Life as physics and chemistry:
a system view of biology

Keith Baverstock
Department of Environmental Science
university of Eastern Finland
Kuopio Campus
Kuopio
Finland

Abstract

Cellular life can be viewed as one of many physical natural systems that extract free energy from their environments in the most efficient way, according to fundamental physical laws, and grow until limited by inherent physical constraints. Thus, it can be inferred that it is the efficiency of this process that natural selection acts upon. The consequent emphasis on metabolism, rather than replication, points to a metabolism-first origin of life with the adoption of DNA template replication as a second stage development. This order of events implies a cellular regulatory system that pre-dates the involvement of DNA and might, therefore, be based on the information acquired as peptides fold into proteins, rather than on genetic regulatory networks. Such an epigenetic cell regulatory model, the independent attractor model, has already been proposed to explain the phenomenon of radiation induced genomic instability. Here it is extended to provide an epigenetic basis for the morphological and functional diversity that evolution has yielded, based on natural selection of the most efficient free energy transduction. Empirical evidence which challenges the current genetic basis of cell and molecular biology and which supports the above proposal is discussed.

Keywords: Dissipative system, evolution, cell regulation, self-organisation, 2nd law of thermodynamics, entropy, ecology

Introduction

In his book "*An Introduction to the Physical Chemistry of Biological Organisation*" the Oxford biochemist Arthur Peacock sets out in what he calls the "central problem of biology" two questions to be asked of a living system, "*how does it work*?" and "*how did it arise*?" (Peacock, 1983). These questions, posed in 1983, remain to be answered definitively and the answers are fundamental to the conceptual foundations of system biology. Peacocke gives extensive consideration to the issues for biology raised by the 2nd law of thermodynamics (hereinafter the 2nd law) in relation to the highly organised and ordered nature of living systems and the increase in these qualities over evolutionary time, since the 2nd law, according to Boltzmann's formulation of entropy, appears to predict their inevitable decline to a state of total disorder, the most probable state of any system.

Historically, the thermodynamic implications of the life process have long been of interest: essentially there were two outstanding questions, how is Boltzmann's disorder avoided and how could a non-equilibrium system achieve stability. In 1943 Erwin Schrödinger, in his lecture "*What is Life?*" described living systems as "*feeding on negative entropy*", i.e., decreasing internal entropy at the expense of increasing the entropy in the environment (Schrödinger, 1944). The Brussels Group, led by Ilya Prigogine, explored extensively the ways in which non-linear dissipative systems might achieve stability (Nicolis and Prigogine, 1989). Peacock , in a detailed account, describes this work as "*a very substantial and sustained endeavour*" (Peacock, 1983). Stuart Kauffman (Kauffman, 1993) explored autocatalytic nets comprised initially of peptides with weak catalytic activity, suggesting that these homeostatic systems might be the basis of an origin of life. So, the view on the emergence of life began to shift from an oddity to a natural outcome. Further along these lines Schneider and Kay (Schneider and Kay, 1995) reasoned that complexity in organisms and the ecologies they inhabit are consequences of 2nd law. Finally, in 2007 Arto Annila's group showed by deriving the 2nd law from statistical physics of open systems that for thermodynamically open networks increasing entropy did not preclude increasing order as long as an inward flowing free energy excess was available (Sharma and Annila, 2007). They based their arguments on the principle that disparity in energy levels will be dissipated in the least time, that is, as quickly and efficiently as the system allows. In this way our world-view changed from regarding life and its origin as exceptional, perhaps even miraculous, to a natural process, i.e., an inevitable consequence of the 2nd law.

Cell regulation is almost universally based on so called genetic regulatory networks (GRN) (Babu et al., 2004; Huang, 2009), which in turn are based on the pioneering work of Monod and Jacob (Monod and Jacob, 1961) in 1961. However, while such approaches may be able to predict the primary products of transcription (mRNA) it has become increasingly clear that post-transcriptional regulation may be independent of the primary product concentrations. For example, in mice the transcriptome only partially (50% of the time) matches the proteome (Ghazalpour et al., 2011) and in any case, increasing the sensitivity of measurement of mRNAs suggests that most gene sequences relevant to particular cell types are transcribed at some, *albeit* very low, level (Ptitsyn, 2008). Regulation of transcription is thus necessary, but not sufficient, to describe the

regulation of the cell. Furthermore, the GRN approach is unable to account for the now well established process of genomic instability (Kadhim et al., 1992). To overcome this problem an epigenetic regulatory model, the independent attractor (IA) model (Baverstock, 2010; Baverstock and Rönkkö, 2008) has been developed. It is argued (Baverstock, 2011) that the IA model is superior in explanatory power to the GRN model which, in any case, entails a serious logical flaw.

This paper will address Peacocke's questions as they relate to the cell, which herein is regarded as the building block for all life forms and therefore the critical element of an organism. The answer proposed for the first question entails abandoning two central tenets of the existing dogma, namely that genetic changes underlie evolutionary adaptation and that the information that regulates the activities of cells is not that encoded in the DNA base sequence. Within the length limitations of this paper it will only be possible to present the proposed logical framework in a highly abstracted form and therefore at the expense of detailed arguments. In addition, evidence in support of the proposed framework will be presented.

Cellular organisation and the 2nd law.

Consider the situation where a thermo-dynamically open system (a cell) experiences an excess of free energy (nutrient) at its boundary with its environment. The principles of least action and least time, dictate that the system will strive to equalise the imbalance by diminishing free energy (consume the nutrient) as efficiently as possible. What Sharma and Annila point out is that if organised structures (the cell with its organelles etc.) can do this more efficiently than disordered structures, as is certainly the case, natural selection will favour ordered systems (Sharma and Annila, 2007), *albeit* that the process has been erroneously seen to entail increased internal entropy when entropy has been equated with disorder rather than with bound and free energy. This of course only applies while there is excess free energy in the environment – when that is exhausted the balance has been attained and no new orderly structures will appear or disappear (Pernu and Annila, 2012). This phenomenon is well illustrated at a very simple level by Bérnard cells which form in a liquid with a temperature gradient caused by uniform heating of the base of the container. A situation of disequilibrium is set up with warm liquid overlaid with cooler, denser, liquid, and only thermal conduction by random collisions of the molecules is available to restore equilibrium. At a critical temperature gradient spontaneous spatial order, in the form of a hexagonal grid of rising molecules convecting the heat upwards and descending cooler molecules, emerges as a more efficient energy transduction mechanism than thermal conduction. So, the increased order is not an end in itself but appears in order to consume free energy more effectively. Conversely, when the heat source is removed a randomly ordered molecular configuration is restored.

