Physically Defining Life: A Thermodynamic Systems Analysis of Biology

Dr. Stephen Palmer

ABSTRACT

This paper applies recent advances in analysing systems behaviour based on thermodynamics to investigating the physical basis of biology and so provide a physical definition of life. The definition of life provided also provides a physical basis for Dawkins's selfish gene theory. The systems analysis of prokaryote evolution and development herein also describes an example of Dawkins's extended phenotype emerging during prokaryote evolution. This paper also identifies the biophysical basis for the emergence of ecological 'r' and 'K' reproductive strategies in prokaryotes. Finally, part of this analytical approach is applied to recent work on the origin of life and used to describe its alignment with a particular hypothesis for the origin of life on Earth.

Keywords: thermodynamics, systems analysis, definition of life, extended phenotype.

Classification: LCC Codes: QH331, QH324.2, QH501

Language: English`

Great Britain Journals Press

LJP Copyright ID: 925614
Print ISSN: 2631-8490
Online ISSN: 2631-8504

London Journal of Research in Science: Natural and Formal

Volume 23 | Issue 10 | Compilation 1.0

Physically Defining Life: A Thermodynamic Systems Analysis of Biology

Dr. Stephen Palmer

ABSTRACT

This paper applies recent advances in analysing systems behaviour based on thermodynamics to investigating the physical basis of biology and so provide a physical definition of life. The definition of life provided also provides a physical basis for Dawkins's selfish gene theory. The systems analysis of prokaryote evolution and development herein also describes an example of Dawkins's extended phenotype emerging during prokaryote evolution. This paper also identifies the biophysical basis for the emergence of ecological 'r' and 'K' reproductive strategies in prokaryotes. Finally, part of this analytical approach is applied to recent work on the origin of life and used to describe its alignment with a particular hypothesis for the origin of life on Earth.

Keywords: thermodynamics, systems analysis, definition of life, extended phenotype.

Author: Stantec Ltd., Stantec House, Kilburn Court, Birchwood, Warrington, U.K.

I. INTRODUCTION

Recent advances in theoretical quantum thermodynamics have enabled new insights into how quantum changes can emerge into physical reality via decoherence (Palmer, 2019). They have also clarified what entropy is.

At the quantum level, a change of microstate requires energy redistribution between an object quantum system and those quantum systems constituting its environment.

The physical elements of quantum microstates follow information systems behavior as described by the statistical mechanics of Shannon information theory:

$$H = - \Sigma p_i \log p_i \quad (1)$$

(where H = Shannon measure of information (the missing information equaling uncertainty) and p_i is the probability of any given microstate)

When the energy distribution within the microstates of a simple closed, isolated physical system is described, the Shannon information description of the system then re-emerges in the form:

$$S = k_B \ln \Omega \quad (2)$$

(where S = system entropy (units of energy /temperature E/T), k_B = Boltzmann constant (energy/temperature), Ω = number of microstates possible for macroscopic constraints (i.e. system total energy E).

This is Boltzmann's definition of entropy, significantly predating Shannon's information theory. The relationship between these two descriptions is a consequence of physical quantum microstate information theory's statistical mechanics for energy distribution within the system. Temperature is a metric of the energy distribution within the system in which the entropy is the uncertainty of how energy is distributed within that isolated system.

II. CLARIFYING AND UNDERSTANDING ENTROPY

The significance of these recent advances is that they allow clarification of what entropy is and importantly, what it is not.

Entropy at the microstate level is simply missing information (uncertainty) in respect of the distribution of the agent of physical microstate change, energy.

There is no element of order or disorder implied by entropy. The association of entropy with order and disorder is a common mistake still regularly seen in scientific publications, but it is entirely erroneous (Ben Naim, 2008; Palmer 2022) leading to misinterpretations of the consequences of entropic system behavior.

The quantum level entropy as microstate uncertainty emerges at the classical level as uncertainty about energy distribution in the form of *energy dissipation*, during 'change' in a classical system.

For systems to change they need to consume free energy (usually termed 'exergy') to carry out work. The physical information of a system s is described by:

$$H = B/T_o \tag{3}$$

(where H = Shannon measure of information (=uncertainty = missing information), B = free energy (exergy) of the system and T_o is the temperature of the systems environment)

Consequently, the physical state and physical information regarding the state of a system is related to how free energy (exergy) is distributed in the system compared to the energy distribution (temperature) of the systems' environment.

This is a key concept in thermodynamic systems behavioural analysis - the physical information defining a system is that which distinguishes that system from its environment. Physical system information is relative; it is the information that distinguishes the system from its background (*its physical environment*).

When we analyse the behaviour of biological systems we are dealing with *open, interacting systems which are also chemical systems*:

$$H = E + P_o V - \Sigma\rho_{io} N_i / T_o \tag{4}$$

(where the system uncertainty H is related to the distribution of molecules in an environment with a number of moles N_i, chemical potential $\Sigma\rho_{io}$, distributed in volume V at pressure P_o and temperature T_o, with energy E)

The thermodynamic information for a system change I is defined by the loss of information that occurred when the system changed to state S_o in which it is indistinguishable from its environment (Tribus and McIrvine, 1971)

$$I = S_o - S \tag{5}$$

The interactions producing microstate changes in an object-system are stochastic (probabilistic). The physical information of a particular change of state in a physical system is a record of its last interactions between the object-system and the systems that populate its environment.

The conditional probability distribution for future states of the interacting systems depends on the present state of the systems. Physical and chemical systems follow *Markovian* behavior wherein information about past interactions between a system and its environment is *dissipated*.

III. DEFINING LIFE IN PHYSICAL SYSTEM TERMS

Life and living systems are typically defined in terms of their observed common behaviors across living systems such as growth, reaction to stimuli, metabolism, energy transformation and reproduction.

Living systems are capable of adaptation and evolution through successive generations, which arises from 'imperfect' system replication. Biological systems are distinguishable from physical and chemical systems by their attribute of memory. For life on Earth, the period for which the agent of memory is the gene in the form of nucleic acids (DNA and RNA), extends at least as far back as the Last Universal Common Ancestor of prokaryotes (LUCA) and probably beyond.

System memory for a physical system arises from a biological system being able to replicate itself within its environment subject to the pressures the environment exerts on the systems success in replication, so a 'system' with memory could first emerge as a self-replicating physical structure or as a self-replicating chemistry.

On a physical system behavioural basis, a universal system definition for biological systems and life is:

'A system (chemically based in known biological systems) with memory which is utilized for environmental fitness, with fitness defined by success in replicating the system in its environment'.'

This thermodynamically referenced definition of life emphasizes that the difference between biological systems and natural physical and chemical systems is biological systems' attribute of memory. The key distinction between life and physical or chemical systems is that biological systems are *non-Markovian* in their system behaviour, due to their replication creating a memory which is transferable between generations, which allows them to adapt to their environment. The attribute of memory makes biological systems *learning systems* in physical information terms. This paper focuses on the biophysical aspects of prokaryotes as an example of the foundational physical systems aspects of biology because prokaryotes are the first emergent biotype for life on Earth and because there is far less distance in prokaryotes between genotype and phenotype and form and function.

The information content of a system obviously describes its physical state and any *change* of state in a physical system is accompanied by a change in the system information. A closed system that does not change is at thermodynamic equilibrium and is *indistinguishable* from itself at a previous point in time.

For open interacting systems such as biological systems, a system at equilibrium is indistinguishable from its background. For such systems where the physical information of the system is different to its background when the system is not at equilibrium with its background.

The profound significance of the relationship $H = B/T_o$ (equation 3) is that a physical system's information is the *difference* in information between the physical system being observed and its environment- the information that *distinguishes* it from its environment.

