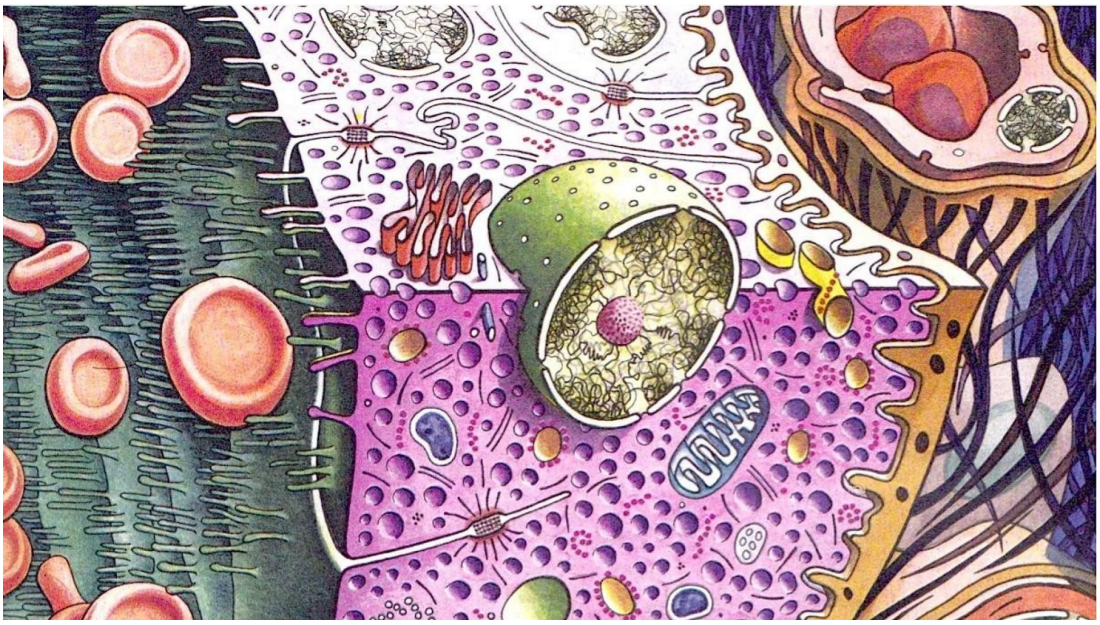
University of Padova

LECTURE NOTES OF BIOLOGICAL PHYSICS

Collection of the lectures notes of professor Fulvio Baldovin.

Edited By

Riccardo Tancredi



Contents

1

Introduction and basic definitions

"Non erit, et lux solis flare nos de tenebris et rerum scientia".
– Lucretius

Lecture 1.
Thursday 2nd March, 2023.
Compiled: Sunday 12th March, 2023.

Physics has always been the combination of experiments and theory. Since Galileo, the observation had always been the starting point to apply the scientific method, followed by hypothesis, deductions and experimental controls. What makes physics an *exact science* is the fact that it is based on mathematical models and principles that are consistently observed in the natural world. These mathematical models allow physicists to make precise predictions about the behavior of physical systems and the results of experiments, which can then be tested through further observations and experiments. In the history of physics many times had happened that theory had predicted reality - think of the Higgs boson - and from this we had learnt that we have a so good comprehension of reality that we are often capable to predict and describe its most inner aspects. On the other hand, biology has not had the same luck: compared to physics, in which, again, both theory and experiments allow us to gain knowledge about reality, there is not a rigorous mathematical theory behind biology. There are not always rigorous models able to describe biology. The study of biology is essentially the study of complexity. Even if *complexity* it's further the goals of this course, we can think of that as the mathematician Peter Dodds would describe it: *"There's no love in a carbon atom, No hurricane in a water molecule, No financial collapse in a dollar bill."* [2]. Biological Physics is basically the study of this complexity with goal the understanding of the physics behind this complexity, behind biology. Contrary to a basic course in physics, in which we could start from the laws of Newton of motion, here there are no Newton's laws in biological physics. By making the assumption that the motion is in absence of friction and that the objects move with constant velocity horizontally and accelerating vertically *"If then the balls of lead, of iron, of stone do not observe that supposed proportion, its damage, we will say that we are not talking about them."* as Evangelista Torricelli would say[1].

The key problem is that every biological particle moves in a dissipative medium, which cannot be neglected: **water**. This actually impose the need to work in a specific framework which is not the traditional one, but it's the stochastic one. We will then deal with stochastic equations. It is important to underline that, since we are working in water, we will work at low **Reynolds-number** so in a regime in which $\vec{F} \propto \vec{v}$ and not $\vec{a}$. Also the $\vec{\nabla}$ operator will be use, expecially in everything concerning fluidodynamics and we will pay attention to tha many representations this operator shows. Moreover, it is relavant to rememeber that the usual time reversal property is no longer possible as we cannot keep track of an Avogadro's number (N_A) of particles.

[1] Se poi le palle di piombo, di ferro, di pietra non osservano qella supposta proporzione, suo danno, noi diremo che non parliamo di esse. [1]

What we wish to show, thanks to statistical mechanics, the *Langevin equantion* and the *Fokker-Planck equations*, is that nature tends to use the same strategies and ideas over and over again. Basically, we will deal with many applications, in order to emphasize that biological physics is an applied subject rather than only a theoretical one. We will start then by considering a single particle moving in water. Then we will exdend this argument to many particles diffusing in water and we will give some possible applications.

