Article

Required Levels of Catalysis for Emergence of Autocatalytic Sets in Models of Chemical Reaction Systems

Wim Hordijk [1,*], **Stuart A. Kauffman** [2,3,4] and **Mike Steel** [5]

[1] Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland

[2] Tampere University of Technology, Finland; E-Mail: stukauffman@gmail.com

[3] University of Vermont, 85 South Prospect Street Burlington, VT 05405, USA

[4] Santa Fe Institute, 1399 Hyde Park Road Santa Fe, NM 87501, USA

[5] Department of Mathematics and Statistics, University of Canterbury, Private Bag 4800, Christchurch, New Zealand; E-Mail: mike.steel@canterbury.ac.nz

[*] Author to whom correspondence should be addressed; E-Mail: wim@santafe.edu; Tel.: +41-21-692-4268.

Abstract: The formation of a self-sustaining autocatalytic chemical network is a necessary but not sufficient condition for the origin of life. The question of whether such a network could form "by chance" within a sufficiently complex suite of molecules and reactions is one that we have investigated for a simple chemical reaction model based on polymer ligation and cleavage. In this paper, we extend this work in several further directions. In particular, we investigate in more detail the levels of catalysis required for a self-sustaining autocatalytic network to form. We study the size of chemical networks within which we might expect to find such an autocatalytic subset, and we extend the theoretical and computational analyses to models in which catalysis requires template matching.

Keywords: autocatalytic sets; chemical reaction networks; template-based catalysis; origin of life

1. Introduction

In previous work we introduced and investigated a mathematical model of catalytic reaction systems and autocatalytic sets [1,2]. It was shown, both theoretically and computationally, that a linear growth rate in the level of catalysis (with increasing length n of the largest molecules in the system) is sufficient for autocatalytic sets to arise spontaneously [2,3] in a well-known binary polymer model of catalytic reaction systems [4,5].

In this paper we take a closer and more detailed look at our model and its results. First, we introduce a small modification to our mathematical definition of autocatalytic sets and the corresponding algorithm for finding them in general catalytic reaction systems (Section 3). This modification makes both the definition and the algorithm slightly simpler, and includes some specific (although probably rare) cases of autocatalytic sets which were previously left out. However, we show (formally, and in simulations) that this modified algorithm does not invalidate any previous results or conclusions.

Second, we show that there is a discrepancy between the theoretical and simulation results (Section 4). Both results show that a linear growth rate in level of catalysis is sufficient for the emergence of autocatalytic sets. However, there is a difference in the parameter values of these linear relations. Here, we recalculate and compare the required levels of catalysis in more detail and under different scenarios.

Third, we show how our model and algorithm can be used to answer other interesting questions relating to the emergence of autocatalytic sets (Section 5). In particular: What is the minimum required size of the molecule set for autocatalytic sets to emerge given a fixed (known) probability of a molecule catalyzing an arbitrary reaction?

Fourth, we show how more chemical realism can be included in our model, for example by considering template-based catalysis (Section 6). Even though this makes the model harder to analyze, it still generates interesting and useful results.

The next section briefly reviews our previously introduced model and definitions. The four sections following it will present the model modifications, extensions, and additional results mentioned above. The final section summarizes the main conclusions and discusses future directions. Mathematical proofs are provided in an Appendix.

Our study fits within a large and growing body of work that aims to formally model how self-sustaining biochemical systems necessary for life might have emerged. This is an area that has been investigated from many angles over the last three decades. Some approaches that are similar in scope but different in their specific details from the one we study here include models based on Petri-nets [6], algebraic approaches based on metabolic closure (such as Rosen's "(M,R) systems", [7–9]), computer simulations of autocatalytic networks involving artificial chemistry [10] or metabolic networks [11], differential equation modeling [12–15], and Erdös-Renyi style random graph theory [16]. The idea of autocatalytic sets as a precursor to life has certainly not been without criticism [17–19], but recent (exciting) experimental evidence shows that they are a real possibility [20–24]. Thus, we believe that our theoretical and computational studies and results are of direct relevance in the larger context of the origin of life [5,25,26].

2. RAF Sets

Autocatalytic sets may have played an important role in the origin of life [5,15,26,27], and are a necessary, although not sufficient, condition for life. Here, *autocatalytic sets* are defined more formally as *RAF sets* as follows (see [1,2] for the full mathematical definition and notations). Given a *catalytic reaction system* (CRS), *i.e.*, a network of molecule types and catalyzed chemical reactions, a (sub)set $\mathcal{R}$ of such reactions (plus the molecules involved in the reactions in $\mathcal{R}$) is called:

1. *Reflexively autocatalytic* (RA) if every reaction in R is catalyzed by at least one molecule involved in any of the reactions in $\mathcal{R}$;

2. *F-generated* (F) if every reactant in $\mathcal{R}$ can be constructed from a small "food set" F by successive applications of reactions from $\mathcal{R}$;

3. *Reflexively autocatalytic and F-generated* (RAF) if it is both RA and *F*.

The food set F contains molecules that are assumed to be freely available in the environment. Thus, an RAF set formally captures the notion of "catalytic closure", *i.e.*, a self-sustaining set supported by a steady supply of (simple) molecules from some food set. Figure 1 shows a simple example. In [2], we also introduced a polynomial-time algorithm to find RAF sets in general catalytic reaction systems.

Note (as already stated earlier [26]) that this notion of an autocatalytic *set* is somewhat different from the (chemical) term autocatalytic *reaction* in which a molecule directly catalyzes its own production. With an autocatalytic set we do *not* mean a set of autocatalytic reactions, but rather a set of molecules and reactions which is collectively autocatalytic in the sense that all molecules help in producing each other (through mutual catalysis, and supported by a food set). Because of this confusion in terminology, we prefer to use the term "RAF set".

