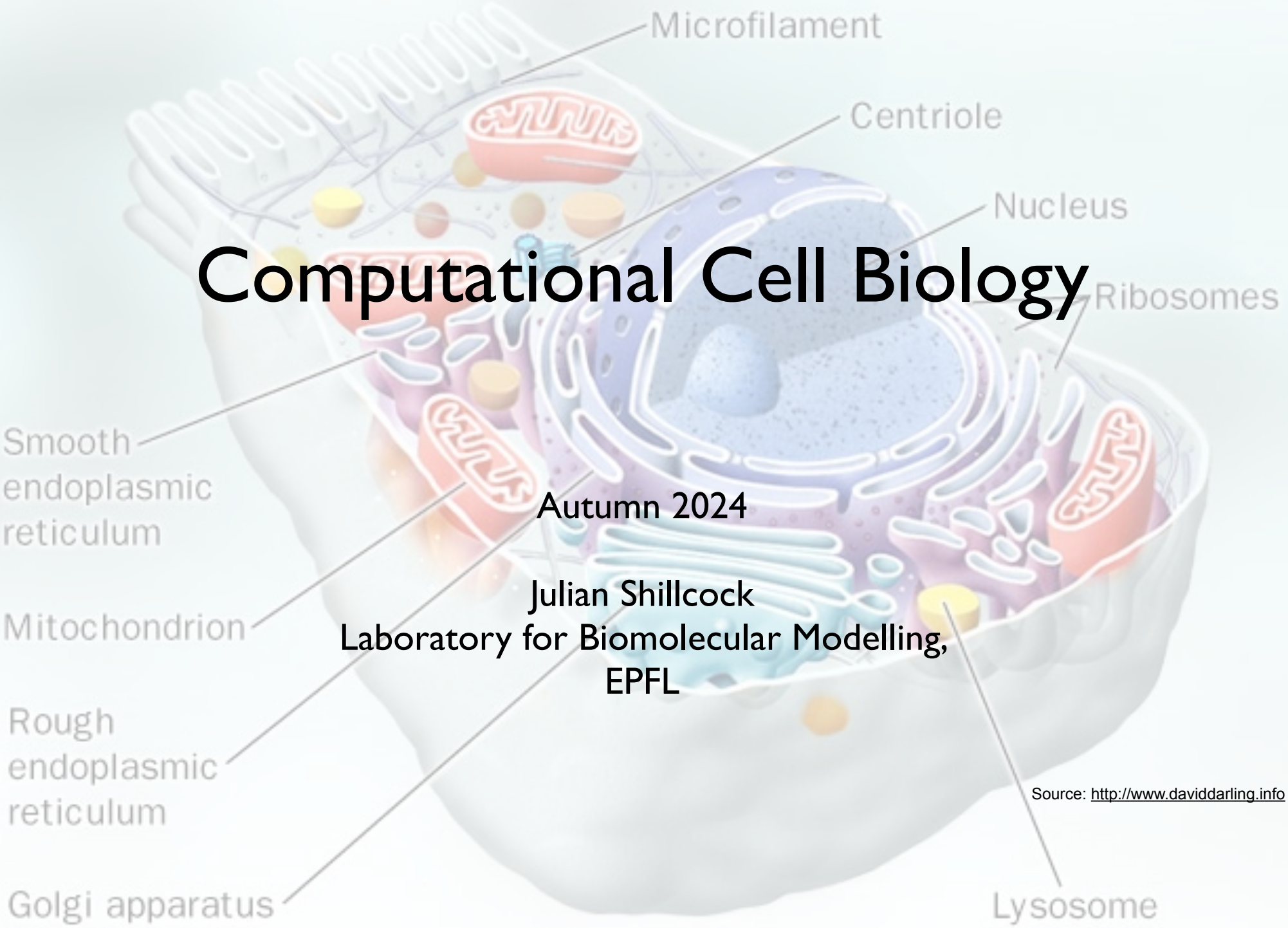


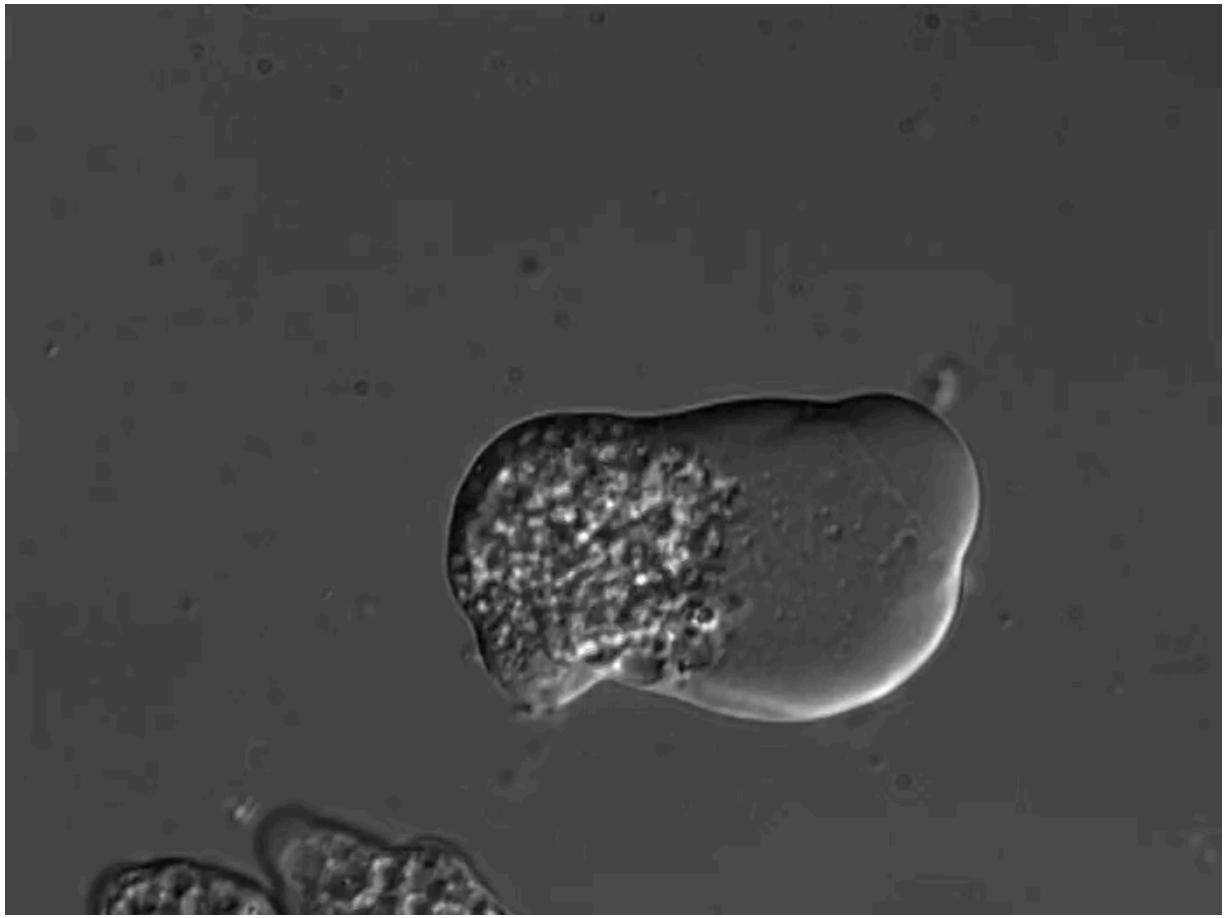
Computational Cell Biology

Autumn 2024

Julian Shillcock
Laboratory for Biomolecular Modelling,
EPFL

Source: http://www.daviddarling.info

Is the cell a machine?



Entamoeba histolytica - anaerobic protozoan

Differential interference microscopy, 5x speedup

hyaline - clear cytosol

vesicle-filled - granular cytosol

leading edge - lobopod

The Roberto Stock group at IBt-UNAM

How can we build a model of this?

We need to know **what** is there, and **what** does it do?

If it is a machine, it's not like a car

EPFL

Common design principles of artificial machines:

- whole is made of precisely-arranged parts that execute ***stable/periodic*** functions ***independently*** of the others (chassis, wheels, transmission, engine, electrics, washer, windscreen, seats, etc. : if brakes fail, the lights still work)
- almost nothing is in equilibrium (by design)
- functions are ***independent*** of the environment (temperature, pressure, etc.)

Common design principles of cells:

- cellular cytosol, proteins, all molecules continually move and *interact* (diffusion, filaments assemble/disassemble, mechanical forces, gradients drive flow)
- many functions operate close to equilibrium because leaving equilibrium is expensive *
- cellular functions often *require randomness* derived from environment (T, P), e.g., diffusion; they are *strongly coupled* to their (changing) environment internally and externally

* cell is often said to be non-eq., but it uses eq. for stability, e.g., [ATP]

Journal of Theoretical Biology 477 (2019) 108–126

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/jtb

Is the cell *really* a machine?

Daniel J. Nicholson[a,b,*]

[a] Konrad Lorenz Institute for Evolution and Cognition Research, Martinstraße 12, Klosterneuburg, 3400, Austria
[b] Centre for the Study of Life Sciences (Egenis), University of Exeter, Byrne House, St. German's Road, Exeter, EX4 4PJ, United Kingdom

See moodle page for today

Quiz - What is your background?

This is NOT an exam! It's supposed to be a fun exercise to see how familiar the language I use to describe cell biology, physics and numerical analysis is.

There are 36 questions to be answered with 1 for **T** or 0 for **F**.

You have ~30 minutes, but can take a break if you finish early and send me your answers.

You can ask questions during the test
but no checking with google or conferring with neighbours!