The principle of *least time* (related to de Maupertuis' principle of least action and Fermat's principle in optics) will apply to any free energy excess in the environment of a cell and thus the most efficient ordered energy transduction processes (that which increases entropy most rapidly) will be favoured. It can, therefore, be concluded the primary basis for natural selection has been and is, variation based on securing the most entropy (Sharma and Annila, 2007). In the view of Sharma and Annila living systems are

simply natural chemical systems which, in principle, cannot be distinguished from inanimate natural systems (Annila, 2010). The over-riding objective of the system is to consume free energy, e.g., by growing and diversifying. Particularly in the early stages of development, growth is limited since powerful energy transduction mechanisms have not yet emerged and organised. Accordingly in the late stages of development, growth decelerates since nearly all free energy has been already consumed. Thus, the overall process gives rise to a "S" shaped evolution of growth with time.

The independent attractor model of cell regulation

The two key features of this model are that a) regulation is an epigenetic process based on information independent of the genomic DNA sequence information encoded on gene products, mainly proteins and b) the concept of attractors is deployed to represent phenotype and therefore stability of the system. The first use of the concept upon which an attractor representing phenotype is based was by Max Delbrück in 1949 when he pointed out in the discussion after a presentation by Sonneborn that ''*many systems in flux equilibrium are capable of several different equilibria under identical conditions. They can pass from one state to another under the influence of transient perturbations.'' (Delbruck, 1949).* Today, the terminology would be different and "flux equilibrium" would be "dynamic steady state". Nevertheless, this statement encapsulates the concept of an attractor that switches from one state to another in response to transient perturbations and in the case of the IA model is supported by a profile of active, that is, interacting, gene products. The interactions among gene products referred to are highly specific, constituting what are called the *rules of engagement* and they comprise the information content of the regulatory process, which is essentially a self-organising phenomenon. From this it can be inferred that gene products carry information but it is not that which is encoded in the DNA base sequences (see accompanying paper) but rather information acquired in the peptide folding process. In summary, the model envisages the cell as a high dimensional (one dimension for each active gene product so a few thousand dimensions for each cell type) complex dissipative system in which the stable states of the system (as dictated by dynamical system theory (Glendinning, 1994)) are discrete. Thus, transitions between such states (phenotypes) are "jumps" or "saltations". This concept of the cell is in marked contrast to the prevailing machine metaphor.

The concept of an attractor to represent phenotype is complicated by the generic nature of the term "attractor" and consequently a degree of looseness in its use. For example, the bi-stable switch that formed the ground work for the modern genetic regulatory theory (Monod and Jacob, 1961) in which the product of one of the genes represses the transcription of another, giving rise to two steady states, is not an attractor in the sense in which the term applies in the IA model. This is because the bi-stable switch is the result of feedback and not dynamic steady states in the system. Ludwig von Bertalanffy in his General System Theory (Bertalanffy, 1969) draws a clear distinction between systems open to information, such as thermostats, that achieve homeostasis, and those open to energy and matter, which are essentially dynamic, where what von Bertalanffy (p46) terms the principle of equifinality applies, that is, where a given final state is reached from several initial starting states

and by different routes. In the former case in biology feedback gives rise to homeostasis (temperature control, for example) and in the latter case the dynamics of the system gives rise to attractors representing phenotype. It is important to keep a clear distinction between these two separate phenomena.

The attractor concept has been used by Stuart Kauffman and Sui Huang , to represent cell types or fates. Kauffman shows that Random Boolean Networks can exhibit state cycle attractors under conditions where the nodes of the network are connected by Boolean functions (rules) (Kauffman, 1993). Huang invokes attractors representing cell fates in genetic networks. In essence the GRN, through directional gene to gene interactions and the application of ordinary differential equations (ODEs), gives rise to an "architecture" that influences transitions between phenotypes, for example, differentiation within a lineage, which "flow" naturally in one direction (stem to terminally differentiated) but can be "pushed" in the other to reprogramme differentiated cells to stem cells (Huang, 2009; Huang, 2012). The concept, as acknowledged by Huang, was explored by Conrad Waddington nearly 50 years ago (Waddington, 1961). As Huang states "*since the interaction specifications of the GRN architecture are determined by the structure of proteins and target DNA sequence, the GRN architecture is "hardwired" into the genome*." (Huang, 2009).

However, there is a clear distinction to be drawn between cell type/fate and phenotype. The term cell fate, where it is derived from a GRN, implies a profile of primary gene products, i.e., mRNAs, whereas phenotype (as deployed in the IA model) refers to the totality of all the features of, and functions being performed by, a cell, at any given point in time and is a function of the active proteins, the relative activities of which is not necessarily reflected in the transcriptome. Thus, the nature of the attractors involved in each case is quite different. A detailed comparison between these two approaches to cell regulation is given in (Baverstock, 2011).

To avoid ambiguities the attractor concept deployed in the IA model has been formalised (Baverstock and Rönkkö, 2008). Essentially, it is hypothesised that peptides, upon folding into proteins, acquire information (as, for example, enzymic activity is acquired or enhanced when peptides fold into proteins) that constitutes the rules of engagement (see accompanying paper) and provides the basis for the self-organisation that appears in the stable form of the attractor. These rules or relations, dictate that if a specific gene product is active in an attractor within a certain range of activities at a time t1, then any other gene product with which it has rules of engagement will be active within certain ranges of activity at time t2, where t2 > t1. This interaction process yields the protein interaction maps such as that determined for yeast (Schwikowski et al., 2000). Thus, the phenotypic state at any given time is represented by a profile of active gene products, each within specific activity ranges, which are related to the size of the basin of attraction surrounding the attractor and thus its robustness in terms of resistance to perturbation. For a cell from a stably replicating species this attractor is termed the *home* attractor (Baverstock, 2000). Central to the home attractor is a set of essential gene products that deal with damage to the genomic DNA. These processes counter the ongoing damage to DNA under physiological conditions (e.g., hydrolysis and oxidative damage) and set-up a dynamic steady state that is fundamental to the life process (Baverstock, 1991), the efficient operation of which ensures replication of gene products

through cell division and fusion. Violations of the rules of engagement can result in the collapse of the home attractor and the adoption of a *variant* attractor or phenotype. Such transitions are discontinuous, stochastic in nature and in practice irreversible: they should be distinguished from the attractor transitions involved in the normal processes of cell differentiation, which conserve the essential properties of the home attractor. Furthermore, it is postulated that the home attractor has been evolutionarily conditioned to provide optimum integrity in DNA replication and optimum robustness in terms of the ranges of activity qualifying a gene product to be part of the attractor. Variant attractors have not been so conditioned and thus are more error prone in replication of the DNA and with smaller basins of attraction and thus less robust to perturbations – the variant cell is genomically unstable and a mutator phenotype (Baverstock, 2000; Baverstock and Rönkkö, 2008; Karotki and Baverstock, 2012).

