PROSPECTS OF A COMPUTATIONAL ORIGIN OF LIFE ENDEAVOR

BARAK SHENHAV and DORON LANCET*
Department of Molecular Genetics and the Crown Human Genome Center, the Weizmann Institute of Science, Rehovot 76100, Israel
(author for correspondence, e-mail: doron.lancet@weizmann.ac.il, phone: 972 8 9343683 or 972 8 9344121, fax: 972 8 9344487)*

(Received 22 January 2003; accepted in revised form 1 April 2003)

Abstract. While the last century brought an exquisite understanding of the molecular basis of life, very little is known about the detailed chemical mechanisms that afforded the emergence of life on early earth. There is a broad agreement that the problem lies in the realm of chemistry, and likely resides in the formation and mutual interactions of carbon-based molecules in aqueous medium. Yet, present-day experimental approaches can only capture the synthesis and behavior of a few molecule types at a time. On the other hand, experimental simulations of prebiotic syntheses, as well as chemical analyses of carbonaceous meteorites, suggest that the early prebiotic hydrosphere contained many thousands of different compounds. The present paper explores the idea that given the limitations of test-tube approaches with regards to such a 'random chemistry' scenario, an alternative mode of analysis should be pursued. It is argued that as computational tools for the reconstruction of molecular interactions improve rapidly, it may soon become possible to perform adequate computer-based simulations of prebiotic evolution. We thus propose to launch a computational origin of life endeavor (http://ool.weizmann.ac.il/CORE), involving computer simulations of realistic complex prebiotic chemical networks. In the present paper we provide specific examples, based on a novel algorithmic approach, which constitutes a hybrid of molecular dynamics and stochastic chemistry. As one potential solution for the immense hardware requirements dictated by this approach, we have begun to implement an idle CPU harvesting scheme, under the title ool@home.

Keywords: molecular dynamics, stochastic chemistry, ool@home, chemical network, computational origin of life, random chemistry

1. Introduction

1.1. PREBIOTIC RANDOM CHEMISTRY

For half a century, since the breakthrough work on prebiotic synthesis (Miller, 1953; Miller and Urey, 1959), chemists have compiled dozens of possible primordial reaction schemes, yielding a diversity of organic compounds, proposed to have seeded life processes (Basile *et al.*, 1984; Ferris and Hagan, 1984; Miller, 1986; Wills and Bada, 2001; Miyakawa *et al.*, 2002). Chemical analyses of carbonaceous meteorites (Briggs and Mamikunian, 1964; Lawless, 1980; Anders, 1989; Deamer and Pashley, 1989; Maurette, 1998; Sephton, 2002) support the suggestion that on primitive earth a large repertoire of organic molecules existed. We have therefore argued in the past that prebiotic scenarios may be adequately described in

Origins of Life and Evolution of the Biosphere **34:** 181–194, 2004.

terms of planet-scale random chemistry reactions (Segre and Lancet, 2000). Based on a count of about 1 million molecule types for alkanes with up to 20 carbon atoms, (Morowitz, 1992; Bytautas and Klein, 1999), it may be estimated that for compounds containing all CHNOPS atoms (Carbon, Hydrogen, Nitrogen, Oxygen, Phosphorus and Sulfur) the number would be much larger than one billion. Some of these may be unstable, and others biased against, but still the relevant state of affairs may be more adequately described through multiple interactions within complex mixtures than in terms of ordinary 'school chemistry' (Cohen and Stewart, 2002).

The prebiotic molecules were relatively small and simple compared to the macromolecules that govern contemporary life. Moreover, as prebiotic molecules did not undergo the refinement and filtering of evolutionary processes, they lacked specificity and tuning towards a defined biochemical process. Thus, the prebiotic milieu could best be characterized by a dense network of weak interactions among relatively small molecules. This situation may have prevailed prior to the emergence of an RNA world (Lazcano and Miller, 1996; De Duve, 1998; Wills and Bada, 2001).

An important open question is whether investigating such a random chemistry setting in the test tube is within the capabilities of current chemistry techniques. What is required is a capacity to handle the high number of molecular species, each potentially at low concentration, in very large volumes and for extremely long periods of time. We pursue here the possibility that certain computational methods could serve to complement the experimental approaches for this task. This concept is based on present-day trends that have permeated all of biology and chemistry, namely, an increasing role for computational tools.

1.2. THE BIOINFORMATICS REVOLUTION

Life as we know it today is heavily based on the existence of molecular networks as integral elements within individual cells and organisms. A pivotal theme in modern cell biology is deciphering complex systems such as transcription schemes, metabolic networks and signaling pathways. The analogous conceptual framework in origin of life research is the metabolism-first mutual-catalysis scenario. This is in contrast to genome-first RNA-world scenarios, where prebiotic entities are often perceived as single molecular replicators, typically biopolymers that undergo mutation-like inter-conversions. In the present paper emphasis will be placed on computational origin of life realizations of the former type, namely mutually catalytic networks, though computational realizations of biopolymer-based origin of life should be considered as well.

The last two decades have witnessed a revolution: a merger of computing and biology has appeared, known as bioinformatics or computational biology. This new synthesis has affected all of the life sciences, and has particularly gained momentum in the wake of the world-wide genome sequencing projects (Lander *et al.*, 2001; Venter *et al.*, 2001).