Physical and chemical systems compete for exergy (free energy) in a given environment to be able to change state but the interaction being observed is Markovian. There is no information carried over to the next interaction for the interacting physical or chemical system. *Markovian behaviour dissipates information.*

In contrast, the defining memory attribute of biological systems allows information on the systems interactions with its background (environment) to be passed down generations of the self-replicating system. There is a second critical system behavioural characteristic being exhibited here: system

self-replication creates the basis for memory, but memory itself is essential for a system to be able to *utilise* information.

The successive (generational) development of information in an observed biological system also has profound implications for how physical structural change can accrue over time. The mathematical behavior of simple systems under persistent, successive positive feedback (in this case feedback on success of replication arising from fitness for replication in the environment) is one that leads to emergence of complex systems (Holland, 1992, Holland 1996, Holland 1998).

This particular characteristic is a key consideration in analysis of emergence of life on Earth (or elsewhere, as these are universal physical principles). System self-replication is the basis for memory, so the earliest emergence of life must be associated with some form of self-replication of physical -chemical systems.

Another key attribute required for a successful self-replicating chemistry to exhibit utilization of information of the information its stores is for the self-replicating system to be incrementally modifiable (e.g. the system needs to have a means for variation to arise its memory). In systems terms, this could be provided by generic mechanisms such as errors in the systems' self-replication or from exchange of information between systems.

These are the physical systems behavioral constraints which the classical biological elements of genotype and phenotype have to act within. A physical structural type of biological system has its genotype as its memory and the agent for system learning fitness to replicate in different environments is the gene (Dawkins, 1976).

The definition of life provided in this paper implies that there is only one physical systems agency for feedback on object- system replicative success. It is the observed systems environment (which is itself a population of other physical, chemical plus biological systems, after the emergence of life).

Physical systems' analysis provides the physical basis for the biologically universal utility of Dawkins' selfish gene' concept. Biological systems are not possible without memory and the success of a gene is simply defined by its ability to self-replicate. On a physical basis, each gene represents an agent of memory competing for the resources needed to optimize its generational self-replication. A physical system approach also allows us to clarify *phenotype* in physical terms. The phenotype is simply the physical structural vehicle for the system memory (gene's) intergenerational replication.

The physical structure of that vehicle (organism structure) encodes its history of actionable information relative to the environment, up to the point of historic structures still retaining present-day utility. If a biological system has memory capacity constraints, it is the environment (*via* its resources for replication and competition for them) that will dictate how that memory is most efficiency utilized to secure replicative fitness (i.e. determine whether genes are retained or lost from the overall available capacity).

This definition of biological systems implies that adaptation to the environment defined by attaining replicative success is the principal system feedback confirming biology. It also implies that a biological system does not to inherently need to metabolize itself: if a biological system can utilize (parasitize) a biological host-system to carry out the work of replication then it needs to sustain and seek far less resources from the environment (but is then of course dependent on host availability in its environment). This paper will explore how critical management of resources to replicative success within biology and how deeply those strategies and tactics for it are conserved, given that they emerge and are generic in prokaryotes.

From the definition of life provided by this paper, viruses are definitively biological systems. My previous reference to parasitism, even applied to a viral biological system strategy, does not imply that only the virus gains. Deeply interactive biological system relationships often have some degree of return to both participants and it should be noted that a virus-host cell interaction also introduces a basis for genetic recombination (horizontal gene transfer), which represents information transfer and acquisition for the host organism.

From consideration of the implications of physical information and how redistribution of energy is the agency for physical state change, we can now also consider Dawkins's 'selfish gene' concept (Dawkins, 1976)and how it aligns with a physical definition of life.

At whatever point early in the emergence of life where a specific form of chemical memory emerges, it will dominate information retention in a biological system and intergenerational information utilization. *From a system viewpoint, the gene is the agent of intergenerational memory of interactions with the environment and the organism is a vehicle for it.*

However, a gene being selfish in terms of its purpose of self-replication should not be confused with the system behaviour needing to be selfish. The purpose of adapting system behavior in relation to the environment is replicative success and the growth strategies and resource acquisition tactics needed to meet that goal will be selfish under some environmental conditions but cooperative under others.

A genome will evolve symbiotic behaviour under generations spent in certain environmental conditions including resource limitation and variability in resources and competition for resources, as will be demonstrated below in a case study of prokaryote evolution.

The critical difference between biological systems and physical or chemical systems is biological systems' capability for utilization of information. Biological systems are learning systems, whose reference point for distinguishing survival-information from survival-misinformation (noise) is its *environment* (Palmer, 2019).

The biological cell and the non-equilibrium thermodynamics of its growth.

Energy redistribution including consumption of free energy is required for systems to change state. The relationship between form and function for successful information (gene) replication of biological systems is a continuous, dynamic process. It is further complicated by the fact that biological systems in turn affect their environment while maximising their potential to replicate within it. The process requires a continuous energy flux through a biological system to maintain an individual system (organism) until it has successfully replicated, in order to propagate the genome through successive physical *vehicle* (phenotype) generations (Dawkins, 1982). On a simple systems basis, once the genome is successfully propagated the individual phenotype (organism) itself is redundant and the individual phenotype can fall back into equilibrium with its environment (its death) without jeopardizing the success of the genotype in replicative terms.

In system terms, the utility of the information system also needs to provide variation in information content in order to provide a basis for creating new genotypic variants relative to its environment. The sources of variation in the information system we are familiar with in the emergent phenotypic prototypes for life on Earth (i.e. LUCA and the prokaryotes) are; mutation, gene loss, error and even recombination via Horizontal Gene Transfer (HGT).

A biological system's most universal structural unit form is the cell. When LUCA emerges in terrestrial evolution, genetic code is provided by DNA and RNA. LUCA's cell structure consisted of a membrane and wall, with the structures for energy generation being part of the membrane (Martin & Russell,

2003; Lane, 2015). Energy generation for biosynthetic activity supporting anabolism (including growth and genome reproduction) is provided by catabolism.

Catabolism is coupled to anabolism to drive anabolism and growth (Fig. 1).