1.1 Elements of thermodynamics: statistical mechanics

Since it is not possible to study $N_A \approx 10^{23}$ particles statistical mechanics is the practical solution to use in order to connect biology and physics.
In thermodynamics we have two representations: **entropic** and **energetic** representation. Equilibrium thermodynamics in terms of entropy representation express the entropy S as a function of linearly additive (extensive) variables: the internal energy, the volume and the number of particles:

$$S = S(U, V, N).$$

Instead, in terms of energy representation, it is possible to invert the the above equation and express the internal energy U as a function of $U = U(S, V, N)$. This equation is less fundamental than the previous one, but it is more used. From this we can generate the *free energies* which are fundamental statical quantities that are equivalent to internal energy with the difference that, using *Legendre transformations*, an extensive parameter is replaced with an intensive one. For example, by replacing the entropy S with the temperature T we get the Helmholtz free energy $F = F(T, V, N)$ or by replacing also the volume V with the pressure P and we get the Gibbs free energy $G = G(T, P, N)$ and so on. These frameworks will be useful to understand ions pumps, thanks to which neurons are able to communicate and work properly in a non-equilibrium environment, or molecular motors, proteins folding, two-states systems (like the folding and unfolding process of RNA using *optical tweezers*), how polymers behave, how they interact and transform, what are their shapes and properties and many many other cases.

To investigate the link with biology let's consider this example: A particle starts

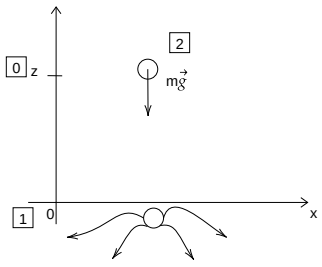


Figure 1.1: Heat and conservation of energy.

at rest at height z is subjected to an external field force $\vec{g}$ pointing downward. The particle has total internal energy $U = \frac{1}{2}m\vec{v}^2(t) + mgz(t)$. By neglecting the action of air

$$\frac{d}{dt}U(t) = \underbrace{\frac{1}{2}m2\vec{v}\cdot\frac{d}{dt}\vec{v}}_{\frac{d\vec{v}}{dt}=(0,0,-g)=\vec{a}} + mg \underbrace{v_z(t)}_{\vec{v}=(0,0,v_z(t))}$$

$$= \underbrace{-mgv_z(t)}_{>0} + \underbrace{mgv_z(t)}_{<0} \stackrel{!}{=} 0$$

This proves that the internal total energy is conserved, as we know from this simple physics problem. We can consider then $\Delta U = U_1 - U_0 = Q$ which represents the total heat absorbed by the system.
Supposing that the ground is made of mud, then energy is transformed to the mud as heat Q, so there is a dissipation of energy. If we consider also step $\boxed{2}$ we can write, since we are considering a cyclic transformation

$$\begin{aligned} \Delta U = U_2 - U_0 &= \underbrace{U_2 - U_1}_{W} + \underbrace{U_1 - U_0}_{Q} \\ &= W + Q \end{aligned} \tag{1.1}$$

where W is the work absorbed by the system. In the case in which there is no mud $\Delta U = 0$, but since we generally do not deal with cyclic transformations the above $\Delta U = W + Q$ represents the conservation of the total energy, also known as the 1^{st} law of thermodynamics. We underline the fact that the Δ is present due to the fact that U is a function of state, while W and Q are not.

The link to biology with this example can be found by considering the case in which the mud is warmed up when the particle hit it: in this case there is a dissipation of energy and the particle is not able to come back to the starting point, due to the fact that we have to take into account the 2^{nd} law of thermodynamics too. In the Helmholtz framework, the Helmholtz free energy is given by $F = F(T, V, N) = U - TS$ where entropy is used to estimate the fraction of *low quality energy*.
The key idea is that the entropic term depicts the disorder of the system and it is linked directly to the energy that the system can no longer use. The amount of *good quality energy* the system loses depends on the temperature: the higher the temperature the more the *good energy* is lost. For biology this is of utmost importance because, since there is no molecular machine that works a 0 temperature, nature uses disorder to produce work, by further continuously find a trade-off between temperature level and quantity of disorder. In this particular range of temperatures - $T = T_r$ is the temperatureof the larger system, called *thermal reservoir* - biology works and it is able to commute outside good energy into low quality one and also to take advantage of the disorder-entropic component in someway. Since every biological system is not a closed system but it is always in touch with a greater system a *reservoir*, it is possible to have $\Delta S < 0$. For an open system we can the restate the second law of thermodynamics as

$$\frac{\Delta U}{T} \leq \Delta S \implies \Delta F \leq 0 \tag{1.2}$$

to underline the possibility of the system to control and reuse entropy, in order to improve the order of the system, as far as the good quality energy, coming from outside, is used and reduced by a proper amount (2^{nd} law) and converted into low quality (1^{st} law) which is ejected in the surrounding. The free energy, so the energy available to perform work, decreases if the open system realizes a spontaneous process.

1.2 How life generates order

Lecture 2.
Friday 3rd March, 2023.
Compiled: Sunday 12th March, 2023.

Biological machines are then completely different from traditional macroscopic ones. The thermodynamic paradigm of living systems is to seek for an interface of an energy flux and to take advantage of order (input), to degradate this order into disorder and, in this process, being able to control and organize the entropy of the organism. To give a practical example, we can consider the bio-chimical reactions that happen in plants thanks to which they are able to store part of the energy coming from the sun, so high quality energy $\Delta U > 0$ carried by photons hitting the

the surface perpendicularly, and then to convert it into chemical energy that can later be used to fuel the organism's activities. In this process - *cellular respiration process* - the energy is employed to preserve or diminisch the entropy of the organism itself and part of this energy is released in the environment as a form of low quality energy in any directions.

