Theory: **Biological systems organize to maximize entropy production subject to information and biophysicochemical constraints**.

Project Summary

Recently in the field of nonequilibrium thermodynamics, Dewar (2003, 2005b) presented a provisional proof on the theory of maximum entropy production (MEP), which posits that steady state systems with sufficient degrees of freedom will organize to maximize the rate of entropy production. While organized structures decrease the entropy of the system, they are maintained by external entropy production and have a higher probability of persistence if their presence increases overall entropy production. However, the configuration of structures that generate entropy, and dissipate energy, are constrained by system resources from which the structures must be synthesized from. Hence, biophysicochemical constraints limit the complexity of dissipative structures. Hurricanes that dissipate thermal energy between the atmosphere and ocean are examples of such dissipative structures. We propose that evolution by natural selection produce biological systems that tend to follow a pathway of maximum entropy production by dissipating high temperature radiation and chemical potential. Consequently, an ecosystem composed of organisms that produce entropy at a high rate has a greater probability of persistence and occupation than an ecosystem under the same constraints that produces entropy at a lower rate. While MEP theory does not distinguish between abiotic and biotic systems, biological systems differ from abiotic ones in one key way: biological systems store information within their metagenome. Therefore, we propose that abiotic systems maximize entropy production by a steepest descent pathway, while information stored within the metagenome allows biological systems to produce entropy along pathways that can increase entropy production when averaged over time. For instance, by storing internal energy, biological systems can maintain entropy production and persist during periods when external energy inputs cease. Based on our theory, we hypothesize that biological systems with greater information content will have higher entropy production rates than biological systems will lower information content.

To test our hypotheses, we will employ flow through microcosms (i.e., chemostats) as experimental systems inoculated with natural microbial communities. Changes in chemical composition will be used to determine entropy production and massively parallel 454 pyrosequencing applied to hypervariable regions in rRNA genes will provide a direct measure of the information content of complex microbial communities. In addition to experimental tests, we will develop a mathematical framework based on our theory to model biogeochemistry orchestrated by biological systems using a distributed metabolic network representation.

Intellectual merit: The current paradigm that governs ecosystem processes is not constrained other than by conservation of mass and energy. Evolution by natural selection provides a mechanism by which complex structures can self organize, but confers no directionality to system evolution because the fitness landscape is as dynamic as the organisms themselves. The theory of MEP under biophysicochemical and information constraints provides directionality, which would greatly advance our understanding of system evolution that is necessary understand energy and mass flow through biological systems and to predict how living systems respond to change. In addition to our main research question, our experiments will significantly contribute to understanding of ecosystem function and community structure.

Broader Impacts: To facilitate application of MEP theory across fields, computational models developed during the project and all experimental results will be distributed on a web site, where we will also develop content for K-12 and higher educational outreach. Feedback from the PIE-LTER Schoolyard program participants will guide web page content and development. Experimental equipment and methods developed will also be used to expand an undergraduate course that the PIs currently CO-teach. The project will support one postdoctoral student, who will be mentored in an interdisciplinary field that crosses biogeochemistry, thermodynamics, information theory, and genomics.

TABLE OF CONTENTS

For font size and page formatting specifications, see GPG section II.C.

	Total No. of Pages	Page No.* (Optional)*
Cover Sheet for Proposal to the National Science Foundation		
Project Summary (not to exceed 1 page)	1	
Table of Contents	1	
Project Description (Including Results from Prior NSF Support) (not to exceed 15 pages) **(Exceed only if allowed by a specific program announcement/solicitation or if approved in advance by the appropriate NSF Assistant Director or designee)**	15	
References Cited	10	
Biographical Sketches (Not to exceed 2 pages each)	4	
Budget (Plus up to 3 pages of budget justification)	7	
Current and Pending Support	3	
Facilities, Equipment and Other Resources	3	
Special Information/Supplementary Documentation	0	
Appendix (List below.) **(Include only if allowed by a specific program announcement/ solicitation or if approved in advance by the appropriate NSF Assistant Director or designee)**		
Appendix Items:		

*Proposers may select any numbering mechanism for the proposal. The entire proposal however, must be paginated. Complete both columns only if the proposal is numbered consecutively.

1. Response to Previous Panel Reviews

This is the third submission of the proposal to the ATB competition. Both previous submissions reviewed very well (E, E, G, and E, E, E, V, V, V), but neither made the funding cut. In the second submission, we changed from DGGE community surveying to 454-tag sequencing, which reviewed favorably, so we focus our comments now on the last review that mostly concerned the proposed experiments. The panel felt that the experiments need to demonstrate that community entropy production is at a maximum. To address this, we have added new mathematical analysis (Section 4.5) and have introduced a new experiment (Exp. 1), which is intended to demonstrate optimality based on community structure. We now use a methanotrophic-based community, as energy consumption can be readily monitored via CH_4 uptake, and the fraction of resources allocated to its uptake can be measured using quantitative PCR of methane monooxygenase. The panel expressed concerns that more experimental data points would be useful, and in response we have increased the number of perturbations we will conduct. However, as articulated in the ATB RFP, theory is typically not proven or falsified with just one set of experiments. In addition to our main research question, our experiments will significantly contribute to understanding of ecosystem function and community structure [109]. We have recently addressed the concerns of translating 454-tag sequences into species diversity [58]. Finally, we are unaware of techniques other than serial dilution to change community structure in complex microbial systems without introducing significant artifacts.

2. Introduction

The development of theory presented in this proposal derives from the basic question: What are the governing principles that determine how energy and matter flow through biological systems composed of independent but interacting individual organisms, such as occurs in ecosystems? Surprisingly, no predictive theory exists for such a fundamental question. While the theory of evolution by natural selection provides a mechanism for self-organization of complex biological structures, the theory is indeterminate in regards to the emergent properties biological systems follow, if any [88, 104]. To address this long-standing question, we will extend a theory from nonequilibrium thermodynamics to biological systems, and will also qualitatively draw upon information theory [2, 117]. Recently, Dewar [29, 31] provided a provisional proof that indicates steady state systems with many degrees of freedom will organize to a state of maximum entropy production (MEP). As will be described in more detail below, entropy production in this proposal can be considered equivalent to energy dissipation. Consequently, self-organization in a system that enhances energy dissipation will have a higher probability of persistence. Based on the MEP principle, the organized structure of a hurricane develops and persists because it facilitates dissipation of energy between the ocean and atmosphere [114]. By analogy, self-organization in a biological system that enhances the degradation of chemical potential, such as oxidation of methane, will have an increased probability of persistence. While the MEP principle (MEPP) will form the basis of our extended theory, many questions must be addressed in order for MEPP to be useful in applications regarding energy and mass flow through biological systems.

As derived by Dewar, MEPP does not distinguish between abiotic and biotic systems. While this allows general application of MEPP, most would agree that hurricanes differ from ecosystems, but what are the important differences in regards to MEPP? How does evolution by natural selection result in a MEP state? Thermodynamics applies to macroscopic properties of systems, as the ideal gas law describes, so what is the proper scale for application of MEPP to biological systems? Does MEPP apply to biochemistry, trophic interactions, or only to whole ecosystems? Can the rate of entropy production at the MEP state be predicted *a priori* for biological systems? Do systems always operate at the MEP state? How do resources available to biological systems constrain biological structure synthesis? How does the evolution of the metagenome [49] affect the MEP state? In this proposal we will address these questions and in the process derive the theory that biological systems organize to maximize entropy production subject to information and biophysicochemical constraints. Experimental systems will be employed to test hypotheses derived from our theory and a mathematical framework for a predictive model will be presented and tested against experiments.

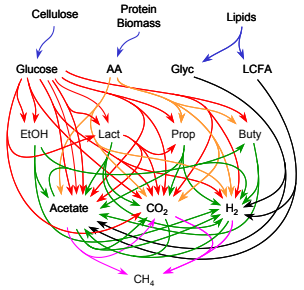

Fig. 1. A subset of possible methanogenic pathways [47, 124].

While the concepts discussed in this proposal apply across all levels of biological organization, we will focus entirely on microbial systems because the concepts are easier to demonstrate due to the short characteristic

timescales and high levels of system organization microbial communities exhibit. For instance, there are surprising regularities that result from competitive and cooperative interactions amongst organisms that are often independent of which organisms are present. Consider methanogenic microbial communities [47, 124]. In the absence of electron acceptors (O_2, NO_3^-, Mn^{4+}, etc.) microbial communities convert glucose to methane and CO_2. While no single organism conducts the overall catalysis, the number of microbial community configurations and metabolic pathways producing the same overall result is staggering (Fig. 1). Microbial systems are fascinating from an evolutionary perspective because the fitness landscape–chemical matrix–is intimately coupled to organismal evolution. Any successful mutation affecting a single organism will result in immediate alteration in system chemistry, which affects all organisms including syntrophic connections that the mutated organism depends upon. This tight coupling has led some to speculate that selection operates across a continuum of hierarchies [20, 21] as well as renewed theories on how such selection may occur [18, 45, 46, 91, 111, 112, 129]. While elucidating the mechanistic details that underlie community level cooperation is important, our focus is at a higher level that does not depend on these details. We posit that competition and cooperation among organisms is sufficiently nonlinear that emergent properties, such as the MEPP, result that are more predictable than the internal dynamics of the underlying community [15, 68]. In fact, predictability of internal community dynamics is likely not possible, except over short time scales [13]; hence the necessity to focus at the whole community scale.

3. Background

3.1 Entropy and Information Definitions Even though entropy is precisely defined in thermodynamics [11], entropy's association with statistical inference and information theory leads to much confusion. In this proposal, entropy will refer to Clausius' [26] original definition; the energy lost when converting internal energy into work. In this case, entropy arises from the atomic structure of matter, in which energy spontaneously disperses into lower energy states due to translational, rotational and vibrational interactions of atoms and molecules. Under constant temperature and pressure, the energy that can be used for work is Gibb's free energy, which accounts for losses due to entropy. Most importantly, if a chemical reaction occurs, such as combustion of methane, but all the energy liberated is dissipated as heat to the surroundings, *then all the free energy is converted to entropy*. Unlike energy, free energy is not conserved. In fact, free energy is more accurately a measure of the energy that can be dissipated as entropy [77]. Hence, for processes of interest here, entropy production will be equated to chemical or radiative energy dissipation. Thermodynamic entropy is associated with the microscopic state of matter; it explains why a sugar cube spontaneously dissolves in water, but *not* the disordering of objects on a desk, which requires external energy dissipation [76]. Shannon entropy describes the latter.

Due to mathematical similarities to statistical mechanics derivations of entropy of gasses by Boltzmann and Gibbs, Shannon [117] also used the term entropy in his original work on information theory. Shannon entropy refers to the amount of information that can be encoded into a message or sent over a communication channel. It is also referred to as algorithmic complexity [74]. In this context, entropy is associated with uncertainty. The ability to predict the next symbol in a message goes down as the information content (Shannon entropy) of the message goes up. A message of only 1's has Shannon entropy of 0; it is quite predictable and contains no information. Interestingly, a message with the greatest information content will appear as a sequence of random symbols. As discussed by Adami et al. [2], in order for information to be useful it must be correlated with the physical world. A random sequence of amino acids, while having high Shannon entropy, would produce a functionless enzyme, as it contains no useful information relative to the environment. Likewise, an enzyme that processes CH_4 is equally irrelevant if the environment lacks CH_4. We will use the term information, then, as Adami [1] defines it, which is also known as structural complexity or useful information.

Finally, we note that Bayesian maximum entropy (MaxEnt) formulation [62], that has garnered interest in population ecology [32, 50, 120], is used to derive MEP [29], but should not be confused with MEP.

3.2 Maximum Entropy Production Principle (MEPP) Dating back to at least Lotka [80], who proposed that ecosystems organize toward a state of maximum power, there has been a significant amount of work in theoretical ecology that focuses on understanding the governing principles that organize ecosystems. In classic equilibrium thermodynamics, a state of maximum entropy and zero entropy production defines the resting state of a system [75], but biological systems are far from equilibrium and contain low entropy ordered structures that are maintained by external energy dissipation [85, 116]. Consequently, there has

been a great deal of interest in nonequilibrium thermodynamic (or thermodynamically inspired) applications to living systems involving: power [80, 94], biomass to maintenance [81], minimum entropy production [108], exergy [84], ascendancy [130], emergy [95], energy dissipation [115], respiration [23, 144], thermodynamic efficiency [89], constructal theory [12], as well as others [67, 128, 133, 145]. However, the theories have not gained wide acceptance and are seldom employed to understand biogeochemistry. While many of the theories have a similar basis, they differ sufficiently to cause confusion, even though many have been shown to be more similar than different [36, 66]. Furthermore, many of the theories are based more on observational intuition than fundamental principles. Consequently, the biological community is faced with the conundrum of which theory is "correct" and which one should be employed, which has stifled advancement.

