FROM INANIMATE MOLECULES TO LIVING CELLS: THE INFORMATIONAL SCAFFOLDING OF LIFE

PEDRO C. MARIJUÁN
Fundación CIRCE, CPS-Universidad de Zaragoza, Zaragoza 50018, Spain
E-mail: marijuan@posta.unizar.es

Both the traditional lack of a unifying approach to informational phenomena in biology and the recent growth of numerous bioinformatic subdisciplines at the interface between biology and computer sciences pave the way for reconsidering the special relationship between information and life. Bioinformation, it will be argued here, should be developed as an integrative scheme in order to characterize the scaffolding of molecular-informational processes on which these new scientific-technological fields are focusing, and on which the organization of the living cell itself is based. In this work, starting out from concepts of the 'molecular recognition' field, which provides a unitary framework to analyze any further molecular-biological processes, the fundamental distinction between sequential and amorphous informational architectures is emphasized –DNA and RNA worlds versus diluted enzymes and proteins. The dynamic, functional overlapping of both classes of informational architectures, particularly taking into account the protein degradation phenomenon, appears at the very center of cellular functioning. The whole structures of the living cell participate in an orchestrated 'evanescent permanence', actually striking an adjustable balance between production and degradation, from which biological adaptability emerges. Subsequently, along the eukaryotic evolution of multicellularity, a scaffolding of informational inventions –i.e., selector genes and cellular signaling systems– has made possible the development and differentiation of a variety of adaptable tissues. In the light of adaptability theory developed by Michael Conrad, a special interrelationship between some of the bioinformatic functional 'omes' would underly the adaptability of each tissue. The resulting adaptive tradeoff between biological functionalities may clarify the intriguing physiological 'shielding' of neuronal tissues in advanced nervous systems. The adaptability tradeoff also provides hints on the background of molecular-informational processes underlying the biological evolution of consciousness.

1. Introduction: beyond the current bioinformatic revolution – bioinformation

The purpose of this paper is the advancement of an integrative perspective of the living cell based on informational conceptualizations. The much vexed question on the attributes which allow distinction between a collection of inanimate biochemical molecules and a living cell will be answered en passant:

it is the informational architecture (or organization, or scaffolding) what implies the basic differences [11]. Biomolecules would remind the case of Silicon compounds, which can also be found either in an unorganized, nonfunctional state in rocks and sands or in a highly organized architecture within modern computers. The structural and dynamic properties of such informational architecture(s) characterizing the living state, and empowering it with its extraordinary capabilities, will be the focus of the present analysis.

Another thread of thought conducing to this informational approach is the contemporary need of, and opportunity for, integrative views [61]. Scientific-technological action in the bioinformatic fields has gone too far away from current theoretical molds. We have to make sense of a growing avalanche of data coming from a number of bioinformatic omes –genome, proteome, transcriptome, metabolome, interactome, plus many other 'omes' in various stages of gestation [28]. Anecdotally, the term *signalome* will be proposed here as comprising the complete set of signaling molecules present in a cell, tissue, or organ –a special relationship between the signalome and some of the other omes will be postulated.

Actually, the relationships among bioinformatic fields are amazingly tangled [61,80]. Rather than a neat downward decomposition of cellular structures and processes into subsystems and families of components, the upward and downward directions of causation criss-cross in a fractal type of interaction. For instance, genomics has to be complemented by transcriptomics and proteomics, and also by protein folding predictions, catalysis and molecular dynamics simulations, structures of multimolecular complexes; biochemical, metabolic and signaling networks analysis; phylogeny, homology and comparative approaches; morphogenesis, developmental and organismal physiology; population and ecological models, etc. [80]. The arrows of causation go all around in multiple directions. In fact, all these heterogeneous fields become linked together in a dense conglomerate where almost any patch or territory may overlap with any other. Even the simplest metabolic question such as "How is isoleucine produced in *Escherichia coli*?" [68] may generate unending quests among the sequences, molecules, pathways, nets, and biochemical literature packed in the Web databases —equivalent to searching in a gigantic dictionary for multifarious linguistic relationships [61].

Although bioinformatics, the 'shotgun marriage' between molecular biology and computer science and engineering during the 1970's, has been hailed as the sine qua non of 21st-century biology [80], and as the definite triumph of reductionism in biological sciences [44], it may also be understood under a very different light. It can be argued that the erratic contour of the

'semiotic' like interactions that actually appear corresponds with emerging, natural properties of biological information [76]. Caught into its own experimental needs, the new computational biology demands not only an exercise of molecular reductionism but a complementary advancement of integrative, systemic biology as well [47].

A side aspect is that the most common denomination for this spattering of new disciplines, *bioinformatics*, literally denotes the application of computer technologies or methodologies to biology, while a more cogent interpretation would be that such computer techniques are allowing us to tap on a whole range of informational properties at all scales of life. *Bioinformation*, we will argue here, seems a more adequate label for this whole new scientific enterprise [61]. Actually, advancing a new integrative approach grounded on the natural properties of biological information (rather than merely computation) would represent an ideal roadmap for making sense of all these new scientific-technological fields.

The complexity of multidisplinary interactions associated to the informational analysis of life does not need much emphasis. Significantly, the pioneers of molecular biology had already couched most of their new conceptualizations under the information metaphor (codes, messengers, transcription, translation, signaling, receptors, etc.); and later on a number of theoretical biology studies have been based on Shannon's information theory, physical information theory, control theory, dynamic systems theory, formal logic, category theory, cellular automata, Boolean networks, genetic algorithms, or the 'proteins as computers' metaphor [9,69]. However, a rigorous and coherent approach to the whole phenomenon of biological information has not been developed yet –here, like in most of its fields, biology has lacked unifying theories, for obvious complexity reasons. What we will attempt in the present work is but a preliminary exploration of the main themes that could be integrated into such bioinformation synthesis.

Starting out from concepts of the 'molecular recognition' field, which provides a unitary background to analyze any further specialized informational architectures, we will basically consider the distinction between *sequential* and *amorphous* architectures; i.e. DNA and RNA worlds versus diluted enzymes and proteins. We will study them separately; afterwards the dynamic, functional overlapping of both informational architectures will be considered, particularly taking into account the new knowledge recently gained on protein degradation. An integrated cellular functioning looms, based on the coupling between protein synthesis and degradation –as we will put it, the whole structures of the living participate in an orchestrated 'evanescent permanence.' It is a generalized

molecular condition of being 'always in the making' from which biological *adaptability* emerges.

Subsequently, along the eukaryotic evolution of multicellularity, a scaffolding of informational inventions has made possible the development and differentiation of a variety of adaptable tissues (in particular, the cellular signaling system appears as the towering developmental tool of eukaryots). In the light of adaptability theory developed by Michael Conrad [18-21] a special interrelationship between three of the previous omes –genome, proteome, signalome– underlies the adaptability of the different tissues. The resulting tradeoff between these three omes may clarify the physiological 'shielding' peculiarities of neuronal tissues in advanced nervous systems (mammals). The adaptability tradeoff also provides hints on the background of molecular-informational processes underlying the evolution of consciousness [60].

2. Molecular Recognition: a unitary background

"Biological Homing" is for W.J. Meggs the genuine force that organizes life [66]. A very long list of recognition phenomena in living cells should be characterized as pairs of molecules that mutually recognize their complementary forms and get attracted together, combining with great specificity and rapidity ('homing'). This homing aspect, crucial to life, spans across vast realms of biomolecules, applying to a number of recognition interactions such as:

enzymes/substrates,
enzymes/effectors,
enzymes/cofactors
antibodies/antigens,
receptors/transmitters,
receptors/hormones,
channels/ions,
channels/ligands
nucleotides/DNA-RNA chains,
RNA/RNA pairing,
RNA/DNA pairing,
DNA/DNA pairing,
DNA/promoters,
DNA/histones,
RNA/ribosomes,
amino acids/protein chains,
proteins/chaperons,

proteins/proteasomes,
proteins/protein multimers,
proteins/protein complexes,
proteins/protein kinases-phosphatases,
tubulins/microtubules,
actins/microfilaments
carbohidrates/glycoproteins,
lipids/lipoproteins,
phospholipids/membranes...