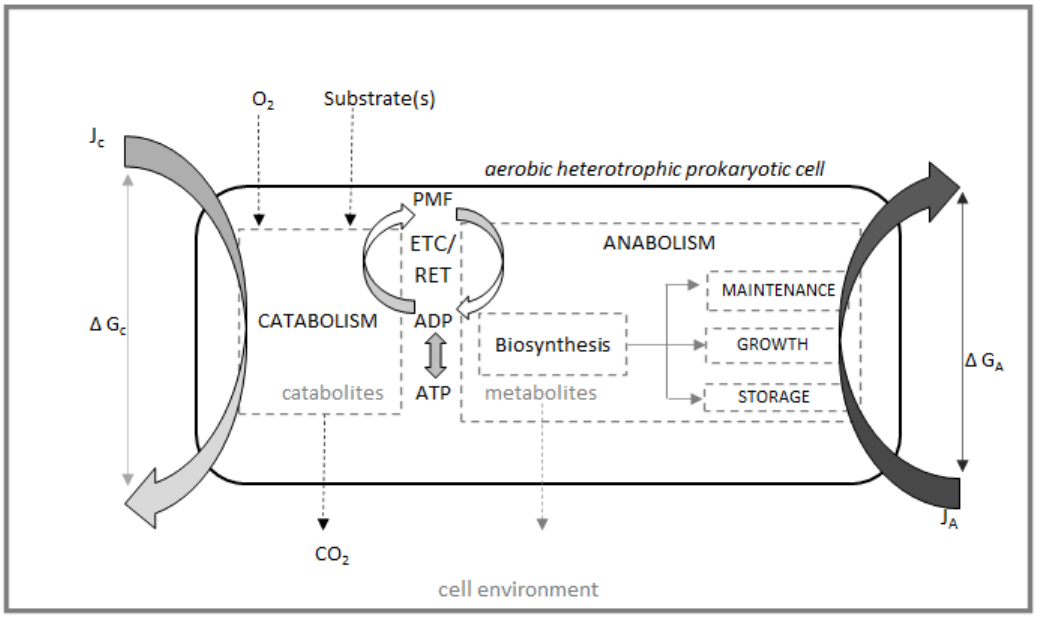

Figure 1: Systems analysis diagram for a prokaryotic cell

A prokaryotic cell system diagram for an aerobic heterotrophic prokaryote. The cell can sustain itself far from thermodynamic equilibrium by coupling the energy generated from catabolism to anabolic work (growth, cell maintenance, substrate storage). This coupling of catabolism to anabolism sustains a driving force J_A for anabolic activity from the energy sink of force J_c. Force J_c is maintained by the exergy provided by Gibbs energy released by catabolism (ΔGc) which in turn creates a driving force J_A for cell reproduction (growth). The origin of force J_c counterbalancing force J_A is the system entropy balance and the system attempting to return to equilibrium with its environment. Two fundamental reproductive (growth) strategies then arise from thermodynamic constraints: growth at maximum reproductive rate, in which anabolic energy and resources are diverted mainly to fast reproduction of a basic viable cell, or growth based on reproducing more resilient cells with greater longevity potential based on a higher proportion of anabolic energy into maintenance, regulation and storage in which cell mass (yield) is higher. These two reproductive strategies can be identified with the ecological theory of 'r' and 'K' strategies for reproduction (Reznick *et al,* 2002) with 'r' strategy for maximum rate of cell reproduction being competitive when resources are not limited and the 'K' strategy of production of higher mass, more resilient, cells.

The entropy balance for reacting chemicals *j* in a mixed flow through environment at constant temperature and pressure is:

$$S_{prod} = \Sigma j\ -\Delta r_j G/T\ \xi j + W/T + \Sigma j\ (\mu i_{in}/T - {}_i\ \mu i_{out}/T)\ \ nj,_{in} \qquad (6)$$

(where S is the entropy produced from the Gibbs reaction energy $-\Delta rjG$, W work done, 'T' is temperature, μi the chemical potential of the 'i'th component, ξj = rate of the jth chemical reaction and $n_{j\ in}$=- the influx flow) (von Stockar, 2013)

For two flux forces exerted across a (prokaryotic) cell both are proportional to their conjugate force 'Zi'**:**

$$J_i = L_i . Z_i \quad (7)$$

(where L*i* is a constant for the flux) (Von Stockar, 2013)

For the flux forces exerted across a (prokaryotic) cell between a bioenergetic catabolic reaction coupled to an anabolic process, assuming the flux for catabolism is designated J_c and that for anabolism designated J_A, the anabolic process can be driven against its driving force by it being *coupled* to the catabolic reaction as shown in Fig.1 below.

Growth-coupling is not fully complete *i.e.* 100% between catabolism and growth-anabolism because the cell also uses other biosynthetic processes in addition to cell replication (growth) to manage entropy (cell maintenance processes) and manage resources (manage starvation risk) which is described in the Herbert -Pirt equation:

$$1/Y = 1/Y_{max} + m/\mu \text{ (Von Stockar, 2013)} \quad (8)$$

(where Y = biomass yield from catabolism, Y_{max} = maximum biomass yield from catabolism, μ is the growth rate and m is the energy used in cell maintenance processes)

Open system thermodynamics for a cell as illustrated in Fig 1 allow derivation of a cell replication (growth) rate in relation to the rate of catabolic substrate consumption:

$$r = L\, RT \ln c + C \text{ (Von Stockar, 2013)} \quad (9)$$

(where r = cell growth rate, L = coupling coefficient for catabolism and anabolism, C = catabolic substrate concentration, c = catabolic substrate saturation concentration, R = universal gas constant, T = temperature)

Equation (9) provides a relationship between growth and catabolic substrate removal kinetics which closely fits that of Monod growth kinetics;

$$\text{Monod kinetic relationship } \mu = \mu_{max}(S/(K_s + S)) \quad (10)$$

for which Substrate removal = $- r_s = \mu(X/Y_s) . (S/(K_s + S))$

(where -r = catabolic substrate removal rate , μ = specific growth, μ_{max} = maximum growth rate, X = cell biomass, Y = biomass yield, S = catabolic substrate concentration, K_s = substrate half-saturation coefficient)

The Monod relationship arose from Monod's observation (Monod, 1950) that in the exponential growth phase, prokaryote biomass formation increased in proportion to substrate consumption. The growth yield 'Y_s' is defined by the catabolic substrate electron consumption per amount of biomass produced. This varies for different substrates with their associated Gibbs free energy and the efficiency of energy transfer directly into cell growth, balanced by any demands for cell maintenance (Equation 8). Monod's work illustrated the function of the prokaryotic gene in providing the feedback between the environment and prokaryote with gene regulation then giving the prokaryote options from the genome in its response to changing environmental conditions.

Evolutionary feedback on cell bioenergetic constraints

Prokaryote bioenergetic outcomes are a balance between irreversible thermodynamic constraints, resource availability and variation in the environment, all of which in turn shapes the genome through natural selection. A high degree of physical structural complexity emerges from a self-replicating information system learning to adapt to its environment even at the level of the prokaryotic cell. Even

this requires its living systems to maintain themselves continuously at a significant distance from equilibrium from their environment while alive.

When there is little or no resource limitation, including competition for substrates and resources, if the cell and its genome can sense this is the state of its environment it can express a phenotype focused on the fastest possible reproduction of the prokaryotic cell. Alternatively, if sensing determines that the ratio of resources to cells for the species is below a certain threshold value, anabolic activity and reproduction can more broadly disperse energy and resources between cell reproduction and increased cell resilience from more investment of resources into maintenance and storage (Bachmann *et al*, 2016).

Figure 2 below illustrates these reproductive alternatives. The sensing system that provides the shift between maximum growth rate (termed 'r' strategy in ecological 'r' and 'K' theory) and more resilient cells (K strategy; higher biomass yield) is *quorum sensing*.

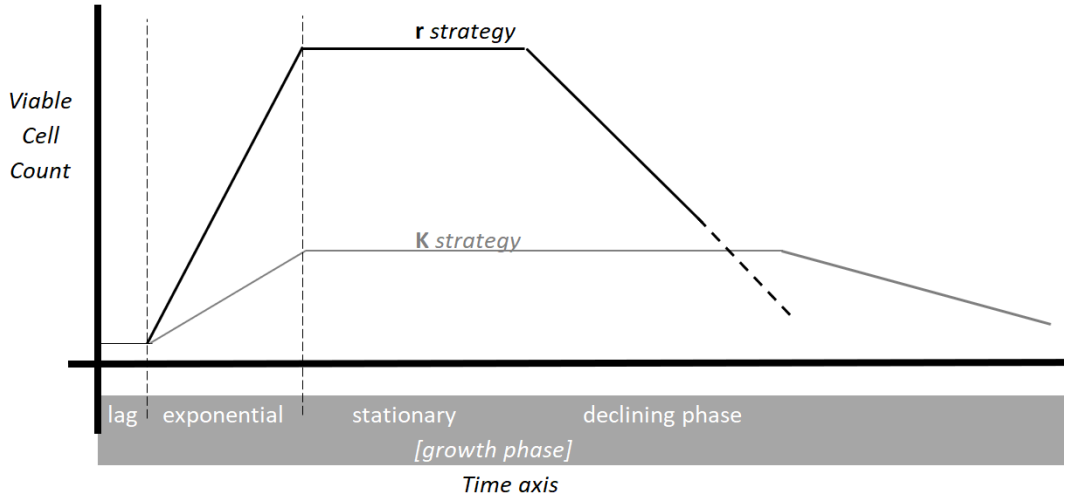

Figure 2: Prokaryote principal growth (reproduction) strategies.