Recently, there has been a renewed interest in the principle of maximum entropy production for nonequilibrium systems due to both theoretical and observational research. The original application of MEP dates to Paltridge [100], who demonstrated that if global heat transport between the tropics and the poles follows MEPP, one can accurately predict meridional heat flux, latitudinal temperatures, and fractional cloud cover. At the time, Paltridge's work was discounted and believed to be coincidental [101]. However, it was later shown that the MEPP also accurately describes the climatology of Mars and Titan where the conventional assumption regarding heat transport fails [78]. Others have also speculated that ecosystems follow MEPP [126, 132], but without theoretical and experimental support, it garnered little attention. Dewar's [29, 31] analysis now provides a theoretical basis for the MEPP, as he derives a provisional proof for MEP for nonequilibrium steady state systems with sufficient degrees of freedom. The general conclusion from the MEPP is that systems will organize, within constraints, so as to maximize the rate of entropy production. MEP is consistent with most, if not all, of the intuitively based ecological concepts referenced above [36, 66]. MEP is an appealing extension to classic, equilibrium thermodynamics that dictates systems will move to a state of maximum entropy at equilibrium. In essence, MEP indicates systems will attempt to achieve equilibrium via the fastest allowable pathway. Equally desirable, the MEP principle does not distinguish between biotic and abiotic systems so can be applied generally.

Dewar's approach is based on the maximum entropy postulate of Bayesian inference developed by Jaynes [61] and uses statistical mechanics reminiscent of Gibbs original work on entropy for equilibrium systems [41]. To arrive at the MEP principle, Dewar uses a probabilistic argument involving microstates and macrostates for steady state nonequilibrium systems; the premise being that a system is most likely to be found in the macrostate that has the most number of microstates. Dewar shows that this occurs with the macrostate of maximum entropy production [cf. 79]. Unlike Hamiltonian systems, there is no requirement that systems operate at MEP, but MEP does serve as a dynamical systems attractor. The MEP principle provides a useful means of describing how complex systems function, and there now exist several examples where this appears to be the case. When ocean circulation models are perturbed, they transition to new steady states with higher entropy production rates [119]. The laminar to turbulent flow transition in smooth pipes can be predicted from MEPP [82]. Beach profiles organize so as to maximize dissipation of incoming wave energy [63]. Entropy production governs growth morphologies of crystals [52]. Kirchhoff's loop law for electronic circuits can be derived from MEP [152] and MEP also explains other properties of electrical systems [24]. Several applications of MEP exist for climate systems [71, 72, 97, 98, 102, 141], including plant transpiration [142]. These examples provide a strong foundation for our use of MEPP as a guide to determine how and why biological systems organize.

4. Proposed New Theoretical Directions

While MEPP has been gaining acceptance in the physical sciences, its validity and application to biological systems remains uncertain. Much of the uncertainly results from the organismal perspective used to understand ecosystem processes, which we believe is an inappropriate scale for considering MEPP. Below we will discuss new ideas for conceptualizing biological systems that are consistent with MEPP and will form the base of this proposal.

4.1 Microstates and Macrostates The MEPP requires that a system must have many degrees of freedom; that is, many different microstate configurations that produce the same entropy producing macrostate. Gibbs' (and Boltzmann's) analysis of gasses using statistical mechanics provides an example [41]. In this case, pressure, volume, and temperature represent macrostate properties, while velocities and positions of molecules represent the microstates. Of course, there are an enormous number of different microstates

that produce the same macrostate. In this proposal, we extend the concept of microstates to biological systems. ***Here, a microstate refers to the organisms present and their connectivity in an ecosystem***. Based on MEPP, there should exist many different species configurations and connectivities that give rise to the same entropy producing macrostate. Over appropriate timescales, species composition and their trophic relationships dramatically change, as has been observed in methanogenic [38], nitrifying [48] and planktonic communities [13] that exhibit dynamics primarily at the species level while maintaining functional stability; a kind of dynamic degeneracy [33]. Indeed, multiplicity of biological microstates appears consistent with neutral theory [110], provided the different configurations are functionally complementary. Due to the large number of microbial species and their high abundances [40, 122], the microstate analogy is most applicable to microbial systems, which will be our primary focus.

4.2 Biophysicochemical Constraints In any implementation of MEPP, system constraints are critical in determining the relevant solution and the rate of entropy production at MEP [27]. In fact, Prigogine's original work on *minimum* entropy production [107] is actually a special case of MEP where the degrees of freedom (the boundary conditions) are constrained [24, 30]. As example, consider a flammable gas mixture of methane and air. If there is a continuous supply of CH_4 and O_2 and combustion is initiated, then a continuous flame will be produced. If no work is done, then all the free energy of combustion is converted into entropy. In this case, the process is operating at the MEP state, where the magnitude of entropy production is constrained by the kinetic theory for gasses, which we could calculate *a priori*. We consider this type of abiotic process operating in a *steepest descent mode*, where the rate of entropy production is maximized at any instance in time, subject to constraints. For analogy, water flowing downhill follows a steepest descent trajectory. Hurricanes also dissipate energy in a steepest descent mode. What happens, however, if the CH_4-O_2 mixture is outside its flammability limits?

Outside the flammability limits, CH_4 and O_2 still react in a steepest descent mode, but the reaction proceeds extremely slow. The reaction rate can be accelerated considerably by addition of a catalyst, but for self organizing systems, the catalyst must be created from CH_4 and O_2 and/or materials contained within the system. The reaction rate, and entropy produced, depends on the effectiveness and quantity of the catalyst. An insufficient amount of catalyst limits the reaction rate, but too much catalyst wastes resources. If the catalyst contains internal energy, which it is likely to, then over-synthesis of catalyst is inconsistent with MEP, because entropy could have been produced instead of catalyst. The catalysts effectiveness (i.e., reaction rate per unit mass) depends on its molecular structure, which in turn depends on the resources available to construct it. Given the common elements (C, N, O, H, S, P, Fe, Mn, Cu, Co, etc), what is the most effective catalyst that could be constructed? Obviously, this is an extremely difficult question to answer given the tremendous number of degrees of freedom. We know the extreme upper bound would be the complete removal of the reactions activation energy, but how much the activation energy can be decreased is constrained by the elemental resources used to build the catalyst. Consequently, unlike the relatively simple kinetics of combustion, we cannot determine *a prior* what the maximum reaction rate is nor what the maximum entropy production rate could be. Experiments are necessary, as well as including our understanding on the stoichiometric limits imposed on biological structures [34, 125].

4.3 MEP by Individuals or Ecosystems? Of course, the catalysts of interest here are enzymes, but we include the organisms that synthesize enzymes as part of the catalyst. Continuing our example, we know from experiments that methanotrophs catalyze methane oxidation as follows,

$$CH_4 + O_2 \xrightarrow{M} \varepsilon_M M + (1-\varepsilon_M)CO_2 + \cdots \quad (1)$$

where methanotrophs, M, both catalyze the reaction and are produced by it (the reaction is autocatalytic [131]) as a function of their growth efficiency, ε_M. The question we address here is how might evolution by natural selection lead to a system that maximizes entropy production? The competitive exclusion principle [7] tells us that the organism with the fastest growth will dominate at the exclusion of all others. To achieve fast growth an organism must balance efficiency versus speed. In (1) above, if ε_M is close to unity (high efficiency), the reaction will proceed slowly due to thermodynamic constraints (see Section 7.1 below) and methanotroph productivity will be low. As ε_M approaches 0 (low efficiency), the reaction will proceed rapidly, but once again with low methanotroph productivity due to low efficiency. Between 0 and 1 lies an optimum ε_M that maximizes methanotroph productivity that is selected for by evolution, but this value does not maximize entropy production because too much methane contributes to

methanotroph biomass. Consequently, we conclude that growth of individual organisms as selected for by evolution does *not* follow MEPP. However, natural ecosystems are not composed of a single species, as grazers, G, are always present to consume prey, such as methanotrophs, as given by,

$$\mathrm{M} + \mathrm{O}_2 \xrightarrow{\mathrm{G}} \varepsilon_{\mathrm{G}}\mathrm{G} + (1-\varepsilon_{\mathrm{G}})\mathrm{CO}_2 + \cdots \tag{2}$$

which is also an autocatalytic reaction that produces exponential growth. Of course, grazers of the methanotroph-grazers are often present, as well as detritivores. Extending this logic naturally leads to food webs observed in nature. A net result of all the predation is that the total biomass in the system is constantly turned over. It is possible to have all organisms growing at their maximum rates without any biomass accumulation, which could lead to an MEP state. ***Hence, it is necessary to have a food web of prey and predators in order to achieve MEP***. If we do not distinguish species, then the system exhibits large-scale cannibalism. The predator-prey interactions also insure biological structures are continuously and dynamically reallocated to those reactions that could assure MEP under changing conditions.

4.4 MEP, Information and Timescales Interestingly, the effectiveness of methanotrophs to catalyze (1) depends on information stored in their genome, which specifies how to construct staggeringly complex organic structures from resources available in the environment. Likewise, predators and other organisms contain information on consuming live and dead biomass. Over evolutionary time, enzymes associated with methanotrophy would presumably improve [1], so that less protein is needed by modern methanotrophs than ancient methanotrophs to attain the same reaction rate. Hence, the amount of entropy produced for a given amount of protein changes with evolution, as well as with the introduction of new enzymes all together. For instance, MEPP predicts the evolution of oxygenic photosynthesis, because entropy production rate per unit biomass is higher with O_2/H_2O redox reactions than anaerobic photosynthesis based on SO_4^{2-}/ H_2S. But information embedded in the metagenome makes it difficult to predict maximum entropy production in biological systems from first principles as can be done with physical systems and demonstrated by Paltridge [100]. As metagenomic sequencing and annotation capabilities increase [151], this information can be used as further constraints. However, metagenome sequencing is still prohibitively costly, so in this proposal we will use a proxy for information content based on short ribosomal "tag" sequences as discussed in Section 6.5, as well as rely on simple metabolic networks for constraints. Information also permits biological systems to integrate entropy production over time and circumvent the steepest descent pathway of abiotic systems.

As discussed above, maximum entropy production rate for a flammable mixture of CH_4 and air is dictated by gas kinetics. While combustion is the MEP solution, it tends to destroy order structures, so has only short persistence. If a perturbation extinguishes the flame, the CH_4 and air mixture will accumulate until a serendipitous spark is reintroduced. Consequently, there can be periods of massive entropy production followed by long periods of no entropy production (and combustion never occurs if the mixture falls outside its flammability limits). If a catalyst is introduced, such as methanotrophs, then entropy can be continuously produced over substantial transients. Even though the instantaneous entropy production will be lower with methanotrophs, the average rate of entropy production will exceed that of the sporadic steepest descent route.

If MEP follows only steepest descent pathways, then not only should CH_4 be oxidized, but all biomass as well, as this would produce the greatest instantaneous entropy proudction. However, if time-averaged entropy production is maximized [86], then allocating some CH_4 to methanotrophs and grazers increases entropy production over the integration interval. Unlike physical systems, biological systems have the capability to predict the future, where the information to do so is contained within the system metagenome. For instance, deciduous forests store some resources during the growing season that allows them to maintain dormancy over the winter or dry period, which requires expectations of future conditions. Microbes also exhibit temporal strategies such as spore formation and luxury uptake mechanisms [14, 69]. When considering biological systems, strategies are not instantaneous, but are integrated over time based on a prediction of the future that has been selected for by evolution. Consequently, by avoiding the steepest descent pathway, information can allow a system to produce entropy even when confronted with perturbations. Of course, perturbations of sufficient magnitude will disrupt biological systems as well. Nevertheless, we postulate ***the difference between abiotic and biotic processes is that the former always follows a pathway of steepest descent, while the later follows a pathway dictated by information that leads to greater entropy production when averaged over time***. We

speculate the MEP asymptote occurs when entropy production is averaged over infinite time and space. Of course, the two pathways are always competing, such as occurs when fire consumes a forest. Pathways of averaged entropy production may be flanked by pathways of steepest descent.

4.5 MEP Independent Variable(s) Entropy production (EP) is the dependent variable in MEP, but what is the independent variable(s) of the optimization? In the original work of Paltridge [100], the independent variable was heat transport from tropics to pole. As modeled heat transport was varied between 0 and ∞, entropy production exhibited a maximum, where the predicted heat transport at MEP corresponded to that observed. For biological systems, we propose that organization of biological structure (i.e., species composition, trophic configuration, etc.) is the independent variable in that different species configurations produce entropy at different rates, with only some configurations producing MEP. We will illustrate this concept with a numerical experiment involving a methanotrophic community in a chemostat sparged with CH_4 and air with a mineral nutrient feed.