More recently, bioinformatics has seen a further change. As a result of the tremendous boost in computing capabilities and the increase in detailed quantitative knowledge of cellular mechanisms, computer simulations of entire biological systems have become more of a reality. Examples are simulations of transcription control networks (Banerjee and Zhang, 2002) and metabolic pathways (Christensen and Nielsen, 2000), in the field of systems biology (Ideker *et al.*, 2001). Some researchers have begun to contemplate a school of *in silico* biology, whereby detailed computer models would capture an entire simple cell (Loew and Schaff, 2001; Tomita, 2001) or even a whole complex organism (Kam *et al.*, 2003).

An important feature of such daring attempts is scale separation. Thus, when an attempt is made to generate a computer model of a bacterium, a scientist might write equations for metabolic pathways or transcription control components in terms of protein interactions. However, such a model would not necessitate computer prediction of the folding of each protein or of the docking of each enzyme-substrate pair.

1.3. Computational chemistry and molecular dynamics

Ever since the advent of digital computers half a century ago, computing power was used to simulate chemical reactions and molecular structure. Quantum chemical calculations (Sponer *et al.*, 1996) are routinely used to predict the lengths, angles, and energies of covalent bonds within molecules, as well as the rate of their formation. One of the glaring paradoxes of biological chemistry, however, is the difficulties encountered when attempting to predict the way by which sets of weaker, non-covalent interactions govern processes such as protein folding and ligand-protein recognition (Bures and Martin, 1998; Schneider and Bohm, 2002). Similar difficulties are encountered when attempting to perform ab-initio calculations of enzyme catalysis. Thus, when it comes to predicting the detailed behavior of a molecular network, it is practically impossible to do so based on first principles of quantum chemistry.

Despite these limitations, computer tools can be found in almost any chemical area (Dessy, 1996). For simulations of the time-dependent behavior of a large number of particles molecular dynamics is employed, that is force fields are used in conjunction with Newtonian dynamics. As an example, simulations of lipids in membranes have shown that such an approach may result in a complete time course description of a complex system (Pohorille and Benjamin, 1993; Tieleman *et al.*, 1997; Goetz and Lipowsky, 1998; Lindahl and Edholm, 2000; Marrink and Mark, 2002). Chemistry based simulations improve due to developments in the underlying algorithms such as in molecular dynamics (Tuckerman and Martyna, 2000) or in stochastic chemistry (Gibson and Bruck, 2000; Gillespie, 2001). Constant improvement is gained also from the continuous progress of hardware. Thus, contemporary tools are capable of rather exact simulation, albeit of relatively small

chemical systems, of up to thousands of molecules, and for relatively short periods of time.

More recently a mixed approach combining Quantum Mechanics and Molecular Mechanics (QM/MM) was developed (Maseras and Morokuma, 1995; Svensson *et al.*, 1996). This approach is now generally accepted and included in commercial computational packages such as Gaussian (Frisch *et al.*, 2001). It allows getting a good accuracy in quantum mechanic calculations applied to large molecules.

Small chemical networks are often described and simulated by means of sets of differential equations (Smolen *et al.*, 2000). However, when analyzing a complex network containing thousands of molecular species, and therefore having millions of potential interactions, differential equations become less useful (Reder, 1990; Hasegawa *et al.*, 1992; Jemmer, 1999). This is particularly true when a very small volume is analyzed, containing only a few copies of each molecular kind, rendering a description in terms of continuous concentration variables untenable (Gueron, 2001). Under such conditions, it is advisable to utilize an alternative approach, namely stochastic chemistry (Gillespie, 1976; Vereecken *et al.*, 1997; Lukkien *et al.*, 1998; Jansen and Lukkien, 1999; Gibson and Bruck, 2000; Gillespie, 2001). In a simple embodiment, a pair of molecules is randomly selected, and a Monte Carlo procedure is employed to decide whether a reaction between them will occur. Rate constants are translated to reaction probabilities, and serve as a basis for such a 'lottery' decision. A more elaborate procedure exists based on the master equation associated with the reacting system (Gillespie, 1976; Gibson and Bruck, 2000), which has a considerably higher computational efficiency. In this procedure the rate constants serve as weights in a weighted 'lottery' for picking a reaction that would take place in the simulation. The stochastic chemistry approach is distinct from molecular dynamics in that no spatial details are obtained and only the population time evolution of the different molecular species in the system is obtained. This allows simulation of much larger chemical systems for much longer durations. The novel prebiotic chemistry approach we suggest here constitutes a merger of molecular dynamics and stochastic chemistry.

1.4. Computer simulations of prebiotic scenarios

One point of view in origin of life research, as in other realms of the life sciences and chemistry, is that computer analyses and simulations should be regarded mainly as an auxiliary tool aimed to help in the comprehension of experimental results. Still, computing tools have long been used in the prebiotic evolution scene for simulation of theoretical models (Eigen, 1971; Eigen and Schuster, 1979; Farmer *et al.*, 1986; Bagley and Farmer, 1991; Kauffman, 1993; Fontana and Buss, 1994; Nir and Lahav, 1997; Alves *et al.*, 2001; de Oliveira and Fontanari, 2001). Computer-based analyses also serve as the basis for a closely related discipline, Artificial Life (Adami, 1998; Sipper and Reggia, 2001; Sipper, 2002; Wilke and Adami, 2002), which aims to recreate life phenomenology within artificial embodiments.

Both origin of life modelers and Artificial Life researchers sometimes sacrifice the coherence of chemistry in order to gain insight on what life is and how it could possibly have evolved.