Figure 1. A simple example of a *catalytic reaction system* (CRS) with seven molecule types $\{a, b, c, d, e, f, g\}$ (solid nodes) and four reactions $\{r_1, r_2, r_3, r_4\}$ (open nodes). The food set is $F = \{a, b\}$. Solid arrows indicate reactants going into and products coming out of a reaction, dashed arrows indicate catalysis. The subset $\mathcal{R} = \{r_1, r_2\}$ (shown with bold arrows) is RAF.
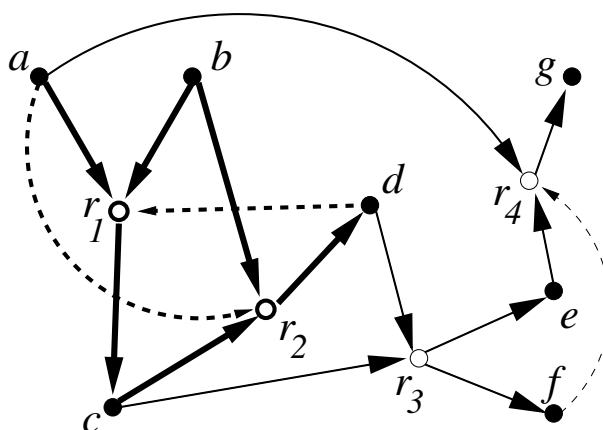
A Model of Catalytic Reaction Systems

In [4,5], a random CRS model was introduced using binary polymers (bit strings) of length at most n as molecule types, ligation and cleavage reactions, and a probability p of any molecule catalyzing any reaction (n and p are parameters of the model). It was argued that in a "sufficiently complex" system (*i.e.*, large enough diversity of molecule types), autocatalytic sets would appear "spontaneously". This was later criticized for requiring an exponential increase (with n) in the (average) number of reactions catalyzed by any one molecule [17], which is chemically unrealistic. However, in [2] it was shown computationally (by applying the RAF algorithm to instances of the random CRS model), and then confirmed theoretically in [3], that only a *linear* growth rate (with n) in the level of catalysis is sufficient for RAF sets to appear with high probability. Furthermore, with these results it is possible to quantify "sufficiently complex" in terms of the size of the molecule set (or maximum molecule length n) and the level of catalysis (average number of reactions catalyzed by any molecule). Here, we continue our investigations of this random CRS model, which we will refer to throughout as the *binary polymer model*.

3. Modification of the Original Definition and Algorithm

In the original (mathematical) definition of RAF sets [2], the "RA" part is stated in terms of the *support* $\mathrm{supp}(\mathcal{R})$ of a set of reactions $\mathcal{R}$, which is defined as all the molecules (reactants and products) that are involved in at least one of the reactions in $\mathcal{R}$. For a set of reactions $\mathcal{R}$ to be reflexively autocatalytic (RA), it is required that every reaction in $\mathcal{R}$ is catalyzed by at least one molecule in $\mathrm{supp}(\mathcal{R})$.

Similarly, the "F" part of the definition is stated in terms of the *closure* $\mathrm{cl}_{\mathcal{R}}(F)$ of the food set F relative to the set of reactions $\mathcal{R}$, which is defined as F together with all molecules that can be constructed from F by repeated applications of reactions in $\mathcal{R}$. For a set of reactions $\mathcal{R}$ to be *F-generated* (F), it is required that the reactants of every reaction in $\mathcal{R}$ are in $\mathrm{cl}_{\mathcal{R}}(F)$.

Now imagine a situation where all reactions in $\mathcal{R}$ conform to the above two requirements, except for one reaction $r \in \mathcal{R}$ which is catalyzed by a molecule $x \in F$ which is not involved in any other way in any reaction in $\mathcal{R}$. So, $x \in \mathrm{cl}_{\mathcal{R}}(F)$, but $x \notin \mathrm{supp}(\mathcal{R})$. In words, the molecule x is in the closure of the food set (by default, as it is part of the food set), but it is not in the support of the reaction set. So, according to the original definition, this set $\mathcal{R}$ is not RAF (it does not conform to the RA part of the definition), even though logically one would consider this case to be a proper autocatalytic set.

To remedy this, and make sure these (probably rare but relevant) cases are also included, we propose a slight modification to our original definition of RAF sets as follows:

3.1. RAF Definition

Given a catalytic reaction system $\mathcal{Q} = (X, \mathcal{R}, C)$, with molecule set X, reaction set $\mathcal{R}$, catalyzation C (a set of molecule-reaction pairs indicating which molecules catalyze which reactions), a nonempty subset $\mathcal{R}'$ of $\mathcal{R}$ is said to be:

- *Reflexively autocatalytic* (RA) if for all reactions $r \in \mathcal{R}'$ there exists a molecule $x \in \mathrm{cl}_{\mathcal{R}'}(F)$ such that $(x, r) \in C$;
- *F-generated* if $\rho(\mathcal{R}') \subseteq \mathrm{cl}_{\mathcal{R}'}(F)$, where $\rho(\mathcal{R}')$ is the set of all reactants in $\mathcal{R}'$;
- *Reflexively autocatalytic and F-generated (RAF)* if $\mathcal{R}'$ is both RA and F.

In other words, a set of reactions $\mathcal{R}'$ is RAF if each reaction in $\mathcal{R}'$ is catalyzed by at least one molecule that can be produced by the set itself (starting from the food molecules), and all reactants (of all reactions in $\mathcal{R}'$) can also be produced by the set itself. So, instead of using $\mathrm{supp}(\mathcal{R})$ in the RA part of the definition and $\mathrm{cl}_{\mathcal{R}}(F)$ in the F part, as in the original definition, the modified definition uses $\mathrm{cl}_{\mathcal{R}}(F)$ in both the RA and the F part, thus simplifying it slightly. Note that an RAF in the original setting of [2] is still an RAF under this new definition by virtue of the following result, a proof of which is provided in the Appendix.

Lemma 3.1 *Any set of reactions that forms an RAF under the earlier definition of [2] is an RAF in the modified definition above.*

This may seem like a minor point, but it could be an important one. Consider, for example, the (reverse) citric acid cycle, which has been argued to have (possibly) been a major step in the origin of life by synthesizing the basic building blocks of organic molecules [28,29]. Even though at present the catalysts that drive the reactions in this cycle are enzymes (proteins), at the early stages these reactions could very well have been catalyzed by much simpler, naturally occurring elements, or ones that are easily synthesized (by basic chemistry) from freely available inorganic molecules [30]. Thus, these original catalysts can be considered food molecules, but they are not involved (as reactants or products) in any of the reactions in the (reverse) citric acid cycle (*i.e.*, they are not in the support). As a consequence, this important metabolic network of reactions would not be classified as an RAF with our original definition, but it is indeed a true RAF according to our modified definition.