Prokaryotes can regulate their cells to grow under two different reproductive strategies which arise from prokaryote cell biomechanical constraints. Under totally unlimited growth conditions (such as unlimited substrate chemostat experiments in which prokaryotes will grow as single cells), prokaryote growth can follow a strategy of maximum rate of minimum viable cell reproduction in order to maximize its numbers and outcompete other species for space. This **'r'** strategy supports maximum consumption of resources and maximum rate of individual cell reproduction (equivalent to Krefts 'ego' strategy (Kreft, 2004)). Prokaryotes can also use *quorum sensing*, an assessment of cell to resource density, to switch to a more *resilient phenotype* in which the cell has increased energy and resource input into maintenance, regulation and storage, which produces cells with higher biomass yield. This 'K' strategy produces more resilient cells providing the prokaryote population (species) with a higher probability of surviving for longer.

However, when resources become limited, the optimum survival strategy shifts to maximising the longevity of the population of the phenotype, making more efficient use of resources to conserve them for longer. The longevity of the population of that phenotype, optimizes yield towards that outcome via the 'K' strategy, where catabolic Gibbs free energy can be redistributed across work on maintenance and storage as well as cell growth (reproduction).

$$\Delta_r G_s = (1\text{-}Y) \,.\, \Delta_r G_{cat} + Y \,.\, \Delta_r G_{an} \text{ (Von Stockar, 2013)} \qquad (11)$$

(where $\Delta_r G_s$ = catabolic substrate Gibbs energy, Y = biomass yield from anabolic process growth coupled to catabolism, $\Delta_r G_{cat}$ = catabolic energy, $\Delta_r G_{an}$ = anabolic energy).

The Gibbs free energy driving cell activity has both an enthalpic and entropic component with the balance differing according to the form of catabolism or energy input i.e. organic carbon aerobic or anaerobic catabolism, autotrophic catabolism, fermentation or photosynthesis):

$$\Delta_r G = \Delta_r H - T\Delta_r . S \text{ (Von Stockar, 2013)} \quad (12)$$

(where $\Delta_r G$ = catabolic Gibbs free energy, $\Delta_r H$ = system enthalpy part of entropy (heat energy dissipated to environment) and $T\Delta_r . S$ represents system entropy exported to the environment as high entropy metabolites (e.g. CO_2 in example given in Fig 1, *etc*)

The outcomes shown in Fig.2 arise from both high bioenergetic thermodynamic driving force and high metabolic rates in maximising the rate of cell replication in the 'r' strategy. However, in a resource limited this high level of resource dissipation per individual cell can lead to competitors with a lower metabolic rate (K strategy) and associated higher resource efficiency persisting for longer as a species.

This means that almost all prokaryotes' $\Delta_r G$ runs in a zone that represents a *balance* between biomass yield efficiency and individual cell replication rate (von Stockar, 2013). Optimizing genome population survival under resource limitation can lead to other tactics in addition to selection of a Y(ield)-based growth strategy (de Mattos &Neijissel, 1997; Wang & Post, 2012, Gonzalez *et al*, 2009).

Prokaryotic cells need not fully couple catabolism to reproductive anabolism, as the reproduction K strategy illustrates. Under a K strategy a prokaryotes' anabolic activity can include more investment of energy and resources in cell maintenance, regulation and storage which results in population longevity being more probable.

Extreme examples of the latter diversion of energy into maintenance and storage that support a 'K' reproductive strategy are now known, in which there is no anabolic energy passed into reproduction. For example, some prokaryotes from deep Earth environments which have very few resources available in their environment have *minimal* metabolism such that even single cells are very long-lived (Bradley *et al*, 2018). Such extreme-environment prokaryotes are likely to represent evolution from a 'K' growth strategy which has geared cell metabolism towards minimal maintenance until the environment again becomes rich in catabolic substrates.

Further bioenergetic biomechanical limitations affecting replication rate arise from the efficiency of energy transfer through the cell's Electron Transport Chain (ETC)and Reverse Electron Transport (RET) where a prokaryote metabolism requires RET. The length of the cell ETC contributes to its inefficiency in energy transfer as described by Heijnen (2013:

$$\text{max. ETC} = 3 \exp [(-69000/R)(1/T - 1/298K)] \quad (13)$$

(where R = universal gas constant, T = temperature (K))

The maximum electron flow rate (mol e/hour) along an ETC is a function of the catabolic substrate 'C' consumption in reaction with its electron donor or acceptor. This leads to a relationship to maximum growth rate in relation to ETC efficiency:

$$\mu_{max} = 3 ((\Delta G_{cat} / y_D^- - m_G)/a_G) \exp[(-69000/R)(1/T - 1/298K)] \quad (14)$$

(where μ_{max} = prokaryote maximum growth rate, y_D = electron supply from catabolic substrate which releases ΔG_{cat} Gibbs free energy from the prokaryote catabolism (per mol electron donor), m_G = Gibbs energy diverted to maintenance a_G = Gibbs energy required to biosynthesize 1 mol Carbon, R = universal gas constant, T = temperature (K)) (Heijnen, 2013)

The bioenergetic returns from catabolism are a function of the carbon source (as in Fig 1) or the energy source used and its electron donor for the catabolic reactions.

The cost of anabolic biosynthesis 'a_G' is affected by the growth conditions (i.e. anaerobic versus aerobic) and the electron donor. Where the electron donor results in an anabolic Gibbs energy demand 'a_G' which is significantly more than zero, such as in autotrophic carbon growth (autotrophic methanogenesis), there is a need for reversed electron transport (RET) along the ETC.

The 'biomechanics' of catabolism bioenergetics also include efficiency considerations related to the length of the catabolic pathway.

For extended pathways, kinetic inefficiencies can arise from leakage or other inefficiencies including operating at elevated concentrations, many enzymatic steps, and their energy losses.

Kinetic theory of catabolism (Costa *et al*, 2006) assumes that natural selection on energy returns from catabolism is driven towards maximal ATP generation, which is favored by optimum catabolic substrate conversion and minimum path length. Costa *et al* (2006) show how the limitations from catabolism in nitrification arising from kinetic theory of optimal pathway length. This explains why prokaryote nitrification has evolved into a two-stage process (with ammonia oxidising prokaryotes feeding nitrite oxidising prokaryotes), due to the shortened catabolic pathways of the cross-feeding prokaryotes offering ATP path efficiency returns on a basis of economic division of labour.

The most significant biophysical structural limitation for the prokaryotic cell arises from how its bioenergetics are integrated with the cell membrane (see Fig 1).

This creates a limit to the bioenergetic capacity of a prokaryotic cell due to the limits imposed by the surface area of the cell membrane. This bioenergetic limitation has resulted in practical limits on the genome size and hence information capacity of prokaryotes.

In comparison, eukaryotes whose possession of mitochondria with their folded membranes have a much higher surface area for their bioenergetics and hence can sustain a larger genome. This emerged as the evolutionary 'solution' to restricted bioenergetic capacity and restricted information capacity in prokaryotes (Lane and Martin, 2007; Lane and Martin, 2012, Lane, 2015).

Lane and Martin's hypothesis has been questioned notably by Lynch and Marinov (2015,2016,2017), with Lynch subscribing to the view that natural selection is overemphasized and is_not the main evolutionary driver of emergent physical structural complexity in biology. He instead, argues it arises from genetic drift and mutation (Lynch, 2007). Schavemaker and Muñoz-Gómez (2022) reviewed Lynch and Marinov's data in the context of cell form and function,) reporting that it supports Martin and Lane's hypothesis in the case of larger cells.