Send answers to me in an email with the format:

To: julian.shillcock@epfl.ch
Subject: BIOENG 455
Body:
1 0
2 1
3 0
etc

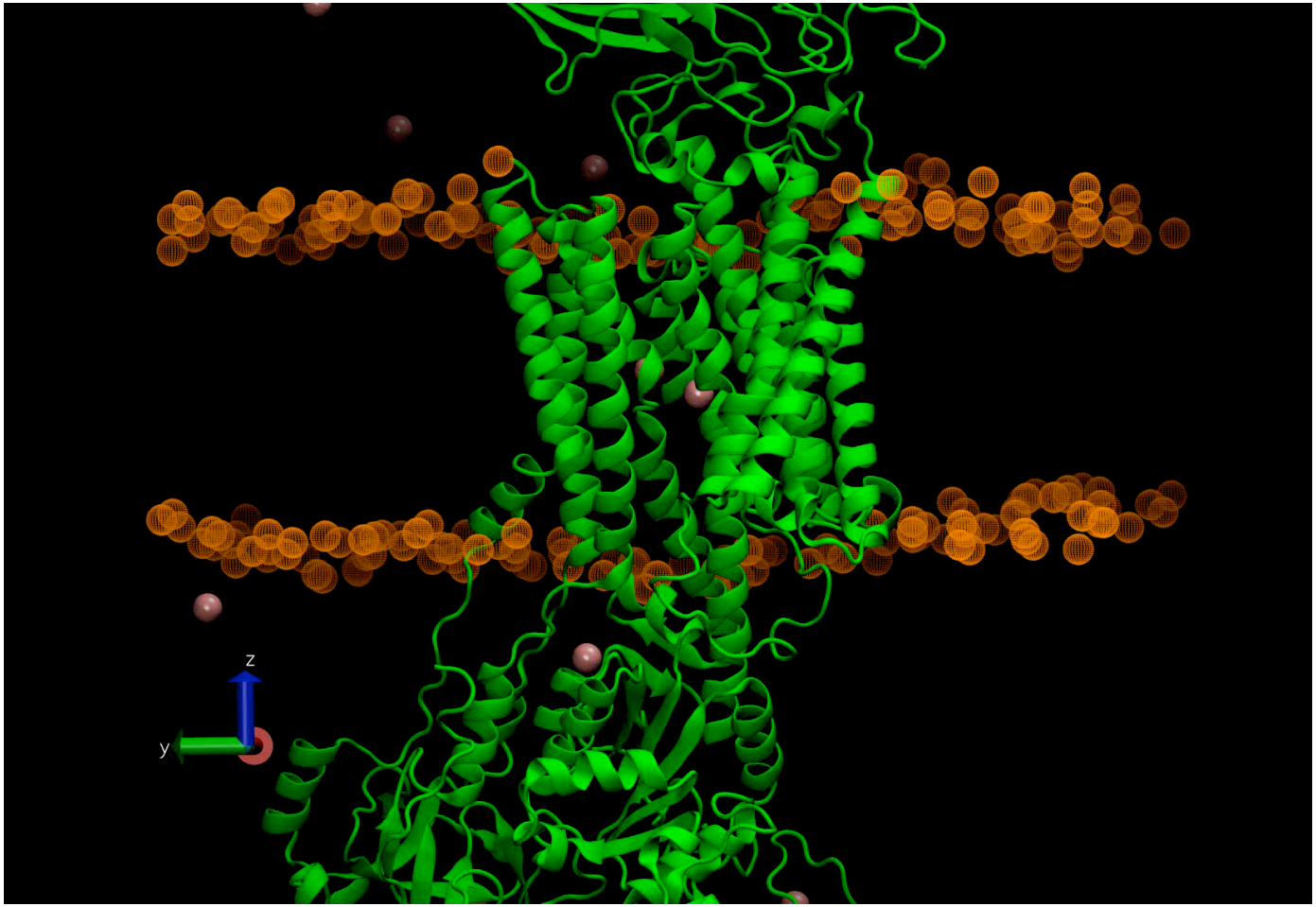
Core Concepts

Scales, ratios, dimensional analysis

Membranes, cells, and models

What is a "good" model?

How does cellular complexity arise out of molecular interactions?



250,000 atoms

50 ns

360 procs.

35 ns/day

www.gromacs.org

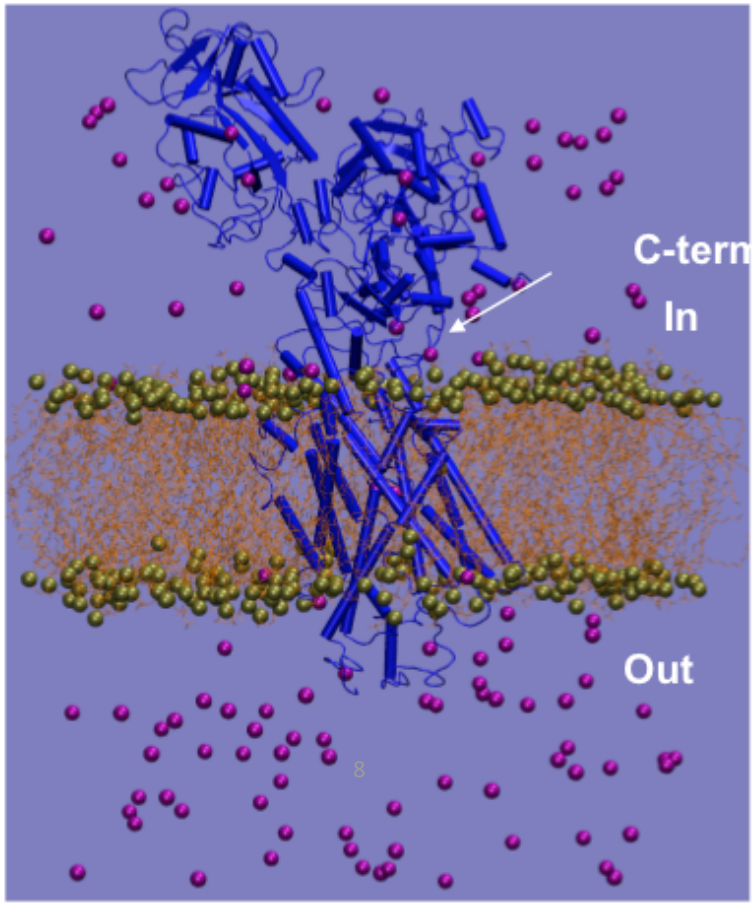
W. Kopec and H. Khandalia, U. Southern Denmark

Why not just solve F = ma for all the atoms in a cell on a big computer?

Cost of atomistic molecular dynamics

Solving Newton's second law to evolve the positions and velocities of the system in time and space

- N ~ 170000 atoms
  - 337 POPC Berger lipids
  - Protein
  - ~ 45000 water
  - Counterions and electrolyte
- NPT ensemble, GROMACS
- Temperature: 310 K
- 115 x 115 x 160 Å
- Time step: 2 x 10^{-15} s



Himanshu Khandalia, U. Southern Denmark

Computing time: 54 cpu years, ~ 80 simulations
6-8 ns a day on 64 cpus for ~ 200,000 atoms

NB. Special HW allows **ms** simulations of 1 protein in days, Shaw et al Science, 330:341-346 (2010)

20,000 lipids can be simulated for **~40 µs** on 20 nodes in 2 weeks, Ingolfsson, JACS 136:14554 (2014)

No obvious insight

Goals of this course

Biology is increasingly computational - an ability to translate biological phenomena into solvable mathematical models (or at least understand the models) is essential.

This course is about making models, computing things

Biology - what is there?

Mathematics - how do we find equations to quantitatively predict behaviour?

Numerical analysis - how to programme a computer to solve the equations

We will construct models, solve equations, plot graphs, do simulations, and explore behaviour as parameters are varied. This requires numerical analysis, approximations, expansion in small parameters, relative magnitudes, etc.

You will become familiar with:

- the *scale* of physical quantities in a cell;

- how to construct their *equations of motion;*

- how to *numerically solve* the equations, particularly using molecular simulations

After this course

Have questions in your head always:

- what is this made of?
- what structures or molecules are the active or relevant ones?
- what can be ignored?
- what are the important interactions?
- how does it move? does it consume energy?
- how can it be perturbed?
- is this similar to another situation I already know?
- where can I re-use this equation?
- what do I do next if I cannot solve it?

Be fluent at *making* models in your head, on paper, in matlab, etc; *breaking* complex systems into simpler parts; *recognising* the important relations or correlations that define the simpler parts.

Transfer techniques from here to other fields

History of cells

History of cell biology is largely the history of the microscope: if we cannot see something, we cannot model it (cp. crawling cell movie)

1660 - Robert Hooke looked at cork and named the (dead) cells *cellulae*

1675 - Leeuwenhoek built microscopes and drew protozoa, bacteria

1931-38 - Ruska and Knoll built first electron microscope

1960s - 1980s - Confocal microscope, Mojmir Petran (invented by Minsky 1955)

1982 - IBM invented the Atomic Force Microscope

1994 - Stefan Hell and Jan Wichman invented STED microscopy

Cell hypothesis:

1. All organisms consist of 1 or more cells
2. Cell is the basic unit of structure of all living things
3. All cells derive from pre-existing cells

A typical cell?

There is no *typical cell* - they are all specialised for their function, and have different shapes, sizes, contents:

- RBC = a small bag of haemoglobin that transports oxygen
- WBC = motile cell that chases bacteria and engulfs them
- Neuron = electrically active cell with long processes (axon, dendrites) that connect at electrochemical synapses to process information

But cells do contain *typical structures or materials*: large proteins, filaments, membranes, vesicles, fluids, … and these have *typical properties*: soft, hard, flexible, stiff, rigid rod, floppy string, sticky blobs, bags, …

Cells exploit the *physico-chemical* properties of materials (water, lipids, polymers, ATP, etc) to carry out their functions because they are always present.

These properties - and the forces they give rise to - are as important for cellular life as the more familiar lock-and-key ligand/receptor binding, direct protein-protein interactions, electrostatic interactions.

