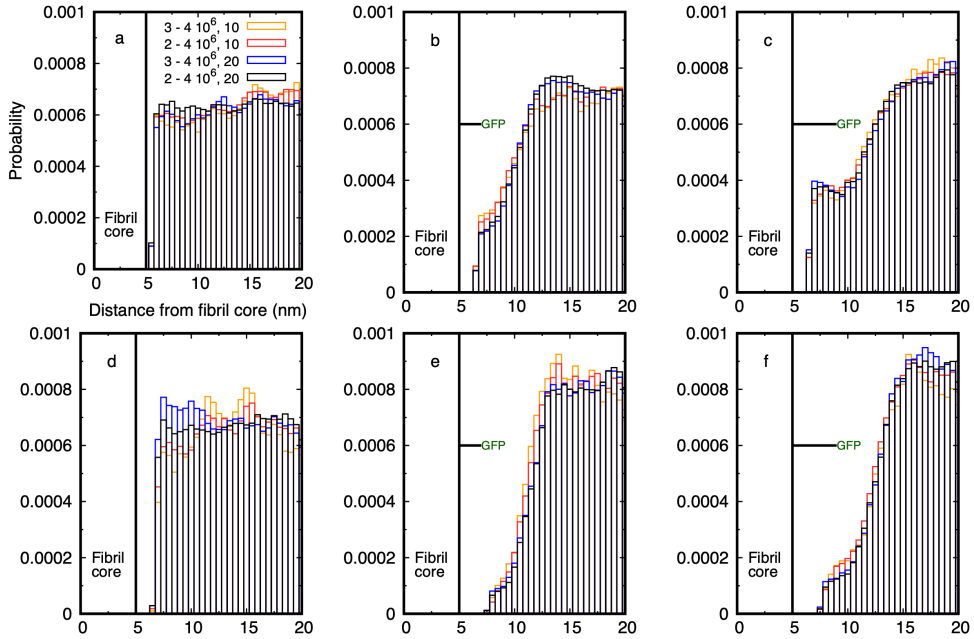


with GFP tag
no high M.wt
proteins in Htt inclusions

no GFP tag

Shillcock et al. Comp. Struct. Bio. J 20:309 (2022)