Lemma 3.1 has several other desirable properties. First, it ensures that the recent result that Rosen's (M,R) systems can be viewed and studied within the RAF framework [9], still holds. Second, it tells us that if a system has no RAF under the new definition, then it clearly has none under the original one, nor, indeed, under more embellished definitions of an RAF that impose further conditions so as to avoid "trivialities". One such condition (from [1]) would be to require that not all reactions in an RAF are catalyzed by molecules in F (or in some larger subset S of molecules that can be generated from F by catalysis from F and resulting molecules). However, we can accommodate this additional condition within the new RAF framework as follows. The algorithm we will describes below constructs a unique maximal RAF (under the new definition), and from this one can easily check whether an RAF exists that satisfies the additional non-triviality condition (simply check whether this maximal RAF contains a reaction that does not have all its catalysts in F, or the larger set S). Third, spontaneous reactions (those that can proceed without any catalyst) can be allowed (and form part of an RAF) if we formally extend F by an extra "token" element with directed (catalysis) arrows from that token to all such spontaneous reactions. And there are various other extensions, restrictions, or variations that can be imposed on our more general framework.

Of course there are many more properties relevant to the origin of self-sustaining biochemistry beyond such ad-hoc conditions to exclude trivialities—for example, the dynamics of the reactions (the quantity of reagents and products (stoichiometry) along with thermodynamic considerations) and the effect of inhibition and degrading side reactions [18,19]. However, our view here is that RAF sets should be regarded very much as minimal *necessary* conditions for such systems, rather than in any sense *sufficient*. Identifying RAF sets, and establishing conditions for their existence is thus a natural question on the road to finding viable candidates for the origin of early biochemistry.

An additional benefit of our modified RAF definition is that it significantly simplifies the corresponding algorithm for finding RAF sets in general catalytic reaction systems, and its correctness proof. The original algorithm is based on repeatedly (and alternately) applying two reduction steps (starting from the full reaction set) [2]:

1. Remove all reactions that do not conform to the RA requirement;

2. Remove all reactions that do not conform to the *F* requirement.

However, since in our modified definition both the RA part and the *F* part are stated in terms of the closure, these two reduction steps can now be merged into one. As a consequence, the algorithm can be simplified to the following:

3.2. RAF Algorithm

1. Start with the complete set of reactions $\mathcal{R}$ and the food set F;

2. Compute the closure of the food set $\mathrm{cl}_{\mathcal{R}}(F)$ relative to the current set of reactions $\mathcal{R}$;

3. For each reaction $r \in \mathcal{R}$ for which (1) all catalysts, or (2) one or more reactants, are not in $\mathrm{cl}_{\mathcal{R}}(F)$, remove r from $\mathcal{R}$;

4. Repeat steps 2 and 3 until no more reactions can be removed.

The resulting (reduced) reaction set $\mathcal{R}$ is either the (maximal) RAF set contained in the given catalytic reaction system, or it is empty, in which case there is no RAF. This was already proved for the original algorithm in [2], and a (simpler) proof for the modified version is provided in the appendix. Note, however, that the overall running time of this modified algorithm is still the same as in the original case ($\mathcal{O}(|\mathcal{R}|^2 \log |\mathcal{R}|)$ worst-case, but shown to be sub-quadratic on average in practice [2]), as it is dominated by step 2 (computing the closure of the food set). All results presented in this paper are generated with this new version of the RAF algorithm.

4. Required Levels of Catalysis

In [2], we showed through computer simulations that a linear growth rate in the level of catalysis (with n, the length of the largest molecules in the system) appears to be sufficient for RAF sets to occur with high probability in the binary polymer model with ligation and cleavage reactions and random catalysis. This was subsequently confirmed theoretically in [3]. However, even though both the computational and theoretical results give a linear relation for the required level of catalysis, there appears to be a significant difference in the values of the parameters of these linear functions.

This discrepancy can partly be explained by the fact that the theoretical analysis in [3] actually assumes RAF sets that involve *all* molecule types in the system, *i.e.*, an RAF that contains the entire molecule set X (but not necessarily all reactions). This, of course, is a much stronger assumption than used in the simulation studies in [2], where the RAF algorithm was used to find *any* RAF set, regardless of how many molecules or reactions it contains. But the question remains whether the discrepancy can be explained entirely by this difference.

To answer this, we repeated the original simulations, using the RAF algorithm to find *any* RAF set, but this time with the modified RAF algorithm. Then we also applied the RAF algorithm to look for RAF sets that involve *all* molecule types in X. In both cases, we collected statistics for the average number $f(n)$ of reactions catalyzed by any molecule (*i.e.*, the level of catalysis) for which there is a probability $P_n = 0.50$ (or close to 0.50) of finding an RAF set in a number of instances of the random catalytic reaction model [31]. From these statistics, we then estimated a linear function $f(n) = a + bn$ using an ordinary least squares regression. We compare these results with the theoretical linear relation which can be calculated from Theorem 4.1 (ii) in [3] (using $P_n = 0.50$ and $k = t = 2$).

So, in short, we compare the required levels of catalysis for RAF sets to occur with high probability for three cases:

A: Computational case for *any* RAF;
B: Computational case for *all-molecule* RAFs;
C: Theoretical case for *all-molecule* RAFs.

The computational values were calculated over 100 to 1000 (depending on the value of n) instances of the random catalytic reaction model, for $n = 7, \ldots, 20$ (because of an exponential increase in the number of molecules $|X|$ and reactions $|\mathcal{R}|$ with n, we are computationally limited to about $n = 20$ in these simulations).

Table 1 presents the linear relations estimated from the simulation data or calculated from the theoretical analysis. There is a difference of almost 2 orders of magnitude between the observed (simulation) slope of the required growth rate (case **A**) and the theoretical one (case **C**). However, as case **B** shows, this difference cannot be fully explained by the fact that the theoretical analysis assumes RAFs involving *all* molecules. Even with this stronger assumption, the theoretical slope (case **C**) is still more than twice as large as the one from the simulations (case **B**).

Table 1. The empirical (cases **A** and **B**) and theoretical (case **C**) linear relations.

A	$f_A(n) =$	$1.0970 + 0.0189n$
B	$f_B(n) =$	$-0.4736 + 0.7012n$
C	$f_C(n) =$	$1.6339n$

Figure 2 shows the complete data (dots for simulation values, lines for estimated or calculated linear relations). Clearly, even though the theoretically predicted value for $f(n)$ grows quite fast with n (slope $= 1.6339$), the actual value grows only very slowly (slope $= 0.0189$). So, for example, for $n = 20$ the theoretically expected value for the average number of reactions that need to be catalyzed by any molecule to have RAF sets occurring with probability at least $P_n = 0.50$ is $f_C(20) = 32.678$, which seems unrealistically high. However, the actual level of catalysis required is only $f_A(20) = 1.475$, which is chemically much more plausible.