As a matter of fact, current explorations of molecular self-assembly in artificial systems are also based on harnessing this very phenomenon of molecular recognition [48].

In spite of the ubiquity and universality of the phenomenon, it is not quite well focused in its generality yet. For instance, following Meggs [66] we can assume that the exclusive homing force between complementary molecular surfaces is electromagnetic (partially including the role of water and hydrophobic forces) and that it arises from complementary distributions of electric charges on the surface of molecules. Thereafter, simplified straight calculations show that for a system with N charges there will be a quantum enhancement factor of N2 over coulomb scattering, versus an enhancement factor of just N in classical mechanics. This enhancement can be very large for molecules with many charges in the complementarity sites and it will greatly enhance reaction rates [66]. Or conversely, just a few coulombian complementary motifs may be enough in order to provide appreciable specificity and affinity.

In what extent other –more complex– electromagnetic forces (hydrogen bonds, permanent dipole, induced dipole, London and van der Waals forces) would yield a similar quantum enhancement in molecular recognition? It goes beyond the competence of this author. But in the field of 'molecular machines', which analyzes the specificity of DNA and RNA sequences recognized by promoters and other proteins, T. Schneider has applied the Shannonian informational analysis to the recognition items involved. Just a few conserved bases, e.g., in the TATA box of operator regions, may be specifically recognized either by promoter or by repressor proteins [77]. Somehow, the comparative information-entropy analysis of binding sites in different intracellular subsystems, checked experimentally in terms of kinetic parameters and Gibb's free energy, could provide a tentative response to the above question on recognition quantum-enhancement in the sequential architecture of the 'DNA world'

Schneider's approach to DNA sequences as recognition targets by proteins is one of the clearest cases of analytical success based on a simplified view of DNA as a symbolic playground [77]. Most of contemporary genomics is actually based on computationally approaching DNA under the metaphor of just four different symbolic items treated as beads-on-a-string. In this regard, the pioneering study of DNA symbolic patterns by means of information theory performed by Lila Gatlin during the 60's [34] has known a remarkable expansion during the intervening years [25]. Probabilistic analysis, recognition algorithms, and long-term correlations are nowadays well studied in prokaryotic and eukaryotic genomes –as information patterns and recognition motifs may be distributed in DNA strings separated by very vast distances too (see next section). Besides, different studies about recognition in the minor groove of the double helix and about other possible DNA configurations, and also about the origins of the genetic code and the symmetry interrelationships between properties of triplets and of amino acids (hydrophobic, polar, non polar, acidic, basic) have contributed to charting out further information properties and recognition motifs of the sequential architecture.

The most stringent molecular-recognition problems, however, are in what we may call the 'amorphous architecture' of diluted enzymes and proteins. Here, the primary sequence of amino acids is scarcely informative about further molecular recognition properties of the folded molecule, and the computational approach has to rely on classical biochemical information about the concerned molecules. In protein science, the classical view on enzyme and protein recognition was the 'lock and key' model proposed by E. Fischer [83]. Then, the 'transition state' theory, the 'induced fit', and other various refinements such as the 'shifting specificity' approach have been proposed later on [10]. However, this type of conceptualizations, in spite of their intrinsic interest, cannot be easily extended to yield informational considerations at higher systemic levels.

Interestingly, a new approach to the intramolecular and intermolecular chemical processes underlying moieties' recognition, self-assembly, self-organization, and folding has been developed by Shu-Kun Lin [51]. His principles of similarity and complementarity relate stability with the overall symmetry status of the molecular system. Along the change process of molecular properties, the entropy of mixing increases monotonically with the similarity of the concerned property among the components involved in the chemical transformation. In cases of complementarity (e.g. templates), the combinations of different properties which offset each other can also be related consistently with changes in the entropy of mixing and in the symmetries of the

final system. This approach sidesteps the Gibbs' paradox on the discontinuity of the entropy of mixing: it now increases continuously –monotonically– with the similarity increase, see Figure 1, in a different way than von Neumann's proposal of continuous decrease of the entropy of mixing with similarity increase [51].

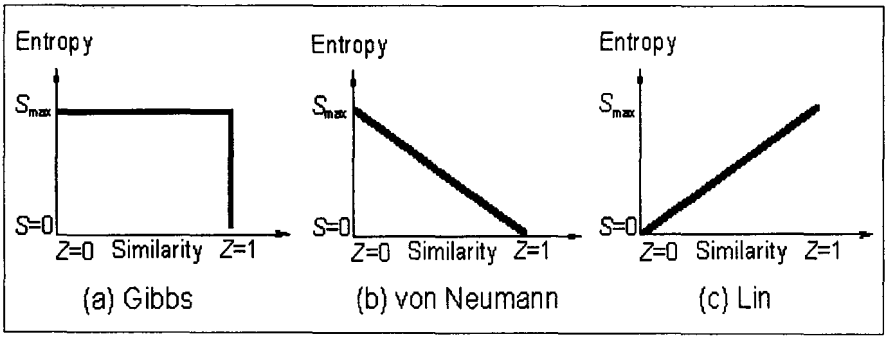


Figure 1. (a) Correlation of entropy (ordinate) of mixing with similarity (abscissa) according to conventional statistical physics, where entropy of mixing suddenly becomes zero if the components are indistinguishable according to the Gibbs paradox. (b) von Neumann revised the Gibbs paradox statement and argued that the entropy of mixing decreases continuously with the increase in the property similarity of the individual components. (c) Entropy increases continuously according to S.K. Lin (not necessarily a straight line because similarity can be defined in different ways). From [51], with permission.

The entropy of mixing relates to the information change of the system through a 'Lewis function':

L = Max. Entropy = S + I

Where the values of entropy S and Information I add to give a new logarithmic function L, which remains unchanged for an isolated system.

Three further laws of information and entropy change are proposed in this new approach to molecular processes [51]; together with the principles of similarity and complementarity, and the symmetry considerations, they represent an exciting new avenue to conceptualize molecular recognition phenomena within the 'amorphous' architecture of the cell.

Let us conclude this brief exploration around biological 'homing' by putting together the two branches of the entire problem. Obviously, we need an approach covering the sequential and the amorphous informational architectures, not in isolation, but in their biological criss-crossing of processes.

How can we make sense of the dynamic information patterns emerging from systematic recognition encounters between molecular partners embedded in sequential versus amorphous architectures, the functional elements of the latter being coded or represented within the patterns of the former?

3. Dynamics of the Sequential Architecture ('DNA World')

The coding relationship was the greatest biological discovery of the 20th century, and the source of the new informational biology of our times. It can be equated with a 'representation' upon DNA-RNA symbols (triplets) of the elementary components (amino acids) integrating the molecular agents (enzymes and proteins) of the amorphous architecture. Besides, the successive reproductive cycles performed by the living cell under environmental pressures enact a continuous 'genetic algorithm,' based on the coding relationship, which optimizes the performances of the molecular agents [57,61].

The above is the core of the Darwinian doctrine. It means, in other words, that by playing around with point mutations and translocations of nucleotide sequences, which imply very simple manipulations upon the nucleotide chains, the DNA-dedicated enzymes of the cell (and many other mutational and 'fluid genome' agents) can alter the functioning of any enzyme, and develop new ones too. Therefore, because of the reproductive capability (amplification) and the process of natural selection (filtering), the 'DNA world' instantiates a transgenerational genetic algorithm devoted to cellular optimization and enzyme specialization following the physicochemical demands posed by the environment.

