Computational Discovery of the Origins of Life

Jan Meisner,†,‡ Xiaolei Zhu,†,‡ and Todd J. Martinez†,‡

†Department of Chemistry and The PULSE Institute, Stanford University, Stanford, California 94305, United States
‡SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, California 94025, United States

In silico reaction discovery using ab initio molecular dynamics shows that the chemistry of life could have originated from two simple inorganic molecules (HCN and water).

Abiogenesis, the emergence of life, is an everlasting and intriguing question in biology and related fields. Compared to the long-time scale of evolution, it appears likely that life sprung up almost instantaneously when conditions permitted it. Indeed, a recent study\(^1\) reported fossils with claimed biological origins dating back to the very oldest evidence of liquid water on Earth, not long after the planet formed 4.5 billion years ago. Despite the apparent effortlessness with which nature achieved abiogenesis, scientists did not know how any of the simple building blocks of life could have formed under early Earth conditions until Urey and Miller\(^2\)'s seminal experiments in 1952. Urey and Miller applied heat and electric sparks to a mixture of simple molecules believed to be abundant on early Earth and obtained various different amino acids. This discovery kicked off the quest to find the chemical origins of life, and many plausible pathways for nonbiological synthesis of all basic building blocks of life have now been proposed, assuming different scenarios including UV radiation, lightning, or hydrothermal vents. Now, Das and co-workers\(^3\) use novel computational methods to show that the chemistry of life could be generated with only two simple inorganic starting materials—water and HCN.

As more and more plausible reactions are proposed, it becomes increasingly evident that the role of individual reactions cannot be understood without considering the kinetics of a complex and strongly interwoven reaction network. Even very simple reaction networks can exhibit complex behaviors, with implications on how complexity could derive from simple molecules. However, studying such large reaction networks presents an almost insurmountable task for traditional tools of investigation, both theoretical and experimental. The number of reactions is extremely large, the conditions are diverse, unfamiliar, or even unknown, and the time scales range from picoseconds for the creation and decay of reactive intermediates to years for high barrier reactions. The traditional hypothesis-driven theoretical approach postulates reactions one at a time and then characterizes minimal energy paths and rates for these reactions. This approach falls flat on its face when confronted with large reaction networks and widely varying time scales. Thus, investigations of prebiotic chemistry demand the use of new techniques to map out large-scale reaction networks.

A recent contribution\(^3\) by Das et al. tackles this problem by using ab initio molecular dynamics (AIMD) simulations, which model bond rearrangement by explicit solution of the electronic Schrödinger equation. As shown schematically in Figure 1, they investigated the reactivity of hydrogen cyanide (HCN) and water under the conditions predominating in the Hadean Earth by means of the ab initio nanoreactor (AINR). Not only did they observe the formation of urea, formaldehyde, and other prebiotic molecules such as formaldehyde and glycolonitrile, they also detected oxazoles, cyanimide, and other important precursors for the synthesis of RNA. Interestingly, they observed many pathways that do not require a strongly reducing environment, in accordance with best estimates of the early Earth environment.\(^3\) They found water and ammonia (NH\(_3\)) acting ubiquitously as catalytic

These results clearly demonstrate how little we understand about the immense chemical space of prebiotic chemistry and confirmed ab initio simulations as a powerful tool for the exploration of complex reaction networks that complements experimental methods.
proton shuttles and conclude that HCN and water might be sufficient to generate a plethora of building blocks of life, even when limiting the considered reactions to those with barrier heights consistent with reasonably rapid reaction rates at $80-100 \, ^\circ \text{C}$ (they use 40 kcal/mol as a plausible upper bound). Under their presumption of a rather high HCN concentration, they could indicate further possible formation pathways to RNA, supporting the RNA world hypothesis. Perhaps more importantly, this work discovered a large number of hitherto unknown yet plausible reactions starting from only two components and without any catalyst. These results clearly demonstrate how little we understand about the immense chemical space of prebiotic chemistry and confirmed ab initio simulations as a powerful tool for the exploration of complex reaction networks that complements experimental methods.

The AINR used in the study of Das et al. is a new computational tool that accelerates AIMD simulations to construct chemical reaction networks in an automated fashion, free of preconceived notions of chemical reactivity. A variety of techniques can be employed to accelerate reactions, including temperature and a virtual piston that periodically pushes molecules together in order to enhance collisions. The computational bottleneck in the AINR is the solution of the electronic Schrödinger equation, which must be repeated for hundreds of atoms and at least millions of time steps. Fortunately, this obstacle can be overcome by exploiting new algorithms that leverage the graphics processing units (GPUs) originally developed for the video game industry. Using GPUs, molecular systems of a few hundred atoms can be studied on the nanosecond time scale to realize sufficient sampling of the space of chemical transformations, resulting in a dense reaction network. In the initial application of the nanoreactor, chemical reactions inspired by the Urey–Miller experiment were performed leading to similar molecules such as glucaldehyde, cyanimide, and glycine.

The AINR and related automated reaction discovery methods will have a lasting impact not only in prebiotic chemistry but also in a much broader swath of chemistry extending to the discovery of novel synthetic strategies and catalysts.

The whole area of automated discovery of chemical reactions and automated generation of reaction mechanisms has evolved over the past few years and shows clear signs of rapid expansion in the near future. Different concepts have been developed ranging from methods based on chemical transformation rules, connectivity graphs, and modified molecular dynamics or combinations thereof. Automated reaction discovery has been shown to be helpful to find new and unexpected reaction pathways in complex chemical systems such as the degradation of biomolecules and metal-catalyzed hydroformylation. The broad applicability of the AINR was also demonstrated by studies of the formation of terpene and related species.
inside nanocapsules and the aggregation of small molecules to form graphene. This suggests that the nanoreactor concept of reaction discovery should also be applied to other fields of chemistry, such as atmospheric chemistry, molecular processes in combustion engines, and complex organic reactions, as shown in Figure 2. We are convinced that the AINR and related automated reaction discovery methods will have a lasting impact not only in prebiotic chemistry but also in a much broader swath of chemistry extending to the discovery of novel synthetic strategies and catalysts. The contribution by Das et al. is a promising early step toward realizing that dream.

Author Information
E-mail: Todd.Martinez@stanford.edu.
ORCID
Jan Meisner: 0000-0002-1301-2612
Xiaolei Zhu: 0000-0002-7825-4815
Todd J. Martinez: 0000-0002-4798-8947
Notes
The authors declare the following competing financial interest(s): TJM is a founder of PetaChem, LLC.

Acknowledgments
Support of the Office of Naval Research (N00014-17-1-2875 and N00014-18-1-2659), Toyota Motor Engineering and Manufacturing North America, Inc., BASF, and the Deutsche Forschungsgemeinschaft (DFG 419817859) is gratefully acknowledged.

References
(1) Dodd, M. S.; Papineau, D.; Grenne, T.; Slack, J. F.; Rittner, M.; Pirajno, F.; O’Neil, J.; Little, C. T. S. Evidence for early life in Earth’s oldest hydrothermal vent precipitates. Nature 2017, 543, 60.
(2) McCollom, T. M. Miller-Urey and Beyond: What have We Learned about Prebiotic Organic Synthesis in the Past 60 Years? Annu. Rev. Earth Planet. Sci. 2013, 41, 207.