Figure 2. The linear relations for the three cases.
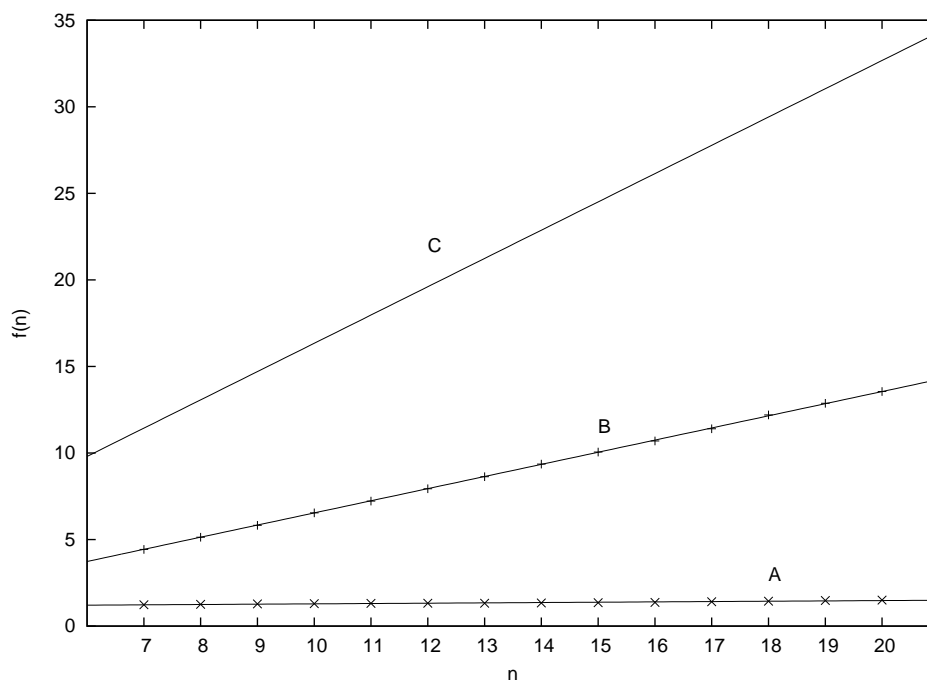
## 5. Required Size of the Molecule Set

In addition to the level of catalysis required for RAF sets to emerge, we can use the RAF algorithm to answer other interesting, and related, questions. For example, one could ask what the minimum required size of the molecule set is (or, in the binary polymer model, the minimum size n of the largest molecules in the system) to get RAF sets with high probability, given a fixed (known) probability p of a molecule catalyzing an (arbitrary) reaction.

Suppose we fix the probability of catalysis at $p = 0.00001$, or perhaps even $p = 0.000001$ (one in a million), which is, for example, roughly the probability in phage display of a random peptide binding an arbitrary ligand [5,32], and "the ease of polypeptide evolution with a small number of arbitrary sequences indicates that a significant fraction of all possible sequences may have functions, at least binding activity in correlation with catalytic activity" [33]. In the binary polymer model, with these probabilities of catalysis, what value of n is required to get a probability of, say, $P_n \geq 0.50$ of RAF sets occurring?

Figure 3 shows the results of applying the (modified) RAF algorithm to instances of the random binary polymer model (Section 2) for the two given values of p and for $n = 5, \ldots, 20$. With $p = 0.00001$ (left curve in Figure 3), $P_n = 0$ for $n \leq 12$, but $P_{13} = 0.982$, and $P_n = 1$ for $n \geq 14$. So, in this case a value of at least $n = 13$ is required to get RAF sets with high probability. For only a slightly higher probability of catalysis ($p = 0.00002$), a value of $n = 12$ would be sufficient (results not shown).

Similarly, with $p = 0.000001$ (right curve in Figure 3), $P_n = 0$ for $n \leq 15$, $P_{16} = 0.939$, and $P_n = 1$ for $n \geq 17$. So, in this case a value of at least $n = 16$ is required to get RAF sets with high probability. Again, for only a slightly higher probability of catalysis ($p = 0.000002$), a value of $n = 15$ would be sufficient (results not shown). The size of the molecule set in this case would be $|X| = 65534$.

Figure 3. The probability P_n of finding RAF sets for different (fixed) catalysis probabilities p and values of n.
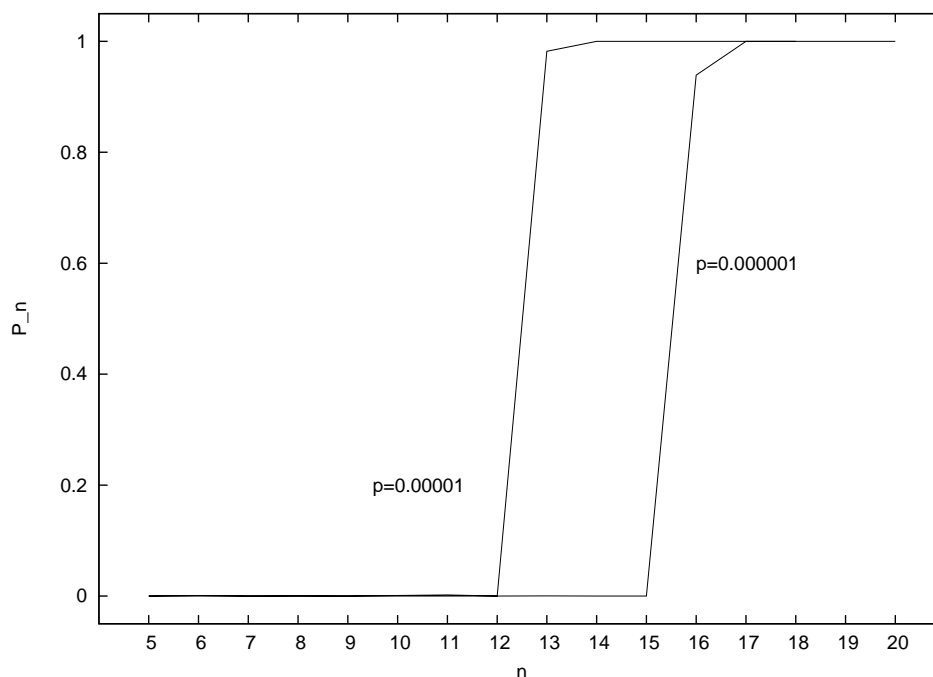
6. Template-Based Catalysis

One could argue that the binary polymer model used in our studies so far is perhaps somewhat oversimplified to be biologically or chemically realistic. However, the model serves as a useful starting point with which precise mathematical statements can be formulated and proved, or at least tested computationally. Furthermore, our RAF definition and algorithm are independent of the particular model that is used, and can in principle also be applied to real catalytic reaction systems (for example metabolic networks, of which the already mentioned citric acid cycle is a core element). And, equally importantly, it is actually not very difficult to add more chemical realism into our mathematical models.