Lipid membrane

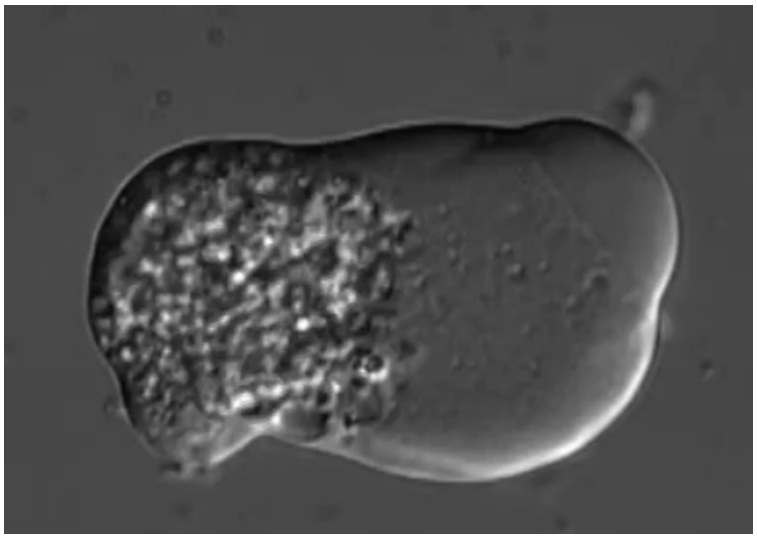
J. Ipsen, MEMPHYS, SDU

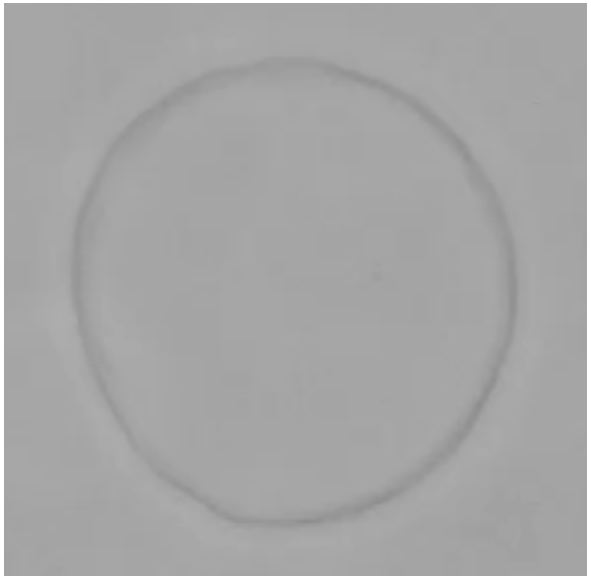
This is a Giant Unilamellar Vesicle (GUV) ~ 20 µm diameter.

Membrane is much thinner than the vesicle diameter and it's a fluid: it fluctuates because of thermal motion of lipids; it is stable because of the hydrophobic effect and reseals on piercing; undulations give rise to a repulsive force between nearby membranes.

What is there? a thin, elastic sheet
What does it do? it gently undulates because of thermal motion of the lipids

Think - Pair -Share





Is the cell diffusing?

Is the vesicle diffusing?

3 mins. One person argue for/against, support your answers with observations from the movies.

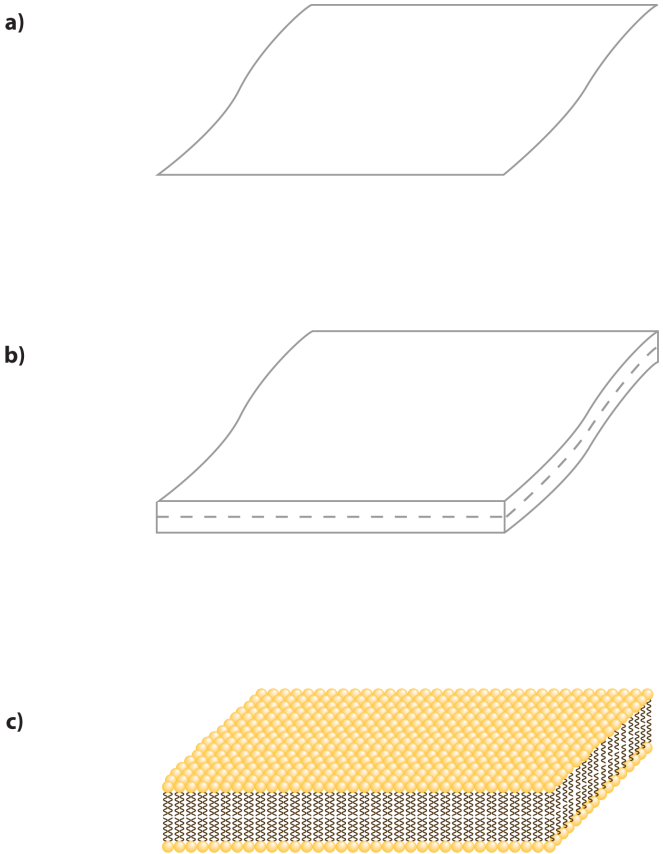
Membranes on different scales

How we describe a membrane mathematically depends on the scale of the question of interest:

Macroscopic - 2d surface, elasticity theory, shape equations, (sufficient for red blood cells), triangulated network "fish net", attach a field (e.g., lipid tilt or different lipid types) to points in the membrane, pore formation

Mesoscopic - adds physical properties involving thickness (e.g. lipid "shape" lateral stress profile/protein conformational changes), geometric asymmetry but no molecular details, pore formation

Molecular - lipids, proteins, protein channel dynamics, permeation, molecular rearrangements, pore formation, fusion



Each scale contributes more insight (and complications) to the still more complex biological membranes.

Lipids on the frontier: a century of cell membrane bilayer
M. Edidin, Nat. Rev. Mol. Cell. Biol. 4:414 (2003)

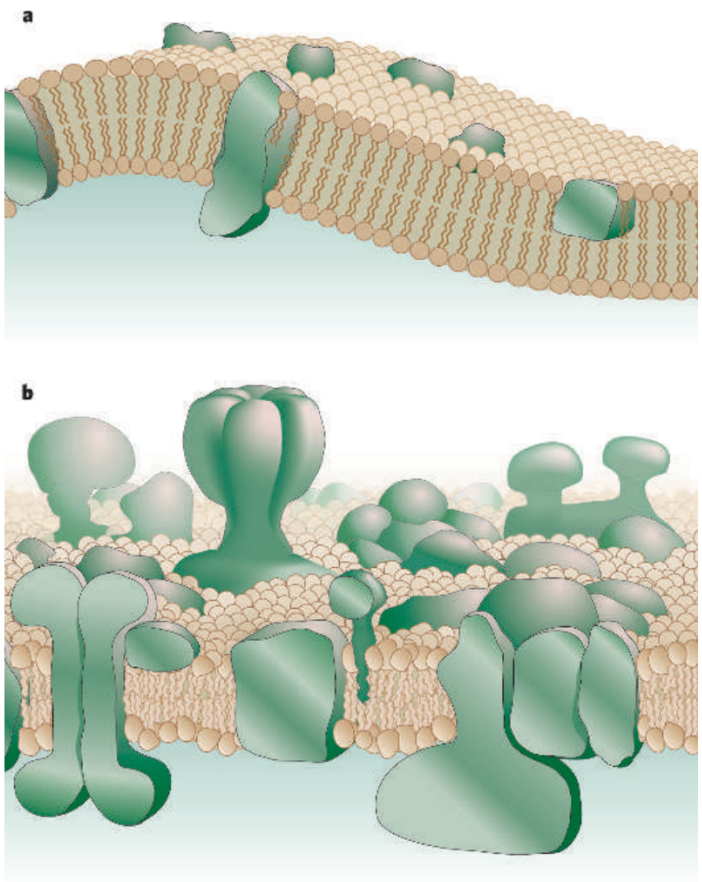
Plasma membrane of a cell

Singer-Nicholson *Fluid Mosaic model* (Science 175:720, 1972)

- PM is a bilayer of lipid molecules held together by the hydrophobic effect of their tails in water
- Proteins are like boats floating on a lipid sea that is unperturbed by them

This explains a lot of data, but it's a rough approximation:

