Identifying molecular selection using Assembly Theory and closed-loop experiments

Amit Kahana[1], Michael Jirasek[1], Silke Asche[1], Abhishek Sharma[1], Stuart Marshall[1] and Lee Cronin[1]

[1]School of Chemistry, University of Glasgow, Glasgow G12 8QQ, United Kingdom

Abstract

Life's origin and chemical evolution requires continuous and substantial selective processes at the molecular scale. However, the spontaneous emergence of selection, its mechanism and system-level influence are still insufficiently explored. To address this, an automated experimental framework has been devised to identify selection in a recursive system of oligomerizing molecules with closed-loop analytics. The approach is based on Assembly Theory, using Molecular Assembly (MA) index as an inherent complexity measure of molecules and molecular networks. A string-based MA model was developed to assist in the efficient analysis of diverse lengthy oligomers and to allow string information procedures. Coupled with smart algorithmic decision-making, the system will attempt to maximize the molecular network's complexity in the reactor over recursive cycles. Following patterns of increasing chemical complexity in the molecular system could reveal definite traces of selection and determine the conditions and agents that promote it. This work elucidates why improbable complex states emerge, pertinent to life's origin and its major evolutionary transitions.

Introduction

The emergence of life is epitomized by a gradual rise in the complexity of sustained simple chemical networks. Biological life is viewed essentially as an autocatalytic chemical network that has undergone evolution, thus acquiring nontrivial function and coordination, both exemplified by higher molecular and system complexity (Xavier et al, 2020). Reaching a living state from a messy mixture of simple monomers requires a substantial complexification process, that is rooted in the fundamentals of chemical selection. Selection is the driving force that enables molecular systems to attain improbable heritable compositional states (Ameta et al, 2021; Kahana et al, 2023). Through a process of selection, a system of reactive molecules transitions from an undirected exploration in chemical assembly space towards a more directed exploration, resulting in a higher degree of complexity (Sharma et al, 2022).

To this day, the spontaneous emergence of selection at the molecular scale has only been partially observed and thus insufficiently understood. Researchers resort to describe selective traits in well-defined systems (Colomer et al, 2020), or otherwise report all the limitedly-annotated changes they observe over time as markers of chemical evolution (Vincent et al, 2019; Frenkel-Pinter et al, 2022). This somewhat-blind approach is smartly employed to study intricate evolving chemical systems; however, it is missing a notion of molecular complexity that can be used to uncover network-wide selective processes.

In recent years Assembly Theory has been developed, which offers an experimentally-verified method of calculating and measuring the innate complexity of molecules, termed Molecular Assembly (MA) index (Marshall et al, 2021). The MA of a molecule is determined by its shortest construction pathway, quantifying the amount of contingency embedded in its graphical representation (Figure 1). It has been reported that MA of molecules can be measured experimentally by multiple analytical techniques (Jirasek et al, 2023), and could be employed as a biosignature in life detection and assessment (Marshall et at, 2021). This enumeration is also applicable to molecular networks, and can be utilized to evaluate their assembly spaces and flow of information (Liu et al, 2021).

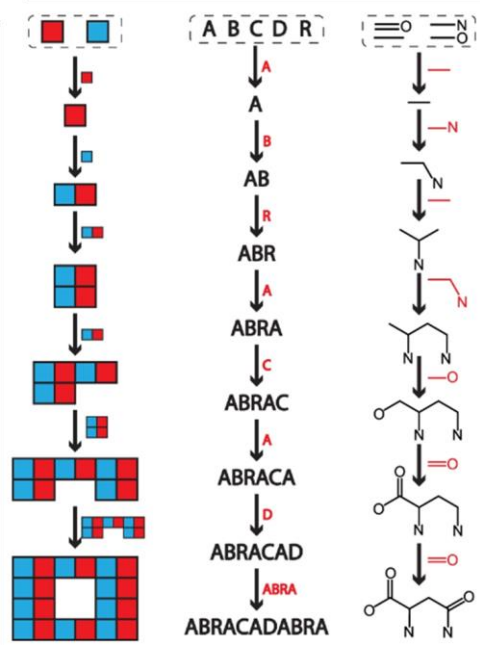

Figure 1: Three graphical representations of assembly index calculations. Arrows are join operations connecting states in the assembly pathway. Molecular assembly refers to the assembly index of molecules (right). The figure was reprinted with permission (Marshall et al, 2021).