Transcription of the genotype is an essential component of the overall regulatory process but it is not sufficient: as noted above evidence shows that the mRNA transcriptome does not equate to the proteome as measured by mass-spectrography (Ghazalpour et al., 2011), let alone the active proteome, which is directly responsible for phenotype. Thus, the IA model can be seen as defining how the cell utilises the available transcribed products to produce phenotype, while the GRN model stipulates which sequences shall be transcribed and be available to the regulatory process as represented by the attractor.

Finally, on cell division and fusion, the attractor is inherited along with the genomic DNA. This is equivalent to saying that the active gene products in the cell (the active proteome) at the time of division are inherited, which has to be the case or else there would be no continuity to the differentiation process and cancer cells would not become progressively more malignant. But the attractor and not the genotype, is the primary "vehicle" of inheritance because the genotype does not carry the information necessary to "reconstruct" the cellular phenotype in terms of the constituent gene products, only the information to replicate the peptides necessary for cellular function. The simplicity of the attractor concept is deceptive: the attractor is far more than the sum of its component gene products; it is a multi-channel parallel processor that entails the future development of the cell and has been conditioned by its history.

Proposed answers to Peacocke's questions

In respect of the first question, how do cells work, from the above it can be concluded that a cell is essentially a transductor of free energy in the surrounding environment and thus its primary function is to metabolise. For three billion years they did this as discrete single entities multiplying in number when free energy was available. During that time they acquired most of the features that characterise multicellular organisms, such as circadian rhythm, sexual reproduction, cooperativity in cell growth, etc.. Bacteria in particular are able to form complex colony structures, exchange genetic material, differentiate among cell types, re-structure their genomes and even appear to have cognitive abilities (Shapiro, 2007; Shapiro, 2011). Thus, the transition from microbial to multicellular life about 500 million years ago was mainly a case of deploying already evolved cellular functions in a new context, namely the "building" of a diverse range of multi-cellular structures that could grow as a single entity. Some 200 human gut bacterial

genomes yielded more than 500,000 distinct bacterial genes (Yang et al., 2009). From an estimate of the number of bacterial species in the human, the total number of bacterial genes is estimated to be about nine million. Out of these, metazoans, for example, have made use of only a few percent so it can be assumed that it was not genetic diversity that constrained the appearance of multicellular life. From the evolutionary record, however, it would appear that once multi-cellular growth had been achieved morphological diversity was able to increase dramatically. Major steps were, for example, the internalisation of free energy sources in a gastrointestinal tract and the development of a vascular system to distribute nutrients.

Regarding the second question concerning the origin of cellular life, natural selection based on metabolic efficiency, rather than genetic variation, points to a metabolism-first origin and the possibility that protein only life forms existed before the acquisition of replication using RNA and DNA as templates. Metabolism-first was initially proposed by Oparin in 1926 and the idea further developed by Dyson as a "toy model" in which catalytic polymerisation occurred in droplets of semi-permeable "oil" suspended in a soup of monomers, without specification of the initial chemical nature of the reaction system(Dyson, 1982; Dyson, 1999). Kauffman explored the possible role of autocatalytic nets based initially on the weak catalytic activity of some peptides in bulk solution. In such systems, feedback loops produce homeostasis which is seen as a life-like property (Kauffman, 1993). The purpose of this paper is not to explore the "origin" process in detail, but simply to assess the plausibility of metabolism-first relative to the alternative replication-first proposals, such as RNA World (Cech, 2011).

Amino acids, the basic components of peptides, would have been relatively abundant on pre-biotic Earth (Miller, S. L. 1992; cited by (De Duve, 1995)). Proteins are essential to RNA and DNA synthesis and replication, and evidence from meteors indicates that nuclecobases could have been in the environment 4 billion years ago (Martins et al., 2008). An early and nearly universal cellular phenotypic function is circadian rhythm (CR) known in cyano-bacteria to be regulated by three proteins, *Kai 1, Kai2* and *Kai3*. Nakajima and colleagues have demonstrated that if these proteins are extracted, purified and incubated with ATP, the phosphorylation of *Kai 3* cycles, relatively independently of changes in temperature, with a period of 24 hours (Nakajima et al., 2005). O'Neill and colleagues have demonstrated in a eukaryote that CR is able to operate after blocking the transcription of the responsible proteins (O'Neill et al., 2011) and shown that erythrocytes, which have no nucleus, also exhibit CR (O'Neill and Reddy, 2011). Finally, *Tardigrades*, eutelic (born with their full complement of somatic cells) metazoans are able to tolerate extreme environmental conditions including very high doses of ionising radiation (~6000Gy) (Horikawa et al., 2006) which would severely damage DNA. Erythrocytes can survive doses in this range but eventually succumb to membrane damage. It might be concluded, therefore, that somatically *Tardigrades* are an example of pure protein life. Metabolism-first with subsequent adoption of replication using DNA as a template is therefore plausible as a two stage origin of cellular life.