Further levels of complexity creep in, as multiple syntactic structures and multiple dynamics of change impinge upon the DNA encoding. For instance, we have an intricate regionalization of the 'chromosome territories', as Lima de Faria put it [49,50], expanding from the single base, or codon, up to the entire chromosome (millions of bases) and whole genome: base, codon, restriction site, operator, promoter, exon, intron, enhancer, gene, split gene, transposon, operon, repetitive DNA, gene family, gene territory, centromere, telomere, chromosome arm, entire chromosome, and genome. Subsequently, many fundamental evolutionary aspects related to global cellular processes such as apoptosis, mitosis, meiosis, sex, aging, evolvability, speciation, etc., may be altered by performing very *simple* changes upon such supraregional levels of DNA structure (e.g., mobile elements, telomere shortening, chromosome fissions, duplications, etc.) [49,58].

The physiological or transcriptional use of DNA initially consists in its expression as RNA –the 'representation' or translation being implemented at the ribosome later on. There is an important difference between the eukaryotic and prokaryotic representational grammars: the modular arrangement of the former. Most of eukaryotic enzyme and protein families are built as serial collections of modules –of folding domains actually– defined throughout the shuffling of exons and introns. The maturation of the mRNA, seemingly carried out at the very gene expression 'factories' attached to pores of the nuclear membrane, implies the selection of a definite series of exon modules [54]. The splicing process of introns conveys the possibility of differential splicing and protein multifunctionality (by domain shuffling), enabling a very large eukaryotic transcriptome around two orders of magnitude above prokaryots [7,16,55].

As a consequence of the coding relationship, triplets corresponding to amino acids lining the active site of enzymes will get identified with a concrete subset of the whole DNA sequence. Such subset will be considered here as the *functional address* of the molecular agent. Changing these particular bases, or triplets, occasionally separated across important distances or situated in different modules, may significantly alter the functioning of the active site. However, there is also a plurality of further sequences or addresses which (distributed or not) are coding for translation, transportation, modification, complex formation, and total or partial degradation of the molecular agent [61]. They do so by being recognized specifically as targets by active sites of other molecular agents. This type of 'receptive' sequences may already be at work in the mRNA stage, but they enter into action more often under the enzyme-protein form. We will consider them as *secondary addresses.* Actually, these secondary addresses orchestrate the boundary conditions surrounding the triggering of the main function: the when, where, how many, how much, for how long, with whom – concerning the action of that specific enzyme or protein. Among the most extensive and sophisticate secondary addresses let us point at those devoted to transcription.

Evolutionarily, such DNA secondary addresses appear as non-functional ones (at least from the strict point of view of enzyme catalysis and protein function), but they have become a fundamental tool to elaborate further molecular networking and to facultatively increase the complexity of the biological system. The whole eukaryotic exploration of tissue differentiation, beauplan, and morphology is fundamentally based on secondary addresses (e.g., controlling cis-sequences for transcription, plus many other addresses for protein transportation, modification, degradation, and formation of complexes). The *selector gene* concept, in particular, becomes a paradigm on life's exploitation of both DNA related complexity and computation by protein

complexes [14,72]. Up to three or four dozen enzymes and proteins may pile up in the really elaborate sets of cis-sequences controlling *selector genes* –those mandating development of complex processes and organs, *Endo 16* and *Pax 6* for instance, the latter initiating eye formation [14,84,87]. There is a host of new views in the recently framed 'evo–devo' discipline, contributing to make an integrated sense of expression and signaling events, Hox batteries, mobile elements, retrotranspossons, gene and genome duplications, repetitive DNA, etc. [1,3,13,72] (one cannot help but remembering the myopic views on *junk* or *selfish* DNA during the 70's and 80's).

Genomes are inherently fluid [1,5]. There is an unending rearrangement of the 'DNA text', roughly in parallel with the successive branching and diversification of the evolutionary tree. The sources of DNA change may be classified into several categories that imply increasing level of 'tinkering' [1,12,15,41,53]:

—first, standard mutational rates derived from the genes involved in DNA metabolism (polymerase editing, mismatch correction, chemical lesion processing),

—second, physiological modulation in the generation of variation (SOS system, Carbon starvation, transient mutators, meiosis),

—third, processes differentially focusing change in the genome (contingency loci, hot and cold recombination regions, duplication and translocation events),

—and fourth, those rare systemic alterations that generate the deepest speciation events (chromosome instabilities, chromosome number alteration, genome duplication, and genome 'ingestion' –symbiosis).

Actually, the arcane phenomenon of symbiosis appears as the central event that has separated the five kingdoms of life at the very roots of the evolutionary tree [56]. The whole levels of DNA dynamics become the internal engine of change in the genetic algorithms of life. They make possible stumbling upon DNA sequences encoding appropriate molecular agents that will keep the system adaptive in its interaction with the ever-changing environment. But let us realize that the DNA text corresponding to an appropriate molecular agent contains, as already said, not only the description of its specialized function, but also multiple secondary addresses providing spatial, temporal and quantitative constraints. Somehow, the *function* and the space-time localizing *addresses* of the function have been put together within the same memory bank and under the same coding scheme. We will discuss later about the evolutionary consequences and augmented problem-solving capabilities implicit in this DNA arrangement of processes. Now it is easy to see the intriguing similarity it has with the von

Neumann architecture of electronic computers, which also encodes operations and addresses together in the central memory.

To summarize, the really elaborate structure and dynamics of the DNA text (*hypertext*, indeed) is the source of its quasi-universality as a problem-solving instrument in the evolutionary context [59,61]. Both the variable coding of physicochemical events and the semiotic-like interaction between heterogeneous networking it provides, greatly enhance the evolvability of biological systems and confer an immense strength and flexibility to its informational, physiological operations. Like the human texts interconnected in the web, this Book of Life inside each living cell piles up vast realms of experiential knowledge encoded into a sequential architecture of symbols.

4. The Amorphous Architecture of Molecular Agents

Enzymes and proteins are flexi-molecular machines of very special energetic properties [31,39]; they have been built, element by element, translated out from the genetic code. Although it is usually neglected, each enzyme or protein has a life cycle of its own. Its existence flows unidirectionally between two especial cavities, birth at the *ribosome* and death or degradation into the *proteasome*. It is curious that research on the latter had to wait almost 30 years once the former was worked out [46].

4.1. *Enzyme and protein function*

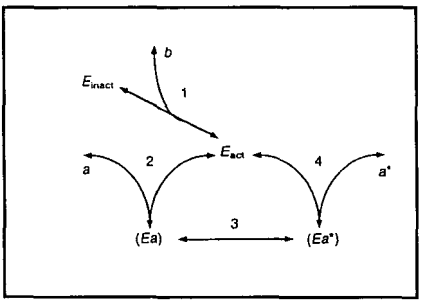
During their variable lifetime, enzymes and proteins display a functioning of fascinating characteristics. It derives from the folding process itself (after exit from the ribosome, the folding of the multi-modular eukaryotic enzyme often needs completion into the protected ambience of chaperons). The mature or 'native state' appears once the enzyme has arrived at the bottom of the *energy landscape* along its folding process [30,31]. In this native state, the enzyme relapses into oscillating motions at different scales, coherent ones included due to its structural property of *tensegrity* [40]; then the enzyme's folding may be destabilized by the arrival of some specific substrate and also by other controlling molecules known as effectors. Once the enzyme is destabilized by the substrate, a new energy landscape emerges, which provides ground for the transformation substrate-product and for completion of a genuine work-cycle with return to the initial state [30,39,57]. The crucial role of water has to be emphasized, both during the folding process and along the work cycle of the enzyme [82,86]. Indeed *water* has to be considered as the first biomolecule, the initial substance that has selected all the other molecules participating in the game of life, including the origins of the genetic code itself [49,50].