Results

Using the aforementioned principles, we devised an automated experimental framework to identify selection in a recursive system of oligomerizing molecules with closed-loop analytics. Such open-ended oligomerization reactions of a defined set of monomers, in a recursive setting, could produce very long and intricate oligomers in high abundance. Since it is difficult to directly calculate their MA, we plan to reduce their graphical structure to a string of monomers, akin to proteins or DNA, as it has been shown that direct MA quantification is at least as hard as NP-complete (Liu et al, 2021). This reduction to a string representation will allow us to generate quicker and more reliable MA calculations and enact string information procedures to monitor the complexity of the sampled system states.

The automated experimental framework, coupled with inline mass spectrometry, enables a long-term recursive process where fresh starting materials are introduced to the reactor every cycle, thus maintaining the reactivity of the system while retaining a *memory* of previous cycles (Asche et al, 2021). The automated set-up provides a high degree of reproducibility and objectivity. In order to drive the complexification of the reactive monomers towards selective states, we will implement a non-biased algorithmic decision-making to push complexity forward. The algorithm will test different input compositions to the reactor, akin to a gradient decent, and decide the next recursive cycle that yields the highest mixture complexity. This will improve the likelihood of detecting the growing presence of selection in the system evident by the sustained increase of MA (Figure 2).

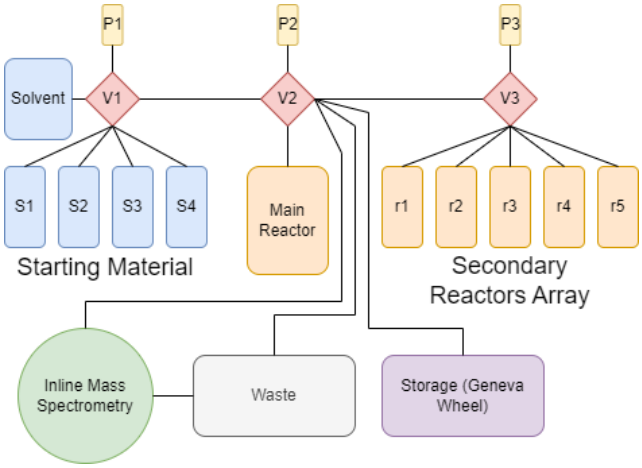

Figure 2: A diagram of the automated experimental framework, based on the chemputer design (Asche et al, 2021). After each recursive cycle in the main reactor, different inputs are tested in the secondary reactors in order to pinpoint the recursive step that will result in optimized complexification. Samples are analysed with inline compressed MS, and also stored for further validation on orbitrap LC-MS/MS. P – pumps, V – valves, S – starting material, r – reactor.

The detection and quantification of selection through dynamic changes in complexity of the molecular system will also allow us to determine conditions and agents that promote selection. For example, our preliminary results indicate that organocatalysts with higher MA have a higher potential for being more specific in their catalytic action (Figure 3). In other words, the more complex the molecule, the more it tends to chemically interact in a meaningful way with a limited set of molecules. This explains the process in which complex molecules could direct the chemical system towards more improbable selective states, which in turn may produce more complex molecules. In an open-ended recursive oligomerization system, where longer oligomers are constantly being generated with enormous diversity, this phenomenon explains why chemical selection may gradually emerge, and possibly how to identify and verify its underlying chemical processes.