A major outstanding question then is "how did multicellular organisms develop the morphological and functional diversity seen in the fossil record and in species alive today?" The current dogma, neo-Darwinism or the modern

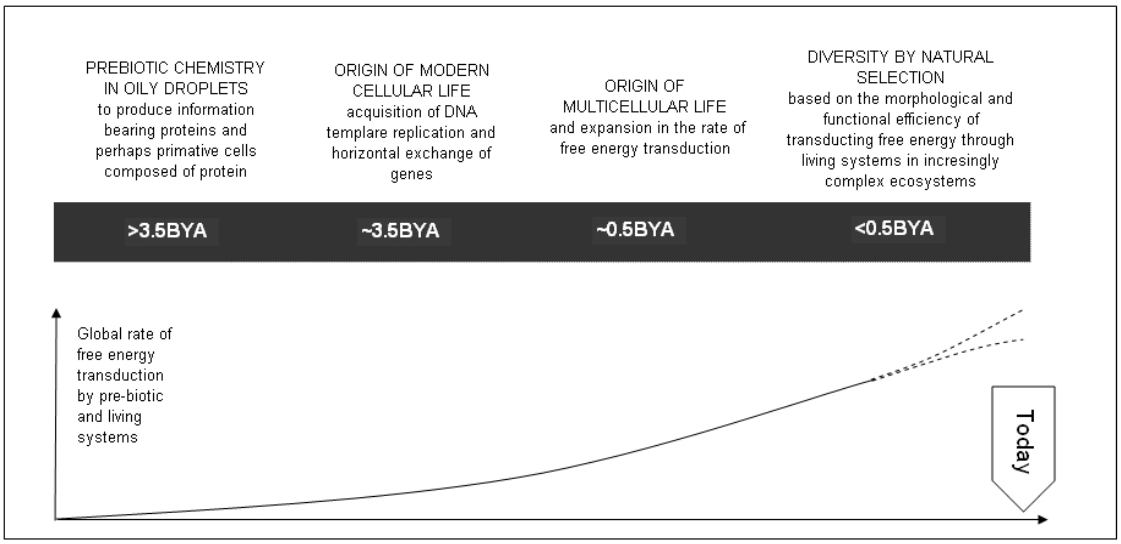
synthesis, attributes this diversity to the vertical inheritance of genetic variation together with natural selection as the creative force. However, all organisms are in close contact with the microbial world and therefore a huge diversity of genetic variants: according to Shapiro , there are no barriers to the horizontal exchange of genetic material between cell types (Shapiro, 2011). Furthermore, a metabolism-first origin implies a functional cellular regulatory system before the acquisition of DNA based replication; indeed, it would be required to acquire that facility. The IA model of cell regulation fulfils this requirement and furthermore relies on exactly the same category of molecules, proteins, that are responsible, as enzymes, for the metabolic processes in the cell.

Diversity in multicellular organisms derives from cellular phenotypic diversity. A cell type, seen as a complex high dimensional dissipative system, with access to some 3,000 active gene products (dimensions) has a huge potential for attractors (discussed in (Baverstock, 2011)) and therefore alternative phenotypes and can, if sufficiently stressed, switch stochastically between them, changing the roles of several gene products in a single irreversible step and thereby generating alternative functions. Thus, the cell to cell signalling functions through which neighbouring cells can influence each other's growth are of particular importance. Morphological diversity can be seen as a matter for cellular cooperativity in a self-similar way to that in which gene products cooperate to produce cellular phenotypic diversity: cooperativity between cells is rooted in the cellular phenotype and its signalling properties. Mauno Rönkkö was able to demonstrate life-like behaviour of virtual worms and beetles in an artificial ecosystem constructed of "atoms" (single particles) with specific rules of interaction between them (Ronkko, 2007). In principle there is no limit on the diversity of structures that could be generated in this way. The life-like dynamics that resulted are attributable to an emergent property of the particle interactions. Thus, it is possible to envisage features at the organism level such as limbs, feathers, sight and hearing, developing purely as result of exploring attractors within a single "genotypic design" that had inherited, from the microbial world, the necessary genetic material, for example, photosensitive proteins. The some 230 cellular phenotypes in the human are derived from some 100,000 gene products, all drawn from an identical genotype. Some 50 million years before *Homo Sapiens* appeared rodents emerged and the genome of the mouse is extremely close to that of the human: slightly fewer chromosomes, a similar number of gene coding sequences many of the same with synteny (the order of sequences on the chromosomes) largely preserved[1]. Additionally, there is now evidence of conservation of synteny for small groups of unrelated genes found in 17 species across several lineages and spanning 600 million years of evolutionary time (Irimia et al., 2012). The clear morphological and functional differences between mice and humans can either be explained by differences in the ways that gene products interact with each other based on nearly identical genotypes, or by minor differences in gene coding sequences in those genotypes. It is argued here that the former is markedly more plausible.

Further support for this contention can be found in the fact that mice and humans scale, along with almost all other living creatures, linearly on a log/log plot for numerous physiological and

[1] See: http://www.evolutionpages.com/Mouse%20genome%20synteny.htm

anatomical features such as metabolic rate with longevity or body mass (Savage et al., 2007). These authors interpret this universal scaling, which incidentally follows the common evolutionary development in time of all natural processes, as derived from physics (Sharma and Annila, 2007), as indicating a common and highly optimised design underlying the multi-cellular life process. Thus, the morphological diversity that is observed is independent of the underlying natural life process, which would be unlikely to be the case if the diversity arose from genetic variation.



A cartoon illustrating the general scheme of the evolution of the natural process of life on Earth. The graphical representation of the rate of free energy transduction by living systems (lower half of the figure) cannot be brought up to date because as the process is indeterminate it cannot be known whether the rate is still rising, levelling off or has already levelled off prior to a collapse of the system.

As illustrated in the figure, it is, thus, a plausible proposition that life arose as a pure protein-chemistry phenomenon exploiting three types of information that emerge as peptides fold into proteins to provide enzymic, structural and regulatory activities. The ability to store in a data base the information required to replicate the necessary proteins by synthesising RNA and DNA was then acquired. This led to the emergence of a diverse range of cellular phenotypes by horizontal DNA exchange that allowed multicellular life to emerge, thereby increasing the complexity and diversity of ecologies (environments). These in turn provided new sources of free energy and thus new challenges to their most efficient exploitation. As stated by Schneider and Kay "*Our study of the energetics of ecosystems treats them as open systems with high quality energy pumped into them. An open system ……. can be moved away from equilibrium. But nature resists*" such movement. "*So ecosystems, as open systems, respond wherever possible with the emergence of organised behaviour that consumes the high quality energy in building and maintaining the newly emerged structure*" (Schneider and Kay, 1995).

