2

Biological Systems, Integrating Information and Entropic Fluxes

Gennaro Auletta
Gregorian University, Rome, Italy
e–mail: auletta@unigre.it

Information acquiring can be considered a three–step or three–component process, in which a processor, a regulator, and a decider are involved. Biological systems are constituted through the integration of these three aspects. In particular, organisms consist of a protein feedback web involving a genetic processor, a regulating metabolic system, and a membrane, separating self and non–self. While a biological system may lack of metabolism, this is the allmarck of organism. These three subsystems can also be regarded as biological systems. The difference between a true organism and other forms of biological systems consists in their ability to integrate two opposite tendencies: the autonomy of the parts and the uniformity of the whole.

KEYWORDS: Infomation, entropy, cybernetics, processor, regulator, decider, genetic system, metabolism, membrane.

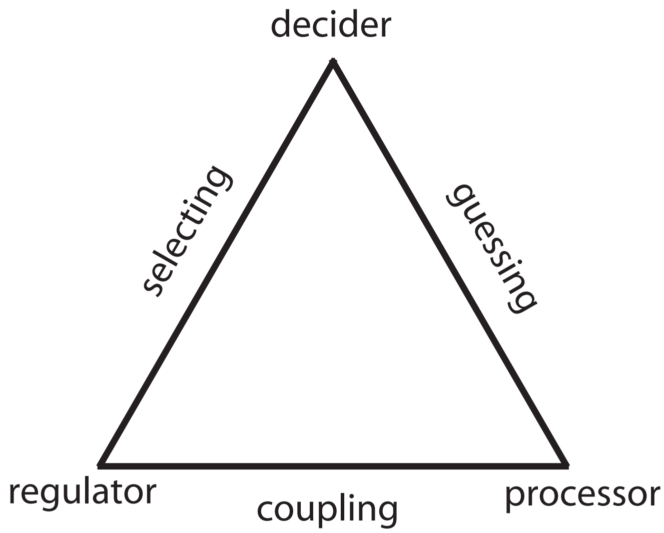

Figure 2.1: Processor–Regulator–Decider.

Introduction

This paper should be understood as an exercise in cybernetics applied to biological system. This means that I try to consider biological systems under a very specific point if view: As systems aiming at acquiring negentropy by controlling environmental information. This means essentially two things:

- This work presupposes molecular biology as the source of data and inputs. A system theory as that proposed here can only be a second–level teory that cannot substitute the fundamental theory.

- For this reason, no new scientific data are presented, but only differently interpreted. This interpretation can be in so far justified, if it provides new and interesting insights about already known problems.

By doing this, I take advantage of the way information is acquired at the physical level, and try to develop some generalizations that could also be applied to biological systems.

2.1 Information–Acquiring From a Physical Point of View

Physically speaking, there are three general aspects or steps in any information–acquisition process [see Fig. 2.1].

- Firstly, a *processor* is necessary, as a source of possible variety. The processor is the component that gives the input so that information can be acquired. It is not necessary that such a processor be random. What is necessary is that the algorithm producing the input be unknown. Otherwise, the acquired information would be valueless.
- The next component is represented by a *regulator*, that is, a system able to work as the interface between the processor and the final detection event. In other words, the regulator provides the necessary coupling, without which we could not speak of information acquiring. I shall return to this point, but let me add here that we never have direct access to any source of variation, we only access its (delayed) effects. This is already true from a relativistic point of view.
- Finally, we need a *decider*, that is, a device that, given a certain coupling, is able to give rise to a decision among a given set (in the simplest case, between two alternatives). In principle, this decision event can have no relation with the initial processor. It is only the coupling (second step) that guarantees that the final event says something about the state of the processor. In this way, we say that the decider has *selected* some information from among the different possibilities to which the processor gives rise.

