Thresholds in Messy Chemistries

Nathaniel Virgo[1]

[1]Earth-Life Science Institute (ELSI), Tokyo, Japan
nathanielvirgo+alifexv@gmail.com

Many of the chemistries studied by chemists as being relevant to the origins of life produce a combinatorial explosion of products. Such chemistries include Miller-Urey chemistry, HCN polymerisation, Fischer-Tropsch synthesis and the Formose reaction, to name just a few. As well as producing a huge variety of products, these reactions can produce some of the basic molecules that comprise living systems, such as amino acids or sugars. The combinatorial explosions occur because the basic building blocks of organic chemistry can be put together in a huge variety of ways; the number of possible molecules that can be made in these systems grows exponentially (or faster) with their size. Moreover, these networks do not have obvious symmetries or other forms of structure that would make them easy to model and analyse.

Researchers in the origins of life have traditionally regarded such messy combinatorial explosions as a problem that needs to be overcome (e.g. Schuster, 2000). However, technology is now progressing to the point where we can explicitly map out the products produced by these chemistries and the networks of reactions that lead to them (e.g. Andersen et al., 2013). The emerging field of systems chemistry has begun to shift the focus away from static descriptions of complex chemical systems and towards an understanding of their dynamics. With this increased interest comes the realisation that the products of messy prebiotic chemistries are not necessarily just an inert "tar" but may be extremely complex dynamic systems in their own right, which we have simply lacked the right tools to study up to now.

This raises a number of important questions regarding the dynamics of such complex, messy chemistries. One such question is whether a sufficiently large and complex reaction network can behave fundamentally differently from what is possible in smaller, cleaner networks. Here I demonstrate an example of this, with the aid of a simple toy example. The example shows the existence of a phase transition, leading to a threshold effect of a kind that cannot occur in a small reaction network. This threshold is closely related to Eigen's error threshold (Eigen, 1971). The difference is that while Eigen's threshold occurs in a "clean" chemistry (template replication) with plenty of symmetry, ours occurs in a model designed to represent a messy chemistry with little structure.

Let us imagine that that there are N chemical species, each of which can react to form some of the other species from some surrounding milleu. The probability that species i causes the formation of species j is assumed to be constant (with value p) and independent of i and j.

This is essentially the same as Kauffman's (1986) model of the formation of autocatalytic sets, except that we have simplified it further to remove the specific structure associated with cleavage and ligation of peptides, since we are interested in a much broader class of chemistries. We interpret our model in terms of small-molecule organic chemistry rather than peptides, as described in (Virgo and Guttenberg, 2015; Virgo et al., 2016). As in Kauffman's model, this model has a percolation transition, meaning that in the large-N limit, if $p > 1/N$ there will be a single giant autocatalytic set consisting of most of the species in the system. In the simulations below we set the kinetic rates of all the catalysis reactions to be equal, with the value k, but one can show analytically that this makes no qualitative difference to the results of the model, as long as the kinetic rates are chosen from a distribution with bounded variance.

We are interested in the case where p is above the percolation threshold. In this regime, if any amount of any species is added to the system, it will eventually produce some amount of almost every species. We are using this simple model to stand in for a much more complex autocatalytic system with a combinatorial explosion. As the amount of matter in the system grows, it samples more and more of the combinatorially huge space of species available to it.

Let us now suppose that within this giant autocatalytic set there exists a smaller subset that can collectively produce its own members at a greater rate. For ease of exposition, we take this set to be a single species that catalyses its own production at a rate k^*, although similar results follow in the more general case. We can ask the following question: can this small subset produce itself faster than the "messy" autocatalytic set that contains it? If so, then if we were to perform this reaction experimentally, we might find that once the "fast" set was discovered it would come to dominate

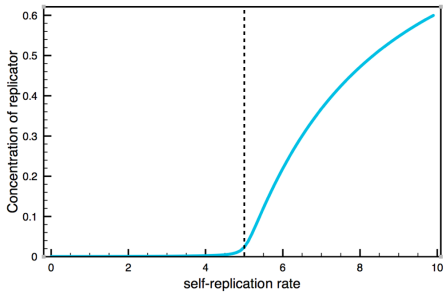

Figure 1: The relative concentration of the self-replicating molecule, plotted as a function of its replication rate k^*. For this plot the values $N = 4000$ and $p = 5/N$ were used. With a greater value of N the threshold becomes sharper.