Discussion

Given the universally accepted fact that life consumes energy it is surprising that mainstream

cell and molecular biology has paid so little attention to thermodynamics, the branch of physics concerned with the utilisation of energy. Seen in the most positive light this could be because it has been recognised that the thermodynamics of closed systems was not relevant to biology[2]. Clausius is credited with recognising and naming the thermodynamic property of entropy, which imposes irreversibility on all energy consuming processes. However, in the interests of computational simplicity thermodynamics has been primarily developed for thermodynamically closed systems, that is, ones that unlike living systems do not exchange energy and material with their environments. In both cases energy can be categorised as either useful or not useful, for performing work: in closed systems useful energy is termed *Gibbs free energy* and the not useful energy, *entropy,* and the 2nd law states that free energy will be irreversibly degraded to entropy, that is, entropy will always increase. In open systems exactly the same applies – life consumes free energy to produce entropy according to the 2nd law.

However, when Boltzmann interpreted entropy in statistical terms at the molecular level he insisted on conserved energy and particle number. Consequently, the description was limited only to stationary systems where only isoenergic processes, order to disorder, were allowed. When the surroundings are disordered the system must move irreversibly towards the most probable state, the state of maximum molecular disorder. At the same time he recognised that living systems "consume entropy" (cited in (Schneider and Kay, 1995)) but apparently remain ordered and even increase in order over evolutionary time. This paradox has dogged biology since: how to explain the apparent violation of the 2nd law.

There is evidence to support the idea that natural selection is based on free energy availability. First, it may be no coincidence that the gastrointestinal tract of most species is close to the centre of gravity of the body, thus, in broad physical terms, optimising the flow of energy within the body. Secondly, for species which are least constrained by their body structure in terms of growth (snakes, for example) body size should correlate with availability of free energy, i.e., be correlated with latitude. The largest known snake is *Titanboa*, weighing more than 1000 kg, the fossil of which have been found in equatorial S America (Head et al., 2009). Thirdly, ecologies should be more complex with greater diversity of species in tropical regions where more than 80% of the Sun's energy falls. It has been recognised since the time of Darwin that there is a latitudinal diversity gradient which declines with increasing latitude. In a model that quantifies the role of energy in generating biodiversity (Allen et al., 2006) this gradient is predicted and it is concluded that metabolic rate is the primary determinant of evolutionary rates. Finally, more complex and mature ecologies would be expected to absorb more of the available free energy and thus have a lower black body temperature of re-emitted radiation. This is in fact observed (Schneider and Kay, 1995). Schneider and Kay see species and the ecologies in which they live, developing together, the former enriching the latter and the latter creating opportunities for new species to evolve due to increased diversity of free energy sources: like Sharma and Annila they regard these natural processes as an inevitable consequence of the Earth being bathed in excess

[2] On the other hand the importance of recognising the thermodynamic openness of living systems was widely recognised in Germany in the 1950s but apparently not elsewhere according to von Bertalanffy (Bertalanffy, 1969).

free energy and the need to dissipate that energy by all *available* means. These points do not definitively prove that natural selection is based on efficiency of entropy acquisition, only that evidence demonstrates that it is plausible.

This interpretation of the life process does not assign any regulatory role to DNA. The evidence that there is at best a very weak correlation between genotype and phenotype is steadily accruing. Perhaps most telling is the failure to find a high degree of correlation between causes of morbidity and mortality in monozygous twin pairs for the most common diseases (Roberts et al., 2012). Evidence that such twin pairs diverge as they age in terms of the patterns of chromatin marking has been available for sometime (Fraga et al., 2005) and attributed, at least in part, to environmental influences, and to errors in copying the methylation pattern at cell division. However, how chromatin marking patterns are determined is not understood and therefore it is not known whether marking is causal or the consequence of another more fundamental process. Consonant with these results is the issue of the "missing heritability" of alleles for common diseases. Single nuclide polymorphisms (SNPs), which are regarded as an important source of genetic variation, have been assayed in thousands of individuals and together with genome wide association studies (GWAS) on some hundreds of individuals, have failed to identify a genetic basis for common diseases such as type I diabetes, hypertension, coronary heart disease, etc. (Sankaranarayanan and Nikjoo, 2011). Similarly, it has not proved possible to account for the strong resemblances observed between parents and offspring on the basis of allele frequency (Gjuvsland et al., 2011). In short, the foundations of population genetics, that mapping the "genotype space" will predict outcomes in the "phenotype space", where certain traits will be selected and thus modify the genotype space of succeeding generations, is currently being challenged by evidence.

At the level of laboratory experimentation a similar situation is emerging. For example, bacteria exposed to a reduced lactose nutrient and grown for 20,000 generations, regained about 50% of the adaptive fitness that was ultimately attained within the first 1000 generations and with the acquisition of only two mutations . The authors concluded that the dogma that genetic variation underpins environmental adaptation was violated (Barrick et al., 2009). The experiments of Kashiwagi and colleagues on a bacterium genetically modified to be able to compensate for the loss of an essential nutrient demonstrated that adaptation to nutrient loss did not involve any pre-existing gene network (there was none) (Kashiwagi et al., 2006) but was rather an example of self-organisation of the transcribed gene products (Baverstock, 2011). Related experiments with yeast demonstrated the recruitment of an essential gene that had been engineered into a different regulatory system within the organism. The authors concluded that for this to happen reprogramming of the regulatory network must have occurred (Stolovicki et al., 2006). However, the alternative possibility is that under stress all genetic resources are transcribed and self-organisation into the state of maximum metabolic activity (maximum rate of free energy consumption) occurs at the gene product level. As well as challenging current dogma based on a regulatory role for DNA these results support the IA model for cell regulation.

Further experimental evidence is provided by the work of Yus and colleagues on *Mycoplasma*

pneumoniae, a bacterium with a reduced genome. By mapping the metabolic network in its entirety the authors were able to predict successfully the response in terms of growth characteristics to a range of nutrient conditions. However, they also found that the reduced bacterium had retained functions for which it did not have the appropriate coding. The authors concluded "*M. pneumoniae shows metabolic responses and adaptation similar to more complex bacteria [presumably its antecedents], providing hints that other, unknown regulatory mechanisms might exist*" (Yus et al., 2009). This result not only provides evidence of post-transcriptional regulation, but seems to indicate a property of the attractor briefly mentioned above, that it entailed not only the future development of the cell but elements of its history.