I note here several things. First of all, there is no irreversibility in the whole process, apart from the final step of decision and selection. According to a theorem of Landauer [Landauer, 1961, Landauer, 1996] and Bennett [Bennett, 1973, Bennett, 1982], it is only information selection that gives rise to irreversibility in any information acquiring. For this reason, information processing can be completely reversible if there is no selection. It follows from this that the processor here can be considered as reversible. It is true that classical information theory [Shannon, 1948] started from the presupposition that information selection is at the source of the process, in terms of the choice of the specific message to be sent. However, this is a circular way to consider the problem: We know nothing about the source if not *through* a regulator (a channel, in information–theory terms), and anything we do know, we know it through a final decision. Why should we predate this final act of decision to an initial one? This would only deprive the whole information–acquiring process of any meaning. At present, this classical formulation can only work in situations where we have technological control over the classical source and over the receiver, but cannot provide an adequate general model of how information is acquired, especially with regard to quantum–mechanical and biological systems. As we shall see, in the latter case the involved systems never have a complete control on the informational sources, and for this reason their information acquiring is affected with some underdetermination, which in turn makes interpretational problems very relevant. Finally, I note that a similar remark was made by the father of cybernetics, Nobert Wiener, about classical systems. According to Wiener [Wiener, 1948], in all classical phenomena where considerations of probability and prediction enter into play, the answers become asymmetrical. One can, in fact, bring a system from the past into the present in such a way that one fixes certain quantities (this is called preparation) and assumes that the other quantities have known statistical distributions. One then observes the statistical distribution of results after a given time (that is, one performs a measurement on the system). This process cannot be reversed, since, in order to do so, one would have to select from the start a fair distribution of systems, which, with no intervention on our part, would end up within certain statistical limits, and thus find out what the antecedent conditions were after a given time. However, for a system starting from an unknown position to end up within a small statistical range is such a rare occurrence that it may be regarded as a miracle. Summing up, one can prepare a system in a certain way and then measure it, but not *vice versa*: In other words, selection comes after prepaation of the initial processor and after regulation. The reason for this is that the very act of selection always consists of an actual reduction from a space of possibilities.

Secondly, for these reasons, the whole information–acquiring process can be understood as a connection between the initial reversible information processing and a final irreversible acquiring event. This means that any process of information acquiring is also an entropic process, through and during which the three components become more disordered than before the process started. One of the biggest mistakes is to mix the concepts of entropy and information (or even to consider information as negentropy). In reality, we have information when there is neither too much order, nor too much disorder [Gatlin, 1972]. Moreover, pure quantum systems are in a zero–entropy state but show a potentially infinite amount of information and at least one bit of information can be obtained from them [Auletta, 2004b, Auletta, 2006].

In all physical processes these three components are separate, that is, each one is instantiated by a different physical system. Even complex systems behave in the same way. Let us take the simple example of Bénard cells. To a certain extent they show a capacity of endogenous self–regulation, provided there are long–term correlations that "coordinate" the motion of the single cells. Moreover, such a network is also able to act as a decider, provided each cell be either levorotatory or dextrorotatory. However, this network has no control at all on the processor, (the source of variety), that lies outside of the system and consists of a source of heat, which here is both an entropic mechanism (allowing the system reach more ordered configurations) and an informational one (able to generate certain specific couplings and decisions of the system). As we shall see later on, biological systems behave differently.

Some General Principles

Here and in the following sections, I try to consider biological systems as systems that emerge from a certain physical structure. In particular, they are characterized by their peculiar ability to integrate the above three components of information acquiring. This is possible if we take into account some general principles indicated by George Ellis so that a system can be considered as truly emergent [Ellis, 2005b, Ellis, 2005a]. They can be summarized as follows:

- There are modular hierarchical structures. These are necessary, since we cannot have the integration of different subsystems without some form of modularity, that is, of (relative) independence of the subsystems, as well as a form of hierarchy, that is, the submission of the systems to the structural and functional constraints of the whole. A modular hierarchy represents the decomposition of a complex problem in constituent parts and processes to handle these parts, each requiring less data and processing, as well as more restricted operations than does the whole.

- There are cybernetic systems, that is a system based on feedback circuits controlling information. In other words, systems in which the dynamics are governed by general formal constraints, and there is a strict connection between informational and entropic aspects. This is true for the whole as well as for the subsystems. Ellis in general stressed [Ellis, 2004] the importance of informational considerations when dealing with true emergent systems.

- There are the dynamics of adaptive modification and evolution. This integration is necessarily a dynamic one. I shall come back later to this point.