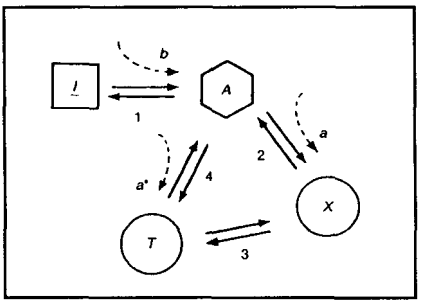
As seen, interactions at the native state with substrates/products and effectors have a machine-like character. There is a clear succession of enzymic states: specific molecular recognition of the substrate, mutual coupling, lowering of the activation energy, interconversion between forms of energy, exit of the substrate transformed into product, and culmination of a regular work cycle [64]. In fact, classical biochemical approaches have described this regular functioning through deterministic rate equations, non-linear ones that are often analyzed in a linear simplified way by means of control theory.

Nevertheless, this functioning may also be approached probabilistically. A stochastic dynamics –*molecular automata*– where enzymes 'fire' their state transitions according to probabilities derived from the free energy differences in between states, can be more realistic than classical equations of control theory [2,57,64]. Moreover, such probabilistic dynamics would be closer to the stochastic nature of transitions in the 'post-folding' energy landscape from which the different states of the enzyme cycle are derived [30]. Figure 2 illustrates the correspondence between the chemical graph and the automata table of states. Needless to say, in quite many intracellular cases of low molecular occupancy, the deterministic equations simply do not hold, while the automata approach keeps its validity and may be handled through the Gillespie's method applied to a stochastic algorithm [2]. Several automata applications, one of them simulating the appreciable complexity of the glycogenetic-gluconeogenetic cycle, have been performed by this author and co-workers [2,32,57,64].

This alternative approach to enzyme dynamics, based on stochastic state transitions, may be put in congruence with Shu-Kun Lin's conceptualization of information change (and entropy of mixing) in molecular recognition. It is intriguing that the inputs-outputs associated to the automata table in Fig.2, correspond with changes in information (and entropy of mixing), but do not correlate automatically with the thermodynamic entropy changes associated to them following the activation energy and the Law of Mass Action. Seemingly, enzymes are able to efficiently decouple their information works from the entropy and free energy changes nominally associated to them. Although the thermodynamic accounts finally have to be settled down, they can be balanced by contributions of other specialized molecular agents participating in the collective division of work. That is the information rationale behind the metabolic network of the cell. Decomposing the vast architecture of thermodynamic cycles into individual reactions carried on by specialized molecular automata creates complex problems [39], but it opens further

possibilities, on when each specific molecular agent has to appear (or be absent) within the whole cellular cycle.





States		Inputs								Outputs	
Previous	Next	a	b	a*	k_{-1}	k_{-2}	k_3	k_{-3}	k_4	a	a*
I	I	/	0	/	/	/	/	/	/	0	0
I	A	/	1	/	/	/	/	/	/	0	0
A	I	0	/	0	1	/	/	/	/	0	0
A	A	0	/	0	0	/	/	/	/	0	0
A	X	1	/	0	0	/	/	/	/	0	0
A	T	0	/	1	0	/	/	/	/	0	0
X	A	/	/	/	/	1	0	/	/	1	0
X	T	/	/	/	/	0	1	/	/	0	0
X	X	/	/	/	/	0	0	/	/	0	0
T	A	/	/	/	/	/	/	0	1	0	1
T	X	/	/	/	/	/	/	1	0	0	0
T	T	/	/	/	/	/	/	0	0	0	0

Figure 2. (Left) Formal mechanism of an isomerase bisubstrate E regulated by the activator b. The substrate and the product are a and a*. (Right) Qualitative representation of the enzyme's action. The states I, A, X and T correspond with Einact, Eact and Ea*. (Below) Logical table of the automaton. The 1 and 0 values represent the occurrence or non-occurrence of the specific phenomenon associated to each variable (e.g., binding of substances, spontaneous dissociation). The sign "/" means that the value of that particular variable is indifferent for the state transition considered. In the most simplified automata approach, probabilities of transition in the Markovian process would be: Pi = 1 – e exp (– Ki.Δt/Δa). See refs. [2,32,57].

4.2. *The complexity of enzyme & protein interactions: nets and complexes*

Like in the case of DNA, further levels of functional complexity creep in, with the important difference that in the dynamics of enzymes and proteins current computational approaches have a rather limited scope. While the abstraction of

DNA as a string of symbols is useful in most biomolecular applications, including bioinformatic ones implying lots of probabilistic analysis and batteries of string algorithms, the 'switching abstraction' that underlies molecular automata and the 'continuous flow' of deterministic equations fail to take into account most of the relevant factors that are interesting, e.g., for proteome, transcriptome, metabolome and signaling science analysis.

For example, a structurally very simple protein, *calreticulin*, which plays an important 'chaperon' role in the intracellular management of Calcium ions, is composed of just three modules or domains: N, P, and C [22]. It has only 46 kDa, around 450 amino acids total. In the module N, one can potentially localize around twelve different activities (a 'secondary address' for ER transportation, a PDI contact and inhibition site, binding sites for Zn, phosphorilation and glycosilation sites, a DNA-binding domain, interactions with perforin, etc.); the module P has around six activities, high affinity Ca^{++} binding sites being the most relevant one; and module C contains another eight, among them high capacity Ca^{++} binding sites [22]. Classical enzymologists working on some varieties of the metabolic enzyme *phosphofructokinase* had also reported an astonishing number of effectors and controlling interactions (7 well authenticated effectors, and up to 23 controlling interactions with other metabolic and signaling molecules), and something similar was happening with other relevant enzymes acting at key metabolic junctions [79].

Obviously, the automata formulation cannot support the accumulation of qualitative and quantitative aspects just described. But one can hardly think that other current computer approaches such as Petri nets, State charts, and Pi calculus would fare much better handling all those 'tactilizing' like interactions [75]. Perhaps combining the basic approaches to enzyme action with the increasing capabilities of computational 'agents' will constitute the most promising strategy of advancement regarding proteomic needs and computer modeling of whole cells [47,75,78].

In many cases –e.g. the calreticulin just seen– enzymes and proteins work as highly 'socialized' agents integrated within functional clusters of modular nature (paradoxically, implying both openness and hierarchical integration); and they may also be incorporated within protein complexes based on specific recognition modules such as SH2, SH3 and others. In these functional modules and protein complexes, the stochastic nature of enzyme-protein function makes a lot of biological sense and is well assumed by molecular biologists working, i.e., in the formation of protein complexes at the nucleus (gene expression machines, and the signaling complexes handling the 'histone code' which regulates the accessibility of DNA to transcription) [67].

As for the functional variety stemming from such massive enzyme-protein networking, in general it is very difficult to handle. Enzyme and protein circuits may display any conceivable class of positive and negative feedback, amplification cascades, feedforward, lineal and parallel processing, robustness and resilience properties, sensitivity, redundancy, graceful degradation, variable interconnection, etc. [23,27,37,47]. At a global scale, power laws clearly emerge in the functional connectedness between enzymes, and also in the formation of protein complexes [42,65]. Precisely there is now a wealth of studies and a growing interest on the connectivity of protein nets, mostly related to the proteomic analysis [35,42,73].

Why should power laws emerge in enzyme and protein connectedness and in formation of complexes? The response should relate to the different classes of DNA tinkering behind the evolution of protein elements –and to the useful properties of nets engaged in power law connectedness (attractor types, robustness, resilience, resistance to damage). Perhaps the 'fractal' arrangement of species in the evolutionary tree, the different DNA tinkering movements, and the power laws in protein connectedness and complex formation are forming part of a global, consistent informational dynamics. It would represent, in other words, the exploration by living matter of the 'natural properties of information' [76]. A *partitional hypothesis* has been inquired by the present author regarding an alternative mathematical background about the emergence of power laws in bioinformational phenomena [62,63].