Another example can be the Belousov-Zhabotinsky reaction: what we observe is the formation of structures, patterns, as over time we can think of this as an increase in order and consequently a dicrease in disorder of the system. The energy flux coming from the reagents of the reaction leaves behind an increase of order and when there is no more external energy the system slowly comes back to its original disordered state, and patterns can not been seen. By taking advantage of order - input energy - and degradating this order into disorder, living systems are so able to control and organize the entropy of the organism the death can be pictured as the maximum possible entropy state a system can reach. When there is no longer the possibility to control this order-disorder mechanism a system is considered dead.

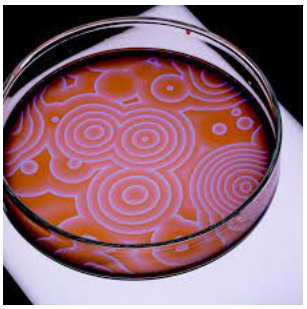


Figure 1.2: B.-Z. reaction [4].

1.2.1 A paradigm for free energy transduction: osmosis flow

A cylinder filled with water is separated into two chambers by a semipermeable membrane which is anchored to the cylinder [3]. Two pistons slide freely allowing the volumes of the two chambers to change as water molecules cross the membrane. Since the water is incompressible, the two pistons must slide together. The open white circles in the picture are sugar molecules and they are confined in the right-hand chamber. As long as the weight is not too heavy **(a)** when the pistons are released we observe an osmotic flow as water crosses the membrane, forcing the both pistons to the right and lifting the weight. The sugar molecules then spread out into the increased volume of water on the right. As the sugar dissolves and spreads through the right-hand chamber, a mysterious force begins to push the pistons to the right: the energy needed to lift the weight came from the outside world, as the system absorbs heat from its surroundings and somehow this thermal energy is converted to mechanicl work. But, it is impossible to completely convert heat into mechanical work, so the order of the sugar molecule is used up: initially each one moves freely, and randomly, throughout the region between the membrane and the right-hand piston.

As water flows through the membrane, forcing the pistons to the right, the sugar molecules lose some of their order (or gain some disorder), being no longer confined to just one half of the total volume of water. When finally the left side has shrunk to zero, the sugar molecules have free run of the entire volume of water between the pistons; their disorder can't increase any more. Osmotic flow sacrifices molecular order, to organize random thermal motion into gross mechanical motion against a load. This can be summarized into the two conditions that must be satisfied $\Delta U > 0$ (the potential energy of the weight increases) and $\Delta S > 0$ (decrease in order), since in the end $\Delta F \leq 0$ as stated by the 2^{nd} principle.
On the contrary, the reverse ormosis can be seen in **(b)** where if the weight is pulled hard enough downward, the pistons will move to the left increasing the concentration of the sugar solution in the right-hand chamber, generating heat. In this case, $\Delta U < 0$ as well as $\Delta S < 0$ such that $\Delta F < 0 \implies \frac{\Delta U}{T} < \Delta S$. This is the perfect example we were looking for: an external input energy - mechanical work in this case

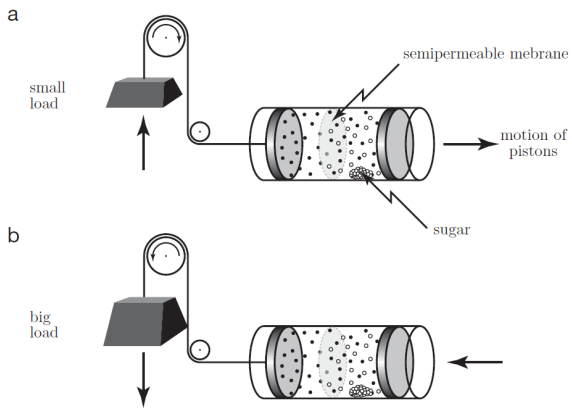


Figure 1.3: A mechanical mechanism trasducting free energy.

- improves the order of our system ($\Delta S < 0$). In agreement with the 1^{st} principle, the energy input must go somewhere and in fact the system gives off heat in the process, to the outside world.
Reverse osmosis is similar to what happens into chloroplasts and plants, where mechanical energy is provided by photons, and the result is the production of order by the production of ATP or more sophisticated molecules. In the same manner, animals too use chemical energy, rather than mechanical or light one, to keep themself organized, by ingesting food rich of bio-chemical compounds.

1.3 Density, current and flux

Doing physics experiments on anthropic scale, we are familiar with common units such as *meters*, *kilos*, *liters* and so on... but, what are the units and the abstract dimensions of things in the biological physics realm? We will say, for instance, that *length* has dimensions $\mathbb{L}$, so writing $[L] = \mathbb{L}$, measured in meters m in **SI** units. We point out that dimensions are distinct from units so, considering more examples, *mass* has dimension $\mathbb{M}$ expressed in kg in SI units; *time* has dimension $\mathbb{T}$ measured in seconds s in SI units; *energy* has dimension $\mathbb{ML}^2\mathbb{T}^{-2}$ expressed in joules J in SI units.
These are just some examples useful to explain very important relations between units: recalling that a *liter* is equal to a cubic decimiter we can write that $\frac{1L}{1dm^3} \equiv 1$ and so, using also $\frac{1m}{10dm} \equiv 1$ we can perform a unite conversion, such as $1L \cdot 1 = 1L \cdot \frac{1dm^3}{1L} \cdot (\frac{1m}{10dm})^3 = 10^{-3}m^3$.