The Most Elementary Components

Let me first consider the most common and clear example of biological systems, the organism, whose most elementary form is the cell when dealing with unicellular organisms.
There are four different building blocks of organisms that are relatively common in our galaxy [De Duve, 2005, 6–9]: nitrogenous bases, sugars, fatty acids, and amino acids. it is not by chance that these compounds are the building blocks of organisms. I will not go as far as to say that any eventual extraterrestrial form of life must necessarily have these compounds exactly, but, it is my opinion that it should show an analogous structure. My reasons are the following:

- *Nitrogenous bases* are chemicals that are especially useful for storing, processing, and transferring information. Let us consider the case of RNA (as is well known, RNA was probably the main information processor in the first steps of life). RNA has four bases, cytosine (C), guanine (G), uracil (U), and adenine (A). Since in the template–directed RNA assembly, C is always connected with G, and U with A, we see that the four bases represent a double binary code that is particularly apt for the transferral of information by creating “negative” images of the original string.

- *Sugars* enter in all the metabolic processes of accumulation and transferral of free energy. Sugars are the principal food compound of many cells. In the course of glucose breakdown through a series of oxidations, energy (in the form of adenosine triphosphate: ATP) and reducing power (in the form of NADH) are saved and stored [Alberts, 1983, 43–45]. A chemical becomes oxidized when it loses electrons, and is reduced when it becomes electron–rich. The net result can be written

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O + \text{energy} . \tag{2.1}$$

  Metabolism is the true regulator of an organism, since it is the mechanism which, thanks to the thermodynamic openness of the organism (i.e. the fact that the organism downloads entropy into the environment), allows for the preservation of the structural order of the organism, that is, of the structures that are built according to the genetic processor.

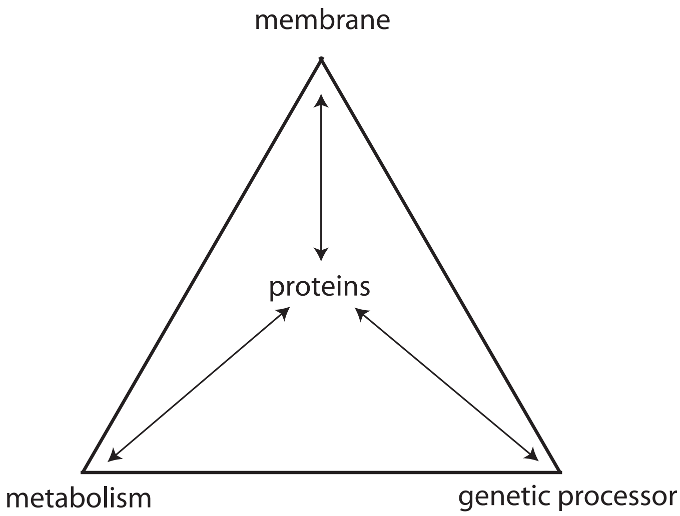

Figure 2.2: Schematic representation of the relationships between the different subsystems of an organism.

- *Fatty acids* pertain to the family of lipids and have originally constituted the compounds that give rise to the cellular membrane. It is the cellular membrane that sharply separates a self from a non–self, and therefore provides a first decider that selects what is allowed to enter into or leave a cell, eventually giving the metabolic system all that is necessary for its working.

- *Amino acids* are the building blocks of proteins. Proteins are often either considered as enzymes or valued for their specific functions. There is, however, a more general function that is of outmost importance: Proteins, toggether with RNA, guarantee all the fundamental passages in the organism's machinery, and constitute, in this way, the vehicles of all feedback interactions within the organism.

I wish to stress that what enters into the cell from the outside is treated at first in informational terms. Any electron that is acquired, for instance, is allowed to enter since it is previously selected as member of an equivalence class (for example, as a high–level electron). This is an informational procedure. The electron that has passed the test *acquires* an entropic meaning only when it enters the cell, since it is inserted into the metabolic web of the system. This also respects the fundamental cybernetic principle according to which cybernetic systems are entropically open and informationally closed [Ashby, 1956][Auletta, 2006]. Moreover, proteins as a whole do not constitute a system apart in the organism, but participate in any work done by the organism: As enzymes in the genetic and metabolic system, and as elements (mostly receptors) helping in the selection procedures of the membrane. Finally, I wish to stress the feedback role of proteins: they are produced according to the instructions of the genetic system and enter into the machinery of the metabolic system and the membrane. However, according to the needs of the latter two systems, they also react back into the genetic system by silencing and expressing parts of the genome [see Fig. 2.2].