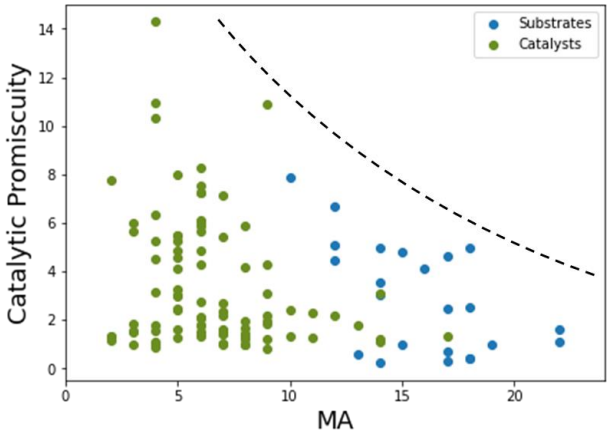

Figure 3: Preliminary results of hydrolytic reactions of nitrophenyl esters with potential catalysts. The project aims to fingerprint the kinetics of over 2300 reactions of substrate-catalyst pairs. The results reveal the relation between the MA of the involved chemical components and their catalytic promiscuity (the coefficient of variation of the catalytic parameters, converse to catalytic specificity).

Conclusions

In sum, our work aims to uncover the spontaneous emergence of chemical selection, by following a directed complexification process of simple molecules in assembly space. Our experiments are conducted in an automated recursive framework, and use a string-based MA model to accurately derive the complexity of lengthy oligomers over time. Coupled with clever algorithmic decision-making, we could drive the chemical system towards increased molecular and network complexity, with a higher potential for emergent selection. This work will shed light on the processes at the root of life and its major evolutionary leaps, and why complexity exists in the world in abundance.

References

Xavier, J. C., Hordijk, W., Kauffman, S., Steel, M., & Martin, W. F. (2020). Autocatalytic chemical networks at the origin of metabolism. Proceedings of the Royal Society B, 287(1922), 20192377.

Ameta, S., Matsubara, Y. J., Chakraborty, N., Krishna, S., & Thutupalli, S. (2021). Self-reproduction and Darwinian evolution in autocatalytic chemical reaction systems. Life, 11(4), 308.

Kahana, A., Segev, L., & Lancet, D. (2022). Protobiotic Network Reproducers Are Compositional Attractors: Enhanced Probability for Life's Origin. Available at *SSRN 4317204*.

Sharma, A., Czégel, D., Lachmann, M., Kempes, C. P., Walker, S. I., & Cronin, L. (2022). Assembly theory explains and quantifies the emergence of selection and evolution. *arXiv* preprint arXiv:2206.02279.

Colomer, I., Borissov, A., & Fletcher, S. P. (2020). Selection from a pool of self-assembling lipid replicators. Nature communications, 11(1), 176.

Vincent, L., Berg, M., Krismer, M., Saghafi, S. T., Cosby, J., Sankari, T., Vetsigian, K., Cleaves, H. J. & Baum, D. A. (2019). Chemical ecosystem selection on mineral surfaces reveals long-term dynamics consistent with the spontaneous emergence of mutual catalysis. *Life*, 9(4), 80.

Frenkel-Pinter, M., Matange, K., Rajaei, V., Costner, J. T., Robertson, A., Kim, J. S., Petrov, A. S., Bowman, J. C. & Williams, L. D. (2022). Chemical Evolution Reimagined. *ChemRxiv*.

Marshall, S. M., Mathis, C., Carrick, E., Keenan, G., Cooper, G. J., Graham, H., Craven, M., Gromski, P. S., Moore, D. G., Walker, S. I. & Cronin, L. (2021). Identifying molecules as biosignatures with assembly theory and mass spectrometry. *Nature communications*, 12(1), 3033.

Jirasek, M., Sharma, A., Bame, J. R., Bell, N., Marshall, S. M., Mathis, C., Macleod, A., Cooper, G. J. T., Swart, M., Mollfulleda, R. & Cronin, L. (2023). Multimodal Techniques for Detecting Alien Life using Assembly Theory and Spectroscopy. *arXiv* preprint arXiv:2302.13753.

Liu, Y., Mathis, C., Bajczyk, M. D., Marshall, S. M., Wilbraham, L., & Cronin, L. (2021). Exploring and mapping chemical space with molecular assembly trees. *Science Advances*, 7(39), eabj2465.

Asche, S., Cooper, G. J., Keenan, G., Mathis, C., & Cronin, L. (2021). A robotic prebiotic chemist probes long term reactions of complexifying mixtures. *Nature Communications*, 12(1), 3547.