Given the long history of the concept of the attractor representing phenotype, dating back to 1949 (Delbruck, 1949), it is surprising how little traction the concept has gained in mainstream cell and molecular biology. The attractor endows phenotypic transitions with a switch like character: as noted by Waddington in 1942 (Waddington, 1942) cells retain their discrete cell type even where tissues are in contact. Of course, the jumps or saltations entailed in phenotypic transitions challenge the Darwinian concept of gradualism, but there is no evidence for a continuum of phenotypic states. The attractor, based on dynamic steady states (rather than feedback) provides the much searched for stability in a thermodynamically open system and according to Annila can be regarded as a free energy minimum in a multi-dimensional state space architecture (Annila, 2010). The interactions between gene products constitute a dissipative process. In the IA model the attractor is in effect performing a parallel computation with some 3000 channels for a human cell type. This has implications for the question of whether the cell can be regarded as computable, a matter of interest to system biology as well as to those concerned with artificial life. Robert Rosen is adamant that organisms are not computable (Rosen, 1991) and Annila would agree: all natural processes have to take the environment into account. The comparatively simple problem of peptide folding is non-deterministic and therefore non-computable, because it is a dissipative process that takes the shortest route down a free energy gradient that is influenced by the peptide's environment in non-deterministic ways (Sharma et al., 2009). Furthermore, Rönkkö points out (private communication) that due to the discontinuities in phenotypic transitions ordinary differential equations would not be applicable as they require continuity.

Discussion of inheritance in mainstream cell and molecular biology focuses almost exclusively on the inheritance of the genotype, but clearly at cell division a specific set of gene products must also be inherited in order that the correct gene products are transcribed in the offspring cells and this is what constitutes the attractor. It is useful to think of this as partitioning the parent cell's non-DNA content (Jablonka and Lamb, 2005). In the IA model it is the attractor, not the genotype that carries the information necessary to define the phenotype of subsequent cell generations (see accompanying paper). The attractor fulfils Mendel's prescription of "units of inheritance" as effectively as does the genotype. It seems that the elegant structure of DNA, its encrypted code and semi-conservative replication made "the gene" an irresistible candidate for the unit of inheritance.

There is a plausible case to be made that the history of biology, for up to 140 years, has been

dogged by two loosely related misconceptions, one about the implications of the 2nd law and the other about the evolution of diversity. The first, dating back to well before 1900 and the second, to 1953, are a result of the strategy of "simplifying the problem to obtain an exact and computable solution", in the first case by concentrating on thermodynamically closed systems and assuming that open systems would behave similarly, and in the second, by basing a general theory on the behaviour of an abnormal type of gene.

In a footnote to Chapter six of his book "*What is Life?*" Schrödinger says that had he been writing for an audience of physicists he would have made his arguments turn on free energy rather than entropy (Schrödinger, 1944). Had he done so he surely would have spotted Boltzmann's misconception, instead he reinforced the mistake. In Annila's view the yearning in physics for predictions that could not be made because of the indeterminate nature of natural systems over-rode common sense (private communication). As noted by Lewontin, Mendel (along with most of twentieth-century molecular geneticists) was careful in his choice of traits to study, choosing, for example, the "drastically dwarfing gene", where variation is easy to see and record. However, much of the phenotypic variation is subtle and therefore much more difficult to detect let alone quantify. This leads to the paradox that "*what we can measure is uninteresting and what is interesting we cannot measure*" (Lewontin, 1974). Now that GWAS and analysis of SNPs enable subtle variation in the *genotype* to be objectively measured it is seen that the behaviour of most genes is, according to the evidence discussed above, not reflected by that of the special subset of traits that has been the object of molecular genetic studies. The generalising of the special case to provide a computable population genetic model has, therefore, been misleading. In the context of the IA model most gene products interact with several others and thus the changes in sequence in one gene, while modifying the gene product, may be of little overall significance to phenotype. This is because small deviations from the "true" phenotype, within the basin of attraction, are accommodated by the attractor, causing "buffering" of small deviations from the norm. On the other hand, gene products that interact with only a single other gene product may have significant consequences for phenotype and if inherited lead to disease. Thus, although there are undoubtedly single locus hereditary effects directly linked to gene mutations they are only a small proportion of the totality of disease and the heritability of common disorders is still missing (Sankaranarayanan and Nikjoo, 2011) and will remain so until inheritance in terms of attractors is addressed.

Finally, the control of transcription needs to be addressed. Its primary purpose is to ensure that the system has access to the correct gene products in order that the phenotype can be expressed. The evidence suggests that in the unstressed state adequate levels of precursors of active gene products are available, as evidenced by the rapid (within minutes) phosphorylation of H2AX site following radiation induced DNA damage (Rogakou et al., 1999). In such experiments a transcriptional response to the damage only occurs after tens of minutes (Watson et al., 2004). It is concluded that such a response is the result of downward causation (Noble, 2012) from the phenotype to the transcriptional processes .

Conclusions

The argument by Annila and colleagues, based on fundamental physical considerations, that it is the selection of the most efficient transductors (metabolisers) of free energy that has driven the process of evolution, rather than the diversity of genomic DNA sequences derived from mutation, is compelling in its simplicity and points to metabolism-first as the likely origin of cellular life based initially purely on proteins and subsequently replication utilising DNA. Logically this order of events implies some form of cell regulation based on proteins pre-dating the deployment of DNA. This is precisely what the IA model provides: a conceptual basis for generating cellular phenotypic diversity from highly specific interactions between proteins based on information acquired as peptides adopt a tertiary structure. In a self-similar way cells acquire inter-cellular interactivity (through signalling) enabling them to form a diverse range of structures as demonstrated by Rönkkö with the life-like properties being an emergent property of these interactions. It is proposed that from a limited range (30,000) of gene coding sequences a broader range (>100,000) of proteins have generated a diverse range of metazoans, all based as West and his colleagues have shown on a single design, a generic genotype. When evolutionists talk about species being adapted to their environments the fundamental issue is "can they exploit the free energy the environment has to offer?" It is proposed here that it is on this criterion that natural selection acts and the beaks of Darwin's finches are a good illustration.

A rapidly accruing body of evidence challenges the DNA centric dogmas that dominate evolution and cell regulation, both of which are predicated on the machine metaphor for the cell and Aristotle's material cause, the genomic DNA sequence. However, when the cell is viewed as a thermodynamically open dissipative system and process underpins function, the efficient cause is the more relevant. While energy transduction processes dominate evolution the efficient cause of cell regulation is the attractor, which entails the future states of the system.