We can consider an additive and trasportable quantity $\mathcal{Q}$, such as linear momentum, energy, or the electric charge or even the number of particles in our system, and given the three dimensional space $\mathcal{V}$, we are able to compute the volume in which this quantity live: $V = \int_{\mathcal{V}} dV$. From this definition we can express other useful physical quantities such as the density $\rho_{\mathcal{Q}}(\vec{r}, t)$ (a field quantity), thanks to which we are able to find $\mathcal{Q}(t) = \int_{\mathcal{V}} dV \rho_{\mathcal{Q}}(\vec{r}, t)$, where $[\rho_{\mathcal{Q}}] = \frac{[\mathcal{Q}]}{\mathbb{L}^3}$. Keeping $\mathcal{Q}$ the more generic as possible we are able to distinguish between charge density ρ_c and mass density ρ_m (?). As well as volumes, we can define areas as $A = \int_{\mathcal{A}} dA$ where $\mathcal{A}$ represents the oriented surface, and from this definition we can finally define the net amount of $\mathcal{Q}$ passing through $\mathcal{A}$ per unit time, so the current $I_{\mathcal{Q}}(t) = \int_{\mathcal{A}} \vec{J}_{\mathcal{Q}}(\vec{r}, t) \cdot d\vec{A}$ where $\vec{J}_{\mathcal{Q}}(\vec{r}, t)$ is the flux - or current density - of particles. Pay attention that since $[A_{\mathcal{Q}}] = \frac{[\mathcal{Q}]}{\mathbb{L}^2\mathbb{T}}$, then $[I_{\mathcal{Q}}] = \frac{[\mathcal{Q}]}{\mathbb{T}}$.

Within $\mathcal{V}$ there could be also the production $\Pi_{\mathcal{Q}}(t)$ of a certain amount of $\mathcal{Q}$ per

unit time. The volumetric production rate is indicated with $\pi_{\mathcal{Q}}(\vec{r},t)$ and it is another field quantity. Then $\Pi_{\mathcal{Q}}(t) = \int_{\mathcal{V}} dV \pi_{\mathcal{Q}}(\vec{r},t)$. In this manner we can define the *integral balance equation* as

$$\frac{d}{dt}\mathcal{Q} = -I_{\mathcal{Q}}(t) + \Pi_{\mathcal{Q}}(t)$$

and using Gauss theorem we can pass to the *differential form*:

$$\frac{\partial}{\partial t}\rho_{\mathcal{Q}}(\vec{r},t) = -\vec{\nabla}\cdot\vec{J}_{\mathcal{Q}}(\vec{r},t) + \pi_{\mathcal{Q}}(\vec{r},t) \tag{1.3}$$

which is known also as **continuity or balance equation**[2]. This equation is fundamental because it expresses how the density is related with the flux of the quantity we are considering and reminds us of the continuity equation in electromagnetism[3].

1.3.1 Negligibility of quantum effects

Consider this simple problem: *Given* $N_{mol} = 6.0 \cdot 10^{23}$ *the number of Avogadro, estimate the size of a water molecule.*

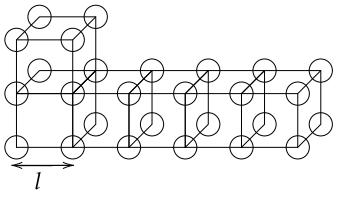


Figure 1.4: Cubic lattice of a water molecule.

To solve this problem recall that $\rho_{m,H_2O} = 1kg/L = \frac{m_N}{V} = \frac{m_N}{NV_c}$. Supposing we are dealing with a cubic lattice of size l, so that each particle has its own cube, except for those at the surface, and supposing we are considering a great size of particles so that we can state that surface effects are negligible with respect to bulk effects, in this way $V_c = l^3 = \frac{m_N}{V} = \frac{m_N}{V\rho_{m,H_2O}} = \frac{\overbrace{n \cdot MM_{H_2O}}}{N\rho_{m,H_2O}} = \frac{18g/mol}{6.0\cdot10^{23}mol^{-1}\cdot\frac{1kg}{L}} = 30 \cdot 10^{-30}m^3$ so, in other words, the size of a water molecule is $l = 3.1 \cdot 10^{-10}\ m = 0.31\ nm$. This result doesn't only show the dimensions of a water molecule and the effectiveness of dimensional analysis, but also allows us to neglect quantum effect, since we will generally work always above this threshold.

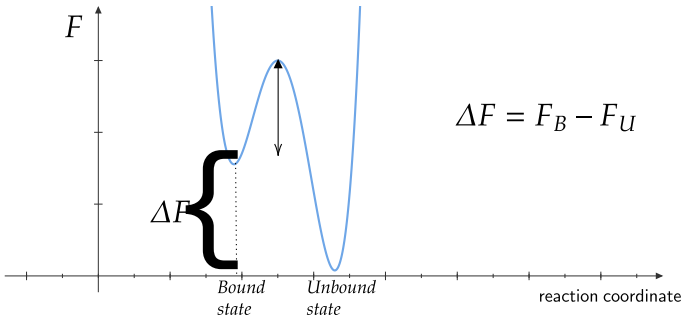


Figure 1.5: Possible free energy landscape of a molecule.