Box 1. Methanotrophic community model.

$$\dot{m}(t) = \varepsilon_m \varphi_m(n(t),c(t))m(t) - \varphi_g(m(t))g(t) - \chi_m m(t) + f_L(m^I - m(t))/V_L$$

$$\dot{g}(t) = \varepsilon_g \varphi_g(n(t))g(t) - \chi_g g(t) + f_L/V_L(g^I - g(t))$$

$$\dot{p}(t) = f_G(p^I - p(t))/V_G + k_L aRT(c(t) - p(t)/h(T))/V_G$$

$$\dot{c}(t) = k_L a(p(t)/h(T) - c(t))/V_L - \varphi_m(n(t),c(t))m(t) + f_L(c^I - c(t))/V_L$$

$$\dot{n}(t) = (-\varepsilon_m \varphi_m(n(t),c(t))m(t) + (1-\varepsilon_g)\varphi_g(n(t))g(t) + \chi_m m(t) + \chi_g g(t))/\rho + f_L(n^I - n(t))/V_L$$

$$\text{where} \quad \varphi_m(n,c) = \varphi_m^{Max} nc/(n+k_{mn})/(c+k_{mc}) \quad \text{and} \quad \varphi_g(m) = \varphi_g^{Max} m^2/(m^2+k_{gm}^2)$$

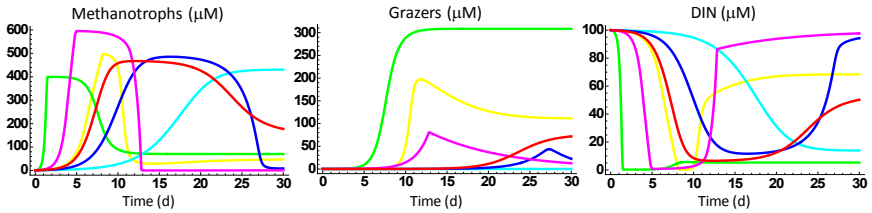

Fig. 2 Transient solutions for eqns. in Box 1 for 6 different parameterizations.

The model (Box 1) uses the standard aggregated compartment approach where methanotrophs, $m(t)$, consume dissolved methane, $c(t)$, and nitrogen, $n(t)$, for growth and are preyed upon by grazers, $g(t)$. Methane is sparged into the chemostat at partial pressure $p(t)$. Ten kinetic growth parameters ($\varepsilon_m, \varepsilon_g, \varphi_m^{Max}, \varphi_g^{Max}, k_{mc}, k_{mn}, k_{gm}, \chi_m, \chi_g, \rho$) involving two Monod-type equations ($\varphi_m(n,c), \varphi_g(m)$) govern the growth of methanotrophs and grazers. A subtle but important aspect of these types of models is that both $m(t)$ and $g(t)$ each represent a diverse community of methanotrophs and grazers, respectively. Consequently, different model parameterizations must be used for different community configurations, and each parameterization will produced different state dynamics (Fig. 2).

From nonequilibrium thermodynamics, it can be shown [35 pp. 131-141] that the rate of entropy production for a chemical reaction is $-r\Delta G_r/T$, where r is reaction rate (mole L^{-1} d^{-1}), ΔG_r is the Gibbs free energy of reaction (J $mole^{-1}$), and T is temperature (ºK). So the rate of entropy production for the model in Box 1 is approximately given by,

$$\dot{S}(t) = -\left[((\Delta G_1(1-\varepsilon_m) + \Delta G_2 \varepsilon_m)\varphi_m + \Delta G_3 \chi_m)m(t) + \Delta G_3((1-\varepsilon_g)\varphi_g + \chi_g)g(t)\right]/T \qquad (3)$$

where ΔG_i i = 1, 2 and 3, are the free energies of CH_4 oxidation to CO_2, CH_4 oxidation to CH_2O, and CH_2O oxidation to CO_2, respectively. Since the free energy of biomass synthesis is close to zero (slightly negative in fact), we treat biomass as glucose here for simplicity [9]. Not surprisingly, different parameterizations of the model (Fig. 2) produce entropy at different rates based on (3), and certain parameter choices maximize entropy production. For example, if we hold all parameters constant except k_{gm}, and plot entropy production for the steady state solution of Box 1 eqns. at a dilution rate of 1.0 d^{-1} with 10 μM DIN feed for differ values of k_{gm}, MEP occurs at k_{gm} = 30 μM (Fig. 3, blue line). If we keep N loading the same, but decrase dilution rate to 0.1 and 0.01 d^{-1} (100 and 1000 μM DIN), MEP occurs at k_{gm} values of 245 and 1080 μM, respectively (Fig. 3). Since k_{gm} affects $g(t)$, Fig 3 shows that too few or two many gazers will result in suboptimal entropy production. Manipulating any

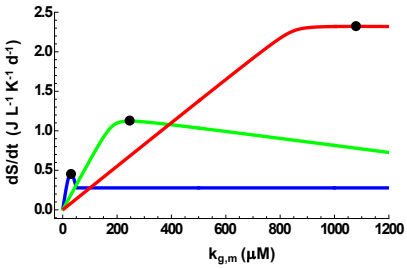

Fig. 3. SS entropy production at different dilution rates but with same N loading. MEP occurs at black dots.

kinetic growth parameter produces similar curves that exhibit maximums. We conclude from this simple example that only certain species and community configurations will maximize entropy production for a specified set of conditions; a random assortment of organisms will not. ***We hypothesize that communities will organize, and give time will evolve, towards the state of MEP.***

4.6 Distributed Metabolic Networks The numerical example above also illustrates the problems with current approaches to modeling biogeochemistry at the organismal level because no theory exists on how community composition changes as drivers change. As a result, current computational models are brittle to changes in communities. Instead of focusing at the organismal level, we will pursue a functional representation using a metabolic abstraction. Ideal MEP constraints for a biological system would arise from complete annotation of the metagenome to establish metabolic capabilities; although, predator preferences and prey avoidance mechanisms would also have to be accounted for. For complex natural communities, this type of information is still many years away; consequently, for this proposal we use a distributed metabolic network to represent community processes.

Although biologists have typically focused at the organism level, this perspective can hide aspects on ecosystem function. For instance, in examining terrestrial below ground processes, emphasis is often placed on the symbiotic relationship between plant and mycorrhizae. While understanding these interactions at the organism level is extremely important, at the functional scale, root, root hair, and mycorrhizae hyphae resemble a space filling fractal network [146] that effectively extracts resources from the soil matrix (Fig. 4). Similarly, the *Riftia*-sulfur bacteria symbiosis found at hydrothermal vents appears to be an efficient design for facilitating transport of H_2S plus O_2 laden water to sulfur bacteria reaction centers, which resemble a packed-bed bioreactor (Fig. 5) that effectively reduces mass transport limitations faced by bacteria [70]. The fact that these metabolic and structural organizations are distributed among species is not relevant to our focus. When looking at metabolic capabilities of microbial systems, we often find that metabolic function is distributed amongst all three domains of life (Bacteria, Archaea and Eukaryote), such as in methanogenic networks [143], microbial mats [139], Winogradsky columns, and almost anywhere redox cascades occur [99]. Since the species that comprise metabolic networks can undergo substantial substitution with only minor impact on functional characteristics [37, 150], we will view microbial systems as metabolic networks that can be distributed in space and time, but resemble multicellular organisms [118, 149]. The abstraction applies to higher organisms as well, but their functional contribution must be identified, such as mastication, filtration, transpiration, etc.

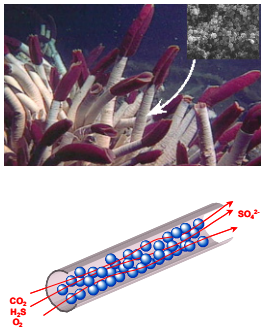

Fig. 5. *Riftia* with sulfur bacteria symbionts.

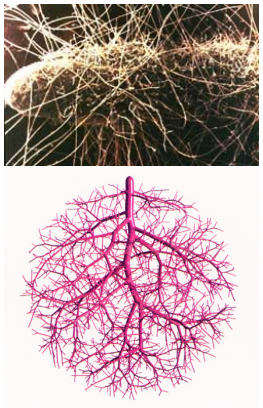
Fig. 4. Mycorrhizae colonizing root hair.

5. Hypotheses

Our overarching theory is that **biological systems organize to maximize entropy production subject to information and biophysicochemical constraints**. Based on theory developed in Section 4 above, we propose several hypotheses below that we will test experimentally. Furthermore, we will also develop a mathematical framework based on our theory to predict how biological systems organize to process energy and matter. We propose the following hypotheses:

H1. *Ecosystems adapt and evolve complementary trophic structures that maximize entropy production.* Perturbations that cause alterations in trophic structure (i.e., too few or too many grazers) temporarily reduces entropy production by the system. We will challenge microbially-based chemostats with alternative microbial communities to demonstrate optimum performance of the adapted community.

H2. *Removal of information from biological systems will cause them to operate at lower entropy production rates.* A biological system with reduced information may lack processes to fully dissipate available energy. To test this hypothesis we will use microbial chemostat experiments where information is removed by serial dilution. Entropy production will be compared to species

richness.

H3. *Information stored within the metagenome allows entropy production to be averaged over long timescales, so that biological systems can maintain entropy production during perturbations and periodic changes in system drivers.* We will test this hypothesis by comparing chemostat experiments with continuous energy input to those with periodic energy input and monitor entropy production and species richness.

The *Null* hypothesis is that biogeochemistry depends solely on what species are present. If species composition is altered, biogeochemistry and entropy production will change. Below we describe the experiments that will be used to test hypotheses H1-H3. In Section 7 we develop a mathematical model based on MEPP that we will use to predict biogeochemistry in our experimental systems.

6. Experimental Approach

6.1 Experimental Apparatus While our MEP-based theory is applicable to natural ecosystems, laboratory microbial systems are ideal for testing of our hypotheses for the following reasons: 1) even dilute microbial systems contain greater than 10^9 organisms with thousands to possibly tens of thousands of "species" per liter [122], thereby insuring many degrees of freedom and extensive metagenomic information; 2) they have fast characteristic timescales which allow us to complete experiments in months rather than decades; 3) they can be maintained in laboratory environments, extensively sampled and manipulated; 4) our new molecular techniques allow us to measure species richness to unprecedented depth [57, 122].

All of our experiments will involve methanotrophic communities and be conducted in replicate 3 L chemostats mixed with cell-culture impellers, outfitted with pH and DO probes, and maintained at 20ºC within dark environmental chambers. Although a similar system could be based on phototrophs, the methanotrophic system has the advantage that energy acquisition by the microbial community can be directly measured. Chemostats will be sparged with a 5% methane in air mixture via mass flow controllers that permit precise measurements of CO_2 production and CH_4 and O_2 consumption rates from on-line gas analyzers that employ laser diode absorption spectroscopy for CO_2 and O_2 (Oxigraf) or NDIR for CH_4 (CAI). A sterilized mineral salts medium (NO_3^-, PO_4^{3-}, salts, and trace metals [17]) of a specified concentration and dilution rate will serve as the input feed. Similar methanotrophic microcosms have been running continuously in our lab for more than three years now (Fig. 6). In addition to standard biogeochemistry measurements (see Section 6.6), our primary measurements will be entropy production rate and community composition.

Since no work is derived from the chemostats, all CH_4 oxidized will be dissipated as heat to the surroundings; consequently, the rate of entropy production can be readily calculated from changes in chemical composition between input and output flows and accounting for accumulation (or loss) in the reactors. Species richness of Bacteria and Archaea will be assessed using massively parallel 454 tag pyrosequencing of hypervariable regions of the rRNA gene (see Section 6.5). Direct microscopic counts will be used to assess species richness of eukaryotes, as well as new 454 approaches. We will also quantify methane monooxygenase genes via quantitative PCR (qPRC) relative to total protein, as our simple model (Box 1 above) shows that to maintain MEP the ratio of methanotroph to total biomass decreases as dilution rate decreases.

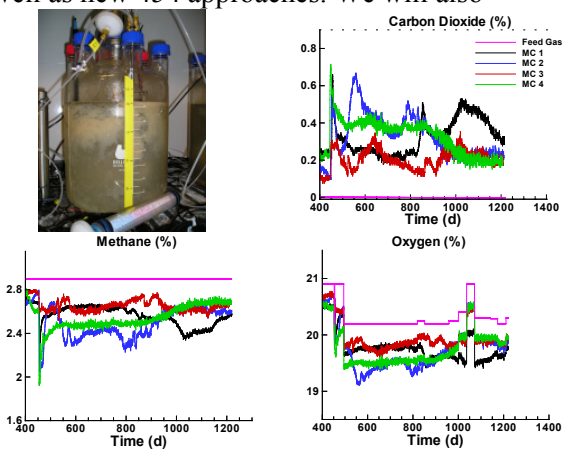

Fig. 6. Gas composition for four long-term semi-closed methanotrophic microcosms over last 2 years.

6.2 Experiment 1, Test of H1 As discussed in Section 4.5, communities organize towards a state of MEP, but the particulars of the community structure depend of external drivers, such as the dilution rate and nutrient loading in a chemostat. This experiment will involve three treatments in duplicate (6 chemostats in all) with the same N loading rate of 10 µmol N L^{-1} d^{-1}, but at the following dilution (D) rates: E1-T1: D=0.03 d^{-1}, 333 µM N; E1-T2: D=0.1 d^{-1}, 100 µM N; E1-T3: D=1.0 d^{-1}, 10 µM N. Each chemostat will be repetitively inoculated (5% v/v) every 2 weeks with

whole water collected from a local pond until addition of inoculum no long affects respiration, which will ensure development of a complex community essential for MEP. Once all chemostats have reached steady state (SS) (ca 3-5 mo.), we will determine entropy production (EP) rate in each treatment and assess community composition via 454-tag sequencing, DAPI counts for eukaryotes and qPCR for *pmoA*. We will also sample community composition once per month during startup phase to assess changes in community structures, as it is possible that more than one community configuration may produce the same amount of entropy. We predict that communities in each treatment will differ significantly but will be functionally stable and operate at MEP. Based on Fig. 3, we also predict EP to follow: T1 > T2 > T1. Following SS operation, perturbation experiments will begin to show that changes in community structure will cause a decrease in EP (e.g. Section 4.5 and Fig. 3).