2. Proposed Strategy

2.1. THE GARD MODEL

The strategy proposed will be presented in the context of specific progress made by our group in the last few years. It involves a quantitative description of a chemically-faithful pre-RNA environment, as featured by the Graded Autocatalysis Replication Domain (GARD) model (Segre and Lancet, 1999, 2000; Segre *et al.*, 2000, 2001). The simplest embodiment of the model is based on the behavior of a non-covalent assembly of diverse lipid-like amphiphilic molecules. This 'lipid world' scenario (Segre *et al.*, 2001) was inspired by previous work on the availability of lipid-like compounds on prebiotic earth (Deamer and Oro, 1980; Deamer and Pashley, 1989) and on the properties of catalytic lipid assemblies (Bachmann *et al.*, 1992; Walde *et al.*, 1994; Kust and Rathman, 1995). In this system, molecules of N_G different types, present in ubiquity in the aqueous environment, may join and leave an assembly, within which they are held together by weak non-specific hydrophobic interactions. The assembly size is restricted by occasional physically-induced fission that keeps it out of equilibrium.

The most crucial assumption of the GARD model is that entry and exit of molecules are governed by interactions with the molecules already present within the assembly. It was demonstrated by computer simulations that specific assembly compositions are capable of propagating themselves while away from thermodynamic equilibrium (Segre *et al.*, 2000). The assemblies are shown to harbor compositional information and complexity (Segre *et al.*, 2000; Shenhav *et al.*, 2003b), which may be transmitted from one generation to another. These GARD simulations employed the stochastic chemistry approach using an efficient approximated algorithm recently devised (Gillespie, 2001).

We have published GARD simulations pertaining to $N_G = 50–200$. More realistic simulations should involve much higher N_G values. Recent work in our laboratory explored such a GARD extension (Figure 1). However, routine studies with these high values of molecular diversity would necessitate considerable augmentation of software and hardware configurations.

2.2. PARAMETER CALCULATION

One of the central issues in simulating chemical or biochemical systems, especially by stochastic chemistry, is the assignment of realistic values to parameters, e.g. kinetic constants or catalytic rate enhancements. Ideally, these should be measured

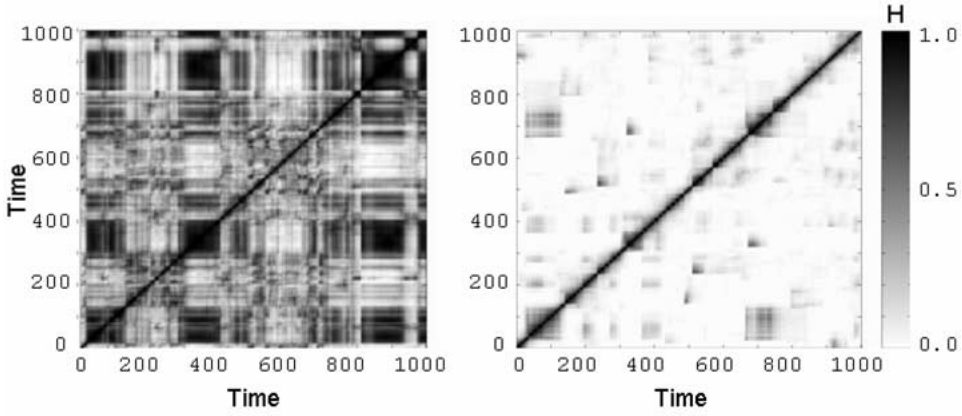

Figure 1. A 'compositional carpet' of a GARD system, as previously described (Segre *et al.*, 2000). Left, with 100 different molecular species ($N_G = 100$), right, with $N_G = 10{,}000$. The drawing depicts a time correlation matrix, where both the ordinate and the abscissa represent the same time scale for the evolution of a particular GARD assembly. Each point in the two dimensional graph is colored by its correlation measure, $H(\mathbf{n_t}, \mathbf{n_{t'}})$, for the compositions at times t and t'. H is defined as

$$H(\mathbf{n_t}, \mathbf{n_{t'}}) = \frac{\mathbf{n_t} \cdot \mathbf{n_{t'}}}{|\mathbf{n_t}||\mathbf{n_{t'}}|},$$

where $\mathbf{n}$ is a compositional vector in N_G-dimensional space. More intuitively defined, each point on the 'carpet' represents how similar are the compositions at two specific time points. Dark colors signify high similarity (close to H = 1) and light colors suggest low similarity (H near 0). The matrix displays dark rectangles, representing quasi-stationary compositions (composomes) and abrupt phase transitions between them. The drawing on the right indicates that at the larger N_G value the composomes are much shorter lived. However, with long simulations and parallelization it might be possible to detect longer-lived composomes in this high dimensionality chemical space.

experimentally, and in fact, this is one of the best opportunities for a fruitful synergy between test tube measurements and computer simulations (Figure 3). Indeed, cooperation of computational and experimental scientists for simulations of biopathways have been established (Regev and Shapiro, 2002). However, for complex systems, which are used to mimic life like behavior, such proposition is less tenable, as it would require millions of individual measurements. Thus computational methods are a viable route.

The published versions of the GARD model employ a unique strategy for obtaining the needed parameters. This strategy derives from the field of statistical chemistry (Segre and Lancet, 1999), and is based on the Receptor Affinity Distribution (RAD) model (Lancet *et al.*, 1993; Rosenwald *et al.*, 2002). Although originally intended for thermodynamic binding parameters, it has now been adapted also to compute kinetic rate-enhancement constants (Lancet *et al.*, 1994; Segre and Lancet, 1999). According to this statistical chemistry approach, the catalytic values are drawn from a realistic distribution of the catalytic potencies, whose phe-

nomenological parameters have been fitted to data obtained from the experimental literature and from databases.