4.3. *Protein degradation*

And finally, degradation, which is the unavoidably fate of every biological enzyme or protein. Although the whole field was forgotten during decades, recent research on protein degradation (and partial proteolysis) has shown as much complexity as in transcription or translation [46]. Protein degradation may be caused by wear and tear, oxidation, electrical fields, mechanical stress, unfolding, misfolding, N rule, Pest sequences, F boxes, proteases, proteasomes, ubiquitination, sumolation, etc. [4,6,38,46,74]

The functional importance of the systemic specializations based on protein degradation and proteolysis is astonishing. In plants, up to three levels of ubiquitin conjugating enzymes, some of them numbering in the hundreds, are superimposed in order to generate endless different functional combinations, occupying up to 5 % of genome in some species (lacking motility and nervous systems, plants rely on a very tight control of their molecular adaptability: adequately and rapidly changing molecular presences and absences) [38]. In animals, research on the cell-cycle control, onset of apoptosis, cancerigenous

metastasis, development, signaling systems, cellular migration, etc. has clearly demonstrated the widespread influence of proteolytic and protein degradation mechanisms [4,46]. A number of illnesses have recently been related to unbalances and systemic malfunctions in protein degradation (particularly neurodegenerative conditions such as Alzheimer, Parkinson, prionic pathologies, etc.).

An orchestrated turnover of components is also a basic metabolic need, otherwise without recycling of protein-degradation constituents into the cellular pool of amino acids, the metabolic burden of protein synthesis would escalate nonsensically [8]. Any living cell (archaebacteria included) has actually integrated into a unified system the work of its ribosomes, chaperons and proteasomes [24]. Undiscovered until late 80's, proteasomes are in fact one of the biggest protein machines, almost half the size of ribosomes (around 2 Mega Da). They ingest into their catalytic cavities any unfolded or misfolded proteins that have been marked by the small peptide ubiquitin – to be digested into very short peptide stretches or into individual amino acids [24]. Eukaryotic evolution has later explored this necessary eliminative function and extended it into a number of functionally diverse processes [4,6,17,58]. The complex of proteasomes, ubiquitin, specialized proteases and conjugating enzymes may be found in gene expression, signaling systems, activation of precursors, transportation, proteases controlling cascades, apoptotic processes, tissue remodelation, development, digestive systems, metabolic balance, complement system, synaptic plasticity, immune surveillance, chemical warfare, venoms, toxins, etc.

The organization rationale behind protein degradation is, thus, highly multiple and heterogeneous. Perhaps protein folding provides an accurate unitary contemplation of the phenomenon. Like in the case of protein function, folding is also an essential factor to understand the whole physical processes behind protein degradation. As long as an individual enzyme or protein keeps intact the folding that characterizes its native state, if none of the 'destruction boxes' that may be contained in its secondary addresses have been activated by ad hoc conjugating factors, then the molecular agent will be unaffected by the ubiquitin/proteasome system. However, a number of environmental factors are conspiring behind the tenuous barrier of activation-energy shielding the proper folding (often this free energy is within the same order of magnitude than substrate/product transformations) [71]. As soon as the combined, cumulative effects of wear and tear, oxidative damage, mechanical stress, thermal agitation, membrane electrical fields, and other chemical injuries irreversibly separate the enzyme from the proper folding of its tertiary structure, the misfolded or

unfolded new molecular surface of the agent becomes a target for dangerous aggregations. So the stringent need of a complex anticipatory system that combines both an immediate quality control of folding at the exit of the ribosome-chaperon system, and a statistical prevention of molecular senescence due to generic accumulation of damage in proteins [81]. As we will discuss in next section, once a functional complex is successfully established along the evolutionary process, it is only a matter of time that it will become co-opted in order to explore and colonize other systemic spaces in the 'transmolecular' regime of the organism.

To conclude: from cradle to grave. The inherent evanescence of enzyme and protein components is at the very center of cellular adaptability. Could any adaptive cell have been organized around 'eternal' enzymes and proteins? The response actually found in nature is, No. There are no working ribosomes without proteasomes. We tend to understand the molecular adaptability of the living as a unidirectional process of protein synthesis, but it implies in actuality the coupling of both synthesis and degradation [58,61]. The cell is always in the making, and in the dismantling. Biological structures are the transient result of a paradoxical evanescent permanence that starts at the molecular level and reappears at all levels of biosocial organization: from molecular, to cellular, tissular, organismic, ecosystemic... Notwithstanding the reductionist reluctance to acknowledge this 'creative destruction' surrounding biological building blocks, weaving and unweaving the own components becomes an unnegotiable trait of life; the source of its organizational flexibility.

5. Integrated functioning of the cell

Returning to our initial scheme based on two informational architectures, now the problem is how the sequential and the amorphous should orchestrate their 'shaking of hands.' Degradation will be taken initially as a guiding theme. The cell appears as a self-producing entity that has to sense the inevitable absences or *functional voids* of its active components [61].

How can any 'absences' be detected? A tentative *detection and measurement* hypothesis has to be formulated. It implies that changes of the symmetries mandated by the Law of Mass Action in the interlinked reactions derived from the functional absences of enzymes-proteins become changes of information / entropy of mixing detectable by the functional modules of the cell at different levels of organization. Once the molecular absences have been detected (and received ad hoc measurement) the continuous evolutionary

optimization process has conduced to a strong economy of action in the cellular organization of responses [58,61].

The detected absences or needs –functional voids– have to be solved at the most appropriate functional level, preferentially 'right in the middle' (effectors acting upon the native function of other enzymes and proteins), or directly by protein synthesis without any specific control at all, like in the case of 'house keeping' genes –relying only on very generic cellular controls associated to ribosomes [85]. Otherwise, different controlling actions may be allocated at the beginning of the protein synthesis process (transcription), or at some intermediate stage (e.g., translation, folding, transportation), and also at the very end (partial proteolysis, protein degradation). Overall, it is a 'sliding' type of control where absences may be solved in a plurality of ways. Either by directly impinging upon other molecular presences, or by resorting to action upon the different fabrication stages conducing to the appearance of enzyme-protein agents. Economically, the controlling actions on earliest stages are favored; i.e., synthesis has to be preferred upon degradation, as the latter implies the degradation work itself plus the previous work invested in the synthesis. Subsequently we may summarize that the cell solves its *functional voids* by *filling* them *in* throughout specific self-production processes [58,59,61].

The above qualitative description represents the crucial 'calculus' scenario of the cell: when and how it has to fill-in its functional voids. Taking into account the multiple detection and response possibilities, the myriad of other concurring internal events, and the possible arrival of external signals mandating (or forbidding) some categories of response, the problem may reach fantastic dimensions. The Cellular Signaling System, *signalome*, appears as the computing apparatus evolved to confront the challenge. Out from prokaryotic origins, the CSS of eukaryots integrates the crucial communication events with the synthesizing/degradation needs [58,59]. It comprises hundreds of different classes of dedicated molecular agents (receptors, channels, transducers, amplification cascades, second messengers, intermediate effectors, final effectors –see Figure 3) that in each tissue are arranged differently. Every cell has tailored its specialized signalome along its developmental trajectory, in dependence of its own history of received signals [1,58,59].

The particular *measurement* role played by second messengers within the general scheme of CSS processes has to be emphasized; it complements the above 'detection and measurement hypothesis' for those cases involving a high complexity of internal/external interactions.