Why are organisms built in this way? Because, it is only by integrating the informational and entropic aspect that they can control environmental information in order to acquire free energy for maintaining and improving their structure. As well as this, the intrinsic formal constraints are the mechanism that determines the dynamic of the organism. The need for free energy is obvious, due to thermodynamic considerations. However, an organism cannot have access to free energy (and pursuing in having that access), without the control of environmental information, especially considering that the environment can change in an unpredictable way. Aristotle already understood that organisms are built through a combination of a informational and metabolic dimensions.

Having integrated the processor inside the system (as a genetic system), the relations between the three systems are deeply modified, since the regulator is no longer coupled with an external unknown source of variations (even if, due to modularity, the genetic system is somehow unknown to the metabolic system). However, the whole is still dependent on external free energy and is therefore also reliant on external information (cellular transduction), so that the whole autarchy of an organism is, in reality, illusory. This is especially evident when considering the organism from an ontogenetic point of view, as we shall see.

I think that this way to consider organisms can be seen as a form of moderate functionalism. As a matter of fact, the different subsystems of an organism have a functional role, and, for this reason, we can assume that, on other planets, there could be forms of life that are characterized by other organic molecules, but playing the same role (for instance, they could present other forms of lipids or even other chemicals, having the function of a membrane, especially in environments without water). On the other hand, the metabolic system necessarily excercise a constraint on a pure functional architecture, since it connects the organism with its specific environment. This is another way to express the dependence of the organism on its environment.

Such a model can account for one of the main problems of present research in explaining the development of the first forms of life: the so–

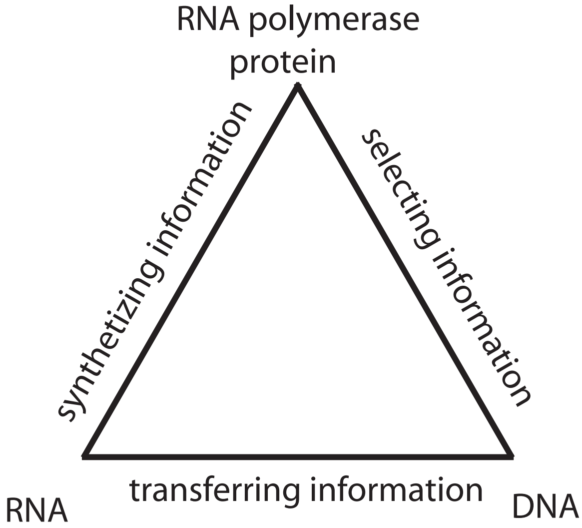

Figure 2.3: Genetic system. It is a true information–transferring system.

called premetabolic phase without RNA [De Duve, 2005, 149–51]. As a matter of fact, the integration process of the building blocks could have lasted a long time, during which proteins of the type known today, and RNA may not necessarily have been formed. More rudimentary and partial forms of these two organic compounds could have worked well, if integrated with sugars and membranes. One of the most commonly made errors is to assume that selection only begins with RNA, whereas that is firstly a consequence of the cell membrane, as I have already pointed out.

The Concept of Biological System

Organisms represent only the most common example of biological systems. In fact, each of the three systems constituting the organism (the genetic, metabolic, and membrane systems) also show general features that are similar to those of the organism. Let me define a *biological system* as any system that is able to integrate a processor, a regulator, and a decider. Let me first consider the genetic system of actual organisms. This can be very schematically depicted as follows:

- The DNA codes the information. This corresponds to a pure processing stage and the random mutation to message variations.
- The RNA (mRNA, tRNA, and rRNA) assures the necessary bridge (the regulator–step) so that this information can be used further.