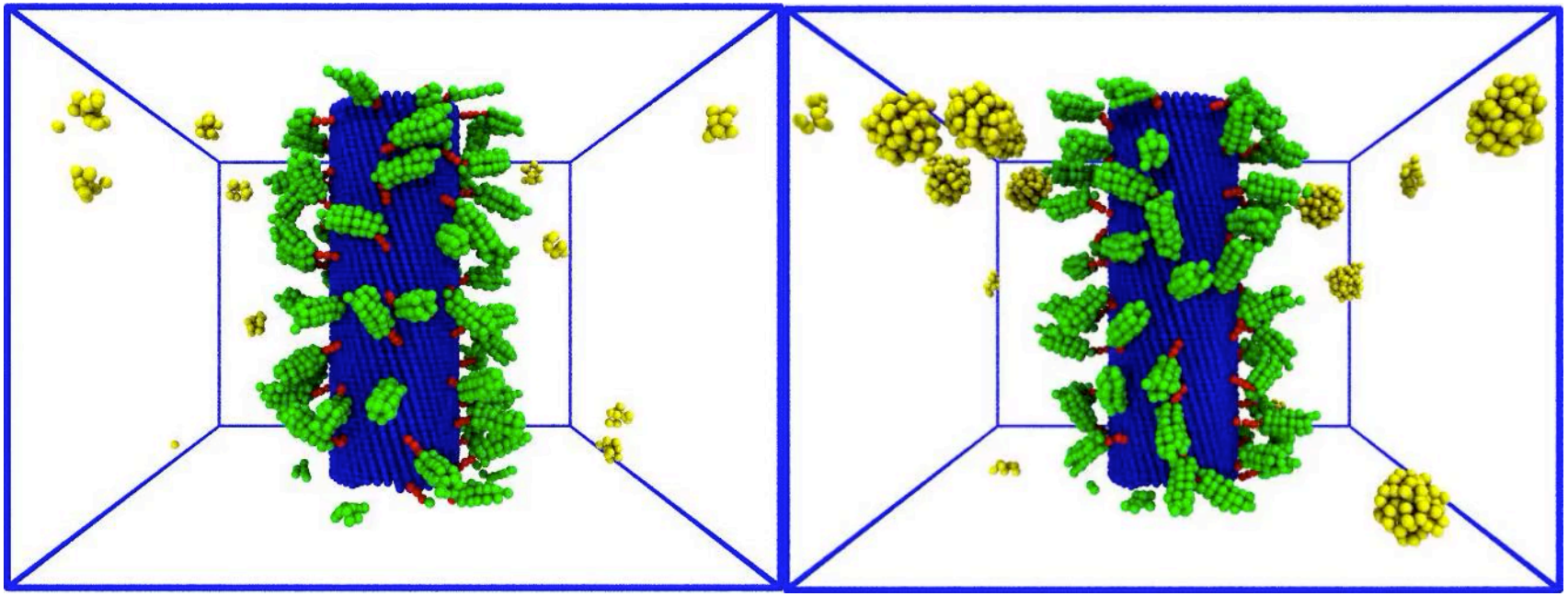




Sphere radius
1 nm

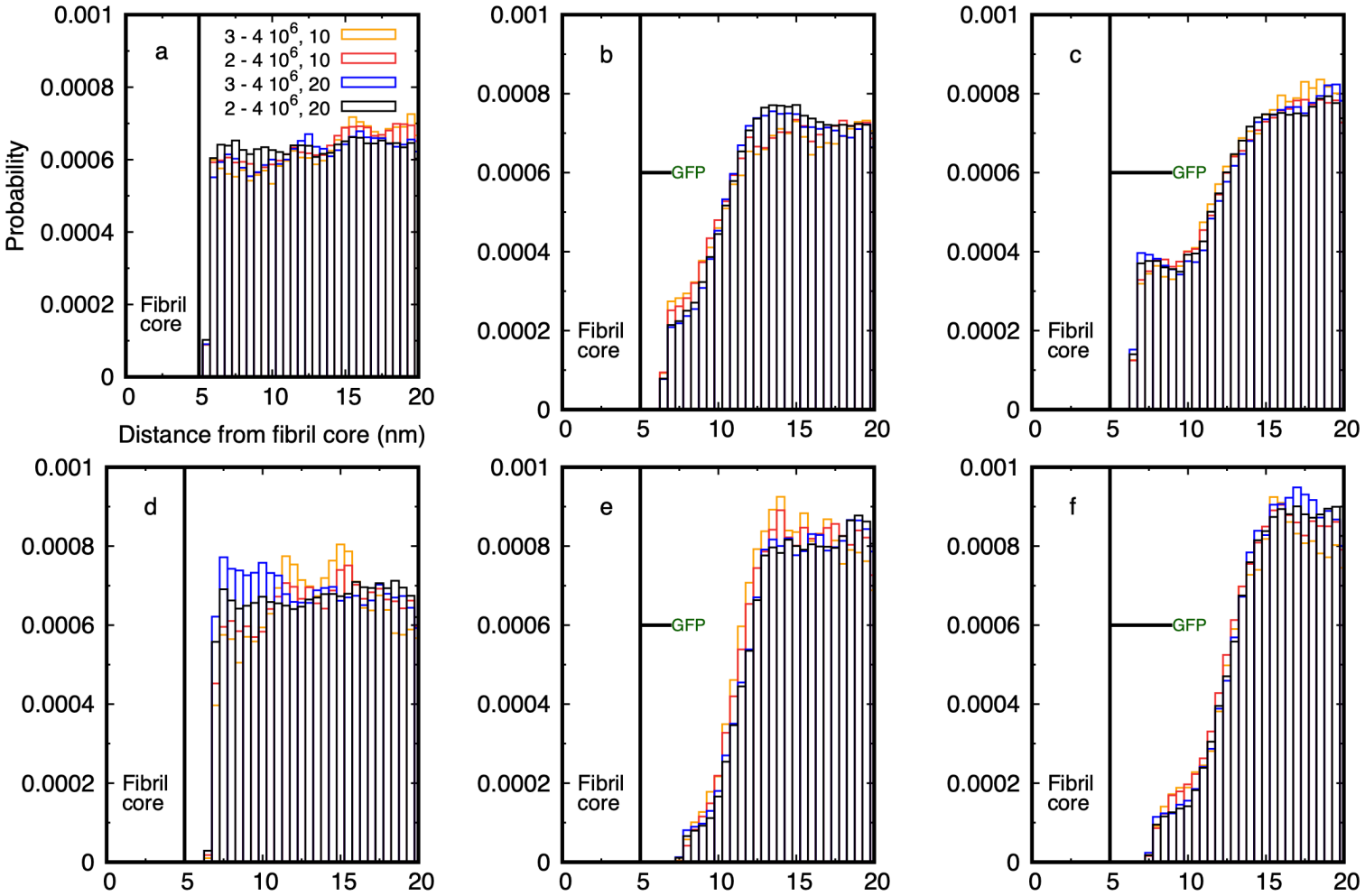
2 nm

Create a rigid fibril with GFP on a linker, and measure the RDF of the nanoparticles as they diffuse around it



RDF shows that the GFP forms a size-dependent "sieve" around the fibril surface — blocking large NPs but letting smaller ones past.

If the linker length > NP size, it **enhances** their residence time at the surface



Unfortunately … it's complicated to set up multiple NPs and a fibril

```
Command CompositeCommandTarget        3000 aggregate
Command CompositeCommandTarget        3000 aggregateBeads

Command CreateCommandGroup            3000 createQD
Command AddCommandToGroup             3000 createQD  SelectBeadTypeInSphere             var1 var2 var3 var4 var5 var6 var7
Command AddCommandToGroup             3000 createQD  SelectPolymerTypeHeadInSphere      var1 var2 var3 var4 var5 var6 var7
Command AddCommandToGroup             3000 createQD  CountBeadTypeInTarget              var1 var2
Command AddCommandToGroup             3000 createQD  PolymerisePolymersInTarget        var1 var2 var3 var4 var5 var6