As one particular example, we have considered *template-based catalysis* [5]. The idea here is that, in order to act as a catalyst, a molecule must match at least a certain area around the reaction site according to some template-based matching rule. Similar to, for example, base-pair complementarity in RNA and DNA, we could require a catalyst to match the complement of at least four positions around the reaction site of a ligation or cleavage (two on either side). Consider the ligation reaction

$$00101 + 0011 \rightarrow 001010011$$
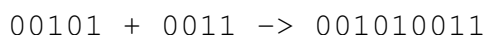
The reaction site template in this case is `0100` (the last two bits of the first binary polymer, plus the first two bits of the second polymer). A given molecule can only act as a catalyst for this reaction if, somewhere along its binary string representation, it contains the complement of this template, *i.e.*, `1011`.

In line with the original random CRS model, and some initial simulations with such template-based catalysis [4,5], we have included this idea as follows. For each combination of a molecule $x \in X$ and a reaction $r \in \mathcal{R}$, if x matches (anywhere along its length) the complement of the reaction site template of length four of r, then with probability p the pair (x, r) is included in the set of catalyzation C. We

present analytical results for this model in Section 6.1 below. However, in our simulations (Section 6.2 below), we used a slightly less constrained version as follows: If a molecule happens to be shorter than length four, then it only has to match part of the complement of the reaction site template, but we require catalysts to be of at least length two. So, the molecule 01 could also catalyze the reaction in the above example. This is mainly done to give some of the food molecules (which are all bit strings up to length two in our simulations) also a chance to act as catalysts.

Note that this template matching requirement is almost the same as in the original simulations [5,34–36], except that we do not allow partial matches here. So, we are considering a slightly "stronger", or more constrained case. Some initial results on similar simulations were reported in [37].

6.1. Theoretical Results

Theorem 4.1 (ii) of Mossel and Steel (2005) shows that for polymers of length up to n over an alphabet of size $\kappa \geq 2$, and with a food set of polymers of length up to t, the probability $P(n)$ that there exists an all-molecule RAF is at least:

$$P(n) \geq 1 - \frac{\kappa(\kappa e^{-\lambda})^t}{1 - \kappa e^{-\lambda}}$$

when each molecule catalyzes on average (at least) λn reactions (provided also that $\lambda > \log_e(\kappa)$). Notice that $P(n)$ can be chosen as close to 1 as we wish by selecting λ large enough (and independently of n!). Thus, this result justifies the statement that the average number of reactions each molecule catalyzes needs to grow only linearly with n in order for there to be a given (high) probability of generating an RAF.

We now describe how this result modifies if catalysis is required to be template-based, as described above. Suppose that a polymer x can catalyze a given reaction only if x contains a substring of length $s = s_1 + s_2$ that is complementary to the end-segment (of length s_1) and the initial segment (of length s_2) of the two molecules involved in the cleavage or ligation. Thus, for the above set-up we have: $s_1 = s_2 = 2$ and so $s = 4$. We also assume that the probability that a molecule x catalyzes a reaction r with a complementary template just depends on x and not on r (this is the analogue of the template-free model assumption (R_2) in [3]).

The following result shows that RAFs will still arise with high probability under linear growth in the average number of reactions each molecule catalyzes, provided the constant involved is increased by a factor of κ^s (this factor would be 16 for the binary model with template size four). More precisely we have the following result, whose proof is provided in the Appendix.

Theorem 6.1 *Let $P_s(n)$ be the probability that there exists an all-molecule RAF under this template matching model. Suppose that each molecule catalyzes (on average) at least $\lambda_s n$ reactions, where $\lambda_s = \kappa^s \lambda$ and $\lambda > \log_e(\kappa)$. Then $P_s(n)$ satisfies the same inequality as $P(n)$, namely*

$$P_s(n) \geq 1 - \frac{\kappa(\kappa e^{-\lambda})^t}{1 - \kappa e^{-\lambda}}$$

6.2. Computational Results

As with the original (random) model (see Section 4), we expect there to be a difference between the theoretically predicted required level of catalysis and the observed level (from simulations) in case

of template-based catalysis. Unfortunately, we are computationally even more restricted with this template-based catalysis case as with the original model. The running time of our RAF algorithm is polynomial in the size of the reaction set $|\mathcal{R}|$, as was shown in [2]. However, since $|\mathcal{R}|$ is *exponential* in the size of the largest molecules n, *i.e.*, $|\mathcal{R}| \propto 2^n$, we can only go up to about $n = 20$ to get computational results in a reasonable amount of time (even on a parallel cluster) for the original model (Section 4). But for the template-based catalysis as described above, it is even worse. We now have to check for each pair of molecule $x \in X$ and reaction $r \in \mathcal{R}$ whether there is a template match between any part of the molecule and the complement of the (4-site) reaction template. Since both $|X|$ and $|\mathcal{R}|$ are $\propto 2^n$, this means that $|X \times \mathcal{R}| \propto 2^{2n}$, and in practice this means that we can only go up to about $n = 16$ with our template-based catalysis simulations.

Figure 4 shows the results of these simulations, using our modified algorithm to find RAF sets (of any size) in the template-based catalysis case, compared to the original model (case **A** in Section 4). As can be expected, for smaller values of n, a higher level of catalysis is needed to find RAF sets with high probability (again, $P_n = 0.50$ is taken as the transition point) in the case of template-based catalysis compared to the purely random model. Since each molecule type $x \in X$ is now restricted, to some extent, in terms of which reactions it can catalyze, the system as a whole is more constrained, and it will be harder to get RAF sets.