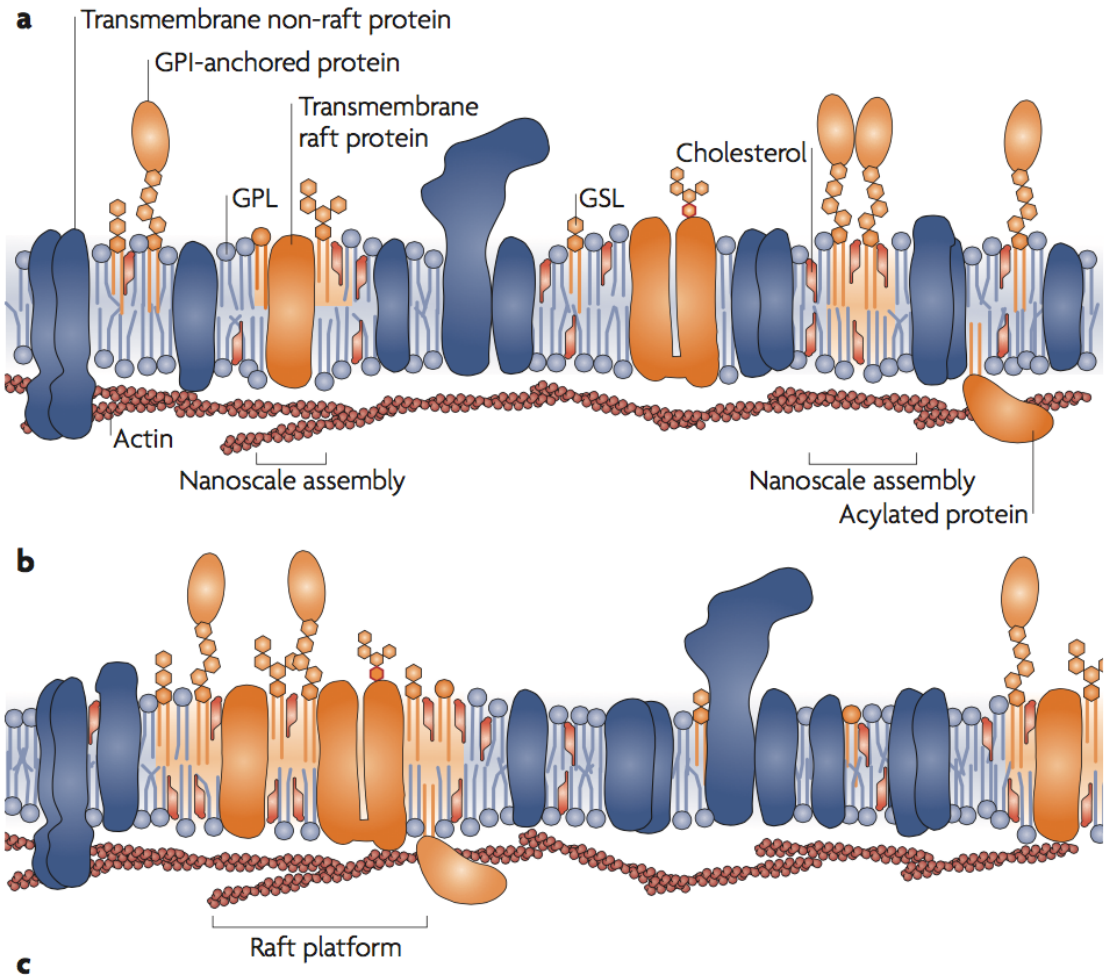
- membrane is patchy (lipid rafts), very dynamic, thickness varies, constrained by actin cytoskeleton
- gains/loses area by vesicle fusion/budding (all lipids in PM turn over in ~ 1 hour *)
- proteins cover a large fraction of the bilayer
- proteins form oligomers and domains with lipids



Engelman DM, Membranes are more mosaic than fluid Nature 438:578 (2005)

* R. M. Steinman et al. J. Cell. Biol. 96:1 (1983)

Lipid rafts



Revitalizing membrane rafts: new tools and insights
K. Simons and M. J. Gerl, Nat. Mol. Cell. Bio. 11:688 (2010)

Functions

Signaling

Endo- and exocytosis

Vesicle trafficking

Virus budding

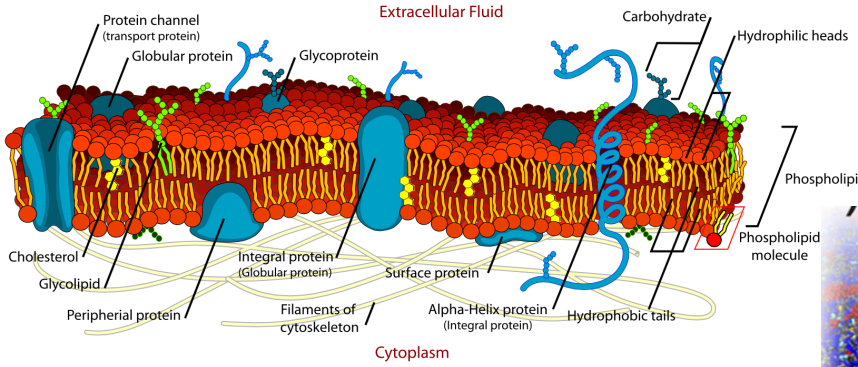
Structure

Protein-lipids-cholesterol

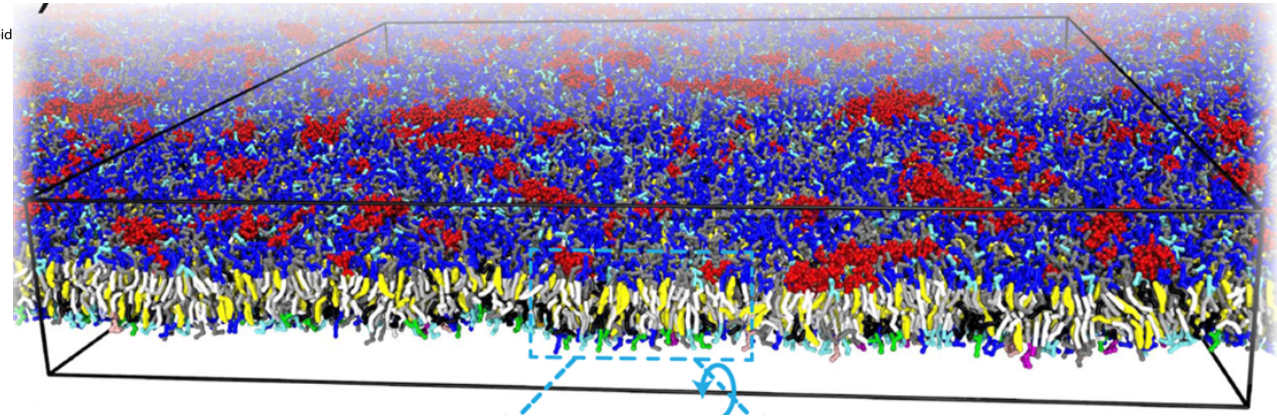
Liquid Ordered phase

Reversible

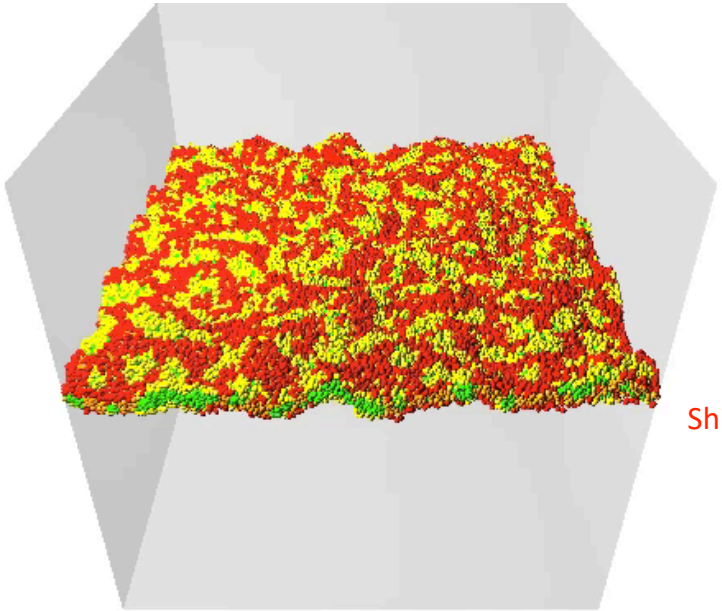
Models of membranes



cgMD: 20,000 lipids of 63 species in 71 x 71 x 11 nm^3 for 40 µs (20 nodes for 2 weeks)



Ingolfsson et al., JACS 136:14554 (2014)