As before mentioned, reactions flows in order to decrease the the value of the free energy F, but to stop biological processes and to keep molecules into *bound states*,

[2] All the $\vec{r}$ coordinates are *eulerian*, but later on in the course we will consider *lagrangian* coordinates and we will introduce total (or material) derivatives.

[3] Continuity equation in electromagnetic theory

enzymes are employed: they play a key role in biological processes since they are able to increase or decrease the activation energy barrier allowing some reactions to happen faster and preventing others to occur[4]. Typical ΔF of chemical bonds are of the order of $\Delta F = q|\Delta V| = 2.4 \cdot 10^{-19}\, J = 240\, pN \cdot nm$, having considered a $\Delta V = 1.5\, V$ and the electric charge q of an electron. The free energy has been expressed both in J and $pN \cdot nm$ to underline the order of magnitude of the quantities involved. We are therefore considering forces that act on the picoNewton scale (trillion times smaller than those we are used to on an anthropic scale) and nanometer distances. Pheraps the most important formula to remember is the **thermal free energy** $\Delta F_{ther} = k_B T_r = 4.1 \cdot 10^{-21}\, J = 4.1\, pN \cdot nm$ which represents the thermal energy of the *reservoir*, so the thermal bath surrounding the system under study. We can see that there is a consistent difference between the thermal free energy and the typical one before mentioned, meaning that we need thermal activation energy for processes to occur, but it is also fundamental to have stable chemical bonds.

1.3.2 Equipartition theorem

Last but not least key element of this introduction on biological physics is the equipartition theorem which states that, at thermal equilibrium, energy is shared equally among all of its possible forms. Traslated into physics, this means that each quadratic-term of the hamiltonian of the system contributes with a $1/2 k_B T$ factor to the average kinetic energy in thermal equilibrium. This means that from the mean kinetic energy $\langle \frac{1}{2} m \vec{v}^{\,2} \rangle = \frac{1}{2} m \langle \underbrace{\vec{v}^{\,2}}_{v_x^2+v_y^2+v_z^2} \rangle = \frac{3}{2} k_B T \implies \sqrt{\langle \vec{v}^{\,2} \rangle} = \sqrt{3 \frac{k_B T}{m}} \simeq 640\, m/s$, T being the room temperature. This last result is of great importance too since expresses the average velocity of water molecules.

[4] By *allosteric control* in biochemistry we refer to the regulation of conformational change in protein dynamics of an enzyme.

2

Molecules are small, but how much?! A journey inside living cells

La vita non è né brutta né bella. È originale.

– Italo Svevo

We are now ready to look a bit closer to the organization of the cells, where same ideas are used and used in many different ways. The goal of this chapter is just to give an idea fo the spatial scales of different organism that we will study in future chapters and to give an insight on the hierarchy of scales in cells. By looking at different size devices, we will try to answer the following question: *how can cells keep track of everything?* The above picture shows many peculiar biological actors:

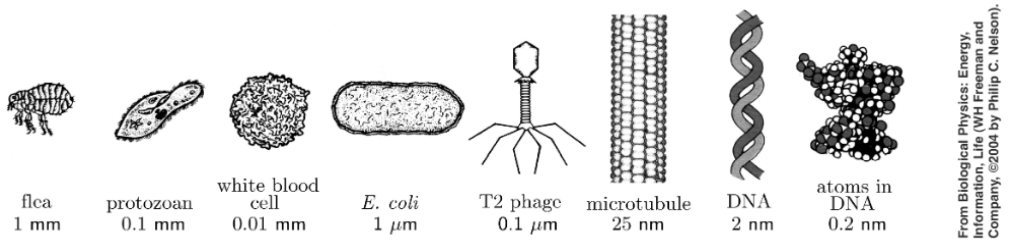


Figure 2.1: Relative sizes of some of the protagonists of this chapter.

starting from a common flea, we are capable to see up to the *Escherichia coli*, which is quite invisible to the naked eye, since the the human eye's resolution limit is of order of 100 mum (*in-vivo*). We know that the objects above pictured have that shape thanks to lenses and visual light employed in electron microscopy (*in-vitro*), which has a resolution limit of $0.1\,nm$.

Setting the scale at $10\mu m$ we can compare (Figure [2.2]) the dimensions of *Escherichia coli* bacteria cells (a) with two cells of baker's yeast (b), whihc are smaller than a human red blood cell (c). Human white blood cell (d) are instead a little bit bigger, but not as much as a human sperm cell (e), or a human skin (*epidermal*) cell (f). On this scale, human photoreptor cells (g), human striared muscle (*myocyte*) cells (i) and also a typical human (*neuron*) nerve cell (i) are the biggest sized actors. Zooming in, on a $0.1\mu m$ scale, we can spot several (a) molecules and macromolecules (such as glucose and ATP molecules, t-RNA, antibodies, DNA and generic enzymes) in comparison with a bacterial cell (b) with its peculiar structure made of flagella, nucleotides and a rigid cell wall. It's easy to distinguish (d) a bacterial virus which has the same dimension of the human immunodeficiency virus (c). In the same car-