Command AddCommandToGroup             3000 createQD  AddTargetToComposite               var1 var2
Command AddCommandToGroup             3000 createQD  AddTargetToComposite               var1 var2


Command SetArgumentToStringSequence   3000  createQD  1 var1 nanoparticle1
Command SetArgumentToStringConstant   3000  createQD  1 var2 W
Command SetArgumentToRealConstant     3000  createQD  1 var6 0.0
Command SetArgumentToRealConstant     3000  createQD  1 var7 1.0
Command SetArgumentsTo3dRectangularLatticeVector 3000  createQD  1 var3 1 var4 1 var5  2 1 6  0.15 0.75 0.1  0.2 0.2 0.15

Command SetArgumentToStringSequence   3000  createQD  2 var1 nano1
Command SetArgumentToStringConstant   3000  createQD  2 var2 Water
Command SetArgumentToArgument         3000  createQD  2 var3 1 var3
Command SetArgumentToArgument         3000  createQD  2 var4 1 var4
Command SetArgumentToArgument         3000  createQD  2 var5 1 var5
Command SetArgumentToArgument         3000  createQD  2 var6 1 var6
Command SetArgumentToArgument         3000  createQD  2 var7 1 var7

Command SetArgumentToArgument         3000  createQD  3 var1 1 var1
Command SetArgumentToIntegerConstant  3000  createQD  3 var2 0
```

these are only ~ 1/2 the commands needed.