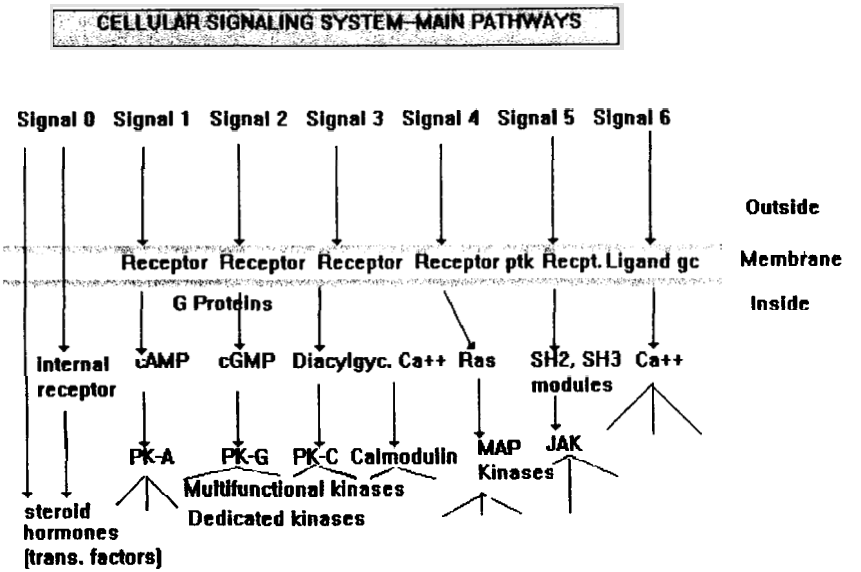


Figure 3. Representation of the principal classes of signaling pathways that operate in eukaryotic cells. The signaling paths in the left (steroids) are the slowest ones, usually associated with cell fate and hormonal effects. Paths 1 to 3, mediated by G proteins, are faster and have a great amplification (ideal ones for sensory receptors), counting with numerous variants. Path 4 corresponds to control of development and cell cycle. Path 5 represents the customary access for neuropeptide action. Path 6, ligand-gated channels, is the genuine cortical path for fast neurotransmitters (GABA, Glutamate). The representation is highly simplified and does not include further effector cascades and cross talking between paths. (Modified from [58].)

These substances (cAMP, cGMP, Ca, InsP3, diacylglicerol, ceramide...) are dramatically modified in their concentrations by the different signaling paths which have been transiently activated, within a generalized cross talking among activated paths –echoing McLuhan, in the cellular system 'the pathway is the message.' Therefore, second messengers provide an integrated perspective (measurement) of the different internal and external influences at play, and are able to pass this integrated influence towards the intermediate and final effectors. At the end of the signaling command chain, the nuclear machinery is going to be fed with ad hoc signals. The nuclear part of the whole signalome apparatus has already been implementing the *histone code,* in order to allow a tight grip upon the euchromatin-heterochromatin states regulating access to transcription –so that the well measured signals from the cytoplasmic signalome may be finally enacted as a new transcription program in relation to the molecular 'absences' to be filled in.

Within this overall 'sliding' system of control, the final transcription process in the nucleus is the usual (preferred) cellular solution to the ongoing cell-cycle problems, including 'artificial' problems generated by external signals impinging on the cell as detected and processed by the CSS apparatus. So to speak, the multicellular organism becomes an extensive 'market' for functional void exchange, adequately endowing each specialized cellular agent with CSS inventions to handle –detect, measure, amplify, solve, export, etc.– the occurring voids (problems) and synthesized solutions in the life cycle of the different cellular communities [58]. It might be argued that a compositional 'language of cells' between tissues emerges, formally played by the signaling pathways of signalomes in support of the standardized exchanges and problem solving interactions of multicellularity [62].

The sliding control we have characterized is in charge of adequating gene expression to cellular needs. It represents the dynamic encounter between the two informational architectures we were distinguishing, the sequential and the amorphous. Let us note that orchestrating the massive cohort of molecular agents and of metabolic and signaling processes it involves is only possible by the flexible concurrence of the enzyme-protein functional modules mentioned in Section 4.2 –their properties of openness and hierarchy, and also the inherent stochasticity of enzyme function.

Evolutionarily, the coupling between the two architectures has been much enriched in eukaryots by exploiting the multiplicity of stages in the 'life' of each enzyme and protein. The evolutionary exploration of molecular solutions on how to build an organism and run its physiology, relies in actuality on an augmented set of trans-molecular interactions to play with. It means that every stage in the 'life' of a molecular agent –transcription, splicing, translation, folding, modification, transportation, modules, complexes, degradation– is amenable to interact with other molecular parties and to play an ad hoc controlling role. If the mutual interaction is appropriately included within the system of secondary addresses, it may be used as a regular element of control and be optimized throughout the evolutionary genetic algorithms. As a general outcome, one can talk about the establishment of a 'transmolecular regime' in the evolution of multicellular organisms, where developmental and physiological solutions may be liberally chosen out from a *transmolecular matrix* of potential interactions between molecular partners situated at quite different stages in their 'life' cycle [61]. The concrete interactions chosen and optimized throughout genomic tinkering will subsequently open new optimization possibilities and evolutionary paths...

This transmolecular rationale might contribute to current evo-devo discussions about the differential characteristics of eukaryotic evolution [13,14,72]. Multicellularity does not exclusively rely on the discovery of new molecular functionalities (throughout functional addresses' tinkering), rather on the exploration of new timings of events, new modular connectivities, and formation of new complexes (throughout secondary addresses' tinkering). This enlarged evolutionary view makes sense in relation with the massive presence of functional loans and the progressive mixing of specialized systems in eukaryotic development and physiology. When inter-addresses conflicts arise (if the optimization of a secondary address conflicts with the main functional address) gene duplications may easily solve the problem. So in plants and in animals there may be a tremendous evolutionary growth of complexity subtended by genomes of relatively modest extension [13,16].

6. Evolutionary coda

Let us recapitulate the basic ideas proposed here. Molecular recognition was our starting point, and making sense of the informational events produced by the dynamic coupling between amorphous and sequential architectures was our main concern. In the extent to which we might establish a rigorous link between a series of theoretical constructs –initially Lin's proposal on information/entropy of mixing [51], afterwards the present author's proposal about molecular automata [64], the emerging functionalities of power-law connectivity in enzyme-protein networks, and the functional void concept as a call to express the sequential architecture and self-modify the population of molecular agents– then an interesting bioinformation scheme of the cell could be drafted.

Presumably, functional voids should be understood as entropy outgrowths. In general, cellular separations (molecular absences) from the regular course of events of the cell's life cycle become losses of dynamic information and approaches to equilibrium in a variety of chemical and cellular contexts [51,59,61]. Adaptively solving its entropy outgrowths is then the main functional concern of the living cell. In the quest for a high level theory to which this peculiar entropy dynamics of the cell could be matched, we have to restrict ourselves to a very simplified scheme. Let us just consider an adaptive structure or *Proteome* –always in the making– selectively decoded out from a *Genome*, following some external instructions plugged in by means of a *Signalome* computing apparatus.

Interestingly, Michael Conrad's *adaptability theory* provides a tradeoff on dynamic (functional) entropies in an ecological setting [18-21]. As recently reviewed by K. Kirby, this is a rather arcane but very flexible characterization based on entropy functions (and not merely entropy values) to be applied to the adaptation of interacting organisms in an ecosystem [45]. Seemingly, it can be extended to the analysis of other adaptive phenomena where an environment, populations of self-modifying entities, and specialized 'computing apparatuses' evolved for the control of the ongoing adaptabilities mutually interact.

Let us assume that within the environment of a multicellular organism each specialized cell adapts by means of a molecular diversity of its own: its proteome derived from its genome, and mediated by the CSS processing. In the light of adaptability theory, the *Prot.* and *Gen.* categories of adaptability indicate a multiplicative relationship, mutually reinforcing the range of available adaptive states, while the CSS or *Sign.* would make a negative contribution. (The negative CSS computational role, congruent with Conrad's tradeoff, is easy to see by paying attention to the signalome's work *in the nucleus*: along the developmental path of the cell towards specialization, signaling actually collapses transcriptomic diversity, as most of the genome is finally caught into a heterochromatin repressed state (see ref. [33] for very illustrative pictures).