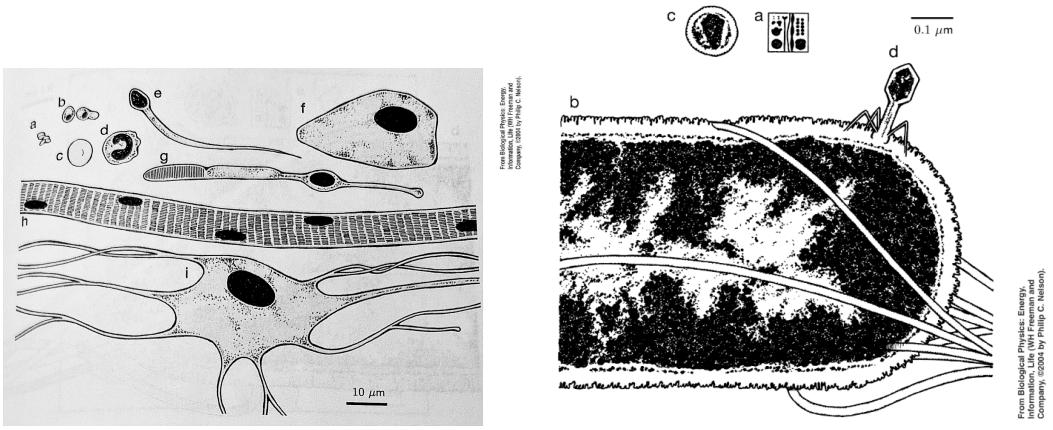


Figure 2.2: Drawing based on light and electron microscopy: relative sizes.

toon extremely different cells are juxtaposed, thinking of neurons capable of receving and outputting electrical signals or bacteria able to live in extreme environmental conditions[1], or further viruses which have a parassitic or symbiotic relationship with the hos. But, they still share similirities such as the way these micro-elements keep themselves organized by using *active transport*, for example, to bring syntesized materials to particular destinations or biochemical reactions to compute highly specific processes.

2.1 Cell physiology

Lecture 3.
Wednesday 8th March, 2023.
Compiled: Sunday 12th March, 2023.

To keep themselves organized, the first thing cells are able to do is to create a clear distinction between the inside and outside world, so between their inner and outer anatomy. Even though there are more than 200 specialized cells in the human body, they all share some common features. Thinking of a cell as separate entities, they are still organism that, in order to work properly, have to take and store energy. A fraction of this energy is used for mechanical activity or for metabolismo - a collection of processes that allows cells to maintein their homeostasis. Energy is further engaged in protein assembly, essential molecules of cells internal structure, or to allow mitosis (duplicating) processes. The inner homeostasis is kept not only by maintaining a fixed interior volume, but also by balancing the concentration of ions there present. Nerves and muscles use this mechanism also to send and receive signals, so to communicate. Communication is of primary importance in order to sense the environmental conditions, so to regulate their interior composition. The presence of needs and food supply or enemy cells has to trigger cells correct response. As part of a feedback loop, cells can also sense their internal conditions and stop particular internal molecular machines. A cell can even destroy itself, in a process called *apoptosis*, in which a specific *protease* is activated because of cell stress, or because of signals from other cells (*cell communication*).

2.2 Internal anatomy

ToDo

[1] In recent years, a number of studies have claimed that ancient bacterial cells are able to survive at very high temperature and to repair their DNA, thanks to an high stress tolerance and resistance to adverse conditions [5]

2.3 Information flow

Lecture 4.
Friday 10th March, 2023.
Compiled: Sunday 12th March, 2023.

The central dogma of molecular biology is the flow of information, starting from the code of life gathered by chromosomes in the DNA of the cell, where it is possible to find *all* the information of an organism. This code is compiled into RNA molecule - specifically *messanger RNA* - which is prepared by *RNA polymerase*, in process called **transcription**, and sent to the endoplasmatic reticulus. This is possible thanks also to regulatory regions and coding regions on the DNA, that specify the amino acid sequences of various needed proteins.

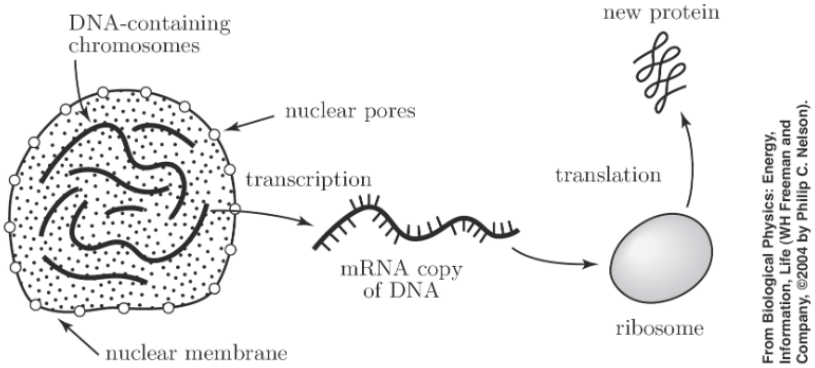


Figure 2.3: General information flow in a cell.