Figure 4. The level of catalysis $f(n)$ required for the template-based catalysis case compared to the original (purely random) case, for different values of n.
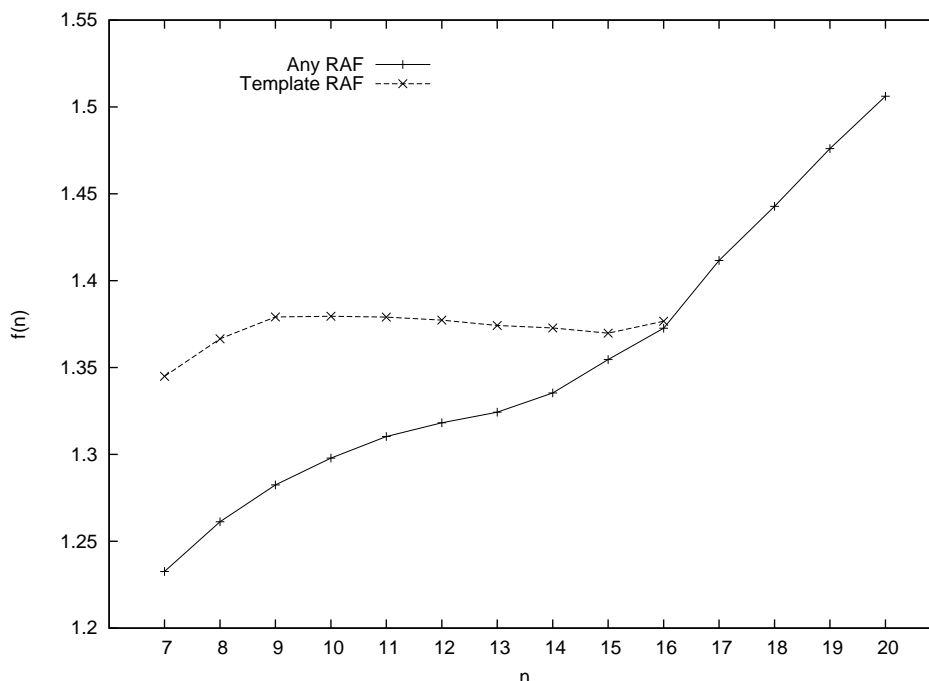
However, for longer and longer molecule types, this restriction becomes less of a problem, as a longer molecule has an increasingly higher probability of matching a given 4-site template somewhere along its length. Indeed, the required level of catalysis for the template-based case tapers off as n increases, and converges to that of the original (purely random) model, reaching the same level at $n = 16$. Given that the purely random model is the "limiting" case for the template-based model, we expect that the level of

catalysis for this template-based model will follow the same pattern as the base model for $n > 16$, and follow a linear relation.

7. Conclusions

Building on our previous work, we have investigated in more detail the levels of catalysis required for the emergence of autocatalytic sets in models of chemical reaction systems. First, we have shown that there is a discrepancy between the theoretically predicted levels and the computationally observed ones. Although both results yield a linear relation between the required level of catalysis and the size n of the largest molecules in the system, in practice this required level is almost two orders of magnitude smaller than the predicted one. However, this discrepancy cannot be fully explained by the fact that the theoretical result is based on a much stronger assumption (requiring the RAF sets to contain *all* molecules in the system). Even for large systems ($n = 20$, containing several millions of molecule types), each molecule only needs to catalyze (on average) between one and two reactions to have RAF sets appear with high probability, which is chemically highly plausible.

Next, we looked at the minimum size of the molecule set (*i.e.*, number of different types of molecules) necessary to get RAF sets given a fixed probability of catalysis. With a (realistic) probability on the order of one in a million of any molecule catalyzing any given reaction, we only need molecules up to length $n = 15$ or $n = 16$, or about 65,000 different molecule types. Again, this number is well within a plausible (experimental) range.

Finally, we studied an extension of the original model, including template-based catalysis. We established formally that in this case a linear growth rate in the level of catalysis also suffices for RAF sets to appear with arbitrary high probability. However, the simulations show that for smaller values of n (the length of the largest molecules in the system), this linear relation is not exact, and that a higher level of catalysis is necessary compared to the original (purely random) model to get RAF sets with a similar probability. But, as n increases, the template matching constraint becomes less of an issue, and the required levels of catalysis converge to those of the original model. This example shows how more chemical realism can be included in our model.

We intend to continue studying the emergence of autocatalytic sets in chemical reaction systems under various scenarios, models, and extensions. However, so far we have mainly studied the (static) underlying graph structures of such systems. One particularly important issue we hope to address in the future is the actual molecular dynamics in a given (catalytic) reaction system. Also, next to studying models of chemical reaction systems, we would like to apply our RAF framework to real (bio)chemical systems such as metabolic networks, or the collection of all known (organic) substrates and reactions. It is our hope that this line of work will help provide more insight into the (possible) origin of life in general.

Appendix

A1. Proof of Lemma 3.1

Let RA denote the original definition of "Reflexively Autocatalytic" from [2], and [RA$_1$] the new one as given here in Section 3. We need to show that [RA] + [F-gen] implies [RA$_1$], so suppose that $r \in R'$. By [RA] there exists $x \in supp(\mathcal{R}') : (x, r) \in C$. Either $x \in \rho(R')$ or $x \in \pi(R')$ (the molecules that are products of at least one reaction in $\mathcal{R}'$). In the first case, by [F-gen], we have $x \in \mathrm{cl}_{\mathcal{R}'}(F)$ so [RA$_1$] holds. In the second case, where $x \in \pi(R')$ there exists $(A, B) \in R'$ for which $x \in B$. By [F-gen], $A \subseteq \mathrm{cl}_{\mathcal{R}'}(F)$ and so $B \subseteq \mathrm{cl}_{\mathcal{R}'}(F)$ and thus $x \in \mathrm{cl}_{\mathcal{R}'}(F)$. So once again [RA$_1$] holds.

A2. Proof of RAF Algorithm Correctness

Let $\mathcal{R}'$ be the output of the RAF algorithm. We need to establish two claims, namely that (i) if $\mathcal{R}'$ is empty then $\mathcal{R}$ contains no RAF, and (ii) if $\mathcal{R}'$ is nonempty then $\mathcal{R}'$ is an RAF, and indeed the unique maximal RAF present in $\mathcal{R}$.