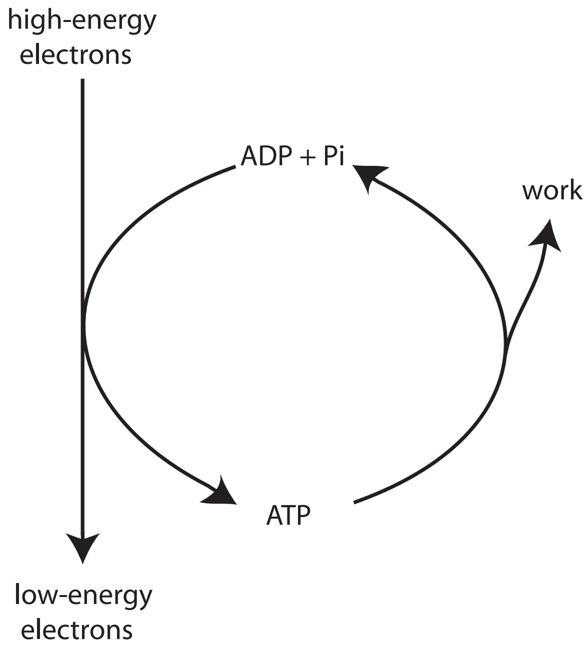

Figure 2.4: The use of high–levels electrons for building ATP. Inspired to [De Duve, 2005, 44].

- The outgoing protein together with the RNA polymerase is the feedback thats select which part of the DNA will continue to be active.

In this way, a single gene can act on many genes, activating and deactivating them. The whole system constitutes a feedback circle [see Fig. 2.3].

As well as this, the whole metabolic process can be considered as a three-system or three-step process:

- First of all, the energy is acquired through some molecular mechanism inside the membrane. One of the most important ways in which this is done is by capturing or producing electrons in excited state. High–energy electrons can be acquired through food for heterotrophic organisms or from mineral donors, in the case of chemotrophy, or can even be powered to excited levels thanks to sun energy in phototrophy (allowing, in this way, the same electrons to be used cyclically).

- Secondly, this energy is stored in the ATP molecules. As we have seen, high–energy electrons can somehow be acquired. By bringing

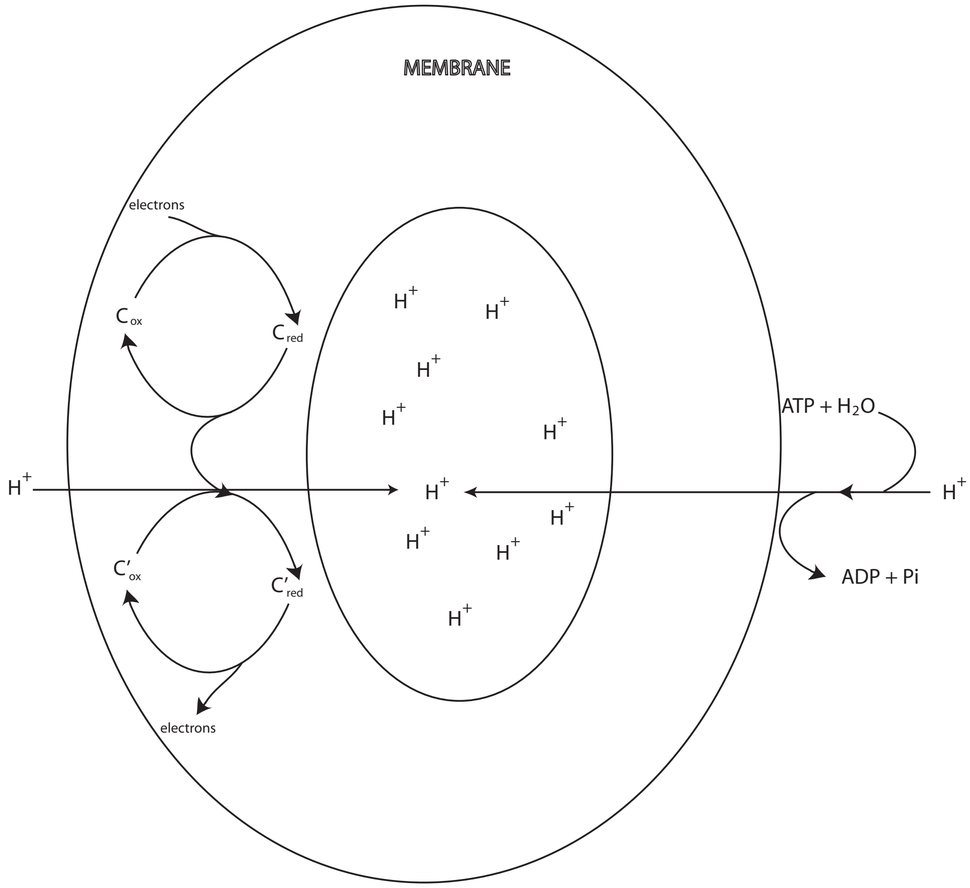

Figure 2.5: Mechanism of the protonmotive force. Inspired to [De Duve, 2005, 134].