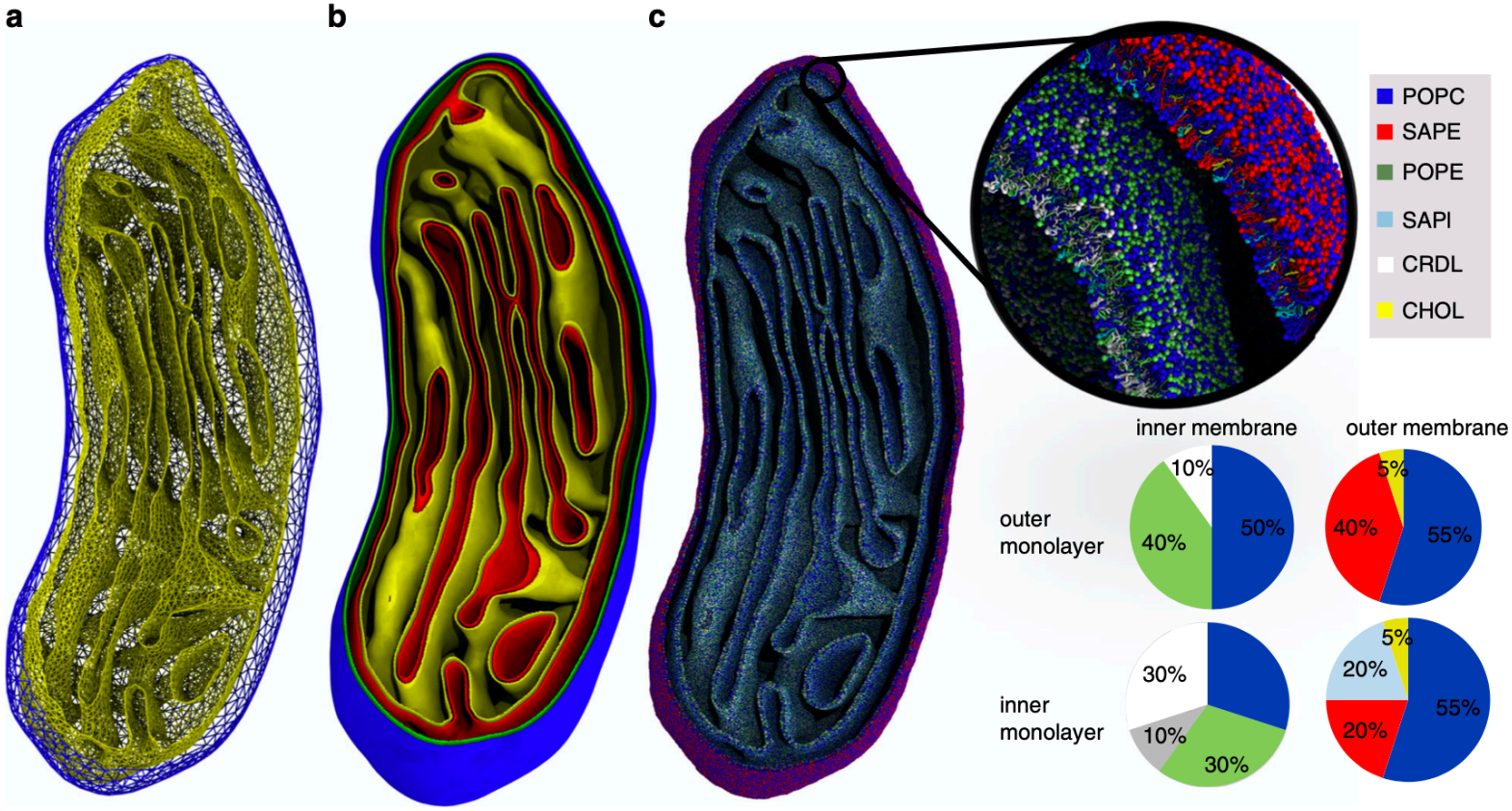


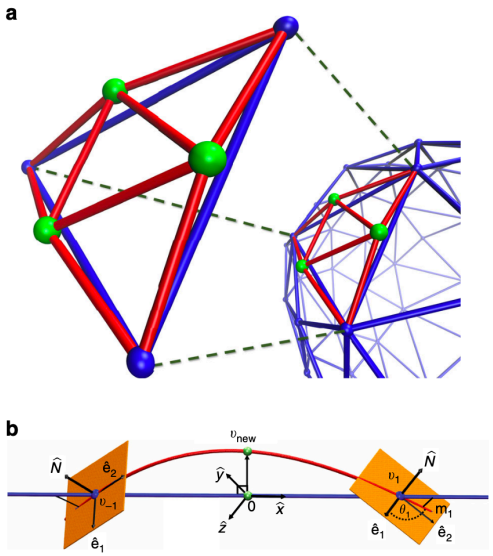
Shillcock , Langmuir 28:541 (2012)

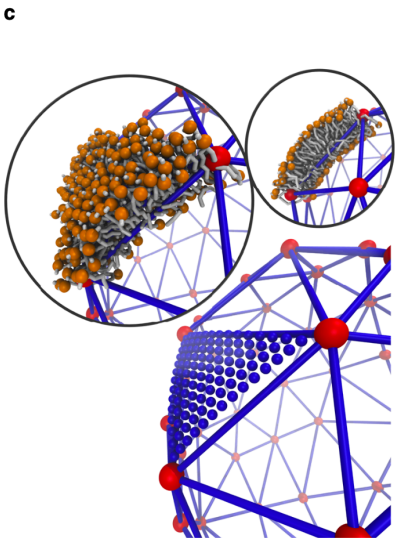
DPD: 9200 + 6128 lipids in 70^3 nm^3 for 80 µs (8 nodes for 1 week)

Simulations are powerful but slow, and limited in space and time scales

Cells contain structures of different dimensions - organelles / 3d, filaments / 1d, membranes / 2d, proteins / 1-3d. We want models that capture essential properties and ignore the irrelevant ones (how do we know what is irrelevant?)



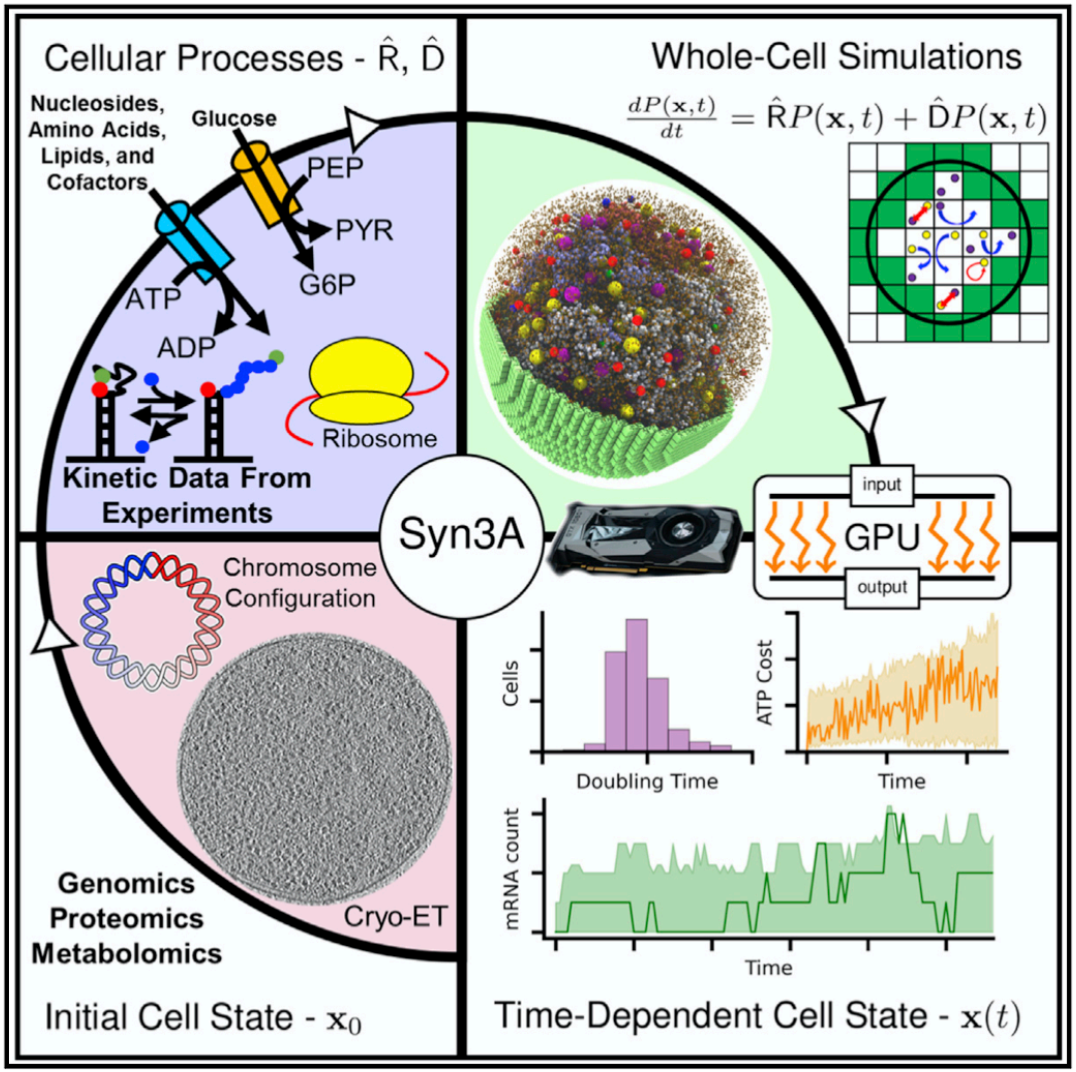




What's next?

Pezeshkian et al. , Nature Communications 11:2296 (2020)

Fundamental behaviors emerge from simulations of a living minimal cell



Thornburg et al., Cell 185: 345-360 (2022)

Quantitative description of a cell

To talk quantitatively about a cell we must measure its properties:

- Light microscope
- EM microscope
- Fluorescence of single particles (Single particle tracking), FCS
- AFM
- others

We collect data, and (at least mentally) construct a (usually mathematical) model that, in its operation, (ideally) produces precisely the data we measure. This requires familiarity with:

- Biology
- Mathematics
- Numerical analysis

Problem: cellular world is not very familiar to us; how can we get intuition?

How is cellular world different from ours?

EPFL

Human scale ~ 1 m

Surrounded by air

Inertia is important

Momentum is conserved

No action at a distance

Gravity dominates

Energy is wasted easily - heat

Temperature is often irrelevant (within a range)

Large gradients easy (T, p, density)

Need glue to bind

Nothing (> 1mm) diffuses

Randomness is often ignorable

You have to do work to move things

Cellular scale ~ 1 micron

Surrounded by (incompressible) water

Nothing accelerates much

Few consequences of mom. conservation

E-fields extend over space

Gravity is minor: charge/dipole/VdW forces

Energy is conserved - T constant

Temperature is crucial and drives diffusion

Gradients are costly to maintain

Proteins are sticky

Diffusion dominates motion

Lot of randomness and fluctuations

You have to do work to keep things still

Think - Pair -Share

Q. Do Newton's laws apply at cellular scale?

3 mins. Ask yourselves what does it mean to "apply"?

Do Newton's laws apply at cellular scale?

EPFL

Are Newton's laws useful in modelling a cell?

Throwing a ball: 1 ball, 6 *degrees of freedom* = 3 position and 3 velocity.

Game of snooker: 15 red balls + 8 coloured balls, all the same material, F = ma, and elastic collisions. Only 4 degrees of freedom per ball as motion is confined to a plane.

Which properties or coordinates are relevant?

But a cell has ~ > 10^{12} water molecules as well as proteins, lipids, sugars, which are enormously complex compared to a water molecule. We can solve Newton's laws for billiard balls, but not for a cell.

Models must have as few parameters as possible - faster to solve, easier to understand.

Degrees of Freedom

A d.o.f is a coordinate of a system that can take a range of values.

It is often obvious which d.o.f are important/relevant in a situation:

- Snooker ball
- NaCl crystal
- Brick in a wall

But not always…. concrete, ice cream, car, city, an economy, earth, sun, solar system.

Not all d.o.f are equivalent, some are tightly connected together (or *correlated*) and do not contribute separately to the dynamics of a system, e.g., atoms in a brick.

It is the density of independently-variable (or *uncorrelated*) *d.o.f* that is important.

How do you identify relevant degrees of freedom in cell biology?