2.3. A COMPUTATIONAL ORIGIN OF LIFE ENDEAVOR

It is rather obvious that the present state of computing technologies does not allow one to contemplate a full digital simulation of the entire path that led from inanimate chemistry to even the simplest protocell. However, it is anticipated that the power of hardware components such as memory capacity and processor speed will continue to increase according to Moore's Law (Moore, 1965; Voller and Porte-Agel, 2002), i.e. expand two-fold every 18 months. It is thus anticipated that 15 years ahead single computers will be 1000 times more powerful. Advances in computer parallelization would enhance this capacity much further. We would like to make a case here for initiating a computational origin of life endeavor. This would entail a process of software development and pursuit of better hardware utilization so as to seed a capacity to study life's origin *in silico*.

The GARD model provides a fertile ground for such an endeavor. Its conceptual framework is general and flexible, and it can be extended in ways that could instill additional details and insight (Shenhav *et al.*, 2003b). An extreme possibility is a complete molecular dynamics simulation of an entire GARD assembly. Here, the detailed motions and interactions of every molecule within the assembly will be simulated in much the same way that protein conformational dynamics is simulated nowadays. However, this will require immense computing resources, expected to be available only decades from now (Figure 2).

With contemporary technology, however, it should be possible to obtain just the parameter values based on first principles, using molecular dynamics tools. Because the most widely explored chemical realization of the model is in the realm of amphiphile assemblies, an example may be drawn from relevant studies (Pohorille and Wilson, 1995; Schuler *et al.*, 2001; Marrink and Mark, 2002) in which the permeability of a lipid bilayer or the dynamics of mixed lipid micelles is simulated by molecular dynamics. Thus, it is feasible to replace the statistical chemistry approach used currently for obtaining the mutual catalytic parameters by computing each such pairwise interaction value based on defined atomic attributes of the molecules involved.

We have therefore begun to develop a hybrid scheme that merges molecular dynamics with stochastic chemistry. Two alternative approaches are considered. In the first, many short molecular dynamics computations are conducted, in which the exchange reactions are followed for a single molecule of type j with respect to a micelle homogeneously composed of molecular species of type i. After a sufficient amount of statistics is gathered, rate parameters for joining and leaving can be established, with which fast stochastic chemistry simulations may then be carried out. In the second alternative, molecular dynamics is used to calculate the rate constants for a specific composition of a heterogeneous assembly, followed by stochastic

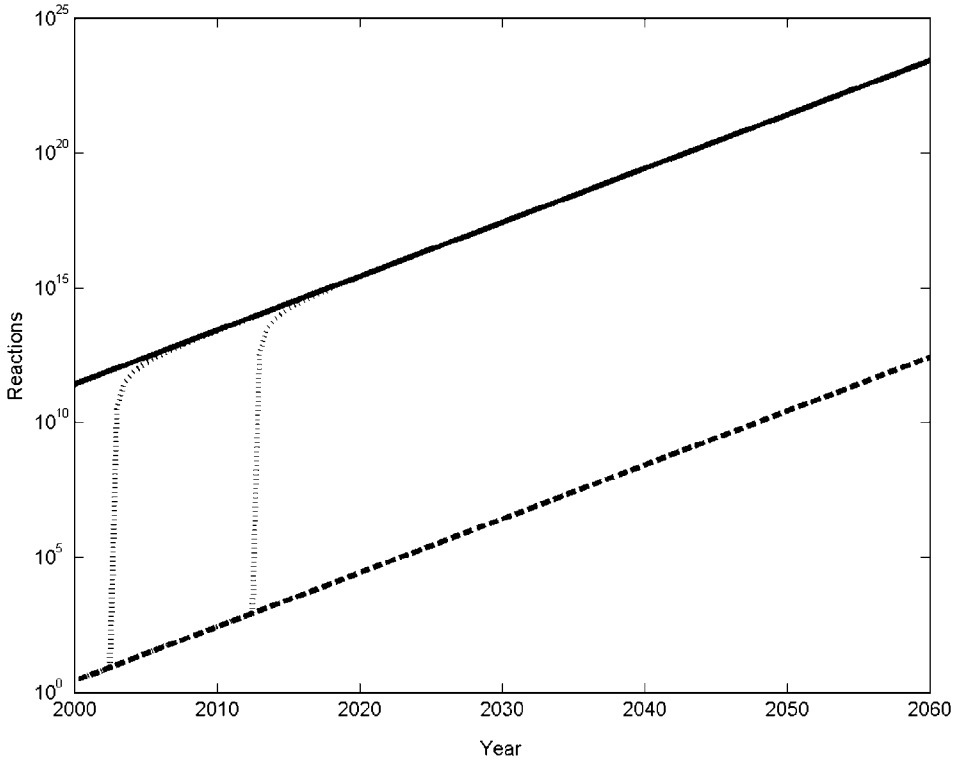

Figure 2. A schematic depiction of the capacity of computers in different years (shown on the abscissa) to compute join/leave reactions of GARD in a month of computation (shown on the ordinate). Solid – stochastic random chemistry in which a single reaction requires approximately 10,000 Flop (Floating point operations) where contemporary desktop computers provide order of 10^9 Flop/sec or 10^{15} Flop/month. The slope of the line reflects Moore's low i.e. doubling of computation power every 18 month. Dashed – pure molecular dynamics, in which a single reaction assumed to require 10^{15} Flops. Dotted – a simple hybrid approach, whereby stochastic simulation is performed following derivation of parameter values by molecular dynamics and interpolation. Light dotted reflects sequential calculation of parameter values while dark dotted assumes the use of 100 computers concurrently for computing the parameters by molecular dynamics in parallel.

simulations. Once the assembly composition changes significantly, a new set of rate constants is similarly computed, either by another round of molecular dynamics or by interpolation of previous runs. In both cases, we explicitly employ an 'on-the-fly' generation of reaction rates, as previously described (Broadbelt *et al.*, 1994). The hybrid approach described here may afford the application of molecular dynamics to the study of prebiotic evolution (Figure 2).