them to ground state, one can use the differential energy for building ATP [see Fig. 2.4] [De Duve, 2005, 41–53]. Another mechanism that is much more widely used is represented by the protonmotive force [De Duve, 2005, 133–48]: two reversible proton (H^+) pumps are coupled, one driven by the transfer of electrons between two carriers and the other by ATP hydrolysis [see Fig. 2.5]. The first pump transfers protons making use of electrons that are given first to a carrier that becomes reduced, i.e. electron rich (C_{red}, which becomes thereafter the oxidized C_{ox}), and from this to another carrier C', subject to an analogous procedure. The second pump transfers protons by the hydrolysis of ATP. When, as is often the case, the electron–driven pump builds a higher proton potential than the ATP–driven pump, the latter functions in

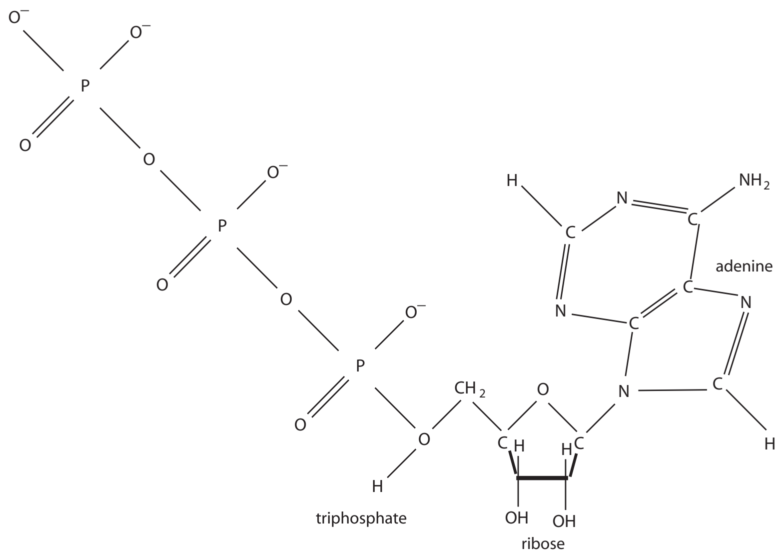

Figure 2.6: Adenosine triphosphate (ATP), the energy storing chemical is a combination of adenine and ribose (which give rise to adenosine) and of triphosphate.

reverse mode and synthesizes ATP [see Fig. 2.6].

- Finally, ATP is used for building polynucleotides, DNA, and RNA, selecting a specific destination from among many possible ones. In general, more complex molecules are built by dehydrating condensation, which, in order to work, is coupled with ATP hydrolysis (a process known as group transfer). ATP undergoes hydrolysis to release inorganic phosphate (P_i) and ADP (constituted by a diphosphate, adenine, and ribose) [see Fig. 2.7]:

$$\begin{aligned} &\text{triphosph.} - \text{adenine} - \text{ribose} \rightarrow \\ &\text{in. phosph.} + \text{diphosph.} - \text{adenine} - \text{ribose}. \end{aligned} \tag{2.2}$$

  This transformation provides energy for work and for chemical synthesis allowing for the storage of structural information. In this way, the organism, in acquiring free energy from the environment and discharging it in entropy, is able to build itself as a structured and ordered system.

It is true feedback, a self–increasing, circle [see Fig. 2.8]. The smallest organisms to present a metabolism are bacteria.