Units and Dimensions

Physical quantities have *dimensions* that are usually expressed in convenient *units* on the human scale:

Mass (**M**) - gm, kg, ton - mass of a person ~ 50-100 kg
Length (**L**) - cm, metre, km - length of a leg ~ 1 m
Time (**T**) - seconds/minutes/hours - heart-beat ~ 1 sec

(and Charge and Temperature which we ignore for now). From these we can derive other units:

Speed = $L\,T^{-1}$ - 1 leg / 1 heart beat = 1 m / sec
Force = Mass x Acceleration = $M\,L\,T^{-2}$ - jump requires 50 kg *9.81 ~ 490 N
Energy = Force x Distance = $M\,L^2\,T^{-2}$ - jump 1 m up requires 490 J

Human-scale units are not useful in cell biology, we need more relevant ones.

Scales in a cell

What are the natural units (or scales) for a cell?

Mass (**M**) - mass of one water molecule, lipid, protein, cell, …

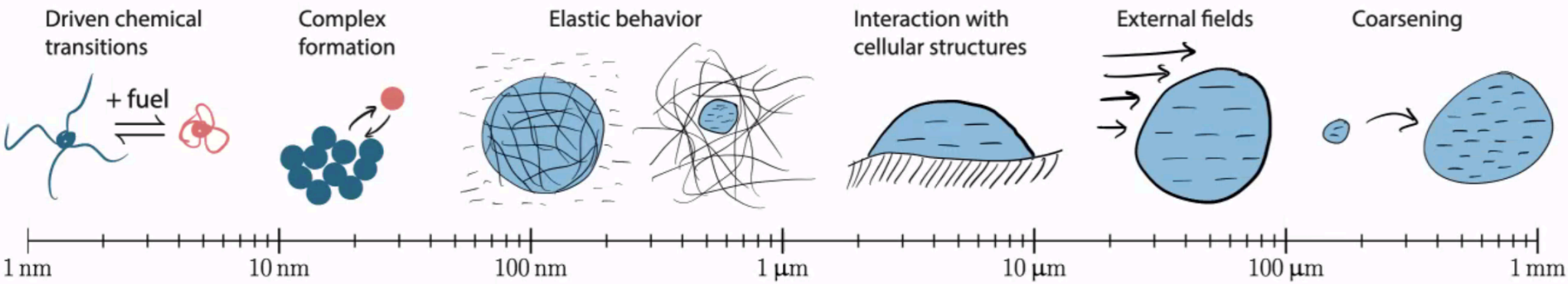
Length (**L**) - size of water, lipid, vesicle, cell

Time (**T**) - lifetime of a cell? Diffusion constant of water or ATP?

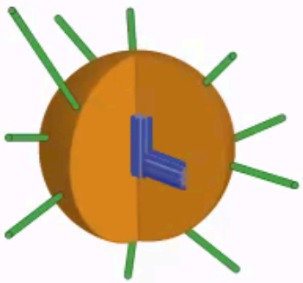
What about speed or force or energy?

You can learn a lot about a problem simply by calculating the ratio of some "relevant" quantities (memorise the table of lengths and times).

Controlling phase separation inside biological cells

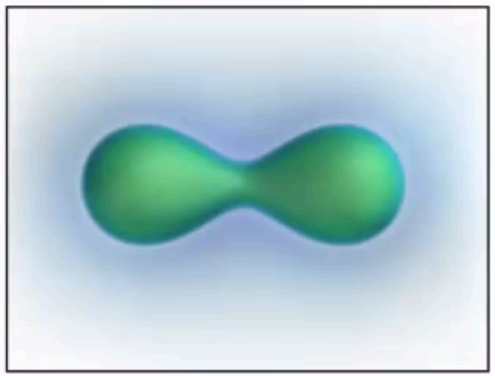


Centrosome organization



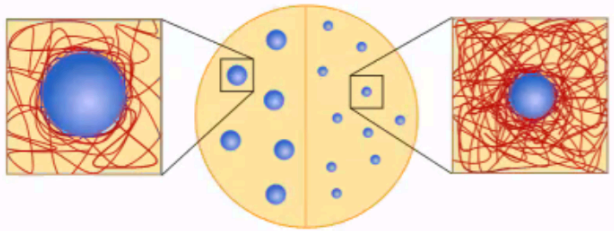
PNAS (2014)
PRL (2018)

Spontaneous division



Nature Physics (2017)

Droplet dynamics in stiffness gradients



Nature Physics (2020)
Soft Matter (2020)
PNAS (2021)

Droplet size control

PRE (2015)
Rep. Prog. Phys. (2019)
Trends Cell Biol. (2019)
J. R. Soc. Interface (2021)

WE'RE HIRING!

More details: www.zwickergroup.org

ZwickerGroup

Scales in a cell - length

Memorise these

Name	Value	Units	Reference
GUV diameter	50	micron (μm)	Mouritsen
RBC diameter	7.5	μm	Guyton
Mammalian cell diameter	20	μm	Alberts
Nuclear diameter	6	μm	Alberts
Lysosome diameter	0.5	μm	Alberts
Synaptic vesicle diam.	60	nanometre (nm)	Alberts
Plasma Membrane thickness	~ 4 (composition?)	nm	Israelachvili
Phospholipid length	~ 2 (lipid?)	nm	Israelachvili
Area per lipid	~ 0.7 (lipid?)	nm^2	Israelachvili
Microtubule width	25	nm	Alberts
Intermediate filament	10	nm	Alberts
Actin filament width	7	nm	Alberts
Tubulin mononer	~5	nm	Alberts
Actin monomer	~5	nm	Alberts
PM area/Total mem. area	0.02	-	Alberts
C-C bond length in lipid tail	0.154 + 0.126*n	nm	Israelachvili

Scales in a cell - time

Memorise these

Name	Value	Units	Reference
Cell division/mitosis time	~30	minutes	Guyton
Vesicle fusion time	~20	ms	Domanska
Clathrin-coated pit formation	~60	sec	Weigel
Actin filament growth rate	3	mono/μM·sec	Fujiwara
Myosin V motor speed	200	nm/sec	book.bionumbers.org
Water diffusion in bulk	2300	$\mu m^2/sec$	Wraight
Water diffusion in gA channel	200	$\mu m^2/sec$	Wraight
Lipid diffusion in membrane	0.1 - 10	$\mu m^2/sec$	Gaede
Lipid flip-flop across membranee	10^2 - 10^5	sec	Israelachvili
Lipid chain equilibration	~1	ns	Roberts

Scales in a cell - energy/force

Name	Value	Units	Reference
k_BT	4.1e-21 J at 300 K ~ 4 pN.nm 1 kJ/mol ~ 0.4 k_BT/molecule	Joules	-
Covalent bond energy	500	kJ/mol	Israelachvili
H-bond energy	20	kJ/mol	Israelachvili
Van der Waals "bond" energy	1	kJ/mol	Israelachvili
Denature a fusion protein	~200	pN	Yersin
Membrane stretch mod. (DMPC)	240 ~ 50 k_BT/nm^2	mN/m	Rawicz
Membrane bending mod. (DMPC)	$0.56.10^{-19}$ J ~ 13.5 k_BT	J	Rawicz
Water-air surface tension	70	mJ/m^2	Wikipedia
Water-oil surface tension	50	mJ/m^2	Israelachvili

References for useful numbers table

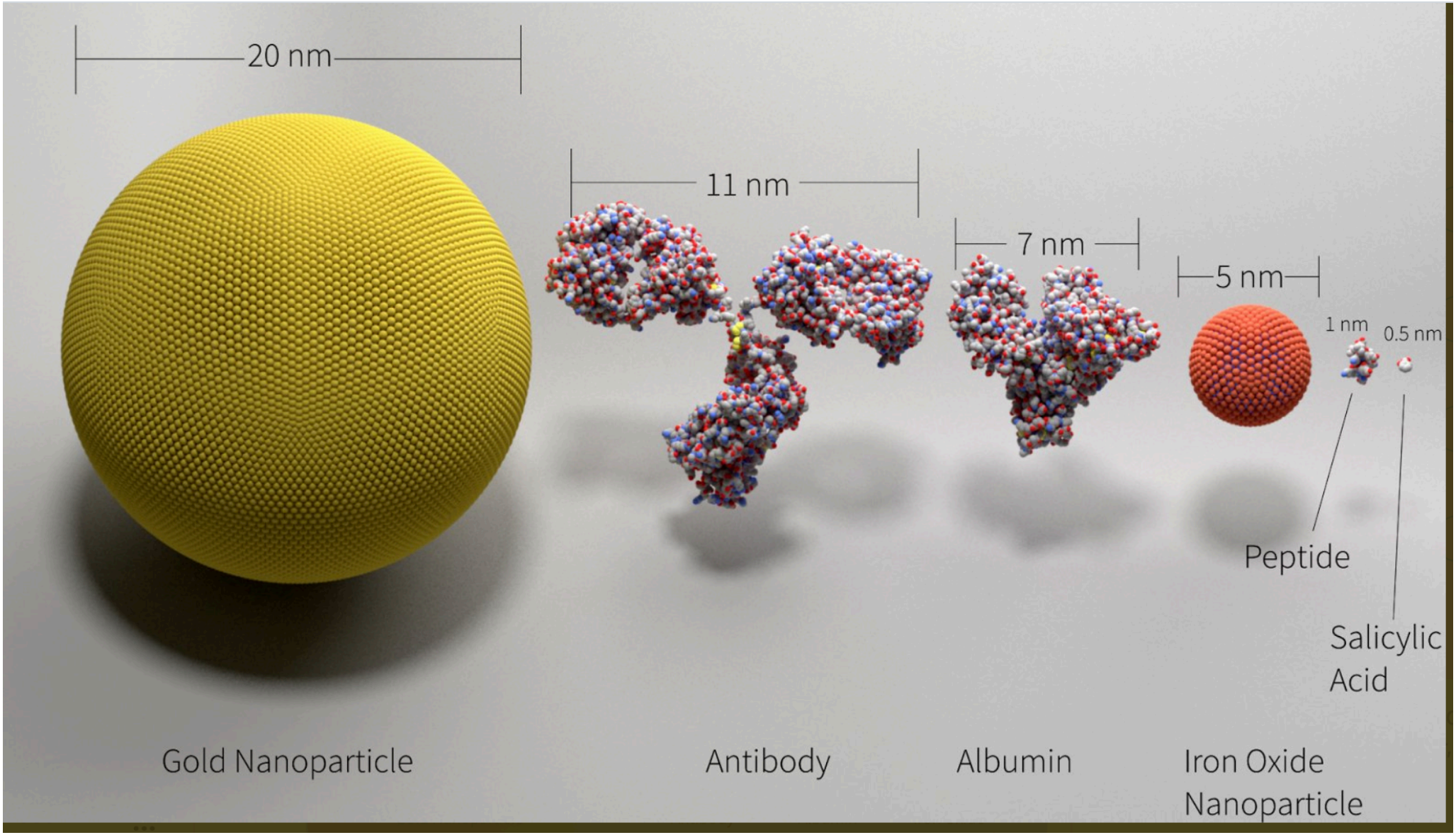
1) O. Mouritsen, **Life as a Matter of Fat** (Springer 2005)
2) A. Guyton, **Textbook of Medical Physiology,** 8th ed. (Harcourt, Brace and Co. 1991)
3) B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson, **Molecular Biology of the Cell**, 2nd ed. (Garland Science, NewYork, 1989)
4) J. Israelachvili, **Intermolecular and Surface Forces** (Academic Press, 1992)
5) M. Domanska et al. **J. Biol. Chem**. 284:32158 (2009)
6) A. Weigel et al. **PNAS** E4591 (published online Nov. 11 ,2013)
7) http://bionumbers.hms.harvard.edu/default.aspx
8) A. Yersin et al. **PNAS** 100:8736 (2003)
9) W. Rawicz et al., **Biophys. J.** 79:328 (2000)
10) H. Gaede an K. Gawrisch, **Biophys. J.** 85:1734 (2003)
11) I. Fujiwara et al. **PNAS** 104:8827 (2007)
12) M. Roberts and A. Redfield, **JACS** 126:13765 (2004)
13) C. Wraight, **Biochim. Biophys. Acta.** 1757:886 (2006)
14) http://book.bionumbers.org/how-fast-do-molecular-motors-move-on-cytoskeletal-filaments/