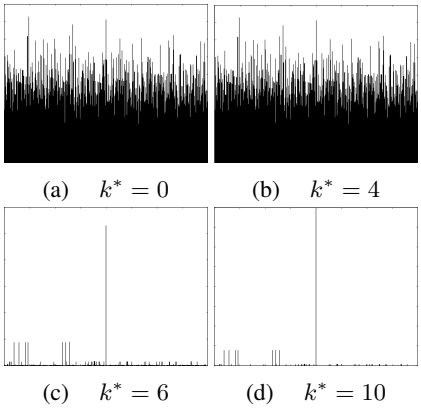

Figure 2: Relative concentration profiles for various values of the replicator's growth rate k^*. In each plot, the species are enumerated along the x axis, with the relative concentration shown on the y axis. The self-replicating species is in the centre of the x axis. Note the stark difference between plots (a) and (b) with k^* below the threshold, versus (c) and (d) above it.

the composition, so that mass spectroscopy would reveal a highly peaked concentration profile, rather than a heterogeneous sampling of all possible products.

Figure 1 shows the relative concentration of a rapidly self-replicating species as a function of its replication rate. (That is, we normalise the vector of concentrations by its sum.) The threshold occurs when $k^* = kp/N$, with k being the rate at which the other species catalyse one another's formation. Below the threshold, all species grow at similar rates. Increasing k^* leads to only a slight relative increase in the amount of the replicator molecule and no change in the relative concentrations of the other species. However, above the threshold the concentration profile changes, becoming dominated by the replicator, and to a lesser extent, by the species whose production can be directly catalysed by it, as shown in figure 2. (This is closely analogous to the "mutational halo" in the case of Eigen's error threshold.)

The significance of this result lies not in the naïve scenario presented, but in showing that such threshold phenomena can exist in messy chemistries at all. Previously two distinct phase transitions were known in prebiotic chemistry models. The error threshold is a phase transition in a chemistry, but not a messy one. Kauffman's model exhibits a phase transition in a messy chemistry, but it is in the opposite direction from ours (i.e. from clean to messy) since once the threshold is passed the system explodes out into the whole combinatorial space. This new model shows a phase transition in which an initially messy chemistry can spontaneously constrain itself to a specific part of its accessible phase space.

We suspect that many more threshold phenomena of this nature will be found, in models much closer to chemical reality than the toy one we have presented. Indeed we suspect that the phase "more is different" (Anderson, 1972) applies to the size of reaction networks as much as it does to the size of a physical system, with phase transitions being a common and generic phenomenon in the dynamics of complex chemistries. Perhaps in time we will find transitions that lead not just to replication of a fixed set of species but to a complex, metabolism-like network of molecular interactions, or to replication with heritable variation, and thence to natural selection.

Acknowledgements.

This publication was supported by the ELSI Origins Network (EON), which is supported by a grant from the John Templeton Foundation. I thank Nicholas Guttenberg, Eric Smith and many others at ELSI for stimulating discussions about a broad selection of ideas, of which this work is only a small part.

References

Andersen, J. L., Andersen, T., Flamm, C., Hanczyc, M., Merkle, D., and Stadler, P. F. (2013). Navigating the chemical space of HCN polymerization and hydrolysis: Guiding graph grammars by mass spectrometry data. *Entropy*, 15:4066–4083.

Anderson, P. W. (1972). More is different. *Science*, 177(4047):393–396.

Eigen, M. (1971). Selforganization of matter and evolution of biological macromolecules. *Naturwissenschaften*, 58(10):465–523.

Kauffman, S. (1986). Autocatalytic sets of proteins. *Journal of Theoretical Biology*, 119:1–24.

Schuster, P. (2000). Taming combinatorial explosion. *PNAS*, 97(14):7678–7680.

Virgo, N. and Guttenberg, N. (2015). Heredity in messy chemistries. In Andrews, P. et al., editors, *Proceedings of the European Conference on Artificial Life 2015*, pages 325–332. MIT Press.

Virgo, N., Ikegami, T., and McGregor, S. (2016). Complex autocatalysis in simple chemistries. *Artificial Life*, early access publication.