One of the most important attributes of the suggested endeavor is that its progress and potential success will be intimately linked to future advances in other fields of molecular sciences. If in the coming decade a highly effective receptor-ligand or enzyme-substrate docking algorithms are perfected, the same algorithms

would be directly applicable to simulate origin of life scenarios. Similarly, as protein folding algorithms become much more efficient, they could also be used to predict the conformations of oligomers, e.g. short peptides, en route to computing their accurate catalytic capacities. This will be essential for extensions of the GARD model which we have begun to develop (Shenhav *et al.*, 2003a, b) in which head groups of lipid-like amphiphiles become more elaborate and begin to resemble simple biopolymers.

2.4. HARDWARE RESOURCES FOR THE ENDEAVOR

The present power of computer hardware is far from being adequate for most simulations of the pre-RNA prebiotic processes, even with the best available algorithms. For a simulation to become a reality, vast hardware improvements are necessary, likely to occur in the foreseeable future. First, the computational endeavor should benefit from the unabated trend of exponentially increasing speed and capacity of single computers (Figure 2). Second, because the typical simulations involve numerous instances of relative independent entities (e.g. multiple instances of GARD assemblies), such systems are known to be 'embarrassingly parallel' i.e. highly concurrent. Thus, the endeavor could benefit maximally from any developments that will entail parallelized CPU utilization. Moreover, the computational endeavor is expected to benefit greatly from schemes for the management and utilization of idle computing resources harvested over the Internet (Shirts and Pande, 2000). This methodology was pioneered by SETI@home for the search of extraterrestrial intelligence (http://setiathome.ssl.berkeley.edu/, Editorial 1998) and is being utilized successfully also for protein folding (Zagrovic *et al.*, 2002a, b) and ligand fitting (Davies *et al.*, 2002). We have recently initiated the OOL@Home venture, which is such a scheme for Origin of Life purpose (Shenhav and Lancet, 2002; Shenhav *et al.*, 2003a).

3. Conclusions

Using computer simulations is widely accepted in cases where laboratory experiments cannot be employed, as exemplified by studies of the origin of stars and galaxies, as well as of the origin of the universe (Bertschinger, 1998; Frenk, 2002). We present here an argument in favor of employing an analogous approach for yet another origin question, that which relates to the complex processes that have led to life's emergence. This may be considered premature by some, who might propose to wait for considerable hardware improvement. We argue for the benefits of a modest beginning, with the prospects of continued progress that will allow gradually a more comprehensive future solution of the problem.

The computational origin of life endeavor is not a replacement for theoretical models or experimental methods (Figure 3). Both will be required to support its

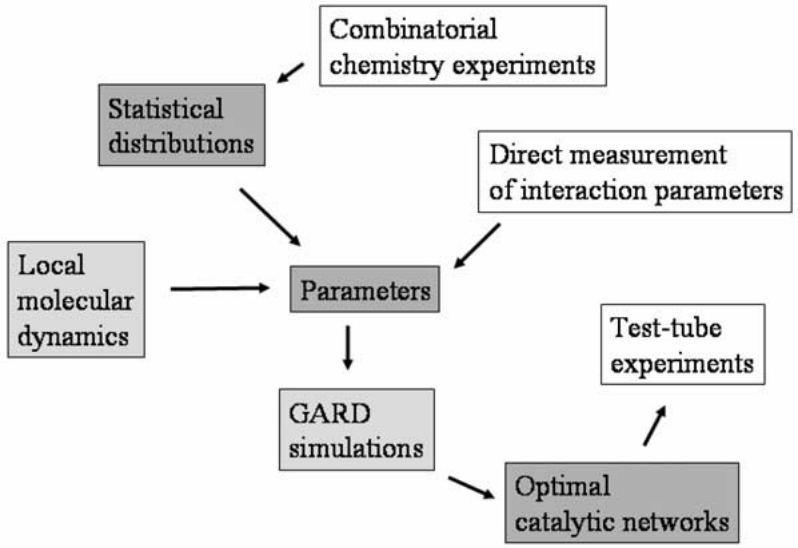

Figure 3. A flowchart showing the interplay between computer simulations and experiments in a computational origin of life project. Theoretical model features (dark gray), such as parameter values, are derived directly or indirectly from bench experiments (white boxes), and in parallel from *in silico* computations (light gray), e.g. molecular dynamics. These parameters are used in a central tool such as programs that simulate the GARD model. The results are used to optimize the set of conditions in which future test tube experiments might be conducted.

parameters and overall design, and to verify or falsify its predictions. Moreover, the computational endeavor may accelerate the development of other methodologies in the realm of origin of life. Computational strategies may not allow, in the immediate future, to fully capture the transition from inanimate components to cellular life. But, it could provide a pertinent set of tools for considerably better understanding of prebiotic evolution.

Acknowledgements

We thank Orna Man for her comments on the manuscript. This research was supported by funds from the Crown Human Genome Center.

References

Adami, C.: 1998, *Introduction to Artificial Life*, Springer Verlag, New York.

Alves, D., Campos, P. R. A., Silva, A. T. C. and Fontanari, J. F.: 2001, Group Selection Models in Prebiotic Evolution, *Physical Review E* **6301**, art. no.-011911.