This is of course where the crux of the information -resource debate rests.

There are bioenergetic constraints for prokaryotic cells below a certain volume to surface area and for faster growing (replicating) prokaryotic cells at larger genome sizes (Schavemaker &Muñoz-Gómez, 2022). Schavemaker and Muñoz-Gómez conclude that larger eukaryote genomes really need the bioenergetic capacity offered by mitochondria (Schavemaker &Muñoz-Gómez, 2022).

This paper also challenges Lynch's view in its presentation of biophysical systems examples of how prokaryote natural selection is driven by competition for resources and space; against biophysical limitations on form and function in the context of survival challenges arising from the environment.

With reference to the case study provided below, if there were such a weak linkage between natural selection of prokaryotes and their environmental adaptation as Lynch implies, the prokaryote extended phenotype case study would not be expected to report such a strong linkage between physical form and function in respect of successful replication and persistence within the case study environment alongside the spread and conservation of those shared traits amongst diverse prokaryote phenotypes within that prokaryote community.

Resource efficiency tactics are coded into many prokaryote genomes and the evolution of resource efficiency has also been a driver for the emergence of *increased information utilization efficiency* in the prokaryote genotype and phenotype.

Information system evidence for evolutionary feedback into development of information utilization efficiency

Prokaryote cell genome size is ultimately limited by prokaryote physical cell form and function constraints. Information utilization is a process towards a goal, which for life is defined as *system replication* (as per definition above). On that basis, regulation of gene expression represents a critical development in the information utilization efficiency of the prokaryote genotype and phenotype in meeting the challenges to replication set by any given environment.

Thermodynamic constraints lead cellular life to reproduce on the basis of fastest rate of individual cell reproduction ('r' reproduction strategy) if the environment is unlimited in resources, or under resource limitation conditions, reproduction on the basis of optimized yield and optimized resource efficiency will tend to be preferred. Thermodynamic constraints can set the threshold for whether a species is likely to establish itself in an environment in which competition for resources exists (Seto, 2014).

The prokaryote genome is structured and regulated on a basis that is inherently resource efficient and information efficient (*e.g.* Struhl, 1999). Metabolic pathway genes are grouped in a cluster – the operon- which has common regulatory control such that gene expression is not just synchronized but rapid. Some environmental stresses induce complex regulation via gene cascades. The genes that encode environment relevant information related to catabolism, metabolism and reproduction are inducible, meaning that presence or absence of a chemical species in the environment leads to their activation.

Within prokaryote metabolism, information system efficiency is increased by catabolism-induced genes being inducible and hence being switched on when the encoded catabolic substrate is present in the cytoplasm and environment. In contrast and further maximizing resource efficiency, anabolic genes are repressed while any anabolic product is still present in the cell cytoplasm so conserving resources. If more than one catabolic substrate is encoded for and is present in the environment and cytoplasm, then the prokaryote genome typically selects the highest survival-potential substrate first (*e.g* E. coli and the LAC operon (Jacob &Monod, 1962)).

Prokaryote transcription level regulation allows for rapid response to critical environmental change and its systems logic 'reads' like a best practice manual for readily and resource efficiently managing risks to reproduction of the prokaryote genotype in any environment.

The sophistication and complexity of how the survival and the reproduction relevant information maintained in prokaryote genomes is managed and utilized is high, as would be expected from a physical-information system point of view for over 3 billion years of prokaryote evolution stochastically exploring different environmental challenges.

The structural sophistication of prokaryotes is ultimately limited by the upper genome size a prokaryotic cell can reproduce. That limitation is likely to have been the driver for 'learning -system exploration' of the catabolic opportunity space – and may have been responsible for the emergence of huge diversity in catabolism in prokaryotes.

Competition for space to reproduce into, translated into material resource use innovation, also leads to niche environment exploitation. Diversity in catabolic substrate and its reactants allows new environmental niches to be explored. Niche environment colonization can itself be internally exploited based on an 'r' reproductive strategy or a 'K' reproductive strategy depending on the bioenergetic returns possible in the 'niche' environment.

Prokaryotes reproduce vegetatively, i.e. daughter cells are clones of parent cells. A key factor in *information utility* for a replicator is a source of information variation to allow system development innovation. Variation in information for prokaryotes arises from mutation and genetic drift (including gene loss) as well as through recombination. Metabolic gene loss mutants are favoured by adaptative fitness benefits when the environment contains the require metabolites (De Souza & Kost, 2016).

Within prokaryotes, the emergence of recombination via various forms of Horizontal Gene Transfer HGT) has allowed for significant information transfer between different prokaryote phenotypes. HGT as a process puts information at risk as transfer operations include viral transduction and conjugation to transfer plasmids where the agents of recombination would be expected to have their own self-interest in being replicated. Negative outcomes range from acquisition of unharmful genetic parasites (which in a prokaryote genome with a limited information capacity represents loss of valuable information space) to acquisition of genes that harm replicative success.

Despite this, HGT is widespread in prokaryotes and some prokaryotes have mechanisms promoting HGT and HGT can influence cooperation and conflict between prokaryote phenotypes (Hall *et al*, 2020) . It seems likely that as a prokaryote habitat becomes more challenging, stress on system replication also increases the value for information system variation within the prokaryote phenotype, notably if the existing genome is not equipped to deal with the emergent stressors.

Resource challenges, from substrate availability and its variation to securing living space, have led to the emergence of significant diversity in catabolic substrate exploitation in prokaryotes, but prokaryote sophistication extends much further than that.

Competition in securing resources from information utilization efficiency is an evolutionary driver for the emergence of sensing in prokaryotes. The significance of the emergence of prokaryote sensing is that *it creates a basis for improved information utilization within an individual cell/individual phenotypes life cycle.*

With sensing coupled to regulation, adaptation is not just occurring between generations but within an individual prokaryote phenotypes' life cycle.

Quorum sensing fulfills a prokaryote resource management purpose in relation to attempting to maximize replicative success through utilization of environmental information (Kreft, 2004; Costa *et al*, 2006, Hense *et al*, 2007). When resources become limited or variable, an anabolic strategy that

maximizes *resilience of the species* is required to maximize prokaryote genotype survival, where symbiotic resource tactics as well as tactics for resource competition play a role.

The limitations on prokaryotes genome size have restricted the physical structural complexity that could be developed within prokaryotic cells. This may be the reason why 'division of labour resource economy 'structures within the prokaryotic cell such as that provided by organelles occurred until the emergence of eukaryotic cells with mitochondria from an archaeal/eubacterial symbiosis.

Physical limitations of the prokaryotic cell phenotype have confined prokaryotic multicellular growth for a single species of prokaryote to one form, 'filamentous' growth, in which linear, linked vegetative reproduction forms a filament.

However, there are also examples in nature of prokaryote cooperative and competitive multicellular growth for mixed species growth, in a prokaryotic example of a common extended phenotype that has emerged in aqueous environments.

Case study: the wastewater prokaryote extended phenotype

An extended phenotype is defined by Dawkins as the wider effect the phenotype of an organism or species can have on its environment, through interactions with its environment that extend beyond the direct mechanical capabilities of its physical phenotype.

Two general forms of extended phenotype are thought possible (Hunter, 2018):

A species physically conforms its environment to favour its survival and reproduction. The classic example used is the beaver and the beaver pond, which other organisms may benefit from, but the beaver's reason for reforming part of its environment to its own specification is based on reproductive self-interest,

Two organisms interact in a relationship where one locally manipulates the behaviour of the other such as a parasite-host interaction; or a subset of individual phenotype interaction where the two organisms influence each other at a distance.