That living systems may have fundamental limits in terms of computability is an important implication for system biology. Attractors figure prominently in applied mathematical research into non-linear dynamical systems of industrial interest (control of chaos): they have been neglected in biological research and should be given greater attention.

Acknowledgements

The author wishes to thank two reviewers for constructive and helpful criticism and Arto Annila, Eliana Beltran, Andrei Karotki, Hooshang Nikjoo, Mauno Rönkkö, and Dillwyn Williams for stimulating and insightful discussion.

References

Allen, A. P., Gillooly, J. F., Savage, V. M. and Brown, J. H. (2006) Kinetic effects of temperature on rates of genetic divergence and speciation. *Proc Natl Acad Sci U S A* **103,** 9130-5.

Annila, A. (2010) The 2nd Law of Thermodynamics Delineates Dispersal of Energy. *Int. Rev.Phys.* **4,** 29 - 34.

Babu, M. M., Luscombe, N. M., Aravind, L., Gerstein, M. and Teichmann, S. A. (2004) Structure and evolution of transcriptional regulatory networks. *Curr Opin Struct Biol* **14,** 283-91.

Barrick, J. E., Yu, D. S., Yoon, S. H., Jeong, H., Oh, T. K., Schneider, D., Lenski, R. E. and Kim, J. F. (2009) Genome evolution and adaptation in a long-term experiment with Escherichia coli. *Nature* **461,** 1243-7.

Baverstock, K. F. (1991) DNA instability, paternal irradiation and leukaemia in children around Sellafield. *Int J Radiat Biol* **60,** 581-95.

Baverstock, K. (2000) Radiation-induced genomic instability: a paradigm-breaking phenomenon and its relevance to environmentally induced cancer. *Mutation Research* **454,** 89-109.

Baverstock, K. (2010) Why do we need a new paradigm in radiobiology? *Mutat Res* **687,** 3-6.

Baverstock, K. (2011) A comparison of two cell regulatory models entailing high dimensional attractors representing phenotype. *Prog Biophys Mol Biol* **106,** 443-9.

Baverstock, K. and Rönkkö, M. (2008) Epigenetic regulation of the mammalian cell. *PloS One* **3,** e2290.

Bertalanffy, L. v. (1969) General system theory; foundations, development, applications, pp. xv, 289 p., G. Braziller, New York,.

Cech, T. R. (2011) The RNA Worlds in Context. *Cold Spring Harb Perspect Biol*.

De Duve, C. (1995) RNA without protein or protein without RNA? In *What is Life? the Next Fifty Years*, (eds. M. P. Murphy and L. A. J. O'Neill), pp. 79 - 82, Cambridge University Press, Cambridge.

Delbruck, M. (1949) translation of discussion following a paper by Sonneborn, T. M. and Beale, G.H. *Unites Biologiques Douees de Continute Genetique***,** 33 - 35.

Dyson, F. J. (1982) A model for the origin of life. *J Mol Evol* **18,** 344-50.

Dyson, F. J. (1999) Origins of life, Cambridge University Press, Cambridge [England] ; New York.

Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., Heine-Suner, D., Cigudosa, J. C., Urioste, M., Benitez, J., Boix-Chornet, M., Sanchez-Aguilera, A., Ling, C., Carlsson, E., Poulsen, P., Vaag, A., Stephan, Z., Spector, T. D., Wu, Y. Z., Plass, C. and Esteller, M. (2005) Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A* **102,** 10604-9.

Ghazalpour, A., Bennett, B., Petyuk, V. A., Orozco, L., Hagopian, R., Mungrue, I. N., Farber, C. R., Sinsheimer, J., Kang, H. M., Furlotte, N., Park, C. C., Wen, P. Z., Brewer, H., Weitz, K., Camp, D. G., 2nd, Pan, C., Yordanova, R., Neuhaus, I., Tilford, C., Siemers, N., Gargalovic, P., Eskin, E., Kirchgessner, T., Smith, D. J., Smith, R. D. and Lusis, A. J. (2011) Comparative analysis of proteome and transcriptome variation in mouse. *PLoS Genet* **7,** e1001393.

Glendinning, P. (1994) Stability, instability, and chaos : an introduction to the theory of nonlinear differential equations. In *Cambridge texts in applied mathematics*, pp. xiii, 388 p., Cambridge University Press, Cambridge [England] ; New York.

Head, J. J., Bloch, J. I., Hastings, A. K., Bourque, J. R., Cadena, E. A., Herrera, F. A., Polly, P. D. and Jaramillo, C. A. (2009) Giant boid snake from the Palaeocene neotropics reveals hotter past equatorial temperatures. *Nature* **457,** 715-7.

Horikawa, D. D., Sakashita, T., Katagiri, C., Watanabe, M., Kikawada, T., Nakahara, Y., Hamada, N., Wada, S., Funayama, T., Higashi, S., Kobayashi, Y., Okuda, T. and Kuwabara, M. (2006) Radiation tolerance in the tardigrade Milnesium tardigradum. *Int J Radiat Biol* **82,** 843-8.

Huang, S. (2009) Reprogramming cell fates: reconciling rarity with robustness. *Bioessays* **31,** 546-560.

Huang, S. (2012) The molecular and mathematical basis of Waddington's epigenetic landscape: a framework for post-Darwinian biology? *Bioessays* **34,** 149-57.

Irimia, M., Tena, J. J., Alexis, M., Fernandez-Minan, A., Maeso, I., Bogdanovic, O., de la Calle-Mustienes, E., Roy, S. W., Gomez-Skarmeta, J. L. and Fraser, H. B. (2012) Extensive conservation of ancient microsynteny across metazoans due to cis-regulatory constraints. *Genome Res*. **22**, 2356-2367

Jablonka, E. and Lamb, M. J. (2005) Evolution in Four Dimensions: Genetic, Epigenetic, Behavioural and Symbiotic Variation in the History of Life. In *Life and Mind: Philosophical Issues in Biology and Psychology*, (eds. K. Sterelny and R. A. Wilson), MIT, Cambridge, Massachusetts.

Kadhim, M. A., Macdonald, D. A., Goodhead, D. T., Lorimore, S. A., Marsden, S. J. and Wright, E. G. (1992) Transmission of chromosomal instability after plutonium alpha-particle irradiation. *Nature* **355,** 738-740.