Anders, E.: 1989, Pre-Biotic Organic Matter from Comets and Asteroids, *Nature* **342**, 255–257.

Bachmann, P. A., Luisi, P. L. and Lang, J.: 1992, Autocatalytic Self-Replicating Micelles as Models for Prebiotic Structures, *Nature* **357**, 57–59.

Bagley, R. J. and Farmer, J. D.: 1991, Spontaneous Emergence of a Metabolism, in Langton, C. G., Tylor, C., Farmer, J. D. and Rasmussen, S. (eds), *Artificial Life II*, Addison-Wesley, pp. 93–140.

Banerjee, N. and Zhang, M. Q.: 2002, Functional Genomics as Applied to Mapping Transcription Regulatory Networks, *Curr. Opin. Microbiol.* **5**, 313–317.

Basile, B., Lazcano, A. and Oro, J.: 1984, Prebiotic Syntheses of Purines and Pyrimidines, *Adv. Space Res.* **4**, 125–131.

Bertschinger, E.: 1998, Simulations of Structure Formation in the Universe, *Annual Review of Astronomy and Astrophysics* **36**, 599–654.

Briggs, M. H. and Mamikunian, G.: 1964, Organic Constituents of Carbonaceous Chondrites, *Life Sci. Space Res.* **2**, 57–85.

Broadbelt, L. J., Stark, S. M. and Klein, M. T.: 1994, Computer-Generated Pyrolysis Modeling – on-the-Fly Generation of Species, Reactions, and Rates, *Industrial & Engineering Chemistry Research* **33**, 790–799.

Bures, M. G. and Martin, Y. C.: 1998, Computational Methods in Molecular Diversity and Combinatorial Chemistry, *Curr. Opin. Chem. Biol.* **2**, 376–380.

Bytautas, L. and Klein, D. J.: 1999, Alkane Isomer Combinatorics: Stereostructure Enumeration and Graph-Invariant and Molecular-Property Distributions, *Journal of Chemical Information and Computer Sciences* **39**, 803–818.

Christensen, B. and Nielsen, J.: 2000, Metabolic Network Analysis. A Powerful Tool in Metabolic Engineering, *Adv. Biochem. Eng. Biotechnol.* **66**, 209–231.

Cohen, J. and Stewart, I.: 2002, *Evolving the Alien*, Ebury Press, London.

Davies, E. K., Glick, M., Harrison, K. N. and Richards, W. G.: 2002, Pattern Recognition and Massively Distributed Computing, *Journal of Computational Chemistry* **23**, 1544–1550.

De Duve, C.: 1998, Constraints on the Origin and Evolution of Life, *Proc. Am. Philos. Soc.* **142**, 525–532.

De Oliveira, V. M. and Fontanari, J. F.: 2001, Extinctions in the Random Replicator Model, *Physical Review E* **6405**, art. no.-051911.

Deamer, D. W. and Oro, J.: 1980, Role of Lipids in Prebiotic Structures, *Biosystems* **12**, 167–175.

Deamer, D. W. and Pashley, R. M.: 1989, Amphiphilic Components of the Murchison Carbonaceous Chondrite: Surface Properties and Membrane Formation, *Orig. Life Evol. Biosph.* **19**, 21–38.

Dessy, R. E.: 1996, New Reagents, New Reactions: Computers in Chemistry, *Technology in Society* **18**, 137–149.

Editorial: 1998, Net News: ... and a Search for Alien Life, *Science* **282**, 853.

Eigen, M.: 1971, Selforganization of Matter and the Evolution of Biological Macromolecules, *Naturwissenschaften* **58**, 465–523.

Eigen, M. and Schuster, P.: 1979, *The Hypercycle*, Springer, Berlin.

Farmer, J. D., Kauffman, S. A. and Packard, N. H.: 1986, Autocatalytic Replication of Polymers, *Physica D* **22**, 50–67.

Ferris, J. P. and Hagan, W. J., Jr.: 1984, HCN and Chemical Evolution: The Possible Role of Cyano Compounds in Prebiotic Synthesis, *Tetrahedron* **40**, 1093–1120.

Fontana, W. and Buss, L. W.: 1994, What Would-Be Conserved If the Tape Were Played Twice, *Proceedings of the National Academy of Sciences of the United States of America* **91**, 757–761.

Frenk, C. S.: 2002, Simulating the Formation of Cosmic Structure, *Philosophical Transactions of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences* **360**, 1277–1294.

Frisch, M. J. and Trucks, G. W., *et al.*: 2001, *Gaussian 98 (Revision A.10)*, Gaussian, Inc., Pittsburgh PA.

Gibson, M. A. and Bruck, J.: 2000, Efficient Exact Stochastic Simulation of Chemical Systems with Many Species and Many Channels, *Journal of Physical Chemistry A* **104**, 1876–1889.

Gillespie, D. T.: 1976, General Method for Numerically Simulating Stochastic Time Evolution of Coupled Chemical-Reactions, *Journal of Computational Physics* **22**, 403–434.

Gillespie, D. T.: 2001, Approximate Accelerated Stochastic Simulation of Chemically Reacting Systems, *Journal of Chemical Physics* **115**, 1716–1733.

Goetz, R. and Lipowsky, R.: 1998, Computer Simulations of Bilayer Membranes: Self-Assembly and Interfacial Tension, *Journal of Chemical Physics* **108**, 7397–7409.

Gueron, S.: 2001, Deterministic Approximations for Stochastic Processes in Population Biology, *Future Generation Computer Systems* **17**, 893–899.