In this wastewater case study, we will examine a prokaryote version of the first 'beaver pond' form of extended phenotype that emerged in prokaryotes in an aquatic type of environment.

The municipal wastewater environment is typically resource limited and rather variable in substrate concentration terms as well as being a flow -through environment. This leads to two predominating natural selection pressures for municipal wastewater prokaryotes: starvation and washout.

Starvation risk management is the most likely purpose for the evolution of quorum sensing in the wastewater prokaryote extended phenotype.

The environment varies in its substrate concentrations, so the heterotrophic bacteria typically associated with the habitat have developed significant catabolic range and diversity in their phenotype and are able to catabolize dissolved organic carbon aerobically and anoxically when oxygen is limited.

This includes an ability to rapidly uptake the most bioenergetically advantageous organic carbon substrates and convert some to intracellular food reserves.

When the primary catabolic substrates are absent, this phenotype includes a capability for shifting resources to production of hydrolytic enzymes to break down particulate volatile solids near the cell. After those resources are exhausted, this heterotrophic phenotype can then shift anabolic activity to maintenance and endogenous respiration.

This diversity in the catabolic phenotype optimizes survival to replication probability in this environment. This cycle of catabolic resource management is seen globally in aerobic municipal wastewater treatment systems daily. (Fig 3).

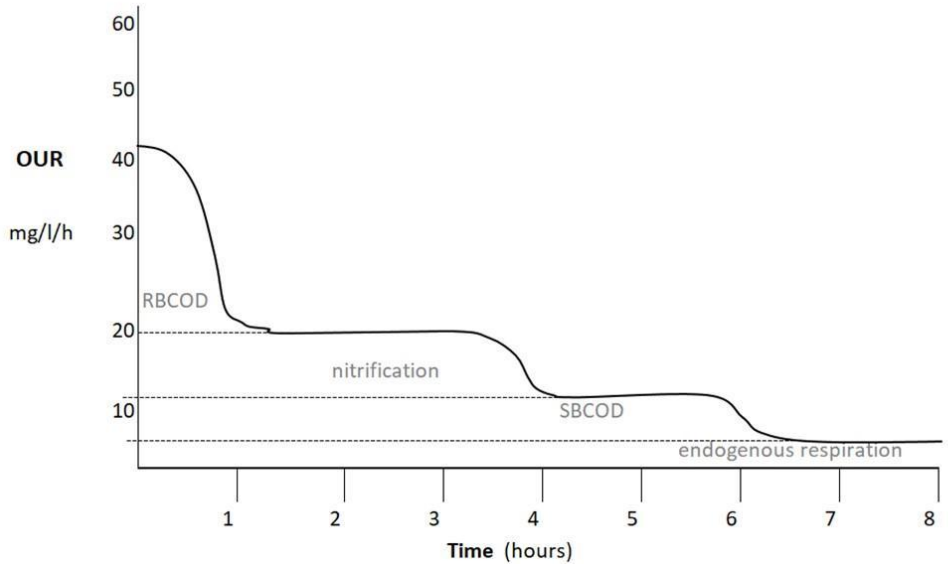

Figure 3: The municipal wastewater prokaryote extended phenotype

The typical municipal wastewater Oxygen Uptake rate for activated sludge systems shows stages of oxygen demand, with three stages in carbonaceous treatment systems and four stages for biotreatment systems designed with a longer mean cell residence time to accommodate slower growing ammonia oxidising and nitrite oxidising prokaryotes. The wastewater extended phenotype heterotrophic prokaryotes dominate oxygen demand due to their better bioenergetic returns on catabolism until their soluble substrates are exhausted. This phenotype then initiates comparatively expensive biosynthesis of hydrolytic enzymes to release more dissolved organic carbon from any local sources of volatile solids. In this period, the slower growing nitrifiers now dominate oxygen removal for ammonia oxidation until ammonia levels reach system equilibrium, at which point the oxygen demand associated with hydrolysed Slowly Biodegradable COD (SBCOD) occurs. After hydrolysed SBCOD has been oxidized, the municipal wastewater extended phenotype shifts to endogenous respiration).

The municipal wastewater extended phenotype includes a quorum sensing capability (Chong *et al,* 2012) that allows the phenotype to sense the local ratio of catabolic substrate to cell.

Below a threshold value which implies reactive resource scarcity, quorum sensing to initiate a K reproductive strategy allows optimized resource efficiency (Hense *et al*, 2007) and optimized resource acquisition capability. The heterotrophic prokaryotes with this phenotype switch to a high uptake rate of acetate and begin to produce External Polymeric Substances (EPS). The ability to produce EPS in this form is likely to have arisen from the genotype for competitive rapid uptake of acetate undergoing mutations and/or gene loss, or recombination, to provide phenotype restructuring towards acetate being shifted into EPS formation outside the cell.

The advantage of this phenotype with external resource storage that would immediately feed back into survival towards replication, is that there are now no space limits within the cell to stop the phenotype accumulating its optimum substrate from the environment while it is available (Palmer *et al*, 2020). Consequently, a structural material (EPS) is formed outside cells of this phenotype. *This minor*

structural change in phenotype also turns out to confer additional significant advantages towards survival directed to replication, including:
EPS facilitates cell aggregation which allows cells of this phenotype to form associations, also creating opportunities for cooperative resource acquisition (Rainey & Rainey, 2003), (Kreft, 2004), (Flemming *et al*, 2016) also reducing the survival risk profile for the phenotype(Boles *et al*, 2004);

- Aggregated growth occurs as biofilm in lower substrate environments and as flocs or granules where aggregated growth can settle under gravity at the upflow rates through that environment.
- What was likely to have originally observed as a starvation management tactic also provides mitigation against phenotype wash-out from the system (Palmer *et al*, 2020),
- The EPS layer entraps particulate material, *reducing the risk* of the phenotype investing in hydrolytic enzyme production once its primary soluble substrates are locally exhausted, as EPS particulate capture makes local particulate VSS highly likely,
- The EPS layer provides a location for growth of other prokaryotes including nitrifiers. Any local nitrate production in a mixed growth system provides this municipal wastewater extended phenotype with a possible diversification of catabolism of their primary substrates via anoxic catabolism,
- The EPS layer traps DNA and increases the probability of information variation via HGT (Merod & Weurtz, 2014) for this phenotype.

This extended phenotype is spread across several species and genera that proliferate in this particular environment.

In order to replicate its genome through successive generations, the phenotype as a vehicle for replication needs to optimize resource acquisition and manage resource paucity. The municipal wastewater extended phenotype presented in this case study is a good example of how minor genetic variation in a resource stressed environment can lead to a range of benefits with complex biological system behavior emerging from a single phenotypic attribute and significantly changing the phenotype's relationship with its environment due to that change locally reconfiguring the phenotype's environment.

Origin of Life: new insights from thermodynamics systems analyses

Physical system analysis is now being applied to the emergence of life on Earth. Hypotheses for the emergence of Life on Earth have for some time included an alkaline hydrothermal vent geological origin for life (Martin & Russell, 2003; Lane, 2015). This hypothesis assumes autotrophic methanogenesis as the basis for the emergence of bioenergetic catabolism:

$$CO_2 + 4H_2 \rightarrow CH_4 + 2H_20$$

Recent research deeper into this hypothesis has included thermodynamic analysis which provides critical contextual insight into what is physically possible.

One paradox for emergence of life is how the agent of memory – genetic code, emerges at the same time as a functional metabolism providing exergy for both code and energy source replication if an 'RNA world hypothesis is followed. More recent hypotheses retain RNA as the initial critical information system but more deeply examine how that is possible for life emerging from a geochemical origin.