Karotki, A. V. and Baverstock, K. (2012) What mechanisms/processes underlie radiation-induced genomic instability? *Cell Mol Life Sci* Online 6 September.

Kashiwagi, A., Urabe, I., Kaneko, K. and Yomo, T. (2006) Adaptive response of a gene network to environmental changes by fitness-induced attractor selection. *PLoS ONE* **1,** e49.

Kauffman, S. A. (1993) The Origins of Order: Self Organisation and Selection in Evolution, Oxford University Press, Oxford.

Lewontin, R. C. (1974) The genetic basis of evolutionary change. In *Columbia biological series no. 25*, pp. xiii, 346 p., Columbia University Press, New York,.

Martins, Z., Botta, O., Fogel, M. L., Sephton, M. A., Glavin, D. P., Watson, J. S., Dworkin, J. P., Schwartz, A. W. and Ehrenfreund, P. (2008) Extraterrestrial nucleobases in the Murchison meteorite. *Earth and Planetary Science Letters* **270,** 130 - 136.

Monod, J. and Jacob, F. (1961) Teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harb Symp Quant Biol* **26,** 389-401.

Nakajima, M., Imai, K., Ito, H., Nishiwaki, T., Murayama, Y., Iwasaki, H., Oyama, T. and Kondo, T. (2005) Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. *Science* **308,** 414-5.

Nicolis, G. and Prigogine, I. (1989) Exploring complexity: an introduction, pp. xi, 313 p., W.H. Freeman, New York.

Noble, D. (2012) A theory of biological relativity: no privileged level of causation. *Interface Focus* **doi:10.1098/rsfs.2011.0067.**

O'Neill, J. S. and Reddy, A. B. (2011) Circadian clocks in human red blood cells. *Nature* **469,** 498-503.

O'Neill, J. S., van Ooijen, G., Dixon, L. E., Troein, C., Corellou, F., Bouget, F. Y., Reddy, A. B. and Millar, A. J. (2011) Circadian rhythms persist without transcription in a eukaryote. *Nature* **469,** 554-8.

Peacocke, A. N. (1983) An Introduction to the Physical Chemistry of Biological Organisation, Oxford Science Publications, Oxford.

Pernu, T. K. and Annila, A. (2012) Natural Emergence. *Complexity* **17,** 44 - 47.

Ptitsyn, A. (2008) Stochastic resonance reveals "pilot light" expression in mammalian genes. *PLoS One* **3,** e1842.

Roberts, N. J., Vogelstein, J. T., Parmigiani, G., Kinzler, K. W., Vogelstein, B. and Velculescu, V. E. (2012) The predictive capacity of personal genome sequencing. *Sci Transl Med* **4,** 133ra58.

Rogakou, E. P., Boon, C., Redon, C. and Bonner, W. M. (1999) Megabase chromatin domains involved in DNA double-strand breaks in vivo. *J Cell Biol* **146,** 905-16.

Ronkko, M. (2007) An artificial ecosystem: emergent dynamics and lifelike properties. *Artif Life* **13,** 159-87.

Rosen, R. (1991) Life Itself: a Comprehensive Inquiry into the Nature, Origin and Fabrication of Life. In *Complexity in Ecological Systems*, Columbia University Press, New York.

Sankaranarayanan, K. and Nikjoo, H. (2011) Ionising radiation and genetic risks. XVI. A genome-based framework for risk estimation in the light of recent advances in genome research. *Int J Radiat Biol* **87,** 161-78.

Savage, V. M., Allen, A. P., Brown, J. H., Gillooly, J. F., Herman, A. B., Woodruff, W. H. and West, G. B. (2007) Scaling of number, size, and metabolic rate of cells with body size in mammals. *Proc Natl Acad Sci U S A* **104,** 4718-23.

Schneider, E. D. and Kay, J. J. (1995) Order from Disorder: the thermodynamics of complexity in biology. In *What is Life? the Next Fifty Years*, (eds. M. P. Murphy and L. A. J. O'Neill), pp. 161 - 174, Cambridge University Press, Cambridge.

Schrödinger, E. (1944) What is life?, Cambridge University Press, Cambridge.

Schwikowski, B., Uetz, P. and Fields, S. (2000) A network of protein-protein interactions in yeast. *Nat Biotechnol* **18,** 1257-61.

Shapiro, J. A. (2007) Bacteria are small but not stupid: cognition, natural genetic engineering and socio-bacteriology. *Stud Hist Philos Biol Biomed Sci* **38,** 807-19.

Shapiro, J. A. (2011) Evolution : a view from the 21st century, pp. xi, 253, FT Press Science, Upper Saddle River, N.J.

Sharma, V. and Annila, A. (2007) Natural process--natural selection. *Biophys Chem* **127,** 123-8.

Sharma, V., Kaile, V. R. I. and Annila, A. (2009) Protein folding as an evolutionary process. *Physica A* **388,** 851 - 862.

Stolovicki, E., Dror, T., Brenner, N. and Braun, E. (2006) Synthetic gene recruitment reveals adaptive reprogramming of gene regulation in yeast. *Genetics* **173,** 75-85.

Waddington, C. H. (1942) Canalisation of development and the inheritance of acquired characters. *Nature* **150,** 563 - ???

Waddington, C. H. (1961) Genetic assimilation. *Adv Genet* **10,** 257-93.

Watson, A., Mata, J., Bahler, J., Carr, A. and Humphrey, T. (2004) Global gene expression responses of fission yeast to ionizing radiation. *Mol Biol Cell* **15,** 851-60.

Yang, X., Xie, L., Li, Y. and Wei, C. (2009) More than 9,000,000 unique genes in human gut bacterial community: estimating gene numbers inside a human body. *PLoS One* **4,** e6074.

Yus, E., Maier, T., Michalodimitrakis, K., van Noort, V., Yamada, T., Chen, W. H., Wodke, J. A., Guell, M., Martinez, S., Bourgeois, R., Kuhner, S., Raineri, E., Letunic, I., Kalinina, O. V., Rode, M., Herrmann, R., Gutierrez-Gallego, R., Russell, R. B., Gavin, A. C., Bork, P. and Serrano, L. (2009) Impact of genome reduction on bacterial metabolism and its regulation. *Science* **326,** 1263-8.