Significant ratios in a cell

Calculate the following ratios:

1. Cell diameter/plasma membrane thickness
2. Cell volume/synaptic vesicle volume
3. How long does a lipid take to diffuse its own size in a membrane?
4. Time for a lipid to flip-flop across the bilayer/lipid hydrocarbon tail equilibration time
5. What is the electric field across a neuron's plasma membrane if $V_{mem} \sim -70$ mV
6. At what separation does the Coulomb potential between Na^+ and Cl^- equal the thermal energy k_BT?
7. If an intrinsically-disordered protein is treated as an phantom chain: what is the ratio of the volume of its backbone to the volume defined by its end-to-end length? How many such polymers could fit inside the volume swept out by one of them?

some of these may turn up on tests …



A White, U Rochester, New York

Dimensional Analysis

Physical quantities have units attached (M, L, T), dimensional analysis helps us in two ways:

- Verifies that equations or relations are physically correct
- Predicts new relations between known physical quantities

If you derive the equation for the Coloumb force between two charges, say, $F = kQ_1Q_2/R^2$, and the RHS does not have units of MLT^{-2} there's a mistake.

How does dimensional analysis create new equations?

Classic example is simple pendulum. (Blackboard calculation)

What happens if we include air density?

What does it mean if a combination of quantities is dimensionless?

Think - Pair -Share

Q. Is the pendulum equation with air density a "good" model?

Q. If not, how could you fix it?

5 mins. What does your intuition tell you about how a pendulum's motion depends on the medium it's in?

Combinatorics

Combinatorics = counting

How many ways can something happen? what is the probability? or statistical weight?

Diversity in *molecular* structure is essential for life:

- Quarks - u, d, s, c, t, b - 6
- Nucleons - p and n - 2
- Elements - H, He, Li, …. ~ 100
- No. of genes in human genome ~ 30,000
- No. of protein types ~ 100,000 - 10^6
- No. of lipid types ~ 100s - 1000s
- No. of nucleic acids with N bases (A, C, G, T) - 4^N
- No. of proteins with N amino acids (20 aa) - 20^N
- No. of C-C molecules with N carbon atoms - 1 if linear, many more if not
- No. of sugar molecules ~ 2 dimensional macromolecular network

Dimensionality is important for combinatorial problems.

Scaling

How does a quantity change as a related mass or size changes?

If the number of small segments of a straight line is doubled, the length doubles

If the linear size of a triangle is doubled, its area is squared

Birthday problem: how does the probability that two people in a group of N share a birthday scale with N?

Neuronal intersection in space: A neuron occupies about 0.02% of the space spanned by its dendrites. If two such neurons are near each other, what is the probability that they intersect?

The physics behind a lot of cell biology often depends on how a quantity scales with the length or size of something else, and the dimension of the quantities involved; which mechanism is used is controlled by scale.

Building mathematical models

Our models are mathematical, first translate the physical/chemical/biological structures into equations:

- Relationships are described by functions, f(x, y, z)
- Differential equations describe their evolution in time
- Partial differential calculus allows us to understand the effects of dimensionality on the system

We solve the equations, sometimes analytically, usually numerically:

- Approximate the functions using Taylor series
- Find maxima or minima, extrapolate to interesting regions
- Integrate ODEs using some scheme with a given accuracy
- Do simulations if too many degrees of freedom or no simplification is possible

Models should have as few parameters as possible - faster to solve, easier to understand.

What do you think is the "simplest" mathematical model of a system in physics?

What is it made of?

How do its molecules interact?

How many degrees of freedom does it have?

How many d.o.f are uncorrelated?

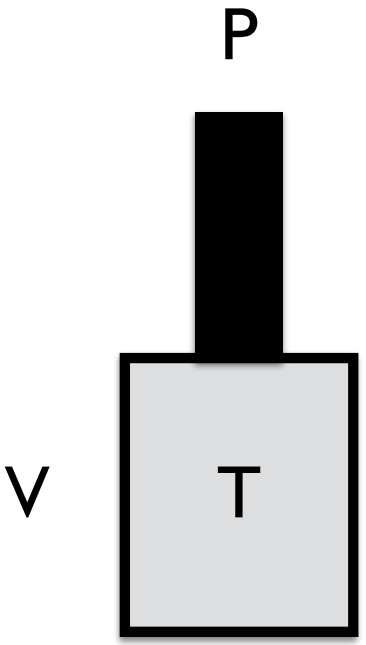
What is the density of the d.o.f in space?

How many measurable physical properties does it have?

Ideal gas in equilibrium

$$pV = Nk_BT$$

p = pressure
V = volume
T = temperature
N = number of molecules in the gas (mass)
k_B = Boltzmann's constant = 1.38 10^{-23} J / K



Why is this simple?

- No interactions, correlations, gradients, history or time evolution
- No size to its parts (molecules are infinitely small)
- Unaffected by container and everything outside except temperature
- Doesn't care about shape, size or material properties of container
- Only 3 measurable properties of which only 2 are independent because of the equation of state

Good models are abstractions

An ideal gas is simple because it is a *generalisation* or *abstraction* from reality.

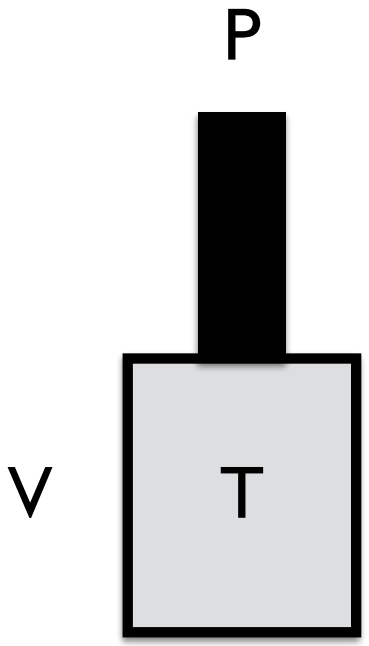
It assumes the molecules:

- are infinitely small
- have infinitely weak interactions but still share energy (infinitely slowly)
- have no time-dependent behaviour (i.e, are in equilibrium, see Lecture 7)

If every atom or molecule had a unique energy dependent on its neighbours' precise position/orientation, we would need to know exactly which atoms interacted, what the strength of the interaction was, and solve complicated equations of motion for every atom in a system.

But this is odd …

No interactions, correlations, gradients, history or time evolution



But … if there are no interactions, there is:

no mean energy,
no energy distribution,
no equilibrium

Everything happens infinitely slowly with infinitesimal exchanges of energy, … equilibrium doesn't care how long it takes to be established.

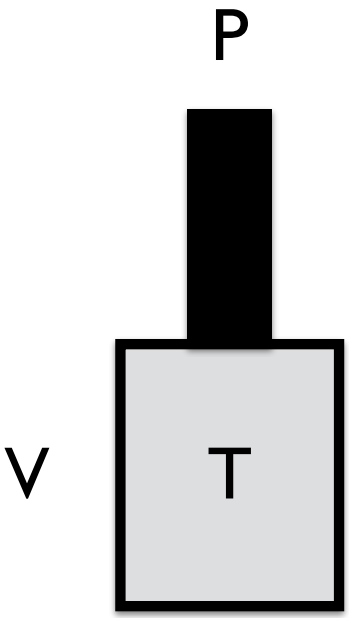
Van der Waal's gas

Two simple modifications to the ideal gas equation of state defines the Van der Waal's gas that has completely different behaviour:

$$(p + aN^2/V^2)(V - Nb) = Nk_BT$$

N*b is effectively the reduction in available volume due to the non-vanishing size of the N molecules

a*N²/V² is the extra pressure that must be added to compensate for inter-molecular attraction that tends to make them clump and reduce the pressure on the walls



VdW equation is one of the simplest EoS that shows a phase transition. It is useful as a guide to more complex models to get a feeling for how they behave.