On a physical system basis, a geochemical system possessing a boundary layer is profoundly different to one with no boundary. A simple fatty acid/lipid boundary layer creates a semi-permeable barrier that can facilitate differentiation between an internal environment for the emergent biological system

and the systems external environment. A semi-permeable boundary structure provides a basis for chemical gradients that can foster metabolise to replace geochemical energy sources and a semi-permeable boundary structure also creates an internal surface for reaction chemistry and initial internal structure development Wimmer *et a*l (2016),(Palmeira *et al*, 2022; Harrison *et al*, 2022).

Consequently, a semipermeable protocell is a structure that can facilitate development of a metabolism to drive replication of the system independent of any geochemistry the system emerged from. A semi-permeable fatty acid/lipid boundary structure also provides a physical structure which could self-replicate in an aqueous system, with volume to surface area of the structure giving a basis for division of the boundary Corominas-Murtra (2019).

These factors are now issues to examine in assessment of the emergence of life from a geochemical source because very recent breakthroughs have been made in understanding the physics (thermodynamics) and chemistry. Wimmer *et a*l (2016) reported on their systematic review of the physical chemistry now assumed for LUCA, which identified that the principal biosynthetic associated with its metabolism could be provided by hydrothermal vent conditions.

Autotrophic methanogenesis creates larger metabolites than its catabolic substrate which means it is entropy decreasing and this results in its driving $\Delta_r G$ being enthalpy ($\Delta_r H$) dominated. The energy dissipation of this reaction is far larger and more enthalpy driven than the aerobic respiration represented in Fig. 1. When the ratio of catabolic enthalpy $\Delta_r H$ to catabolic driving force $\Delta_r G$ is plotted for the known range of types of catabolic reactions and their driving forces ((von Stockar, 2013) all prokaryote metabolic regimes have a ratio of 1 or less, with the lone exception of autotrophic methanogenesis which has a ratio over 4. This significant thermodynamic distance from other metabolism clearly identifies autotrophic methanogenesis as a thermodynamic candidate for the emergence of prokaryotic life from a geochemical 'cradle' as described by Wimmer *et a*l (2016).

More recently, Corominas-Murtra (2019) investigated the thermodynamics of the duplication of lipid cells and determined that duplication required a balance between supply of exergy (free energy) and entropic forces associated with lipid boundary and sphere growth and maintenance up to the point of sufficient resources for its duplication. This work showing that protocell emergence requires a thermodynamic window, provides support for two studies showing how autotrophic protometabolism could develop in an alkaline hydrothermal vent environment and lead to the production of protocells (Palmeira *et al*, 2022; Harrison *et al*, 2022).

These thermodynamic system analyses have identified a development route by which protocells with autotrophic protometabolism, in which autocatalytic nucleotide synthesis and CO_2 fixation drive growth and protocell replication.

This route for the emergence of life starts with a geochemical environment that provides the chemistry and enthalpy required to support autotrophic metabolism, which in turn in that environment produces lipid-bound protometabolic in which nucleotides play a critical role. This provides a replicable protocell within which RNA can form (Palmeira *et al*, 2022; Harrison *et al*, 2022).

IV. CONCLUSIONS

Physical systems analysis of biological systems provides insights into the physical mechanisms that life both acts on and is defined by and the critical role played by the environment. Life and biological systems are distinguished from biological and chemical systems by their attribute of memory, the agent of which is the gene.

Using the physical definition of life provided in this paper:

'A c system with memory which is utilized for environmental fitness, with fitness defined by success in replicating that system.'

Lynch's contention that natural selection plays a minimal role in biological evolution and in the emergence of the complexity observed in biological systems (Lynch, 2007) cannot be reconciled with how biology works in physical terms as explained in this paper.

Complexity arises even in simple systems if they have sustained positive feedback over generations (Kaufmann, 1996; Holland, 1992, Holland 1996, Holland 1998).

In the case of biological systems, the purpose of information storage, i.e. memory (genotype), is for reproduction. The physical context for any use of information is that system information's distinguishability between the system and its environment.

For biological systems their purpose is self-replication and their environment and their alignment to it determines their reproductive success. Extinction of a genotype is the price of the failure of a phenotype to achieve the minimum level of resource management through information utilization required to successfully replicate its genome. For biological systems, the definition of information (compared to noise or misinformation), is accurate environment-relevant data.

Reproduction requires material resources and energy resources which have to be obtained from the environment which will usually include competitors for those resources. Biological system complexity arises from the need to optimize replication through securing resources needed for it. The phenotype is the physical structural information arising from genome replication and expression needed to secure the energy and material resources needed to sustain and reproduce a biological system. There is continuous positive system feedback between a genome and the environment through the phenotype's ability to secure all the resources, including space, needed to continue to replicate the genome.

The prokaryote genome often encodes information for two reproductive strategies that map directly to the ecological theory of 'r' strategy reproduction (maximum rate of reproduction of offspring) and 'K' strategy (lower rate of reproduction of more resilient offspring) and this in turn includes sensing based regulation which represents information utilization by the genome in expressing a phenotype within the organism's lifetime.

Sensing emerged in prokaryotes to satisfy basic physical requirements of reproduction. Quorum sensing emerged as the basis for prokaryotes to shift phenotype between an ''r' reproductive strategy or a 'K' reproductive strategy. Although these reproductive strategies have fallen out of favour for use in large, complex eukaryote ecology, this paper shows how both strategies arise from biophysical constraints in prokaryotes and emerge early in the evolution of life on Earth. In prokaryotes at least, 'r' strategy and 'K' strategy exactly describe how reproduction is tailored to environment resource opportunity or limitation and the characteristics (Reznick *et al*, 2002) attributed to each strategy in ecology fit well to prokaryote use of r' or 'K' strategy for reproduction in their microenvironments. For more complex organisms 'r' strategy and 'K' strategy might be expected to provide a less accurate fit to observed behaviour but for prokaryotes and microorganisms. Secondly, ecological analyses have until recently lacked systems analyses that took into account the significance of non-Markovian processes in biology in forming a view resource appreciation in reproductive strategy formulation, but such

approaches such, as discounted reproductive number (Reluga *et al*, 2009) are now available and may help encourage increased use of r' and 'K' strategy in future.

Resource acquisition strategies and tactics in biology are central to early ecology and are critical physical challenges for all biological systems to manage within their relationship to their environment in order to optimize genome replication. The systems behaviour we describe as 'economics' appears in biological resource acquisition strategies and tactics emerging in prokaryotes described in this paper.

ACKNOWLEDGEMENTS

The author has no conflicts of interest to declare.