The first set of perturbation experiments involve cross inoculation between treatments. One liter from one treatment will be added to a different treatment as follows: T1 $\rightarrow$ T2; T2 $\rightarrow$ T3; T3 $\rightarrow$ T1. Following cross inoculation, EP and community composition will be monitored for several months. Since the initial SS communities will be at MEP, the inoculums will not be competent and will simply be washed out. We will repeat the cross inoculation experiment in the reverse order as well (T1 $\rightarrow$ T3; T3 $\rightarrow$ T2; T2 $\rightarrow$ T1), but expect the same results. For the final perturbation, we will swap dilution rates between chemostats and cross inoculate at the same time. In this perturbation, we expect the inoculum to overtake the existing community as conditions for MEP will favor the inoculum. This experiment should take between 15-20 mo. to complete.

6.3 Experiment 2, Test of H2 In this experimental test of hypothesis H2, we will demonstrate that removal of information from the system by reducing species richness will result in lower entropy production. Two treatments, in duplicate, involving chemostats at 0.1 d^{-1} dilution rate and 10 µmol N L^{-1} d^{-1} loading will be started. Treatment E2-T1 will be inoculated with the full microbial community (we will likely use two chemostats from Exp 1). Treatment E2-T2 will be operated under the same conditions, but the inoculum will be serially diluted by a factor of 10^7. To minimize mortality, we will use filter sterilized waste medium collected from the E1-T2 chemostats from Exp. 1. At 10^7 dilution, there is only a 1% probability that cells at a concentration of 10,000 L^{-1} will be found in the inoculum. This will remove rare bacteria as well as remove large grazers such as ciliates and perhaps nanoflagellates. The inoculum will be grown under batch conditions until cell numbers approach typical values, then chemostat operation will begin. We will measure entropy production and species richness as above. We expect entropy production to be significantly reduced due to the removal of grazers that are necessary to maintain entropy production via resource cycling (Section 4.5). The experiment should run for 3-4 mo.

6.4 Experiment 3, Test of H3 In the last experiment we will make use of our existing methanotrophic microcosms (Fig. 6) to demonstrate that information contained within the metagenome allows entropy production to be averaged over time (Section 4.4). Microcosms will be cross inoculated to insure equal metagenomic content and will be converted to chemostat operation at 0.1 d^{-1} dilution rate with 10 µmole N L^{-1} d^{-1} loading. Methane feed to two of the four microcosms will then be placed on a 50% duty cycle, with two days CH_4 plus air and two days of just air (E3-T2). Based on hypothesis H3, entropy production in both treatments (E3-T1 and E3-T2) should be equal when averaged over a 4 d period once transient dynamics and system adaptation have completed (ca. 3-4 mo.). During the experiments microcosms will again be assessed for EP and community composition on an approximately monthly basis.

6.5 Determination of Microbial Population Structure The diversity of the microbial community (bacteria and archaea) in inocula and chemostats will be assessed using a 454 tag sequencing strategy that allows for extremely sensitive, relatively quick, and cost-effective screening of microbial diversity and generation of taxonomic inventories in individual samples [57, 122]. Sequence tags from hypervariable regions of small subunit ribosomal RNAs (SSU rRNA) provide descriptions of Operational Taxonomic Units (OTUs) in microbial communities. Nearly unique rRNA tag sequences serve as a proxy for individual "species" or OTUs [58]. Enumerating the number of different rRNA tags provides a detailed description of the relative occurrence of specific microbes in a sample. For the development of predictive frameworks that link biogeochemical processes with particular microbial populations, we must have statistically significant assessments of microbial diversity, relative abundance, and community structure. The 454 tag sequencing methodology deeply samples a microbial population, thus allowing robust statistical predictions about the community. It also includes detection of both dominant members and low abundance "rare biosphere" members [103, 122]. These rare organisms may become particularly

important under certain conditions in our chemostat experiments; hence, they may significantly contribute to system information. The strategy exploits the massively parallel, pyrosequencing capability of the Roche Genome Sequencer Titanium (GSTi) system to sample the hypervariable regions from >1,000,000 rRNAs in a single sequencing run without requiring the construction of recombinant clones or preparation of sequencing templates. We will use a "keyed" primer approach [57] that allows us to tailor our experiments to include communities that may need to be unequally sampled, such as the bacterial communities which may demand upwards of 50,000 sequences, in comparison to archaeal communities, which may only need ~5,000 sequences. DNA- and RNA-based methods will be used to assess both the total potential diversity of the system, as well as the diversity of only the active portion of the microbial community. At any single time point, we can extract RNA or DNA, carry out reverse transcription PCR or regular PCR, and obtain community profiles by 454 sequencing. In addition to examining ribosomal genes, we will also use PCR to follow changes in the abundance and diversity of methanotrophs. This strategy uses the key enzyme responsible for the oxidation of methane to methanol, particulate methane monooxygenase [17], which is encoded by the *pmoA* gene. This gene is highly conserved and present in all methanotrophs, therefore PCR primers can be designed for specifically amplifying *pmoA* out of environmental samples to determine their abundance and diversity [53, 83]. When sampling for total community profiles via 454, we will also carry out quantitative PCR of the *pmoA* gene. For select time points, we will clone and sequence the amplicons to examine how changes in abundance of methanotrophs relate to changes in their phylogenetic diversity.

For all samples, fluids will be collected on 0.22 μm Sterivex filter, DNA extracted according to Sinigalliano *et al*. [121], and PCR amplicon libraries built or qPCR carried out. Sequencing on the GSTi, qPCR, and data analysis will all be performed at the Bay Paul Center. Correspondence and other statistical analyses will be used to determine if complementary communities arise and how this relates to time of sampling or biogeochemical measurements using statistical tools in the R package (www.r-project.org). Since communities must recycle resources to attain MEP, we expect whole community shifts as opposed to simply removing one OTU and replacement by another.

6.6 Biogeochemical Measurements Concentration of the following nutrients will be measured to calculated entropy production and to allow model comparisons to observations (see Section 7.2): NO_2^- and NO_3^- (via Lachat QuikChem 8000 autoanalyzer); NH_4^+ [123]; PO_4^{3+} [87]; O_2 (by Winkler titration [73] and electrode); dissolved inorganic carbon (DIC) (via UIC Coulometrics [65] or GC); dissolved organic carbon (DOC) and nitrogen (TDN) (via a Shimadzu TOC-Vcph for carbon and TNM-1 unit for total nitrogen); particulate organic carbon (POC) and nitrogen (PON) (on Perkin Elmer 2400 CHN elemental analyzer); DAPI counts of bacteria and protists [106], and SYBR Green I counts of viruses [90]. Both dissolve (DOM) and particulate (POM) organic matter will be characterized for total protein [96], which is proportional to total biological structure ($\mathbb{S}_T$).

7. Development of a Metabolic Network Model Based on MEP

An abridged description of the modeling approach is discussed below. Space limits the details that we can include, so this section is only intended to provide overview. The basic model framework uses the metabolic network perspective discussed earlier, and a methanotrophic community in a batch reactor that is sparged with methane and air will serve as an example. The formulation is similar to how single cells control metabolism by regulating the synthesis and degradation of enzymes associated with pathways necessary for growth. However, at the ecosystem level, metabolism is distributed amongst many phyla, and expression of metabolic pathways often occurs via changes in microbial species composition that have differing metabolic capabilities.

7.1 Optimized Metabolic Ecosystem Network (OMEN) Model Although a multi-trophic community, which includes Bacteria, Archaea, phage, protists and other microbial grazers, is ultimately responsible for environmental chemical transformations, organisms are not explicitly modeled because of the lack of information and the dynamic and potentially chaotic nature of communities at the organism level [13, 48, 150]. Instead, a metabolic network synthesizes generic biological structure, $\mathbb{S}$, which consists largely of enzymatic protein, but also represents other macromolecules expressed by microbial communities for growth and form. The biological structure can be allocated to any reaction in the network and serves as the reaction's catalyst. The metabolic network also orchestrates energy and mass acquisition necessary to construct biological structure itself, where the MEP principle governs how biological structure is

allocated to any metabolic reaction. The mathematical framework is general and can be extended to any biological system by expanding the distributed metabolic network based on knowledge from classic microbiology and metagenomics. One can also examine how the introduction or evolution of new metabolic functions alters resource allocation and system dynamics as well as examine dynamics when metabolic functions are removed.

Table 1. Half reactions used for methanotroph metabolic network.

Reaction	Structure	Rate
$CH_4 + H_2O \rightarrow CH_2O + 4e^- + 4H^+$	$\mathbb{S}_1$	r_1
$CH_2O + 2H_2O \rightarrow CO_2 + 4e^- + 4H^+$	$\mathbb{S}_2$	r_2
$O_2 + 4e^- + 4H^+ \rightarrow 2H_2O$	$\mathbb{S}_3$	r_3
$NO_3^- + 8e^- + 10H^+ \rightarrow NH_4^+ + 3H_2O$	$\mathbb{S}_4$	r_4
$CH_2O + NH_4^+/\rho_S + 0.05e^- \rightarrow \mathbb{S}$	$\mathbb{S}_5$	r_5
$\mathbb{S} \rightarrow f_{aC}\, CH_2O + (1\text{-}f_{aC})dC + (f_{aN}\, NH_4^+ + (1\text{-}f_{aN})dN)/\rho_S + 0.05e^-$	$\mathbb{S}_6$	r_6
$dC \rightarrow CH_2O$	$\mathbb{S}_7$	r_7
$dN \rightarrow NH_4^+$	$\mathbb{S}_7$*	r_8

* Note, $\mathbb{S}_7$ catalyzes both reactions 7 and 8.

For our aerobic methanotrophic metabolic network example (Fig. 7), eight half reactions account for methane oxidation (r_1, r_2, r_3), including sugar biosynthesis (r_1), nitrate reduction (r_4) for biological structure synthesis from NH_4^+ and sugars (r_5), and biological structure and detritus degradation (r_6, r_7, r_8) (Table 1). Each of the eight reactions has associated biological structure ($\mathbb{S}_{1\ldots7}$), except for reactions 7 and 8, which are both catalyzed by the same structure, $\mathbb{S}_7$, to minimize model degrees of freedom. The use of half reactions (Table 1) increases the flexibility of the network to utilize available electron acceptors or donors, but to insure electron conservation all half reactions are coupled to an electron shuttle reaction as follows,

$$\varepsilon_i (2e^- + NAD^+ + H^+ \rightarrow NADH) \tag{4}$$

where ε_i is the number of electron pairs produced by reaction *i*. Gibbs free energies are calculated for each reaction in Table 1 after coupling to (4). We use the approach of Alberty [4-6] to calculate the standard Gibbs free energy of reaction, which accounts for proton dissociation equilibria between chemical species ($H_2CO_3 \Leftrightarrow H^+ + HCO_3^-$, etc.) at a specified pH and temperature. The overall free energy of reaction, denoted as $\Delta G_r^\dagger(r_i)$, accounts for the concentration of reactants and products ($\mathbf{c}(t)$), and ionic strength, I_S, is used to estimate activity coefficients [4]. For the methanotrophic network simulation, we use the following conditions: pH 7, I_S 1.92 μM, at 20ºC. Free energy, enthalpy and entropy of formation for biological structure is based on Battley [8, 10] for bacteria; however, this is not critical as the free energy of living organisms is similar in value to the substrates they are constructed from.

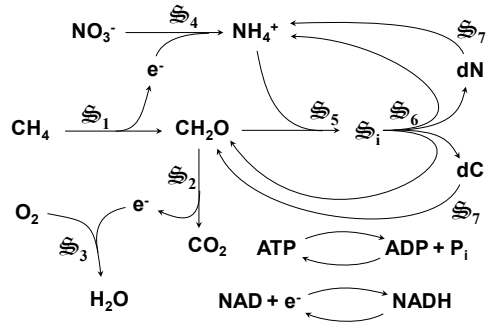

Fig. 7. Simple distributed metabolic network for methanotrophic-base food web. Seven biological structures, $\mathbb{S}_i$, catalyze the eight reactions.