Let us now consider the membrane system. The membrane itself is a lipid bilayer constituted by the hydrophobic tails sticking together while

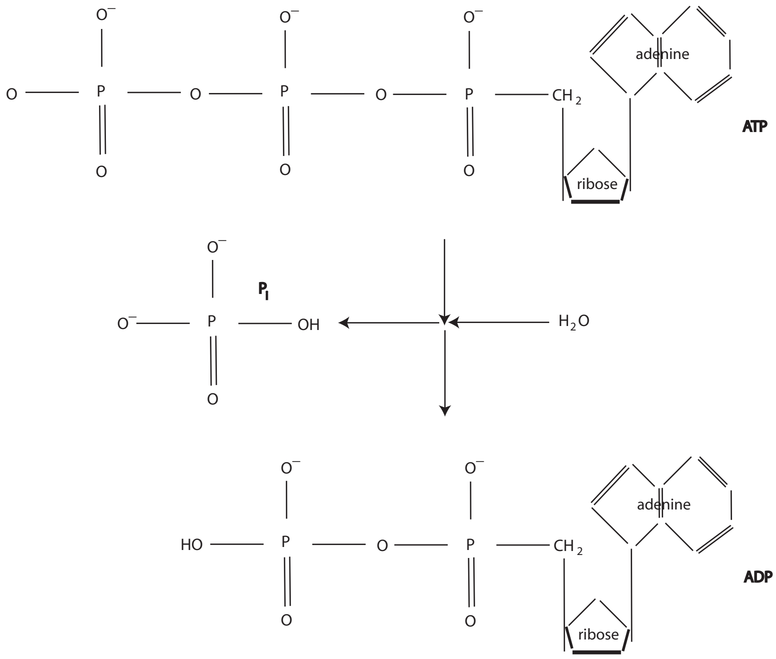

Figure 2.7: The chemical reaction producing ADP from ATP.

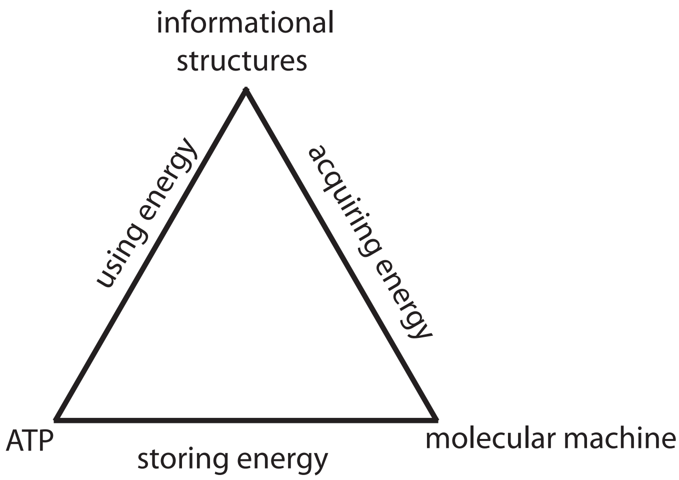

Figure 2.8: Metabolic circle. It is an energetic circle. The direction is here important and cannot be inverted: It runs clock–wise.

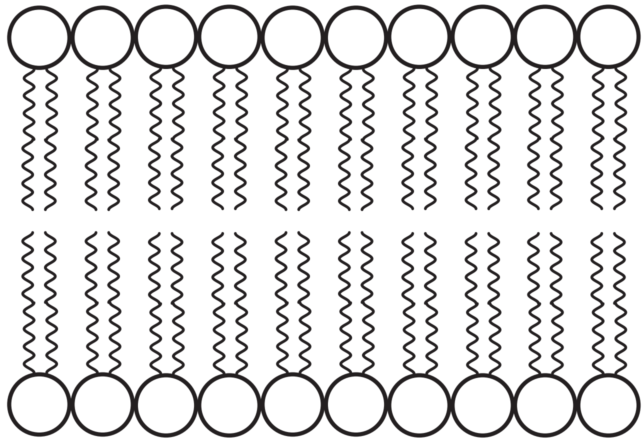

Figure 2.9: Membrane as a lipid bilayer.

the hydrophilic heads remain in contact with water [see Fig. 2.9]. Membranes spontaneously generate sack–like vesicles, that is, (relatively) closed systems. I have mentioned above that this system allows things to enter the cell from outside. Often, however, nothing enters, but an external input, with the help of additional proteins, gives rise to a cascade reaction *inside* the cell. In order that such a mechanism may work, it is first of all necessary that there be a receptor mechanism capable of identifying the external chemical. A comparator is then necessary in order to establish what kind of signal this chemical represents. In the most elementary cases, the receptor is tuned to specific signals, and therefore, is itself a comparator as well [see Fig. 2.10]. Finally, a molecular mechanism acts as a decider by giving rise to the appropriate reaction (in the most elementary case, acceptance or rejection). This is a true informational control system.