To establish claim (i) we first justify a third claim (iii): If a reaction r lies in any RAF (say $\mathcal{R}_1$) that is contained within an arbitrary subset (say $\mathcal{R}_2$) of $\mathcal{R}$, then r is not eliminated if step 3 of the RAF algorithm is applied with $\mathcal{R}_2$ being taken as the "current set of reactions". To see this, the fact that $\mathcal{R}_1$ is an RAF implies that there exists a molecule $x \in \mathrm{cl}_{\mathcal{R}_1}(F)$ that catalyzes r, and that the reactants of r are contained within $\mathrm{cl}_{\mathcal{R}_1}(F)$. Now, $\mathcal{R}_1 \subseteq R_2$ and so $\mathrm{cl}_{\mathcal{R}_1}(F) \subseteq \mathrm{cl}_{R_2}(F)$; thus step 3 will not eliminate r from $\mathcal{R}_2$. This establishes claim (iii) which implies, by induction, the further claim (iv): If a reaction r lies within any RAF that is contained within $\mathcal{R}$, then r is never eliminated at step 3 of the RAF algorithm at any stage starting with $\mathcal{R}$. Claim (i) now follows immediately.

To establish claim (ii) we first note that, since the algorithm has terminated at some non-empty set of reactions $\mathcal{R}'$, no reaction in $\mathcal{R}'$ can be further eliminated by step 3. By definition, this means that each reaction r in $\mathcal{R}'$ has the properties that (1) at least one catalyst of r is in $\mathrm{cl}_{\mathcal{R}'}(F)$ and (2) all reactants of r are in $\mathrm{cl}_{\mathcal{R}'}(F)$. Condition (1) implies that $\mathcal{R}'$ satisfies the (RA) condition, and property (2) shows that $\mathcal{R}'$ satisfies the F-generated condition. Thus $\mathcal{R}'$ is an RAF; moreover it is the unique maximal RAF by virtue of claim (iv) established in the previous paragraph.

A3. Proof of Theorem 6.1

We first recall two quantities (x_m, r_m) from [3]. Let

$$x_m = \frac{\kappa^{m+1} - \kappa}{\kappa - 1} \tag{1}$$

denote the number of polymers of length at most m over an alphabet of size κ and let r_m denote the number of forward (ligation) reactions that produce polymers of length at most m. We have:

$$r_m \leq m x_m \tag{2}$$

For any polymer x of length at most n let $T_s(x)$ denote the set of polymers of length s that occur within x, and for any polymer t of length s let $X_n(t)$ denote the set of polymers of length at most n that contain at least one copy of t. Thus,

$$t \in T_s(x) \Leftrightarrow x \in T_n(t)$$

Clearly,

$$1 \le |T_s(x)| \le \kappa^s \tag{3}$$

Also, for $n \ge s$, and t any polymer of length s we have:

$$|X_n(t)| \ge 1 + \kappa + \kappa^2 + \cdots + \kappa^{n-s} = \frac{\kappa^{n-s+1} - 1}{\kappa - 1} \tag{4}$$

Given a (forward) reaction r, let t_r denote the complement of the template polymer of length $s = s_1 + s_2$ (which is also polymer of length s). If we let $R_n(t)$ denote the set of forward reactions r with $t_r = t$, then:

$$|R_n(t)| = r_{n-s} \tag{5}$$

In the template polymerization model, x can catalyze r only if $x \in X(t_r)$ (or, equivalently, if $r \in R_n(t)$ for some $t \in T_s(x)$). Let $p(x, r)$ be the probability that molecule x catalyzes reaction r. We have assumed that each molecule that could catalyze reaction r (by virtue of template matching) has the same probability of doing so, that is:

$$p(x, r) = \begin{cases} p_x, & \text{if } t_r \in T_n(x) \\ 0, & \text{otherwise} \end{cases}$$

where p_x just depends on x and not r (the analogue of requirement (R2) in [3]).

Thus, if we let μ_x be the expected number of (forward) reactions ($r \in R_+(n)$) that a molecule x catalyzes, then:

$$\mu_x = \sum_{r \in R_+(n)} p(x, r) = p_x \cdot |S(x)| \tag{6}$$

where $S(x) = \{r \in R_+(n) : x \in X(t_r)\}$. Note that $|S(x)|$ is bounded above by the number of pairs (t, r) where $t \in T_s(x)$ and $r \in R_+(n) : r_t = t$. By virtue of Inequality (3) and Equation (5) this set of pairs has size at most $\kappa^s r_{n-s}$ and so

$$|S(x)| \le \kappa^s r_{n-s} \tag{7}$$

Now, Inequality (2) applied to $m = n - s$, and (7) gives:

$$|S(x)| \le \kappa^s r_{(n-s)} \le \kappa^s (n - s) x_{n-s}$$

Thus, by the assumption that $\mu_x \ge \lambda_s n$ and applying (6), we obtain:

$$\lambda_s \cdot n \le \mu_x \le p_x \kappa^s (1 - s/n) x_{n-s} \cdot n$$

Since $\lambda_s = \kappa^s \lambda$ this last inequality can be written as:

$$p_x \ge \frac{\lambda}{(1 - s/n) x_{n-s}} \tag{8}$$

Now, following [3], for a forward reaction $r \in R_+(n)$ let q_r^* be the probability that r is not catalyzed by any molecule. Then

$$q_r^* = \prod_{x \in X(t_r)} (1 - p(x, r)) = \prod_{x \in X(t_r)} (1 - p_x)$$

By Inequalities (4) and (8) and the inequality $(1 - x)^y \leq \exp(-xy)$ for $x, y > 0$ we obtain:

$$q_r^* \leq \exp\left(-\lambda \cdot \frac{\kappa^{n-s+1} - 1}{(1 - s/n)x_{n-s}(\kappa - 1)}\right) < \exp(-\lambda)$$

where the second inequality follows by replacing x_{n-s} by the expression given by Equation (1) for $m = n - s$ and checking that the ratio that post-multiplies λ is greater than 1. The remainder of the proof now follows the argument in [3] for the proof of Theorem 4.1 (ii).