Of course, $\Delta G_r^\dagger(r_i)$ represents the available or needed energy only if the reaction is run sufficiently close to equilibrium (i.e., reversibly), which is clearly not the case for biological systems. To account for inefficient energy transfer or energy production, all reactions are coupled to an energy source (or sink) reaction in the form of ATP hydrolysis (or synthesis) of the form,

$$\eta_i(t)(ATP + H_2O \rightarrow ADP + P_i). \tag{5}$$

Even though we know most ATP reaction couplings from biochemistry, cells have the ability to dissipate ATP [113], and different organisms in a community can have different ATP couplings for the same pathways [51]. Consequently, we treat energy coupling, $\eta_i(t)$, as a control variable to be determined by optimization. The combined whole reaction *i* consists of (4) plus (5) and half reaction *i* in Table 1. The Gibbs free energy of the combined reaction *i*, $\Delta G_c^\dagger(r_i)$, is then given by

$$\Delta G_c^\dagger(r_i, t) = \Delta G_r^\dagger(r_i) + \eta_i(t)\Delta G_r^\dagger(r_{ATP}), \tag{6}$$

where $\Delta G_r^{\dagger}(r_{ATP})$ is the Gibbs free energy of (5). By altering the magnitude and sign of $\eta_i(t)$, an endergonic reaction can be driven forward, and by inefficient coupling of (5) with exergonic reactions, energy can be extracted at less than 100% efficiency, allowing the reaction to proceed at higher rates. This is the standard tradeoff between power and efficiency; operating reactions at high throughput necessitates low efficiency, while attaining high efficiency limits reaction rate, which all biotic and abiotic system must contend with [44, 105].

Metabolic network reaction rates are given by,

$$r_i = \nu_i \mathbb{S}_i(t) F_i^K(\mathbf{c}) F_i^T(\mathbf{c}) \tag{7}$$

where ν_i is an experimentally determined rate constant per unit of biological structure allocated, $\mathbb{S}_i(t)$ is biological structure allocated to reaction *i*, and $F_i^K(\mathbf{c})$ and $F_i^T(\mathbf{c})$ are kinetic and thermodynamic forces, respectively. The kinetic force is given by the general form,

$$F_i^K(\mathbf{c}) = f_{NADH}(\zeta_{NADH}, \varepsilon_i) f_{ATP}(\zeta_{ATP}, \eta_i) \prod_j \left(\frac{c_j}{c_j + K_{i,j}} \right)^{\Phi_{i,j}} \tag{8}$$

where c_j is the concentration of substrate *j*, and the matrix element, $\Phi_{i,j}$, equals 1 if reaction *i* consumes substrate *j*; otherwise $\Phi_{i,j}$ is zero (Table 1). To ensure electron and energy conservation, we assume biological structure has a fixed amount of NAD+NADH and ADP+ATP storage per unit of total biological structure, $\mathbb{S}_T = \sum_i \mathbb{S}_i$, as specified by the constants ζ_{NADH}^{Tot} and ζ_{ATP}^{Tot}, respectively (μmol (mmol $\mathbb{S}_T$)$^{-1}$). The functions $f_{NADH}(\zeta_{NADH}, \varepsilon_i)$ and $f_{ATP}(\zeta_{ATP}, \eta_i)$ are simple Michaelis-Menten-like expressions of the electron, $\zeta_{NADH}(t)$, and energy, $\zeta_{ATP}(t)$, availability that constrain reactions that use ATP, ADP, NADH or NAD as determined by ε_i and $\eta_i(t)$.

Chemical reaction rates are often limited by kinetics, so the thermodynamic force is often ignored because it is usually close to unity. However, as the reaction approaches equilibrium, $F^T(\mathbf{c})$ approaches zero and constrains the net reaction rate no matter how favorable the reaction kinetics. It can be shown [16, 64] that the thermodynamic force is related to the Gibbs free energy of the reaction, $\Delta G_c^{\dagger}(r_i)$, as follows,

$$F_i^T(\mathbf{c}) = 1 - \exp\left(\frac{\Delta G_c^{\dagger}(r_i)}{RT\chi_i} \right) \text{ for } \Delta G_c^{\dagger}(r_i) \le 0, \tag{9}$$

where *R* is the gas constant, *T* is temperature (ºK) χ_i is the average stoichiometric number for net reaction *i* [16, 140]. The thermodynamic force is an important aspect of our model, since altering energy coupling, $\eta_i(t)$, allows the model to select from high-power, low-efficiency to low-power, high-efficiency modes. Because of the complexity of community metabolic networks, χ_i are treated as tunable parameters.

Reactions $r_5(t)$ and $r_6(t)$ represent biological structure synthesis and degradation, respectively. However, which of the seven biological structures to produce or degrade at time *t* is not specified. Consequently, we introduce an additional control variable, $\sigma_i(t)$, that specifies the partitioning of biological structure synthesis (analogous to transcription + translation) (Fig. 8). Degradation of biological structure is nonspecific and only depends on the relative concentration of $\mathbb{S}_i$ to total biological structure, $\mathbb{S}_T$, and the concentration of $\mathbb{S}_6$. Since degradation of biological structure in ecosystems is largely the result of grazing by higher trophic levels, the non specificity of degradation is closer to how natural food webs operate. In addition, biological structure is not perfectly degraded to building block materials (i.e., CH_2O and NH_4^+), but produces some detrital carbon (dC) and nitrogen (dN) (Fig. 7), as specified by the assimilation parameters f_{aC} and f_{aN}, respectively (Table 1). Breakdown of detrital C and N is controlled by reactions $r_7(t)$ and $r_8(t)$, but only one biological structure, $\mathbb{S}_7$, catalyzes both reactions.

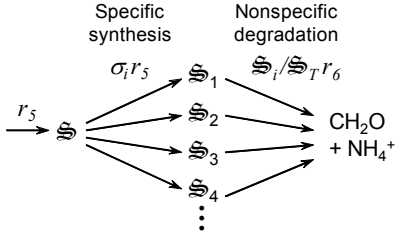

Fig. 8. Illustration of how partition function (σ_i) controls synthesis of biological structures.

Based on the metabolic network and reaction rates, a mass balance model can be constructed for the state variables, which represent the concentrations and partial pressures of chemical species, $\mathbf{c}(t)$, biological structures, $\mathbb{S}(t)$, as well as the system's redox and energy

states (ζ_{NADH} , ζ_{ATP}). In the methanotroph example $\mathbf{c}(t)$ include: [CH_4], [CH_2O], [O_2], [CO_2], [NH_4^+], [NO_3^-], [dC], [dN], p_{CH4}, p_{O2} and p_{CO2},. The general form of the state model is given by,

$$\frac{d[\mathbf{c}(t), \mathbf{s}(t), \boldsymbol{\zeta}(t)]^T}{dt} = \mathbf{f}(\mathbf{r}(t), \mathbf{r}_A(t), \mathbf{q}(t), \mathbf{u}(t), t) \tag{10}$$

where $\mathbf{r}(t)$ is the vector of reaction rates given by (7), $\mathbf{r}_A(t)$ is a vector of abiotic reaction rates, $\mathbf{q}(t)$ is a vector of external sources, such as methane feed, and $\mathbf{u}(t)$ is the vector of the two control functions: $\mathbf{u}(t) = [\boldsymbol{\eta}(t), \boldsymbol{\sigma}(t)]^T$. Consequently, once $\mathbf{u}(t)$ is specified, (10) can be solved for $\mathbf{c}(t)$, $\mathbf{s}(t)$ and $\boldsymbol{\zeta}(t)$.

The control variables, $[\boldsymbol{\eta}(t), \boldsymbol{\sigma}(t)]^T$, are determined by formulating and solving an interval optimization problem in which average entropy production rate, $\langle dS/dt \rangle_j$, is maximized over a specified interval of time, δ_t . Entropy production is given by the negative of a reaction rate times the Gibbs free energy of the reaction divided by temperature, summed over all reactions [35 pp. 131-141], which is readily calculated. The general form of the optimal control problem is as follows:

$$\begin{array}{ll} & \text{Subject to :} \\ \max \quad \left\langle \frac{dS}{dt} \right\rangle_j = \frac{-1}{T\delta_t} \int_{t_j}^{t_j+\delta_t} \sum_i r_i(\tau) \Delta G_c^{\dagger}(r_i(\tau)) d\tau & \boldsymbol{\eta}_L \le \boldsymbol{\eta}(t) \le \boldsymbol{\eta}_U \\ \text{wrt} \quad \boldsymbol{\eta}(t), \boldsymbol{\sigma}(t) & 0 \le \sigma_i(t) \le 1 \\ & \sum_i \sigma_i(t) = 1 \end{array} \tag{11}$$

To obtain a solution over a specified time domain, $[t_0, t_f]$, (11) is solved repetitively over a sufficient number of intervals, $[t_j, t_j + \delta_t]$, to cover the entire domain. There are several ways of solving (11), including linear programming following linearization at time, *t* [137, 138]. However, we have recently implemented an interval optimization method using SNOPT [43] that solves the nonlinear programming problem over the interval, δ_t via sequential quadratic programming coupled with block implicit methods to solve the associated differential equations (10) [19]. In the current implementation the optimal interval, δ_t , is subdivided into n_g grid points, and the control functions $[\boldsymbol{\eta}(t), \boldsymbol{\sigma}(t)]^T$ are discretized over the interval as linear piecewise continuous functions. The length of the optimization interval, δ_t , is an interesting aspect of the model that relates to hypothesis H3, as it reflects the characteristic time scale of environmental variability that a living system has evolved to cope with over a given spatial scale [92]. We expect that microbial systems can be described by short optimal intervals (~days), while in forest ecosystems δ_t would be at least the length of the four seasons (~year). It is also possible to cast the problem as a type of infinite horizon [22] or receding horizon [59] optimal control problem, which we plan to investigate. Figures 9 and 10 illustrate a solution to (11) for our methanotrophic microcosms at zero dilution rate using a 8 d optimization interval.

7.2 Model Comparison with Experiments During all three experiments, we will extensively characterize the chemostats so that we can compare model output to observations. Changes in chemical compositions will also allow us to determine overall reaction rates and consequently entropy production rate in all experiments, which will also be compared to model predictions. The primary model outputs (see Figs. 9 and 10) are 1) concentration and pressures of nutrients and gasses, 2) concentration of biological structures, 3) metabolic network reaction rates (Fig. 7 and Eq. 7), 4) current metabolic expression (σ_i partition coefficients) and 5) reaction efficiencies (η_i variables).

During the first phase of Experiment 1, model parameters will be calibrated based on experimental observations. The perturbation experiments will then test the models ability to capture changes in community structure and biogeochemistry that will occur in the final phase of Exp. 1, without model recalibration. We will compare our MEP-based OMEN model to a standard compartment model (e.g., Box 1 above), which will be similarly calibrated. A model based on fundamental principles, such as MEP, should outperform an empirically-based model under perturbations which neither model has been calibrated for.

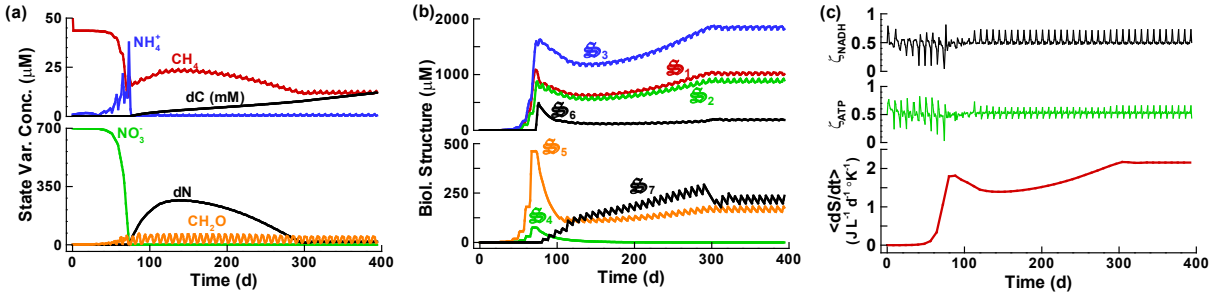

Fig. 9. Concentration of state variables (a), biological structures (b), NADH and ATP storage, and entropy production rate with eight day entropy production averaging.

Model parameters requiring calibration include ν_i, χ_i, and $K_{i,j}$ in (7), (9) and (8), respectively. Data from Experiment 1 will be use to calibrate model parameters using data assimilation techniques we have developed [134]. The 454 tag sequences will be used to estimate OTUs

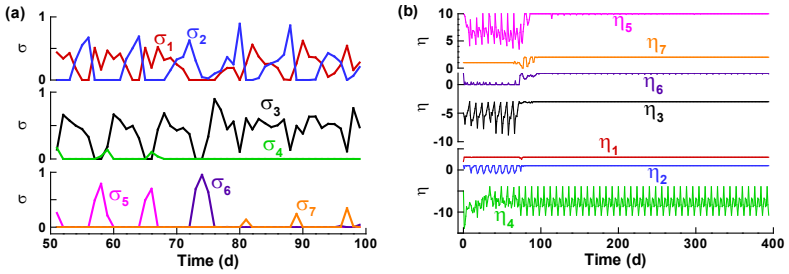

Fig. 10. (a) Biological structure partitioning and (b) reaction energy coupling control variables. Note, σ plotted for t = 60 to 100 d.

and species richness, and queried against a reference database to extract information about taxonomic identity where possible. This information along with *pmoA* activity and total protein will be related to predictions of biological structure allocations, $\mathbb{S}_i(t)$.