We could tentatively apply Conrad's entropy functions to the molecular adaptability of a cell in the organismic context as follows:

Range of adaptive possibilities = *Prot* x *Gen* / *Sign*

Adaptability = Entropy [*Prot*]+ Entropy [*Gen*]– Entropy [*Sign*]

We could examine the adaptability of any tissue under these premises. It is easy to realize, then, the exceptionality of neurons in vertebrate nervous systems. They purport very high values of those three functional entropy elements: a vast proteome endowed with a very high metabolic activity, a number of gene expression events within a very fluid euchromatin state in the nucleus, and also a hypertrophied CSS in the cellular body and in synaptic structures [1,29]. Therefore, neurons have pushed the three functional factors of adaptability up to their very limits –overall resulting in a very poor physiological adaptability [29]. Vertebrate nervous systems require a high degree of tissular shielding (with a separate blood circulation in the brain, upheld by an extra cephaloraquid circulatory system), they also incorporate a retinue of accompanying glial cells (in an approximate ratio 100 / 1 to neurons), and altogether imply a very high metabolic cost (10 times the average of other tissues). At the same time, the functional fragility of neurons is evident: usually

it is the first tissue to metabolically collapse and bring death to the rest of the organism (e.g., intoxications, hypoxia, hypo- and hyper-thermia, etc.).

The neuronal cul-de-sac on adaptability may have forced species endowed with advanced nervous systems –conspicuously mammals and primates– to tap new sources of computing (information processing) without adding to the adaptability overload of neurons. Could that hypothesis contribute to explain the evolutionary co-optation of the intracellular system of microtubules into brain information processing, at the quantum level, as claimed by Hameroff and Penrose? [36,70].

The new functionality accompanying neuronal microtubules could be discussed in an evolutionary context including molecular changes in tubulines and associated proteins (caused by the evolutionary optimization of the new function), development of new anatomical and physiological cerebral characteristics (probably recapitulating the differences between protostome and deuterostome brains [29]), and ecological-environmental factors propelling the whole process of evolutionary change (mammalian seclusion into nocturnal niches during reptilian dominance was proposed by H.J. Jerison as the main factor behind the evolution of consciousness [43]). In any event, consciousness as a biological information-processing phenomenon has a fascinating evolutionary-physiological-molecular history behind it [60].

The information-entropy functions discussed for cellular/organismic adaptability reappear in the regular functioning of central nervous systems. One can easily find 'microscopic' and 'mesoscopic' entropy functions as a natural byproduct of the neuron's *signalome* –sensitization, desensitization, inhibition, etc. They clearly show up in the functional organization of perception, for instance Weber-Fechner's general law of 'sensation' as a *log* of 'stimulus' intensity [52]. Do we perceive but samples of functional entropies about the ongoing environmental changes? It is an intriguing speculation that could extend Conrad's adaptability theory into the very dynamics of central nervous systems. Besides, G.M. Edelman and G. Tononi's recent claim about entropy changes in the discharges of neuronal maps as *bona fide* correlates of consciousness, would be pointing out in a similar direction [26].

Informational inventions are the very 'matter' of life. Starting out from quantum involvement in the 'homing' of biomolecules and culminating in the (quantum?) neuronal dynamics of our own consciousness there is a continuum of bioinformation themes that we have barely sketched here –molecular recognition, sequential and amorphous architectures, prokaryotic and eukaryotic representation-grammars, protein connectedness and degradation, biological self-production, multicellular evolution, neuronal adaptability... They definitely

call for a unifying grand vision. Bioinformation becomes, simply, an emerging synthesis of our times.

Acknowledgments

Many thanks to Professor Franco Musumeci, DMFCI–University of Catania, for his kind invitation to a superb meeting, and to the other organizers of the conference for the great job done.

References

1. Alberts B., Bray D., Lewis J., Raff M., Roberts K. and Watson J.D., Molecular Biology of the Cell (Garland Publishing Inc, NY, 1994).
2. Alves D. and Marijuán P.C., Information and Symmetry in the Cellular System, Symmetry: Culture and Science 6, 1 (1995) pp. 38-41.
3. Arthur W., The emerging conceptual framework of evolutionary developmental biology. Nature 415 (2002) pp. 757-764.
4. Avilés F.X., Guasch A. and Vendrell J., Activación de precursores de proteínas. Investigación y Ciencia Marzo (1994) pp. 74-81.
5. Avise J.C., Evolving Genomic Metaphors: A New Look at the Language of DNA. Science 294 (2001) pp. 86-87
6. Bachmair A., Finley D. and Varshavsky A., In Vivo Half-Life of a Protein Is a Function of its Amino-Terminal Residue, Science 234 (1986) pp. 179-186.
7. Baltimore D., Our genome unveiled. Nature 409 (2001) pp. 814-816.
8. Bradshaw R.A., Protein translocation and turnover in eukaryotic cells. TIBS 14 (1989) pp. 276-279.
9. Bray D., Protein molecules as computational elements in living cells. Nature 376 (1995) pp. 307-312.
10. Britt B.M., A Shifting Specificity Model for Enzyme catalysis. J. Theor. Biol. 104 (1993) pp. 181-190.
11. Brooks R., The relationship between matter and life. Nature 409 (2001) pp. 409-412.
12. Capy P., A plastic genome. Nature 396 (1998) pp. 522-523.
13. Carroll S.B., Chance and necessity: the evolution of morphological complexity and diversity. Nature 409 (2001) pp. 1102-1109.
14. Carroll S.B., Grenier J.K. and Weatherbee S.D., From DNA to Diversity. (Blackwell Science, Malden-MA, 2001).
15. Charlesworth B., Sniegowski P. and Stephan W., The evolutionary dynamics of repetitive DNA in eukaryotes. Nature 371 (1994) pp. 215-220.
16. Claverie J.M., What If There Are Only 30,000 Human Genes? Science 291 (2001) pp. 1255-1257

17. Conaway R.C., Brower C.S. and Conaway J.W., Emerging Roles of Ubiquitin in Transcription Regulation. Science 296 (2002) pp. 1254-1258.
18. Conrad M., Adaptability (Plenum Press, New York, 1983).
19. Conrad M., Physics and Biology: Towards a Unified Model. Applied Mathematics and Computation 32 (1989) pp. 75-102.
20. Conrad M., Cross-scale information processing in evolution, development and intelligence. BioSystems 38 (1996) pp. 97-109.
21. Conrad M., Physical-Biological Roots of Information Processing. Cybernetics and Human Knowing 5, 4 (1998) pp. 46-59.
22. Corbett E.F. and Michalak M., Calcium, a signaling molecule in the endoplasmic reticulum? TIBS 25 (2000) 307-311.
23. Csete M.E. and Doyle J.C., Reverse Engineering of Biological Complexity. Science 295 (2002) pp. 1664-1669.
24. Driscoll J., The Role of the Proteasome in Cellular Protein-Degradation, Histology and Histopathology 9, 1 (1994) pp. 197-202.
25. Ebeling W. and Frömmel C., Entropy and predictability of information carriers. BioSystems 46 (1998) pp. 47-55.
26. Edelman G.M. and Tonino G., A Universe of Consciousness: How Matter Becomes Imagination. (Basic Books, New York, 2000).
27. Endy D. and Brent R., Modelling cellular behaviour. Nature 409 (2001) pp. 391-395.
28. Fields S. and Johnston M., A Crisis in Postgenomic Literature. Science 296 (2002) pp. 671-672.
29. Fox R.F., Energy and the Evolution of Life (W.H. Freeman and Co., New York, 1988).
30. Frauenfelder H., Sligar S.G. and Wolynes P.G., The Energy Landscapes and Motions of Proteins. Science 254 (1991) pp. 1598-1603.
31. Frauenfelder H. and Wolynes P.G., Biomolecules: Where the Physics of Complexity and Simplicity Meet. Physics Today February (1994) pp. 58-64.
32. García-Olivares A., Villarroel M. and Marijuán P.C., Enzymes as molecular automata: a stochastic model of self-oscillatory glycolitic cycles in cellular metabolism. BioSystems 56 (2000) pp. 121-129.
33. Gasser S.M., Visualizing Chromatin Dynamics in Interphase Nuclei. Science 296 (2002) pp. 1412-1416.
34. Gatlin L., Information Theory and the Living System (Columbia University Press, New York, 1972).
35. Gavin A.C. et al., Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature 415 (2002) pp. 141-147.
36. Hameroff S., Consciousness, the Brain, and Spacetime Geometry. Annals of the New York Academy of Sciences 929 (2001) pp. 74-104.
37. Hartwell L.H., Hopfield J.J., Leibler S. and Murray A.W., From molecular to modular cell biology. Nature 402 (1999) pp. C47-C52.