The concept of a biological system is of wide applicability. From the point of view of its reproduction, the organism may be regarded as a system consisting of a genotype (the processor), a rybotype (the coupling which ensures the necessary building blocks (proteins) according to genetic instructions), and a phenotype, (the final output selected) [Barbieri, 2003]. From an ontogenetic point of view, the phenotype is structured as a signal transducer (the processor), the metabolic system (the regulator), and a decision system. Here, dependence on the external environment is evident. Even if the organism tries to control it

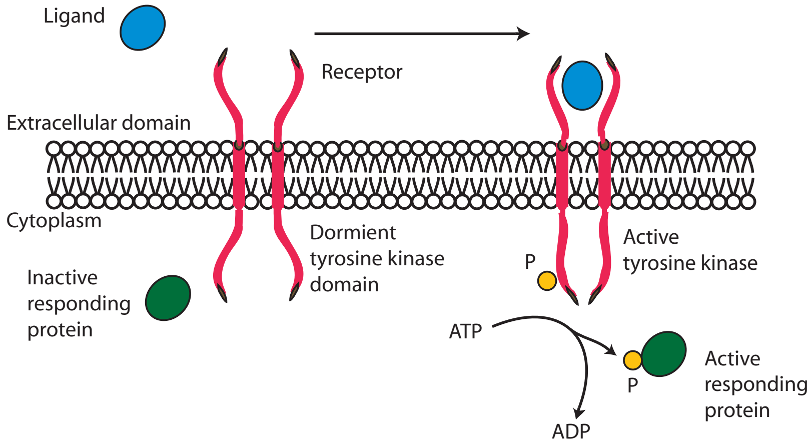

Figure 2.10: The basic mechanism of (paracrine) cell signalling. The external inducer give rise to an enzymatic activity. Usually, this is a kinase activity using ATP to phosophorylate specific kynase residues of cxertain proteins. Inspired to [Gilbert, 1991, 147]. See also [Wolpert, 2002, 299].

completely, it cannot entirely succeed.

2.2 Final Considerations

The structure of biological systems is a true fractal one, since there are no *a priori* limits in building biological systems as parts of previous ones or in collecting biological systems that already exist. An example of the latter is to be found in an ecosystem, in which the different species are the processor (source of variation), the ecological niche or niches, the regulator, (that is, the ground upon which different species encounter each other and somehow establish relations), and natural selection is the decider. In fact, it is probable that biological systems can be defined as scale free systems[1]. As such, population of different organisms are expected to follow general principles that might be the same as those defining the molecular network at the level of one organism. When looking at organisms from that perspective, individualistic and cooperation behaviors are those that are really in balance.

This explains an important aspect of life: biological systems at any level have the tendency to autarchy, that is, to become true organisms. If they are part of a bigger biological organism, this tendency manifests itself as an anarchic one, as happens, for example, with cancer cells. On the other hand, when the whole tries to become an organism, we have

[1] I owe this remark to Luc Jaeger.

the tendency to "totalitarism", as sometimes happens in human society or even with social insects. This tendency cannot succeed, however, as human society, for instance, cannot provide a true metabolism, but is dependent on the metabolism of its members. In general, a biological system can exist in dynamic equilibrium between these two opposite tendencies. This is the reason why biological systems are evolvable and developmentable, that is, are intrinsically dynamic. It is worth mentioning that even the so–called homeostasis, that is, the equilibrium of an organism, should be interpreted instead in dynamical terms as a homeorhesis [Waddington, 1974].

Obviously, as I have already remarked, even the autarchy of an organism is, to a certain extent, illusory, since it always relies on external free energy, and therefore on external sources of information.

From the above considerations, the fact also emerges that viruses are not biological systems, since although they are provided with at least some genetic material and a (non–lipidic) membrane, they have no regulatory mechanism (and this is also the reason why they proliferate *ad infinitum*). An interesting situation is when virus proliferation is kept under control. The case of the bacteriophage λ has been studied [Arber, 1983]. After certain steps, a species of "decision" is done: Either the bacteriophage pursue its reproduction idefinitely (this response is called lytic), and in this case the cell finally dies, or the virus reproduyction is kept under control trough production of the λ repressor cI which is in turn produced by promoters p_{RE} and p_{RM} (a response that is called lysogenic).