8. Broader Impacts, Education and Postdoctoral Student Mentoring

We have proposed that biological systems evolve towards a state of maximum entropy production as constrained by resources, biophysicochemistry and evolved useful information. While we have chosen microbial systems to test this theory, it applies across a broad range of biological organization and has many implications. Our use of information [2] to distinguish biotic from abiotic processes allows predictions on invasive species and introduces the importance of timescales. In order for an exotic species to be successful, it must cause entropy production (EP) to increase over the pre-invasion EP at some timescale. If an exotic flourishes by oxidizing existing biological structure, then EP will spike, followed by a crash in a steepest decent manner that is analogous to a forest fire. Such an exotic will not be persistent, because the original system had greater EP over longer timescales. However, if the exotic increases averaged EP over that of the original system, then it is likely to be persistent and difficult to eradicate. It is also clear that useful information [2], not Shannon information, should be the basis for assessing how biodiversity relates to ecosystem function, because a random collection of organisms, which would have a high Shannon index, is unlikely to be complementary and exhibit high EP.

The MEP principle as described in this proposal provides a definitive direction to evolution, but allows–requires–a multitude of pathways. The MEPP also has interesting connections to the metabolic theory of ecology (MTE) [146]. Since MTE (or similar theories [3]) shows that respiration scales to the 3/4 power of body mass, a kg of bacteria has a much higher respiration than a kg of whale, so why would large organisms evolve, as they appear to waste resources (consider ν in Eq. (7))? One possibility is that they provide stability, as they act as a means of internal energy storage that is necessary to maintain EP during external perturbations. While larger organisms often perish during perturbations, their biomass provides energy to maintain hierarchical structure and EP in the system over the perturbation, which may explain why biomass specific respiration decreases with increases with ecosystem maturity [93]. The MEPP connects to society as well, as the human planetary dominance is solely a result of our ability to dissipate the tiny amounts of energy that were stored, not dissipated, during transient events in Earth's history, such as during the Carboniferous period. Under MEPP, natural systems will never conserve energy, but excel at recycling resources if they limit entropy production. Clearly, if MEPP underlies the organization of biological systems, it will have wide spread application.

To facilitate application of MEP theory across scientific fields and to introduce the concept to non scientists, we plan to dedicate 25% of the half-time RA II to the development of a MEP web site. All computational models developed during the project will be distributed, with documentation, on the web site, as well as all chemostat data and general advancements and applications of MEPP to biology. All sequence data will be submitted to NCBI and made available at our Visualization and Analysis of Microbial Population Structures (VAMPS) website (http://vamps.mbl.edu), which will be cross linked to the MEP web site. The other function of the MEP web site will be educational outreach to K-12 and higher. To provide feedback on the usefulness of the educational content developed, we will seek input from the PIE-LTER schoolyard program participants, as one of us (Vallino) is a CO-PI on the PIE project. We will use Winogradsky columns for in-class discussion and participation, as these systems nicely illustrate the complexity of microbial processes.

This project will also allow us to enhance our teaching in the Semester in Environmental Science (SES) undergraduate program (http://ecosystems.mbl.edu/SES/). The SES program has a strong history of drawing underrepresented groups in science, as classes have been on average 84% women and several minority colleges and universities belong to the SES consortium. Starting in 2006 we began the URGES (Under-Represented Groups in Environmental Science) Program, which supports two minority students together with a faculty member on sabbatical leave to participate jointly in SES. The URGES program is intended to allow faculty members from minority colleges and universities to augment their curriculum at their home institutions based on the SES program. Both PIs currently co-teach a laboratory/lecture course on microbial methods in ecology as part of SES, and we will use the chemostat experiments as a model system in our methods course as well as introduce new methods based on our research. This project would also allow us to bring molecular techniques along with our chemostat experiments to the independent research projects that all SES students conduct during the last five weeks of the program.

8.1 Mentoring of Postdoctoral Student This project will support one postdoctoral student for the duration of the project and will be co-advised by both PIs. The postdoc will have a unique opportunity to engage scientists from both the Ecosystems Center and the Bay Paul Center for Comparative Molecular Biology and Evolution, which we will facilitate. The postdoc will actively be involved in the development of proposals, and will be allowed to gain teaching experience via our SES Methods course. While it is unlikely we will find a student with expertise in both molecular biology and computational thermodynamic modeling, the student will gain such experience. Of course the student will be required to present results at professional meetings and take the lead in manuscript writing.

9. Results of Prior NSF Support.

LTER-Plum Island Sound Comparative Ecosystems Study. OCE–9726921, 8/98-7/04: $4,130,000 and OCE-0423565 8/04-7/09: $4,446,385. PI's: *Hopkinson, Giblin,* ***Vallino****, Hobbie, Peterson, Deegan, Morris, Vörösmarty, Buchsbaum*. During the two projects, we have designed and implemented a comprehensive, long-term study of a major, coupled, land-estuarine system in the Acadian biogeographic province in eastern New England. The PIE LTER seeks to develop a predictive understanding of the long-term dynamics of watershed and estuarine ecosystems at the land-sea interface and to apply this knowledge to the wise management and development of policy to protect the natural resources of the coastal zone. The project has resulted in 175 peer-reviewed publications and 20 theses or dissertations since funding began in 1998. Research at the site has involved over 65 individuals from over a dozen institutions. Vallino leads the modeling effort for the PIE-LTER project and has contributed 16 publications [25, 28, 39, 42, 54-56, 60, 127, 134-138, 147, 148]. **Broader Impacts** – PIE LTER has developed a substantial education and outreach program. Our education program consists of Schoolyard, Undergraduate, Graduate and Post-Graduate components. Our schoolyard program alone involves about 1500 students and 42 teachers in the present school year. Our strategies at the undergraduate level include summer research internships (4-7/yr), developing new, LTER-based curriculum, undergraduate research projects and senior theses, guest lectures and LTER sponsored field trips, and research experiences/collaborations for college faculty. Graduate students from Clark University, UNH and USC have been active LTER participants. We have also hosted several foreign graduate students: 2 Denmark, 1 Portugal.

10. References

1. Adami,C. (2002) Sequence complexity in Darwinian evolution. *Complexity* **8**, (2) 49-56, doi: 10.1002/cplx.10071.

2. Adami,C., Ofria,C., and Collier,T.C. (2000) Evolution of biological complexity. *Proceedings of the National Academy of Sciences of the United States of America* **97**, (9) 4463-4468.

3. Agutter,P. and Wheatley,D. (2004) Metabolic scaling: consensus or controversy? *Theoretical Biology and Medical Modelling* **1**, (1) 13, doi: 10.1186/1742-4682-1-13.

4. Alberty,R.A. (2003) *Thermodynamics of biochemical reactions*, Hoboken, NJ, Wiley & Sons, 397 pp.

5. Alberty,R.A. (2006) *Biochemical thermodynamics: Applications of Mathematica*, Hoboken, NJ, Wiley & Sons, 464 pp.

6. Alberty,R.A. (2006) Thermodynamic Properties of Weak Acids Involved in Enzyme-Catalyzed Reactions. *J.Phys.Chem.B* **110**, (10) 5012-5016.

7. Armstrong,R.A. and McGehee,R. (1980) Competitive Exclusion. *Amer.Nat.* **115**, (2) 151-170.

8. Battley,E. (2003) Absorbed heat and heat of formation of dried microbial biomass: Studies on the thermodynamics of microbial growth. *Journal of Thermal Analysis and Calorimetry* **74**, (3) 709-721, doi: 10.1023/B:JTAN.0000011003.43875.0d.

9. Battley,E.H. (1998) The development of direct and indirect methods for the study of the thermodynamics of microbial growth. *Thermochimica Acta* **309**, (1-2) 17-37.

10. Battley,E.H. (1999) An empirical method for estimating the entropy of formation and the absolute entropy of dried microbial biomass for use in studies on the thermodynamics of microbial growth. *Thermochimica Acta* **326**, (1-2) 7-15.

11. Battley,E.H. (1999) On entropy and absorbed thermal energy in biomass; a biologist's perspective. *Thermochimica Acta* **331**, (1) 1-12.

12. Bejan,A. (2007) Constructal theory of pattern formation. *Hydrology and Earth System Sciences* **11**, (2) 753-768.

13. Beninca,E., Huisman,J., Heerkloss,R., Johnk,K.D., Branco,P., Van Nes,E.H., Scheffer,M., and Ellner,S.P. (2008) Chaos in a long-term experiment with a plankton community. *Nature* **451**, (7180) 822-825, doi: 10.1038/nature06512.

14. Berman-Frank,I. and Dubinsky,Z. (1999) Balanced growth in aquatic plants: Myth or reality? *BioScience* **49**, (1) 29-37.

15. Bhalla,U.S. and Iyengar,R. (1999) Emergent Properties of Networks of Biological Signaling Pathways. *Science* **283**, (5400) 381-387, doi: 10.1126/science.283.5400.381.

16. Boudart,M. (1976) Consistency between kinetics and thermodynamics. *J.Phys.Chem.* **80**, (26) 2869-2870.

17. Bowmen,J. (2000) The Methanotrophs—The Families Methylococcaceae and Methylocystaceae. In: *The Prokaryotes: An evolving electronic resource for the microbiological community*, Edited by Dworkin,M., New York, Springer, http://141.150.157.117:8080/prokPUB/index.htm.

18. Brookfield,J.F.Y. (1998) Quorum Sensing and Group Selection. *Evolution* **52**, (5) 1263-1269.

19. Brugnano,L. and Magherini,C. (2004) The BiM code for the numerical solution of ODEs. *Journal of Computational and Applied Mathematics* **164-165**, 145-158, doi: 10.1016/j.cam.2003.09.004.

20. Caldwell,D.E. (1999) Post-modern ecology - is the environment the organism? *Environmental Microbiology* **1**, (4) 279-281.

21. Caldwell,D.E., Wolfaardt,G.M., Korber,D.R., and Lawrence,J.R. (1997) Do bacterial communities transcend Darwinism? *Adv.Microb.Ecol.* **15**, 105-191.

22. Carlson,D.A., Haurie,A.B., and Leizarowitz,A. (1991) *Infinite horizon optimal control: Deterministic and stochastic systems*, 2 Ed., Berlin, Springer-Verlag, 332 pp.

23. Choi,J.S., Mazumder,A., and Hansell,R.I.C. (1999) Measuring perturbation in a complicated, thermodynamic world. *Ecol.Modelling* **117**, (1) 143-158, doi: 10.1016/S0304-3800(99)00042-3.

24. Christen,T. (2006) Application of the maximum entropy production principle to electrical systems. *Journal of Physics D: Applied Physics* **39**, (20) 4497-4503.

25. Claessens,L., Hopkinson,C., Rastetter,E., and Vallino,J. (2006) Effect of historical changes in land use and climate on the water budget of an urbanizing watershed. *Water Resour.Res.* **42**, (3) 1-13, doi: 10.1029/2005WR004131.

26. Clausius,R. (1850) Ueber die bewegende Kraft der Wärme und die Gesetze, welche sich daraus für die Wärmelehre selbst ableiten lassen. *Annalen der Physik und Chemie* **155**, (3) 368-397, doi: 10.1002/andp.18501550306.

27. Crooks,G.E. (2007) Beyond Boltzmann-Gibbs statistics: Maximum entropy hyperensembles out of equilibrium. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)* **75**, (4 (041119)) 1-5, doi: 10.1103/PhysRevE.75.041119.

28. Deegan,L.A., Bowen,J.L., Drake,D., Fleeger,J.W., Friedrichs,C.T., Kari,A., Hobbie,J.E., Hopkinson,C., Johnson,D.S., Johnson,J.M., LeMay,L.E., Miller,E., Peterson,B.J., Picard,C., Sheldon,S., Sutherland,M., Vallino,J., and Warren,R.S. (2007) Susceptibility of salt marshes to nutrient enrichment and predator removal. *Ecological Applications* **17**, (sp5) S42-S63.

29. Dewar,R. (2003) Information theory explanation of the fluctuation theorem, maximum entropy production and self-organized criticality in non-equilibrium stationary states. *Journal of Physics A: Mathematical and General* **36**, 631-641.

30. Dewar,R.C. (2005) Maximum entropy production and non-equlibrium statistical mechanics. In: *Non-equilibrium thermodynamics and the production of entropy: life, earth and beyond*, 41-55. Edited by Kleidon,A. and Lorenz,R.D., Berlin, Springer-Verlag.

31. Dewar,R.C. (2005) Maximum entropy production and the fluctuation theorem. *Journal of Physics A: Mathematical and General* **38**, (21) L371-L381, doi: 10.1088/0305-4470/38/21/L01.

32. Dewar,R.C. and Porte,A. (2008) Statistical mechanics unifies different ecological patterns. *Journal of Theoretical Biology* **251**, (3) 389-403, doi: 10.1016/j.jtbi.2007.12.007.

33. Edelman,G.M. and Gally,J.A. (2001) Degeneracy and complexity in biological systems. *PNAS* **98**, (24) 13763-13768.

34. Elser,J.J., Sterner,R.W., Gorokhova,E., Fagan,W.F., Markow,T.A., Cotner,J.B., Harrison,J.F., Hobbie,J.E., Odell,G.M., and Weider,L.W. (2000) Biological stoichiometry from genes to ecosystems. *Ecology Letters* **3**, (6) 540-550.