REFERENCES

1. Ben Naim, A. (2008) A Farewell to Entropy: Statistical Thermodynamics based on information. Publ. World Scientific Publishing Co. Pte Ltd UK. ISBN: 978-981-270-707-9.
2. Palmer, S.J. (2019). A review of the physical roles of information in nature. Journal of Biological Physics and Chemistry. Vol 19 p22-40 (2019).
3. Tribus, M. and McIrvine, E.C. (1971). Energy and information. Scientific American, September 225, 179–188.
4. Holland, J.H. (1992) Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence (Complex Adaptive Systems). mergence: from chaos to order. MIT Press. ISBN 978-0262581110.
5. Holland, J.H. (1998) Emergence: from chaos to order. Oxford University Press. ISBN 978-0-19-286211-2.
6. Holland, J.H. (1996) Hidden Order: How Adaptation Builds Complexity. Basic books/Perseus Books Group. ISBN 0-201-44230-2.
7. Dawkins, R. (1976). The Selfish Gene (first edition). Oxford University Press. ISBN 0-19-857519-1
8. Dawkins, R. (1982) The Extended Phenotype: the long each of the gene. Oxford University Press. ISBN 0-19-288051-9.
9. Martin, W. Russell, M.J. (2003) On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. Phil. Trans. R. Soc. Lond. B 2003 358, p59-85.
10. Lane, N. (2015). The Vital Question: Why is life the way it is? Publ. Profile Books Ltd (2015, 2016). ISBN 978 1 78125 037 2.
11. von Stockar, U. (2013) Live cells as open non-equilibrium systems. Chapter 16, 399-421. In: von Stockar, U. (ed)(2013. Biothermodynamics: the role of thermodynamics in Biochemical Engineering. EPFL Press 1st edition (2013).
12. Heijnen, J.J. (2013) A thermodynamic approach to predict black box model parameters for microbial growth. Chapter 18, 443-473. In: von Stockar, U. (ed)(2013. Biothermodynamics: the role of thermodynamics in Biochemical Engineering. EPFL Press 1st edition (2013).
13. Wang, G., Post, W.M. (2012). A theoretical reassessment of microbial maintenance and implications for microbial ecology modelling. FEMS Microbiol. Ecol. 81p610-617.
14. Reznick, D., Bryant, M.J., Bashey, F. (2002) Ecology 83(6) p1509-1520.
15. Monod, J. (1950) La technique de culture continue. Théorie et applications.. Ann. Inst. Pasteur. 79. p390-410.
16. Herwig Bachmann, H., Bruggeman, F.J., Molenaar, D., dos Santos, F.B., Teusink, B. (2016) Public goods and metabolic strategies. Current Opinion in Microbiol. 31p109-115.
17. Bradley, J.A., Amend, J.P., LaRowe, D.E. (2018) Survival of the fewest: Microbial dormancy and maintenance in marine sediments through deep time. Geobiology 17: p43-59.

18. De mattos, M.J.T.., Neijssel, O.M. (1997) Bioenergetic consequences of microbial adaptation to low-nutrient environments. Journ. Biotech. 59 p117-126.
19. Gonzalez,O., Gronau, S., Pfeiffer, F., Mendoza, E., Ralf Zimmer, R., Oesterhelt, D. (2009) Systems analysis of the Bioenergetics and Growth of the Extreme Halophile *Halobacterium salinarum* PLOS Comp. Biol. 5 (4) p1-11.
20. Costa E., Perez J., Kreft ,J.U. (2006) Why is metabolic labour divided in nitrification? Trends Microbiol 14: 213-219.
21. Lane, N and Martin, W. (2007) The energetics of genome complexity. Nature. Vol. 467 p929-934.
22. Lane, N. and Martin, W.F. (2012). The origin of membrane bioenergetics. Cell 151, 1406–1416.
23. Lynch, M. (2007) The frailty of adaptive hypotheses for the origins of organismal complexity. Proc. Nat. Acad. Sci. 104(Suppl 1):8597-8604.
24. Lynch, M., Marinov, G.K. (2015) The bioenergetic costs of a gene. Proc. Nat. Acad. Sci. 112(51):15690-15695.
25. Lynch, M., Marinov, G.K. (2016) Reply to Lane and Martin: Mitochondria do not boost the bioenergetic capacity of eukaryotic cells. PNAS 11113 E667-E668.
26. Lynch, M., Marinov, G.K. (2017) Membranes, energetics and evolution across the : prokaryote-eukaryote divide. eLife, 6 e20437.
27. Schavemaker, P.E., Muñoz-Gómez, S.A. (2022) The role of mitochondrial energetics in the origin and diversification of eukaryotes. Nat Ecol Evol 6, 1307–1317 (2022). https://doi.org/10.1038/s41559-022-01833-9.
28. Jacob, F., & Monod, J. On the regulation of gene activity. Cold Spring Harbor Symp. Quant. Biology 26, 193–211 (1962).
29. Struhl, K. (1999) Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. Cell 98, 1-4.
30. Seto, M. (2014) The Gibbs free energy threshold for the invasion of a microbial population under kinetic constraints. Geomicrobiology Journ. 31, 645-653.
31. De Souza, G., Kost, C. (2016) Experimental evolution of metabolic dependency in bacteria. LPOS genetics DOI: 10.1371/journal.pgen.1006364.
32. Hall, R.J., Whelan, F.J., McInerney J.O., Ou Y,Domingo-Sananes, M.R. (2020) Horizontal gene transfer as a source of conflict and cooperation in prokaryotes. Frontiers in Microbiology, 17 Jul 2020, 11:1569 DOI: 10.3389/fmicb.2020.01569.
33. Hense, B., Kuttler, C., Müller, J., Rothballer, M., Hatmann, A., Kreft ,J.U. (2007) Does efficiency sensing unify diffusion and quorum sensing?. Nat Rev Microbiol 5, 230–239 (2007). https://doi.org/10.1038/nrmicro1600.
34. Hunter, P. (2018) The revival of the extended phenotype. EMBO Reports 19:e46477 DOI: 10.15252/embr.2018464477.
35. Chong, G., Kimyon, O., Rice, S.A., Kjelleberg, S., Manefield, M. (2012). The presence and role of bacterial quorum sensing in activated sludge. Microbial. Biotech. 5(50, p621-633.
36. Palmer, S.J., Noone, G., Hoyland, G. (2020).A review of the physical relationships between form and function in wastewater biotreatment biophysical relationships and a new theory relating activated sludge and biofilm systems behaviour. Water and Env. Journ. Doi:10.1111/wej.12614.
37. Rainey, P.B., Rainey, K. (2003) Nature Letters Vol 425 4 Sept 2003 p72-74.
38. Kreft ,J.U. (2004) Biofilms promote altruism. Microbiology 150: 2751-2760.
39. Flemming, H-C., Wingende, J., Szewzyk, U., Steinberg, P., Rice, S.A., Kjelleberg, S. (2016) Biofilms: an emergent form of bacterial life. Nature Reviews 14, 563-574.
40. Boles, B.R., Thoendel, M., Singh, P.K. (2004) Self-generated diversity produces “insurance effects” in biofilm communities. PNAS 101 (47) 16630-16635 https://doi.org/10.1073/pnas.0407460101

41. Merod R.T., Weurtz, S., (2014)Extracellular Polymeric Substance Architecture Influences Natural Genetic Transformation of *Acinetobacter baylyi* in Biofilms. Appl. Environ. Microbiol. 80 (24) p7752-7757
42. Corominas-Mutra B (2019) Thermodynamics of Duplication Thresholds in Synthetic Protocell Systems. Life 9 (9) DOI:10.3390/life9010009
43. Wimmer, J.L.E., Xavier, J.C., Vieira A.N, Pereira, D.P.H., Leidner, J., , Sousa, F.L., Kleinermanns, K., Martina Preiner., M., Martin, W.F. (2021) Energy at Origins: Favorable Thermodynamics of Biosynthetic Reactions in the Last Universal Common Ancestor (LUCA). Frontier in Microbiol. 12 p1-18 DOI 10.3389/fmicb.2021.793664
44. Nunes Palmeira R, Colnaghi M, Harrison SA, Pomiankowski A, Lane N. The limits of metabolic heredity in protocells. Proc Biol Sci. 2022 Nov 9;289 (1986):20221469. doi: 10.1098/rspb.2022.1469. Epub 2022 Nov 9. PMID: 36350219; PMCID: PMC9653231.
45. Harrison, Stuart A; Palmeira, Raquel Nunes; Halpern, Aaron; Lane, Nick; (2022) A biophysical basis for the emergence of the genetic code in protocells. Biochimica et Biophysica Acta (BBA) - Bioenergetics , Article 148597. 10.1016/j.bbabio.2022.148597.
46. Kauffman, S. (1996) At Home in the Universe: The Search for the Laws of Self-Organization and Complexity. Oxford University Press (USA). ISBN 978-0195111309
47. Reluga, T.C., Medlock, J., Galvani, A. (2009) Math. Biosciences & Eng. 6(2)p377-393