In addition, DNA contains a rich array of regulatory sequences for the binding of regulatory proteins, along with immense stretches with no known function (rewrite this, it's just copied). Then, the *mRNA* is read, in process called **traduction**, where the *transfer RNA* binds to a particular triplet of monomers (*bases*) in the transcript and each brings the corresponding amino acid monomer (*residue*) to be added to the growing polypeptide chain: the biological protein syntesis (*mRNA translation*) is achieved thanks to essential macromolecular machines called *riboromes*. Polypeptide then can create part of the inner cell's structure, in the cytoplasm (or cytosol ??), or can be a regulatory protein, modifying the behaviour of the cell itself. Since this last scenario creates a mechanism for orchestrating the development of the cell, even the DNA in the end can be modified. This "dogma" is therfore not a real one, as

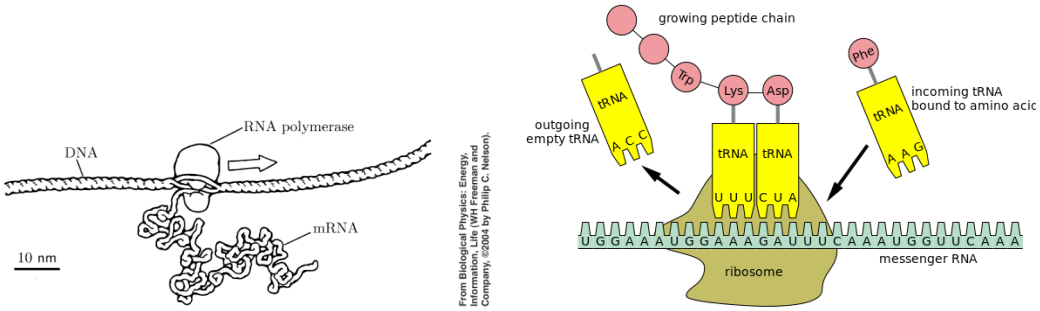


Figure 2.4: Transcription of DNA to *mRNA* by *RNA polymerase* on the left. Peptide systesis on the right.

Francis Crick had defined it. Further discoveries have highlighted this possibility of DNA modification by proteins and macromolecules: the study of such effects is called **epigenetics**, meaning *on top of genetics* from the Greek. Moreover, random point mutation can change either permanently or temporarily the DNA and we have to point out that even multiple clones of the same animal are not completely identical, since they share both the same DNA and also, since the come from a common egg,

its cytoplasm. We can think also of viruses that modify DNA, by manipulating the specif protein production.

Exercise 1.

3

Brownian motion

Molecules are immersed in water and they jiggle all the time due to collisions with H_2O molecules. Even though water molecules (Figure 1.4) are smaller than the particles we are here considering, but 10^{12} is the number of collisions per seconds (?) estimated to which particles are subjected to. This is responsible for the presence of a random force that make particles move randomically in space, making them describing casual patterns.

3.1 Law of diffusion

3.1.1 Random walk model

The paths described by particles are continuous and not differentiable, as well as scale invariant: in fact, zooming in a specific pattern, it is possible to ascertain the fractacl geometry of the trajectories described.

Let's consider a $1 - dim$ problem for semplicity, in which, at each time step:

- the particle make a jump of length $\Delta L > 0$, independent from the previous step, with equal probability to the left or to the right: $P_+ = P_- = \frac{1}{2}$;
- each time interval is deterministic[1]: $\Delta t > 0$;
- $p_i(N)$ is the probability of finding the particle in position $x_i(N) = i\Delta L$, with $i \in \mathbb{Z}$, aftern N steps, so after a time $t = N\Delta t$.

To formally reproduce a possible trajectory, we will state the **Langevin equation**, a *stochastic differential equation* (**SDE**) which uses lagrangian coordinates $\vec{r}(t)$. In eulerian coordinates, instead, $\vec{r}$ is fixed but there is a field $\rho(\vec{r}, t)$ that gathers the stochasticity information of the process, so it expresses the probability of finding the particle somewhere. These considerations will lead us to the *partial derivative equation* (deterministic equation) known as the **Fokker-Planck equation**.

A possible way of formulating $p_i(N)$ is by considering all the paths $M_i(N)$ of N steps ending at position $x_i = i\Delta L$: $p_i(N) = \frac{M_i(N)}{M(N)}$ where $M(N)$ is the total number paths obtainable within N steps. This last variable is easy to compute, since the particle can equally jump to the right or to the left, $M(N) = \underbrace{2 \cdot 2 \cdot 2 \cdot \cdots \cdot 2}_{N\text{-terms}} = 2^N$.

Defining N_r the number of steps the particle takes to the right and $N_l = N - N_r$ the number of steps it takes to the left, then:

$$\begin{aligned} x_i(N) = i\Delta L &= N_r \Delta L + N_l(-\Delta L) \\ &= (N_r - N_l)\Delta L = (2N_r - N)\Delta L \end{aligned}$$

[1] We won't consider the case of random Δt_i, so dealing with a *continuous random walk model.*

From $i = (2N_r - N) \implies N_r = \frac{N+i}{2}$. If we consider exactly N steps, we can have many different patterns, starting at different positions, which stop in i: to count them we have to consider all the possible way to group N objects in N_r subgroups, so basically considering a binomial distribution given by:

$$M_i(N) = \binom{N}{N_r} = \frac{N!}{(N-N_r)!(N_r)!} = \frac{N!}{(\frac{N-i}{2})!(\frac{N+i}{2})!}$$

Note that there is no problem in computing this binomial since N and i must have the same parity: if i is odd (even) there is no possible way of reaching position x_i in an even (odd) number of steps N. In conlusion, the sought probability is

$$p_i(N) = \begin{cases} \frac{1}{2^N} \frac{N!}{(\frac{N+i}{2})!(\frac{N-i}{2})!} & \text{if } i \text{ and } N \text{ share the same parity,} \\ 0 & \text{otherwise.} \end{cases} \tag{3.1}$$