What is a good model?

The ideal gas is simple and illustrates an important point:

No model captures *everything* about a real system

An ideal gas **cannot** have a phase transition: the equation just does not allow it.

But it captures the fundamental properties of any **dilute** gas in such a way that two simple changes give it enough structure to have a liquid-gas phase transition for a **dense** gas.

It is simple, relevant params., extensible, which makes it a good model

Cell biology has come a long way

Optional reading on moodle: Enabling the next 25 years of cell biology
J. Woodgett and D. T. Laughlin, *Trends in Cell Biology* **26**:789 (2016)

25 years ago:

- no PCR machine in the lab
- whole genome sequencing took years
- fluorescent microscopes were expensive and rare
- no gene editing tools

Now we can visualise almost all cellular structures, use single molecule precision to manipulate them, and edit genes for therapeutic benefit, and have cryoEM, AFM, STED, PALM, FRAP, FRET, opto-genetics - we can see and pull single molecules!

The next 25 years will be a time of computational tools.

What will you do when a computer can do your job?

Computers are amazing at solving equations fast

Machine Learning is great for extracting patterns from Big Data

Robots can do electrophysiology untiringly

But this isn't model building

Humans are amazing at:

Creative Thinking = seeing things differently, comparing quantities, making models

Innovation = modifying models to predict new situations.

In the exercises and project, you will make models, code, set up and run simulations, and identify what is important to include in a model - all skills that can be used in any discipline not just cell biology.

Summary

- This course is about building models and making predictions in cell biology (but not only cell biology): you don't have insight into the real system, only a model in your head
- We have to know basic length and time scales for cellular properties to gain *intuition* into cellular behaviour
- How to identify what is important and what to ignore in a model? Experience, and extrapolating from what you already know
- Always ask questions!

Break

Format of the course

Lecture in class (recording is on mediaspace) + exercises (4 hr/week) + tests + journal club + simulation project (at home 4 hr/week)

- Tests are marked but no collaborating (based on lecture material/reading)
- Journal club/Exercises are marked and may collaborate (simulations, calculations)
- Projects are marked and you should work in groups of 3 - 5 so the data and workload can be shared - but everyone has to write a separate report
- Background reading introduces the following week's lecture material
- Is there a class delegate? Anonymous feedback can be directed through them

Marking

50% DPD project and report in the form of a scientific paper
(Introduction, Method, Results, Conclusions, References - sample reports on moodle)
30% 3 x 1 tests (Weeks 4, 6, 11, practise test on moodle -2 weeks)
15% 2 x Homework exercises (due in Weeks 6, 11)
5% Journal club presentation (shall we have this early or late?)

BIOENG-455 Table of Events

Lecture	Theme	Test/Homework/ Journal Club	In-class Derivation
1	The cellular scale	Quiz	
2	Macromolecules in a cell	Announce HW 1	D1 - Freely-jointed chain
3	Forces in a cell		D2 - Equipartition Theorem
4	Simulation types	Test 1	
5	How to coarse-grain	JC ?	
6	Brownian motion	HW 1 due - Test 2 take home	D3 - Random walk, Langevin Eq.
		Semester break / No lecture	
7	Thermodynamics 1	Take-home test 2 due	
8	TD 2/Phase transitions	Announce HW 2	D4 - Entropic spring
9	Membranes 1	JC ?	
10	Membrane 2 pores/fusion	Test 3	
11	Shiga toxin and symmetry	HW 2 due	
12	Membraneless organelles 1		
13	Membraneless organelles 2		
14		Project presentations	

Exercise

Exercise period is for:

- working on the homeworks/project
- running short simulations to get familiar with the code
- anything else you want to ask/do about the course

1) Install DPD simulation code and get the User Guide
2) Check you can run the sample input file (dmpci.001) on your laptop

Laptop poll

How many are using which OS?

windows

mac

linux

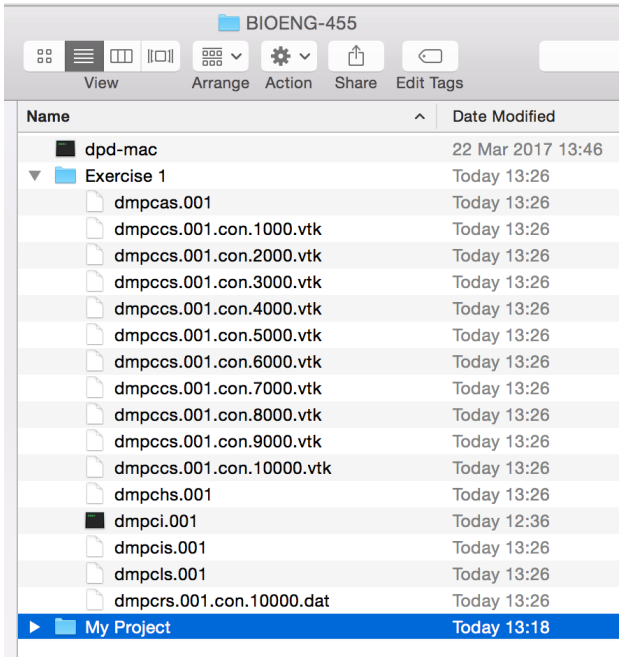
other?

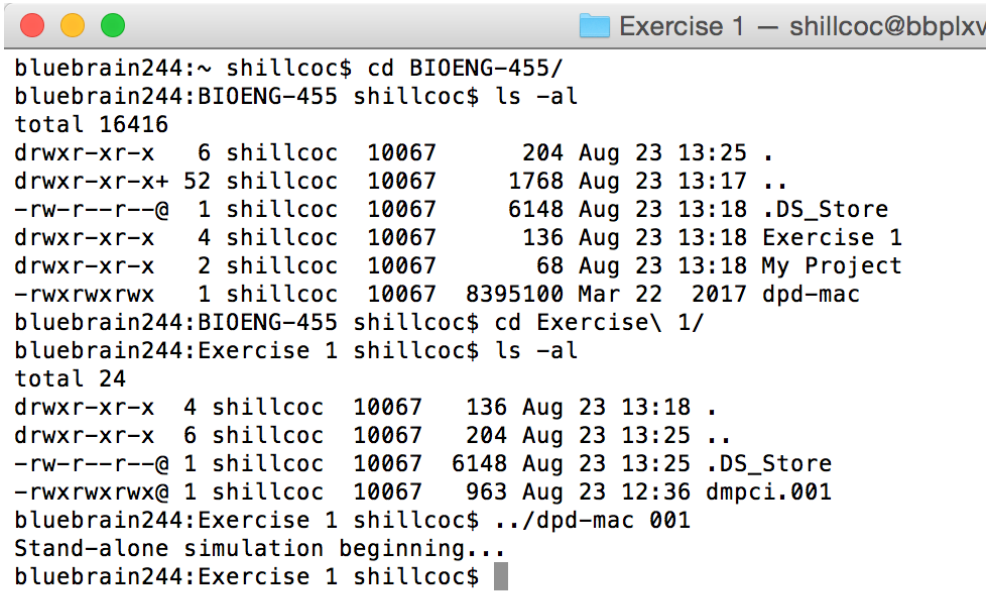
Download and run the DPD code

- Download the executable (Linux, Mac, Windows) and User Guide from moodle (or source code from https://github.com/Osprey-DPD/osprey-dpd)
- Create a directory structure to hold the runs, e.g., ~/BIOENG-455/Exercise 1
- Download the input file: dmpci.001 and run the code:

  > dpd-mac 001

Scan for source code





dmpci.001 input file for "water"

```
dpd

Title   " Water "
Date    19/09/18
Comment " Pure water simulation. Measuring the pressure as a function of the bead
          density parameter (Density 3) allows the equation of state to be determined.
          Ignore the first analysis period (1 - 5000 timesteps) to allow the system
          to equilibrate and then take the value from the second period (5001 - 10000).
|
         Note. If you edit the title above or this comment there must be at least
         one space between the quotes and the text. Blank lines are allowed.   "
|
State   random


Bead  W
      0.5
      25
      4.5


Polymer Water    1.0   " (W) "


Box         10  10  10       1  1  1
Density     3
Temp        1
RNGSeed     -33145
Lambda      0.5
Step        0.02
Time        10000
SamplePeriod     10
AnalysisPeriod   5000
DensityPeriod    10000
DisplayPeriod    1000
RestartPeriod    10000
Grid        1  1  1


Command ToggleBeadDisplay            1    W
Command SetCurrentStateCamera        1    0.5 -1.0 -0.5  0.5 0.5 0.5
Command SetCurrentStateDefaultFormat 1    Paraview
```

The runId can only contain letters, numbers, - and _

Title, Date and description of run - there MUST be space between text and " "

Initial state type

Bead type definitions
(Name, radius, cons. int., diss. int.)

Polymer (or molecule) type definitions
(Name, number fraction, shape) - note spaces between shape and " "

```
Box          10   10   10        1   1   1
Density      3
Temp         1
RNGSeed      -33145
Lambda       0.5
Step         0.02
Time         10000
SamplePeriod      10
AnalysisPeriod    5000
DensityPeriod     10000
DisplayPeriod     1000
RestartPeriod     10000
Grid         1   1   1

Command ToggleBeadDisplay             1    W
Command SetCurrentStateCamera         1    0.5 -1.0 -0.5  0.5 0.5 0.5
Command SetCurrentStateDefaultFormat  1    Paraview
```

Simulation box size/CNT cell size

Bead density, Temperature

RNG Seed and "lambda parameter"

Integration step size

No of time steps, sampling period, etc

Grid size for analysis

Commands to change display

Commands must be time-ordered