Hasegawa, T. I., Herbst, E. and Leung, C. M.: 1992, Models of Gas-Grain Chemistry in Dense Interstellar Clouds with Complex Organic-Molecules, *Astrophysical Journal Supplement Series* **82**, 167–195.

Ideker, T., Galitski, T. and Hood, L.: 2001, A New Approach to Decoding Life: Systems Biology, *Annu. Rev. Genomics Hum. Genet.* **2**, 343–372.

Jansen, A. P. J. and Lukkien, J. J.: 1999, Dynamic Monte-Carlo Simulations of Reactions in Heterogeneous Catalysis, *Catalysis Today* **53**, 259–271.

Jemmer, P.: 1999, Symbolic Algebra in the Analysis of Dynamic Chemical-Kinetic Systems, *Mathematical and Computer Modelling* **30**, 33–47.

Kam, N., Marelly, R., Kugler, H., Pnueli, A., Harel, D., Hubbard, E. J. and Stern, M. J.: 2003, Formal Modeling of C. Elegans Development: A Scenario-Based Approach in Priami, C. (ed.), *Proceedings of First International Workshop on Computational Methods in Systems Biology (CMSB)*, Rovereto, Italy, Lecture Notes in Computer Science **2602**, pp. 4–20.

Kauffman, S. A.: 1993, *The Origin of Order*, Oxford University Press.

Kust, P. R. and Rathman, J. F.: 1995, Synthesis of Surfactants by Micellar Autocatalysis – N,N-Dimethyldodecylamine N-Oxide, *Langmuir* **11**, 3007–3012.

Lancet, D., Kedem, O. and Pilpel, Y.: 1994, Emergence of Order in Small Autocatalytic Sets Maintained Far from Equilibrium – Application of a Probabilistic Receptor Affinity Distribution (RAD) Model, *Berichte Der Bunsen-Gesellschaft-Physical Chemistry Chemical Physics* **98**, 1166–1169.

Lancet, D., Sadovsky, E. and Seidemann, E.: 1993, Probability Model for Molecular Recognition in Biological Receptor Repertoires – Significance to the Olfactory System, *Proceedings of the National Academy of Sciences of the United States of America* **90**, 3715–3719.

Lander, E. S., Linton, L. M., *et al.*: 2001, Initial Sequencing and Analysis of the Human Genome, *Nature* **409**, 860–921.

Lawless, J. G.: 1980, Organic Compounds in Meteorites, *Life Sci. Space Res.* **18**, 19–27.

Lazcano, A. and Miller, S. L.: 1996, The Origin and Early Evolution of Life: Prebiotic Chemistry, the Pre-RNA World, and Time, *Cell* **85**, 793–798.

Lindahl, E. and Edholm, O.: 2000, Mesoscopic Undulations and Thickness Fluctuations in Lipid Bilayers from Molecular Dynamics Simulations, *Biophysical Journal* **79**, 426–433.

Loew, L. M. and Schaff, J. C.: 2001, The Virtual Cell: A Software Environment for Computational Cell Biology, *Trends Biotechnol.* **19**, 401–406.

Lukkien, J. J., Segers, J. P. L., Hilbers, P. A. J., Gelten, R. J. and Jansen, A. P. J.: 1998, Efficient Monte Carlo Methods for the Simulation of Catalytic Surface Reactions, *Physical Review E* **58**, 2598–2610.

Marrink, S. J. and Mark, A. E.: 2002, Molecular Dynamics Simulations of Mixed Micelles Modeling Human Bile, *Biochemistry* **41**, 5375–5382.

Maseras, F. and Morokuma, K.: 1995, Imomm – A New Integrated Ab-Initio Plus Molecular Mechanics Geometry Optimization Scheme of Equilibrium Structures and Transition-States, *Journal of Computational Chemistry* **16**, 1170–1179.

Maurette, M.: 1998, Carbonaceous Micrometeorites and the Origin of Life, *Orig. Life Evol. Biosph.* **28**, 385–412.

Miller, S. L.: 1953, A Production of Amino Acids under Possible Earth Conditions, *Science* **117**, 528–529.

Miller, S. L.: 1986, Current Status of the Prebiotic Synthesis of Small Molecules, *Chem. Scr.* **26B**, 5–11.

Miller, S. L. and Urey, H. C.: 1959, Organic Compond Synthesis on the Primitive Earth, *Science* **130**, 245–251.

Miyakawa, S., Yamanashi, H., Kobayashi, K., Cleaves, H. J. and Miller, S. L.: 2002, Prebiotic Synthesis from CO Atmospheres: Implications for the Origins of Life, *Proc. Natl. Acad. Sci. U.S.A.*.

Moore, G. E.: 1965, Cramming More Components onto Integrated Circuits, *Electronica* **38**.

Morowitz, H. J.: 1992, *Beginnings of Cellular Life*, Yale University Press, London.

Nir, S. and Lahav, N.: 1997, Emergence of Template-and-Sequence-directed (TSD) Syntheses. 2. A Computer Simulation Model, *Orig. Life Evol. Biosphere* **27**, 567–588.

Pohorille, A. and Benjamin, I.: 1993, Structure and Energetics of Model Amphiphilic Molecules at the Water Liquid Vapor Interface – A Molecular-Dynamics Study, *Journal of Physical Chemistry* **97**, 2664–2670.

Pohorille, A. and Wilson, M. A.: 1995, Molecular-Dynamics Studies of Simple Membrane Water Interfaces – Structure and Functions in the Beginnings of Cellular Life, *Orig. Life Evol. Biosphere* **25**, 21–46.

Reder, C.: 1990, Analysis of Multiscale Biochemical Systems: Graph Methods, *Biomed. Biochim. Acta* **49**, 671–680.