38. Hellmann H. and Estelle M., Plant Development: Regulation by Protein degradation. Science 297 (2002) pp. 793-797.
39. Ho M.V., Bioenergetics (The Open University, London, 1995).
40. Ingber D.E., The Architecture of Life. Scientific American January (1998) pp. 30-39.
41. Jacob F., Complexity and Tinkering. Annals of the New York Academy of Sciences 929 (2001) pp. 71-73.
42. Jeong H., Mason S.P., Barabási A.L. and Oltvai Z.N., Lethality and centrality in protein networks. Nature 411 (2001) pp. 41-42.
43. Jerison H.J., Paleoneurology and the Evolution of Mind. Scientific American 234 (1976) pp. 90-101.
44. Kay L.E., The Molecular Vision of Life (Oxford University Press, New York, 1993).
45. Kirby K.G., Biological adaptabilities and quantum entropies. BioSystems 64 (2002) pp. 33-41.
46. Kirschner M., The cell cycle then and now. TIBS 17 (1992) pp. 281-285.
47. Kitano H., Systems Biology: Towards System Level. In Foundations of Systems Biology, ed. by H. Kitano (The MIT Press, Cambridge, 2001).
48. Lehn J.M., Toward Self-Organization and Complex Matter. Science 295 (2002) pp. 2400-2403.
49. Lima-de-Faria A., Evolution without selection: Form and function by autoevolution. (Elsevier, Amsterdam, 1988).
50. Lima-de-Faria A., Biological Periodicity. (JAI Press Inc., London, 1995).
51. Lin S.K., The Nature of the Chemical Process. 1. Symmetry Evolution – Revised Information Theory, Similarity Principle and Ugly Symmetry. Int. J. Mol. Sci. 2 (2001) pp. 10-39.
52. Llinás R.R., I of the Vortex: From Neurons to Self. (A Bradford Book-The MIT Press, Cambridge-MA, 2001).
53. Lynch M., Gene Duplication and Evolution. Science 297 (2002) pp. 945-947.
54. Maniatis T. and Reed R., An extensive network of coupling among gene expression machines. Nature 416 (2002) pp. 499-506.
55. Maniatis T. and Tasic B., Alternative pre-mRNA splicing and proteome expansion in metazoans. Nature 418 (2002) pp. 236-243.
56. Margulis L., The Conscious Cell. Annals of the New York Academy of Sciences 929 (2001) pp. 55-70.
57. Marijuán P.C., Enzymes, automata and artificial cells. In Computing with biological metaphors, ed. by R.C. Paton (Chapman & Hall, London, 1994) pp. 50-68.
58. Marijuán P.C., The Cell as a Problem-solving "Engine". In Computation in Cellular and Molecular Biological Systems, ed. by R. Cuthberson, M. Holcombe and R. Paton (World Scientific, Singapore, 1996) pp. 183-194.

59. Marijuán P.C., Information and symmetry in the biological and social realm: New avenues of inquiry. Symmetry: Culture and Science 7, 3 (1996) pp. 281-294.
60. Marijuán P.C., Cajal and Consciousness: Introduction. Annals of the New York Academy of Sciences 929 (2001) pp. 1-10.
61. Marijuán P.C., Bioinformation: untangling the networks of life. BioSystems 64 (2002) pp. 111-118.
62. Marijuán P.C., Pastor J. and Villarroel M. The Language of Cells: A Compositional Approach to Cellular Communication. Symmetry: Culture and Science 9, 2-4 (1998) pp. 383-392.
63. Marijuán P.C. and Villarroel M., On Information Theory Stumbling-Blocks. Cybernetics & Human Knowledge 5, 4 (1998) pp. 19-29.
64. Marijuán P.C. and Westley J., Enzymes as molecular automata: a reflection on some numerical and philosophical aspects of the hypothesis. BioSystems 27 (1992) 97-113.
65. Maslov S. and Sneppen K., Specificity and Stability Topology of Protein Networks. Science 296 (2002) pp. 910-913.
66. Meggs W.J., Biological homing: hypothesis for a quantum effect that leads to the existence of life. Medical Hypothesis 51 (1998) pp. 503-506.
67. Misteli T., Protein Dynamics: Implications for Nuclear Architecture and Gene Expression. Science 291 (2001) pp. 84-847.
68. Overbeek R., WIT citation in NetWatch (wit.mcs.anl.gov/WIT2). Science 285 (1999) p. 1323.
69. Paton R.C., Computing with biological metaphors-some conceptual issues. In Computing with Biological Metaphors, ed. by R.C. Paton (Chapman & Hall, London, 1994).
70. Penrose R., Consciousness, the Brain, and Spacetime Geometry: An Addendum. Annals of the New York Academy of Sciences 929 (2001) 105-110.
71. Ponnuswamy P.K. and Gromiha M.M., On the Conformational Stability of Folded Proteins. J. Theor. Biol. 166 (1994) pp. 63-74.
72. Raff R.A., The Shape of Life (The University of Chicago Press, Chicago, 1996).
73. Ravasz E., Somera A.L., Mongru D.A., Oltvai Z.N. and Barabási A.L., Hierarchical Organization of Modularity in Metabolic Networks. Science 297 (2002) pp. 1551-155.
74. Rechsteiner M. and Rogers S.W., PEST sequences and regulation by proteolysis. TIBS 21 (1996) pp. 267-271.
75. Regev A. and Shapiro E., Cells as computation. Nature 419 (2002) p. 343.
76. Scarrott G.G., The Formulation of a Science of Information: An Engineering Perspective on the Natural Properties of Information. Cybernetics and Human Knowing, 5, 4, (1998) pp. 7-17.

77. Schneider T.D., Information and Entropy of Patterns in Genetic Switches. In Maximum-Entropy and Bayesian Methods in Science and Engineering (Vol. 2) (Kluwer Academic Publishers, 1988) pp. 147-154.
78. Shimizu T.S. and Bray D., Computational Cell biology –The Stochastic Approach. In Foundations of Systems Biology, ed. by H. Kitano (The MIT Press, Cambridge, 2001).
79. Sols A., Multimodulation of enzyme activity. Curr. Topics Cellul. Regul., 19 (1981) pp. 77-101.
80. Spengler S.J., Bioinformatics in the Information Age. Science 287 (2000) pp. 1221-1223.
81. Stadtman E.R., Protein Oxidation and Aging. Science 257 (1992) pp. 1220-1224.
82. Szent-Györgyi A., Bioelectronics. A study in cellular regulation, defense, and cancer. (Academic Press, New York, 1968) pp. 8-9.
83. Tanford C. and Reynolds J., Nature's Robots (Oxford University Press, Oxford, 2001).
84. Tjian R., Molecular Machines that Control Genes. Scientific American, Febr. (1995) pp. 38-45.
85. Todorov I.N., How Cells Maintain Stability. Scientific American December (1990) pp. 32-39.
86. Urry D.W., Elastic Biomolecular Machines. Scientific American January (1995) pp. 44-49.
87. Yuh C.H., Bolouri H. and Davidson E.H., Genomic Dis-Regulatory Logic: Experimental and Computational Analysis of a Sea Urchin Gene. Science 279 (1998) pp. 1896-1902.