35. Eu,B.C. (1992) *Kinetic theory and irreversible thermodynamics*, Montreal, John Wiley & Sons Canada, Ltd., 752 pp.

36. Fath,B.D., Patten,B.C., and Choi,J.S. (2001) Complementarity of Ecological Goal Functions. *Journal of Theoretical Biology* **208**, (4) 493-506.

37. Fernandez,A.S., Hashsham,S.A., Dollhopf,S.L., Raskin,L., Glagoleva,O., Dazzo,F.B., Hickey,R.F., Criddle,C.S., and Tiedje,J.M. (2000) Flexible community structure correlates with stable community function in methanogenic bioreactor communities perturbed by glucose. *Appl.Microbiol.Biotechnol.* **66**, (9) 4058-4067.

38. Fernandez,A., Huang,S., Seston,S., Xing,J., Hickey,R., Criddle,C., and Tiedje,J. (1999) How Stable Is Stable? Function versus Community Composition. *Applied and Environmental Microbiology* **65**, (8) 3697-3704.

39. Filoso,S., Vallino,J.J., Hopkinson,C., Rastetter,E., and Claessens,L. (2004) Modeling Nitrogen Processing and Transport in the Ipswich River Basin, Massachusetts, Using HSPF: Present Conditions and Future Scenarios. *Journal of the American Water Resources Association* **40**, (5) 1365-1384.

40. Gans,J., Wolinsky,M., and Dunbar,J. (2005) Computational Improvements Reveal Great Bacterial Diversity and High Metal Toxicity in Soil. *Science* **309**, (5739) 1387-1390.

41. Gibbs,J.W. (1902) *Elementary principles of statistical mechanics*, New Haven, C. Scribner's sons, 207 pp.

42. Giblin,A.E. and Vallino,J. (2003) The role of models in addressing coastal eutrophication problems. In: *The Role of Models in Ecosystem Science*, 327-343. Edited by Canham,C.D., Cole,J.J., and Lauenroth,W.K., Princeton, NJ, Princeton University Press.

43. Gill,P.E., Murray,W., and Saunders,M.A. (2005) SNOPT: An SQP Algorithm for Large-Scale Constrained Optimization. *SIAM Review* **47**, (1) 99-131.

44. Gnaiger,E. (1990) Concepts on efficiency in biological calorimetry and metabolic flux control. *Thermochimica Acta* **172**, 31-52.

45. Goodnight,C.J. (1990) Experimental studies of community evolution I: The response to selection at the community level. *Evolution* **44**, (6) 1614-1624.

46. Goodnight,C.J. (1990) Experimental studies of community evolution II: The ecological basis of the response to community selection. *Evolution* **44**, (6) 1625-1636.

47. Gottschalk,G. (1986) *Bacterial metabolism*, 2 Ed., New York, Springer-Verlag

48. Graham,D.W., Knapp,C.W., Van Vleck,E.S., Bloor,K., Lane,T.B., and Graham,C.E. (2007) Experimental demonstration of chaotic instability in biological nitrification. *ISME J* **1**, (5) 385-393, doi: 10.1038/ismej.2007.45.

49. Handelsman,J., Rondon,M.R., Brady,S.F., Clardy,J., and Goodman,R.M. (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chemistry & Biology* **5**, (10) R245-R249.

50. Harte,J., Zillio,T., Conlisk,E., and Smith,A.B. (2008) MAXIMUM ENTROPY AND THE STATE-VARIABLE APPROACH TO MACROECOLOGY. *Ecology* **89**, (10) 2700-2711, doi: 10.1890/07-1369.1.

51. Helling,R.B. (2002) Speed versus Efficiency in Microbial Growth and the Role of Parallel Pathways. *J.Bacteriol.* **184**, (4) 1041-1045, doi: 10.1128/jb.184.4.1041-1045.2002.

52. Hill,A. (1990) Entropy production as the selection rule between different growth morphologies. *Nature* **348**, (Nov 29) 426-428, doi: 10.1038/348426a0.

53. Holmes,A.J., Costello,A., Lidstrom,M.E., and Murrell,J.C. (1995) Evidence that particulate methane monooxygenase and ammonia monooxygenase may be evolutionarily related. *FEMS Microbiol.Lett.* **132**, (3) 203-208.

54. Hopkinson,C.S., Buffam,I., Hobbie,J., Vallino,J.J., Perdue,E.M., Eversmeyer,B., Prahl,F., Covert,J., Hodson,R., Moran,M.A., Smith,E., Baross,J., Crump,B., Findlay,S., and Foreman,D. (1998) Terrestrial inputs of organic matter to coastal ecosystems: an intercomparison of chemical characteristics and bioavailability. *Biogeochemistry* **43**, 211-234.

55. Hopkinson,C.S. and Vallino,J.J. (1994) Toward the development of generally applicable models of the microbial loop in aquatic ecosystems. *Microb.Ecol.* **28**, 321-326.

56. Hopkinson,Jr.C.S. and Vallino,J.J. (1995) The relationships among man's activities in watersheds and estuaries: A model of runoff effects on patterns of estuarine community metabolism. *Estuaries* **18**, (4) 598-621.

57. Huber,J.A., Welch,D.B.M., Morrison,H.G., Huse,S.M., Neal,P.R., Butterfield,D.A., and Sogin,M.L. (2007) Microbial Population Structures in the Deep Marine Biosphere. *Science* **318**, (5847) 97-100, doi: 10.1126/science.1146689.

58. Huse,S.M., Dethlefsen,L., Huber,J.A., Welch,D.M., Relman,D.A., and Sogin,M.L. (2008) Exploring Microbial Diversity and Taxonomy Using SSU rRNA Hypervariable Tag Sequencing. *PLoS Genet* **4**, (11 (e1000255)) 1-10, doi: 10.1371/journal.pgen.1000255.

59. Ito,K. and Kunisch,K. (2002) Receding horizon optimal control for infinite dimensional systems. *ESAIM: Control, Optimisation and Calculus of Variations* **8**, 741-760, doi: 10.1051/cocv:2002032.

60. Jay,D.A., Uncles,R.J., Largier,J., Geyer,W.R., Vallino,J., and Boynton,W.R. (1997) A review of recent developments in estuarine scalar flux estimation. *Estuaries* **20**, (2) 262-280.

61. Jaynes,E.T. (1957) Information theory and statistical mechanics. *Phys.Rev.* **106**, (4) 620-630.

62. Jaynes,E.T. (2003) *Probability theory: The logic of science*, Cambridge, Cambridge University Press, 758 pp.

63. Jenkins,S.A. and Inman,D.L. (2006) Thermodynamic solutions for equilibrium beach profiles. *J.Geophys.Res.* **111**, (C2) 1-21, doi: 10.1029/2005JC002899, 2006.

64. Jin,Q. and Bethke,C.M. (2003) A New Rate Law Describing Microbial Respiration. *Applied and Environmental Microbiology* **69**, (4) 2340-2348.

65. Johnson,K.M., Wills,K.D., Butler,B.D., Johnson,W.K., and Wong,C.S. (1993) Coulometric total carbon dioxide analysis for marine studies: Maximizing the performance of an automated gas extraction system and coulometric detector. *Mar.Chem.* **44**, 167-187.

66. Jorgensen,S.E. (1994) Review and comparison of goal functions in system ecology. *Vie Milieu* **44**, (1) 11-20.

67. Jorgensen,S.E., Patten,B.C., and Straskraba,M. (2000) Ecosystems emerging: 4. growth. *Ecol.Modelling* **126**, (2-3) 249-284.

68. Katare,S. and Venkatasubramanian,V. (2001) An agent-based learning framework for modeling microbial growth. *Engineering Applications of Artificial Intelligence* **14**, (6) 715-726, doi: 10.1016/S0952-1976(02)00015-5.

69. Khoshmanesh,A., Hart,B.T., Duncan,A., and Beckett,R. (2002) Luxury uptake of phosphorus by sediment bacteria. *Water Res.* **36**, (3) 774-778, doi: 10.1016/S0043-1354(01)00272-X.

70. Kiorboe,T., Ploug,H., and Thygesen,U.H. (2001) Fluid motion and solute distribution around sinking aggregates. I. Small-scale fluxes and heterogeneity of nutrients in the pelagic environment. *Mar.Ecol.Prog.Ser.* **211**, 1-13.

71. Kleidon,A. (2006) Quantifying the biologically possible range of steady-state soil and surface climates with climate model simulations. *Biologia* **61**, S234-S239.

72. Kleidon,A., Fraedrich,K., Kunz,T., and Lunkeit,F. (2003) The atmospheric circulation and states of maximum entropy production. *Geophys.Res.Lett.* **30**, (23) 1-4, doi: 10.1029/2003GL018363.

73. Knapp, G. P., Stalcup, M. C., and Stanley, R. J. (1990) Automated Oxygen Titration and Salinity Determination. **Tecnical Report WHOI-90-35**, Woods Hole Oceanographic Institution, Woods Hole, MA, 25 pp.

74. Kolmogorov,A.N. (1965) Three approaches to the quantitative definition of information . *Problems of Information Transmission* **1**, (1) 1-7.

75. Kondepudi,D. and Prigogine,I. (1998) *Modern thermodynamics: From heat engines to dissipative structures*, New York, Wiley & Sons, 486 pp.

76. Lambert,F.L. (1999) Shuffled cards, messy desks, and disorderly dorm rooms - Examples of entropy increase? Nonsense! *Journal of Chemical Education* **76**, 1385-1387.

77. Lineweaver,C.H. and Egan,C.A. (2008) Life, gravity and the second law of thermodynamics. *Physics of Life Reviews* **5**, (4) 225-242, doi: 10.1016/j.plrev.2008.08.002.

78. Lorenz,R.D., Lunine,J.I., and Withers,P.G. (2001) Titan, Mars and Earth : Entropy production by latitudinal heat transport. *Geophys.Res.Lett.* **28**, (3) 415-418.

79. Lorenz,R. (2003) COMPUTATIONAL MATHEMATICS: Full Steam Ahead-Probably. *Science* **299**, (5608) 837-838.

80. Lotka,A.J. (1922) Contribution to the Energetics of Evolution. *PNAS* **8**, (6) 147-151.

81. Margalef,R. (1968) *Perspectives in ecological theory*, Chicago, University of Chicago Press, 112 pp.

82. Martyushev,L. (2007) Some interesting consequences of the maximum entropy production principle. *Journal of Experimental and Theoretical Physics* **104**, (4) 651-654, doi: 10.1134/S1063776107040152.

83. McDonald,I.R. and Murrell,J.C. (1997) The particulate methane monooxygenase gene pmoA and its use as a functional gene probe for methanotrophs. *FEMS Microbiol.Lett.* **156**, (2) 205-210, doi: 10.1111/j.1574-6968.1997.tb12728.x.

84. Mejer,H. and Jorgensen,S.E. (1979) Exergy and ecological buffer capacity. In: *State-of-the-art in ecological modelling*, 829-846. Edited by Jorgensen,S.E., London, Pergamon Press.

85. Morowitz,H.J. (1968) *Energy flow in biology: biological organization as a problem in thermal physics*, New York, Academic Press, 179 pp.

86. Moroz,A. (2008) On a variational formulation of the maximum energy dissipation principle for non-equilibrium chemical thermodynamics. *Chemical Physics Letters* **457**, (4-6) 448-452, doi: 10.1016/j.cplett.2008.04.050.

87. Murphy,J. and Riley,J.P. (1962) A modified single solution method for the determination of phosphate in natural waters. *Anal.Chim.Acta* **27**, 31-36.

88. Murray,B.G., Jr. (2001) Are ecological and evolutionary theories scientific? *Biological Reviews* **76**, (2) 255-289, doi: 10.1017/S146479310100567X.

89. Nielsen,S.N. and Ulanowicz,R.E. (2000) On the consistency between thermodynamical and network approaches to ecosystems. *Ecol.Modelling* **132**, (1-2) 23-31.

90. Noble,R.T. and Fuhrman,J.A. (1998) Use of SYBR Green I for rapid epifluorescence counts of marine viruses and bacteria. *Aquat.Microb.Ecol.* **14**, (2) 113-118.

91. Nowak,M.A. (2006) Five Rules for the Evolution of Cooperation. *Science* **314**, (5805) 1560-1563, doi: 10.1126/science.1133755.

92. O'Neill, R. V., DeAngelis, D. L., Waide, J. B., and Allen, T. F. H. (1986) A hierarchical concept of ecosystems. Monographs in population biology, Princeton, Princeton University Press, **23,** 1-253.

93. Odum,E.P. (1969) The strategy of ecosystem development. *Science* **164**, (3877) 262-270.

94. Odum,H.T. and Pinkerton,R.C. (1955) Time's speed regulator: the optimum efficiency for maximum power output in physical and biological systems. *Am.Sci.* **43**, 321-343.

95. Odum,H.T. (1988) Self-Organization, Transformity, and Information. *Science* **242**, (4882) 1132-1139.

96. Ogunseitan,O.A. (1997) Direct extraction of catalytic proteins from natural microbial communities. *J.Microb.Methods* **28**, (1) 55-63.