References and Notes

1. Steel, M. The emergence of a self-catalysing structure in abstract origin-of-life models. *Appl. Math. Lett.* **2000**, *3*, 91–95.
2. Hordijk, W.; Steel, M. Detecting autocatalytic, self-sustaining sets in chemical reaction systems. *J. Theor. Biol.* **2004**, *227*, 451–461.
3. Mossel, E.; Steel, M. Random biochemical networks: The probability of self-sustaining autocatalysis. *J. Theor. Biol.* **2005**, *233*, 327–336.
4. Kauffman, S.A. Autocatalytic sets of proteins. *J. Theor. Biol.* **1986**, *119*, 1–24.
5. Kauffman, S.A. *The Origins of Order*; Oxford University Press: New York, NY, USA, 1993.
6. Sharov, A. Self-reproducing systems: Structure, niche relations and evolution. *BioSystems* **1991**, *25*, 237–249.
7. Letelier, J.C.; Soto-Andrade, J.; Abarzúa, F.G.; Cornish-Bowden, A.; Cárdenas, M.L. Organizational invariance and metabolic closure: Analysis in terms of (M,R) systems. *J. Theor. Biol.* **2006**, *238*, 949–961.
8. Cornish-Bowden, A.; Cárdenas, M.L.; Letelier, J.C.; Soto-Andrade, J. Beyond reductionism: Metabolic circularity as a guiding vision for a real biology of systems. *Proteomics* **2007**, *7*, 839–845.
9. Jaramillo, S.; Honorato-Zimmer, R.; Pereira, U.; Contreras, D.; Reynaert, B.; Hernández, V.; Soto-Andrade, J.; Cárdenas, M.; Cornish-Bowden, A.; Letelier, J. (M,R) Systems and RAF Sets: Common Ideas, Tools and Projections. In *Proceedings of the Alife XII Conference*, Odense, Denmark, 19–23 August, 2010; pp. 94–100.
10. Flamm, C.; Ullrich, A.; Ekker, H.; Mann, M.; Hoegerl, D.; Rohrschneider, M.; Sauer, S.; Scheuermann, G.; Klemm, K.; Hofacker, I.L.; *et al*. Evolution of metabolic networks: A computational framework. *J. Syst. Chem.* **2010**, *1*, DOI:10.1186/1759-2208-1-4.
11. Kun, A.; Papp, B.; Szathmáry, E. Computational identification of obligatorily autocatalytic replicators embedded in metabolic networks. *Genome Biol.* **2008**, *9*, DOI:10.1186/gb-2008-9-3-r51.
12. Awazu, A.; Kaneko, K. Discretness-induced transition in catalytic reaction networks. *Phys. Rev. E* **2007**, *76*, 041915:1–041915:8.
13. Brogioli, D. Marginally stable chemical systems as precursors of life. *Phys. Rev. Lett.* **2010**, *105*, 058102:1–058102:4.

14. Bartsev, S.I.; Mezhevikin, V.V. On initial steps of chemical prebiotic evolution: Triggering autocatalytic reaction of oligomerization. *Adv. Space Res.* **2008**, *42*, 2008–2013.

15. Dyson, F.J. A model for the origin of life. *J. Mol. Evol.* **1982**, *18*, 344–350.

16. Bollobas, B.; Rasmussen, S. First cycles in random directed graph processes. *Discret. Math.* **1989**, *75*, 55–68.

17. Lifson, S. On the crucial stages in the origin of animate matter. *J. Mol. Evol.* **1997**, *44*, 1–8.

18. Szathmary, E. The evolution of replicators. *Philos. Trans. R. Soc. Lond. B* **2000**, *355*, 1669–1676.

19. Orgel, L.E. The implausibility of metabolic cycles on the prebiotic earth. *PLoS Biol.* **2008**, *6*, 5–13.

20. Sievers, D.; von Kiedrowski, G. Self-replication of complementary nucleotide-based oligomers. *Nature* **1994**, *369*, 221–224.

21. Lee, D.H.; Severin, K.; Ghadiri, M.R. Autocatalytic networks: The transition from molecular self-replication to molecular ecosystems. *Curr. Opin. Chem. Biol.* **1997**, *1*, 491–496.

22. Ashkenasy, G.; Jegasia, R.; Yadav, M.; Ghadiri, M.R. Design of a directed molecular network. *Proc. Nat. Acad. Sci. USA* **2004**, *101*, 10872–10877.

23. Hayden, E.J.; von Kieddrowski, G.; Lehman, N. Systems chemistry on ribozyme self-construction: Evidence for anabolic autocatalysis in a recombination network. *Angew. Chem. Int. Ed.* **2008**, *120*, 8552–8556.

24. Lincoln, T.A.; Joyce, G.F. Self-Sustained Replication of an RNA Enzyme. *Science* **2009**, *323*, 1229–1232.

25. Penny, D. An interpretive review of the origin of life research. *Biol. Philos.* **2005**, *20*, 633–671.

26. Hordijk, W.; Hein, J.; Steel, M. Autocatalytic sets and the origin of life. *Entropy* **2010**, *12*, 1733–1742.

27. Dyson, F.J. *Origins of Life*; Cambridge University Press: Cambridge, UK, 1985.

28. Morowitz, H.J.; Kostelnik, J.D.; Yang, J.; Cody, G.D. The origin of intermediary metabolism. *Proc. Nat. Acad. Sci. USA* **2000**, *97*, 7704–7708.

29. Smith, E.; Morowitz, H.J. Universality in intermediary metabolism. *Proc. Nat. Acad. Sci. USA* **2004**, *101*, 13168–13173.

30. Morowitz, H.J.; Srinivasan, V.; Smith, E. Ligand field theory and the origin of life as an emergent feature of the periodic table of elements. *Biol. Bull.* **2010**, *219*, 1–6.

31. It was shown in [2] that $P_n = 0.50$ can be taken as the "transition point" between RAFs not existing at all and RAFs occurring with high probability.

32. Scott, J.K.; Smith, G.P. Searching for peptide ligands with an epitope library. *Science* **1990**, *249*, 286–390.

33. Yamauchi, A.; Nakashima, T.; Tokuriki, N.; Hosokawa, M.; Nogami, H.; Arioka, S.; Urabe, I.; Yomo, T. Evolvability of random polypeptides through functional selection within a small library. *Protein Eng.* **2002**, *15*, 619–626.

34. Bagley, R.J. *A Model of Functional Self Organization*. Ph.D. Thesis, University of California, San Diego, CA, USA, 1991.

35. Bagley, R.J.; Farmer, J.D. Spontaneous Emergence of a Metabolism. In *Artificial Life II*; Langton, C.G., Taylor, C., Farmer, J.D., Rasmussen, S., Eds.; Addison-Wesley: Upper Saddle River, NJ, USA, 1991; pp. 93–140.

36. Bagley, R.J.; Farmer, J.D.; Fontana, W. Evolution of a metabolism. In *Artificial Life II*; Langton, C.G., Taylor, C., Farmer, J.D., Rasmussen, S., Eds.; Addison-Wesley: Upper Saddle River, NJ, USA, 1991, pp. 141–158.

37. Andersen, I.T.; Nan, L.; Kjaersgaard, M.I.S. Search for life in catalytic reaction systems. Available online: http://www.stats.ox.ac.uk/research/genome/projects/pastprojects (accessed on 5 May 2011).