Rosenwald, S., Kafri, R. and Lancet, D.: 2002, Test of a Statistical Model for Molecular Recognition in Biological Repertoires, *Journal of Theoretical Biology* **216**, 327–336.

Schneider, G. and Bohm, H. J.: 2002, Virtual Screening and Fast Automated Docking Methods, *Drug Discov. Today* **7**, 64–70.

Schuler, L. D., Walde, P., Luisi, P. L. and van Gunsteren, W. F.: 2001, Molecular Dynamics Simulation of n-Dodecyl Phosphate Aggregate Structures, *Eur. Biophys. J.* **30**, 330–343.

Segre, D., Ben-Eli, D., Deamer, D. W. and Lancet, D.: 2001, The Lipid World, *Orig. Life Evol. Biosphere* **31**, 119–145.

Segre, D., Ben-Eli, D. and Lancet, D.: 2000, Compositional Genomes: Prebiotic Information Transfer in Mutually Catalytic Noncovalent Assemblies, *Proceedings of the National Academy of Sciences of the United States of America* **97**, 4112–4117.

Segre, D. and Lancet, D.: 1999, A Statistical Chemistry Approach to the Origin of Life, *Chemtracts – Biochemistry and Molecular Biology* **12**, 382–397.

Segre, D. and Lancet, D.: 2000, Composing Life, *Embo Reports* **1**, 217–222.

Sephton, M. A.: 2002, Organic Compounds in Carbonaceous Meteorites, *Nat. Prod. Rep.* **19**, 292–311.

Shenhav, B. and Lancet, D.: 2002, OOL@Home – Origin of Life at Home!!! *ISSOL Spring News Letter*.

Shenhav, B., Kafri, R. and Lancet, D.: 2003a, The transition from Compositional to Polymer-Based Replication, *Orig. Life Evol. Biosphere* **33**, 267.

Shenhav, B., Segre, D. and Lancet, D.: 2003b, Mesobiotic Emergence: Molecular Assemblies that Self-Replicate without Biopolymers, *Adv. Complex Syst.* **6**, 15–35.

Shirts, M. and Pande, V. S.: 2000, Computing – Screen Savers of the World Unite! *Science* **290**, 1903–1904.

Sipper, M.: 2002, *Machine Nature: The Coming Age of Bio-Inspired Computing*, McGraw-Hill, New York.

Sipper, M. and Reggia, J. A.: 2001, Go Forth and Replicate, *Scientific American* **265**, 26–35.

Smolen, P., Baxter, D. A. and Byrne, J. H.: 2000, Modeling Transcriptional Control in Gene Networks – Methods, Recent Results, and Future Directions, *Bull. Math. Biol.* **62**, 247–292.

Sponer, J., Leszczynski, J. and Hobza, P.: 1996, Hydrogen Bonding and Stacking of DNA Bases: A Review of Quantum-Chemical ab initio Studies, *J. Biomol. Struct. Dyn.* **14**, 117–135.

Svensson, M., Humbel, S., Froese, R. D. J., Matsubara, T., Sieber, S. and Morokuma, K.: 1996, ONIOM: A Multilayered Integrated MO+MM Method for Geometry Optimizations and Single Point Energy Predictions. A Test for Diels-Alder Reactions and Pt(P(t-Bu)(3))(2)+H-2 Oxidative Addition, *Journal of Physical Chemistry* **100**, 19357–19363.

Tieleman, D. P., Marrink, S. J. and Berendsen, H. J. C.: 1997, A Computer Perspective of Membranes: Molecular Dynamics Studies of Lipid Bilayer Systems, *Biochimica et Biophysica Acta (BBA) – Reviews on Biomembranes* **1331**, 235–270.

Tomita, M.: 2001, Whole-Cell Simulation: A Grand Challenge of the 21st Century, *Trends Biotechnol.* **19**, 205–210.

Tuckerman, M. E. and Martyna, G. J.: 2000, Understanding Modern Molecular Dynamics: Techniques and Applications, *Journal of Physical Chemistry B* **104**, 159–178.

Venter, J. C. and Adams, M. D., *et al.*: 2001, The Sequence of the Human Genome, *Science* **291**, 1304–1351.

Vereecken, L., Huyberechts, G. and Peeters, J.: 1997, Stochastic Simulation of Chemically Activated Unimolecular Reactions, *Journal of Chemical Physics* **106**, 6564–6573.

Voller, V. R. and Porte-Agel, F.: 2002, Moore's Law and Numerical Modeling, *Journal of Computational Physics* **179**, 698–703.

Walde, P., Wick, R., Fresta, M., Mangone, A. and Luisi, P. L.: 1994, Autopoietic Self-Reproduction of Fatty-Acid Vesicles, *Journal of the American Chemical Society* **116**, 11649–11654.

Wilke, C. O. and Adami, C.: 2002, The Biology of Digital Organisms, *Trends in Ecology & Evolution* **17**, 528–532.

Wills, C. and Bada, J.: 2001, *The Spark of Life: Darwin and the Primeval Soup*, Perseus Publishing.

Zagrovic, B., Snow, C. D., Shirts, M. R. and Pande, V. S.: 2002a, Simulation of Folding of a Small alpha-Helical Protein in Atomistic Detail Using Worldwide-Distributed Computing, *Journal of Molecular Biology* **323**, 927–937.

Zagrovic, B., Sorin, E. and Pande, V.: 2002b, Distributed Computing Simulations of Folding of beta Hairpin from Protein G and Villin Headpiece, *Biophysical Journal* **82**, 474a-474a.