97. Ou,H.W. (2006) Meridional thermal field of a coupled ocean-atmosphere system: a conceptual model. *Tellus A* **58**, (3) 404-415, doi: doi:10.1111/j.1600-0870.2006.00174.x.

98. Ozawa,H., Ohmura,A., Lorenz,R.D., and Pujol,T. (2003) The second law of thermodynamics and the global climate system: A review of the maximum entropy production principle. *Rev.Geophys.* **41**, (4, Art. No. 1018) 1-24, doi: 10.1029/2002RG000113.

99. Paerl,H.W. and Pinckney,J.L. (1996) A mini-review of microbial consortia: their roles in aquatic production and biogeochemical cycling. *Microb.Ecol.* **31**, 225-247.

100. Paltridge,G.W. (1975) Global dynamics and climate-a system of minimum entropy exchange. *Q.J.Roy.Met.Soc.* **104**, 927-945.

101. Paltridge,G.W. (2005) Stumbling into the MEP racket: An historial perspective. In: *Non-equilibrium thermodynamics and the production of entropy: life, earth and beyond*, 33-40. Edited by Kleidon,A. and Lorenz,R.D., Berlin, Springer-Verlag.

102. Paltridge,G.W., Farquhar,G.D., and Cuntz,M. (2007) Maximum entropy production, cloud feedback, and climate change. *Geophys.Res.Lett.* **34**, (14 (L14708)) 1-6, doi: 10.1029/2007GL029925.

103. Pedros-Alio,C. (2006) Marine microbial diversity: can it be determined? *Trends in Microbiology* **14**, (6) 257-263, doi: 10.1016/j.tim.2006.04.007.

104. Peters,R.H. (1976) Tautology in evolution and ecology. *Am Nat* **110**, (971) 1-12.

105. Pfeiffer,T. and Bonhoeffer,S. (2002) Evolutionary consequences of tradeoffs between yield and rate of ATP production. *Zeitschrift für Physikalische Chemie* **216**, (1) 51-63, doi: 10.1524/zpch.2002.216.1.051.

106. Porter,K.G. and Feig,Y.S. (1980) The use of DAPI for identifying and counting aquatic microflora. *Limnol.Oceanogr.* **25**, 943-948.

107. Prigogine,I. (1978) Time, structure, and fluctuations. *Science* **201**, (4358) 777-785.

108. Prigogine,I. and Nicolis,G. (1971) Biological order, structure and instabilities. *Q.Rev.Biophys.* **4**, 107-148.

109. Ptacnik,R., Solimini,A.G., Andersen,T., Tamminen,T., Brettum,P., Lepisto,L., Willen,E., and Rekolainen,S. (2008) Diversity predicts stability and resource use efficiency in natural phytoplankton communities. *PNAS* **105**, (13) 5134-5138, doi: 10.1073/pnas.0708328105.

110. Pueyo,S., He,F., and Zillio,T. (2007) The maximum entropy formalism and the idiosyncratic theory of biodiversity. *Ecology Letters* **10**, (11) 1017-1028, doi: 10.1111/j.1461-0248.2007.01096.x.

111. Rainey,P.B. and Rainey,K. (2003) Evolution of cooperation and conflict in experimental bacterial populations. *Nature* **425**, (04 Sept) 72-74, doi: 10.1038/nature01906.

112. Reeve,H.K. and Holldobler,B. (2007) The emergence of a superorganism through intergroup competition. *PNAS* **104**, (23) 9736-9740, doi: 10.1073/pnas.0703466104.

113. Russell,J.B. and Cook,G.M. (1995) Energetics of bacterial growth: balance of anabolic and catabolic reactions. *Microbiol.Mol.Biol.Rev.* **59**, (1) 48-62.

114. Schneider,E.D. and Kay,J.J. (1994) Life as a manifestation of the second law of thermodynamics. *Mathematical and Computer Modelling* **19**, (6-8) 25-48.

115. Schneider,E.D. and Kay,J.J. (1994) Complexity and thermodynamics: towards a new ecology. *Futures* **26**, (6) 626-647.

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

117. Shannon,C.E. (1948) A mathematical theory of communication. *Bell System Technical Journal* **27**, 379-423-623-656.

118. Shapiro,J.A. (1998) Thinking about bacterial populations as multicellular organisms. *Annu.Rev.Microbiol.* **52**, 81-104.

119. Shimokawa,S. and Ozawa,H. (2002) On the thermodynamics of the ocean general circulation: Irrerversible transition to a state with higher rate of entropy production. *Q.J.Roy.Met.Soc.* **128**, 2115-2128.

120. Shipley,B., Vile,D., and Garnier,E. (2006) From Plant Traits to Plant Communities: A Statistical Mechanistic Approach to Biodiversity. *Science* **314**, (5800) 812-814, doi: 10.1126/science.1131344.

121. Sinigalliano,C.D., Gidley,M.L., Shibata,T., Whitman,D., Dixon,T.H., Laws,E., Hou,A., Bachoon,D., Brand,L., maral-Zettler,L., Gast,R.J., Steward,G.F., Nigro,O.D., Fujioka,R., Betancourt,W.Q., Vithanage,G., Mathews,J., Fleming,L.E., and Solo-Gabriele,H.M. (2007) Impacts of Hurricanes Katrina and Rita on the microbial landscape of the New Orleans area. *PNAS* **104**, (21) 9029-9034, doi: 10.1073/pnas.0610552104.

122. Sogin,M.L., Morrison,H.G., Huber,J.A., Welch,D.M., Huse,S.M., Neal,P.R., Arrieta,J.M., and Herndl,G.J. (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". *PNAS* 0605127103, doi: 10.1073/pnas.0605127103.

123. Solorzano,L. (1969) Determination of ammonia in natural waters by the phenolhypochlorite method. *Limnol.Oceanogr.* **14**, 799-801.

124. Stams,A.J.M. (1994) Metabolic interactions between anaerobic bacteria in methanogenic environments. *Antonie van Leeuwenhoek* **66**, (1) 271-294, doi: 10.1007/BF00871644.

125. Sterner,R.W. and Elser,J.J. (2002) *Ecological stoichiometry: The biology of elements from molecules to the biosphere*, Princeton, NJ, Princeton University Press, 439 pp.

126. Swenson,R. (1989) Emergent attractors and the law of maximum entropy production: Foundations to a theory of general evolution. *Systems Research* **6**, (3) 187-197.

127. Tobias,C.R., Cieri,M., Peterson,B.J., Deegan,L.A., Vallino,J., and Hughes,J. (2003) Processing watershed-derived nitrogen in a well-flushed New England estuary. *Limnol.Oceanogr.* **48**, (5) 1766-1778.

128. Toussaint,O. and Schneider,E.D. (1998) The thermodynamics and evolution of complexity in biological systems. *Comp.Biochem.Physiol.A* **120**, (1) 3-9.

129. Traulsen,A. and Nowak,M.A. (2006) Evolution of cooperation by multilevel selection. *PNAS* **103**, (29) 10952-10955, doi: 10.1073/pnas.0602530103.

130. Ulanowicz,R.E. (1986) *Growth and development: ecosystems phenomenology*, New York, Springer-Verlag, 203 pp.

131. Ulanowicz,R.E. (1997) *Ecology, the ascendent perspective*, New York, NY, Columbia University Press, 201 pp.

132. Ulanowicz,R.E. and Hannon,B.M. (1987) Life and the production of entropy. *Proc.R.Soc.London, Ser.B* **232**, (1267) 181-192.

133. Ulanowicz,R.E. and Platt,T. (1985) Ecosystem theory for biological oceanography. Can. Bull. Fish. Aquat. Sci., 213, Ottawa, 260 pp.

134. Vallino,J.J. (2000) Improving marine ecosystem models: use of data assimilation and mesocosm experiments. *J.Mar.Res.* **58**, 117-164.

135. Vallino,J.J. and Hopkinson,C.S. (1998) Estimation of dispersion and characteristic mixing times in Plum Island Sound estuary. *Estuarine, Coastal Shelf Sci.* **46**, (3) 333-350.

136. Vallino,J.J., Hopkinson,C.S., and Garritt,R.H. (2005) Estimating estuarine gross production, community respiration and net ecosystem production: a nonlinear inverse technique. *Ecol.Modelling* **187**, (2-3) 281-296, doi: 10.1016/j.ecolmodel.2004.10.018.

137. Vallino,J.J., Hopkinson,C.S., and Hobbie,J.E. (1996) Modeling bacterial utilization of dissolved organic matter: Optimization replaces Monod growth kinetics. *Limnol.Oceanogr.* **41**, (8) 1591-1609.

138. Vallino,J.J. (2003) Modeling Microbial Consortiums as Distributed Metabolic Networks. *Biol Bull* **204**, (2) 174-179.

139. van Gemerden,H. (1993) Microbial mats: A joint venture. *Marine Geology* **113**, (1-2) 3-25, doi: 10.1016/0025-3227(93)90146-M.

140. Vuddagiri,S.R., Hall,K.R., and Eubank,P.T. (2000) Dynamic Modeling of Reaction Pathways on the Gibbs Energy Surface. *Ind.Eng.Chem.Res.* **39**, (2) 508-517, doi: 10.1021/ie9904091 S0888-5885(99)00409-1.

141. Wang,J., Salvucci,G.D., and Bras,R.L. (2004) An extremum principle of evaporation. *Water Resources Res.* **40**, (9) 1-14, doi: 10.1029/2004WR003087.

142. Wang,J., Bras,R.L., Lerdau,M., and Salvucci,G.D. (2007) A maximum hypothesis of transpiration. *JGR Biogeosciences* **112**, (G03010) 1-11, doi: 10.1029/2006JG000255.

143. Warnecke,F., Luginbuhl,P., Ivanova,N., Ghassemian,M., Richardson,T.H., Stege,J.T., Cayouette,M., McHardy,A.C., Djordjevic,G., Aboushadi,N., Sorek,R., Tringe,S.G., Podar,M., Martin,H.G., Kunin,V., Dalevi,D., Madejska,J., Kirton,E., Platt,D., Szeto,E., Salamov,A., Barry,K., Mikhailova,N., Kyrpides,N.C., Matson,E.G., Ottesen,E.A., Zhang,X., Hernandez,M., Murillo,C., Acosta,L.G., Rigoutsos,I., Tamayo,G., Green,B.D., Chang,C., Rubin,E.M., Mathur,E.J., Robertson,D.E., Hugenholtz,P., and Leadbetter,J.R.

(2007) Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* **450**, (7169) 560-565, doi: 10.1038/nature06269.

144. Washida,T. (1995) Ecosystem configurations consequent on the maximum respiration hypothesis. *Ecol.Modelling* **78**, (3) 173-193.

145. Weber,B.H., Depew,D.J., and Smith,J.D. (1988) Entropy, information, and evolution: New perspective on physical and biological evolution. MIT Press, Cambridge, MA, 390 pp.

146. West,G., Brown,J.H., and Enquist,B.J. (1997) A General Model for the Origin of Allometric Scaling Laws in Biology. *Science* **276**, (5309) 122-126.

147. Williams,M., Hopkinson,C., Rastetter,E., and Vallino,J. (2004) N budgets and aquatic uptake in the Ipswich River basin, northeastern Massachusetts. *Water Resources Res.* **40**, (11) 1-12, doi: 10.1029/2004WR003172.

148. Williams,M., Hopkinson,C., Rastetter,E., Vallino,J., and Claessens,L. (2005) Relationships of Land Use and Stream Solute Concentrations in the Ipswich River Basin, Northeastern Massachusetts. *Water, Air, & Soil Pollution* **161**, (1-4) 55-74, doi: 10.1007/s11270-005-2830-0.

149. Wilson,D.S. and Sober,E. (1989) Reviving the superorganism. *Journal of Theoretical Biology* **136**, (3) 337-356.

150. Wittebolle,L., Vervaeren,H., Verstraete,W., and Boon,N. (2008) Quantifying Community Dynamics of Nitrifiers in Functionally Stable Reactors. *Applied and Environmental Microbiology* **74**, (1) 286-293, doi: 10.1128/AEM.01006-07.

151. Yooseph,S., Sutton,G., Rusch,D.B., Halpern,A.L., Williamson,S.J., Remington,K., Eisen,J.A., Heidelberg,K.B., Manning,G., Li,W., Jaroszewski,L., Cieplak,P., Miller,C.S., Li,H., Mashiyama,S.T., Joachimiak,M.P., van Belle,C., Chandonia,J.M., Soergel,D.A., Zhai,Y., Natarajan,K., Lee,S., Raphael,B.J., Bafna,V., Friedman,R., Brenner,S.E., Godzik,A., Eisenberg,D., Dixon,J.E., Taylor,S.S., Strausberg,R.L., Frazier,M., and Venter,J.C. (2007) The Sorcerer II Global Ocean Sampling Expedition: Expanding the Universe of Protein Families. *PLoS Biology* **5**, (3 e16) 432-466, doi: 10.1371/journal.pbio.0050016.

152. Zupanovic,P., Juretic,D., and Botric,S. (2004) Kirchhoff's loop law and the maximum entropy production principle. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)* **70**, (5) 056108-5.