In lagrangian coordinates $x(N)$ is the position of the particle after exactly N steps. Then:

$$\langle x(N) \rangle = 0 \tag{3.2}$$

since, in this model, the particle has the same probability of going either to the left or to the right. Expressing $x(N) = x(N-1) + k(N)$, where $k(N) = \pm \Delta L$, we have that

$$\langle x(N)^2 \rangle \overset{(a)}{=} \langle x(N-1)^2 \rangle + 2\langle x(N-1) \cdot k(N) \rangle + \langle k(N)^2 \rangle$$

where in (a) we have expanded the square and applied the linearity of the mean value. We observe now that $\langle x(N-1) \cdot k(N) \rangle = \langle x(N-1) \rangle \langle k(N) \rangle$ because steps are independent one from the other. Furthermore, as equation [3.2], $\langle x(N-1) \rangle = 0$ and also $\langle k(N) \rangle = 0$ from its definition: these relations actually describe the so called **Markovian process** [6], a stochastic model describing a sequence of possible events where the probability could depend only on the state attained in the previous event - *order 1 Markovian process* for the x variable - or can be completely memoryless - *order 0 Markovian process* for the k variable - so where there is no dependency on previous steps.

In conclusion: $\langle x(N)^2 \rangle = \langle x(N-1)^2 \rangle + \Delta L^2$. Considering $N = 2$, we have that $\langle x(2)^2 \rangle = \underbrace{\langle x(1)^2 \rangle}_{=\Delta L^2} + \Delta L^2 = 2\Delta L^2$. By induction we can finally find the **law of diffusion**:

$$\langle x(N)^2 \rangle = N\Delta L^2 = \frac{t}{\Delta t} \Delta L^2 = t \frac{2\Delta L^2}{2\Delta t} = 2Dt \tag{3.3}$$

$D = \frac{\Delta L^2}{2\Delta t}$ is the diffusion coefficient which, as a *transport coefficient* as dimensions $[D] = \frac{\mathbb{L}^2}{\mathbb{T}}$.
In $3 - dim$ this can be easily extended to:

$$\langle \vec{r}^2 \rangle = \langle x(t)^2 \rangle + \langle y(t)^2 \rangle + \langle z(t)^2 \rangle = 6Dt \tag{3.4}$$

3.1.2 Diffusion processes in biological organism

Now that we have uncovered the main law of diffusion, seeing as the variance of this kind of processes grows linearly in time, we ask ourselves: *is diffusion a good mechanism for transport?*
To answer this question, suppose we have two spherical - for simplicty - cells of radius $R = 1\mu m$ (size of an *E. Coli* approximately) and $R = 10\mu m$ (size of an *eukaryotic* cell approximately) respectively. In the center of these cells there is a specific biological compound such as a glucouse molecule or a globular protein. They have diffusion

coefficients $D_{C_6H_{12}O_6} \simeq 10^{-9} m^2/s = 1 \mu m^2/ms$ and $D_{protein} \simeq 10^{-2} \mu m^2/ms$ respectively[2]. The time τ needed for the molecule to diffuse outside the cell, so to walk a distance R, is given by equation [3.4]: $R^2 = 6D\tau \implies \tau = \frac{R^2}{6D}$.

Table 3.1: Diffusion time.

	Glucose molecule - $C_6H_{12}O_6$	Globular protein
E. coli	$\tau \simeq 0.2\, ms$	$\tau \simeq 20\, ms$
Eukaryotic cell	$\tau \simeq 20\, ms$	$\tau \simeq 2\, s$

As shown in Table [3.1], diffusion is an ideal process for small enough organism, in the case there is the need of transport of tiny molecules. Otherwise, diffusion is not efficient: think, just for example, that brain signals travels on average at $\tau \simeq 2\, ms$: this means that there exist other forms of trasport, of *active transport* specifically, in which cells spend some energy - ATP molecules - to carry out things in specific directions. Diffusion represents a *passive transport* mechanism, perfect for spreading in all directions, mostly because it is energy free, since it is the direct consequence of thermal activity of the environment in which cells are plunged.

[2] Notice how the diffusion coefficient is smaller for bigger cells: the bigger the cell the less molecules tend to diffuse

Bibliography

[1] Evangelista Torricelli a Michelangelo Ricci [in Roma], Firenze, 10 Febbraio 1646, Istituto e Museo di Storia della Scienza, Firenze, Italia.

[2] M. De Domenico, D. Brockmann, C. Camargo, C. Gershenson, D. Goldsmith, S. Jeschonnek, L. Kay, S. Nichele, J.R. Nicolás, T. Schmickl, M. Stella, J. Brandoff, A.J. Martínez Salinas, H. Sayama. Complexity Explained (2019). DOI 10.17605/OSF.IO/TQGNW

[3] Nelson, Philip Charles and Radosavljević, Marko and Bromberg, Sarina and Goodsell, David S, Biological physics: energy, information, life, 2008, WH Freeman New York.

[4] Steve Nadis, Uncovering life's operating code, Science, National Geographic, 2021.

[5] Johnson, Sarah Stewart and Hebsgaard, Martin B and Christensen, Torben R and Mastepanov, Mikhail and Nielsen, Rasmus and Munch, Kasper and Brand, Tina and Gilbert, M Thomas P and Zuber, Maria T and Bunce, Michael and others, Ancient bacteria show evidence of DNA repair, National Academy of Sciences, 2007.

[6] Karlin, Samuel and Taylor, Howard E, A second course in stochastic processes, 1981, Elsevier.