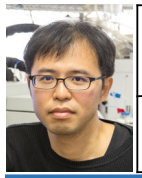
【Grant-in-Aid for Transformative Research Areas（B）】

Birth and Explosive Diversification of Proteins （Protein Big Bang）



Principal Investigator	RIKEN, Center for Biosystems Dynamics Research, Professor TAGAMI Shunsuke　　Researcher Number : 40586939
Project Information	Project Number : 24B305　　Project Period (FY) : 2024-2026 Keywords : the origin of life, structural biology, synthetic biology, evolution

Purpose and Background of the Research

●Outline of the Research

Earth is a planet filled with life. Then, is there life beyond us in this universe? Or are we alone? As our current technology only allow for limited observations of the universe, answers to such questions have yet to be found. In this research project, we will study on one of the key conditions for the emergence of life, the birth and explosive diversification of proteins or ”Protein Big Bang.”

Proteins are polymers made of 20 different amino acids, each folding into a unique shape according to its amino acid sequence. Each protein's unique structure enables various functions. In fact, the majority of our biological reactions are carried out by a wide variety of proteins.

However, the origins of functional proteins on Earth still hold many mysteries. The number of possible amino acid sequences of proteins is immense. For example, even a relatively small protein with 100 amino acids could have an astronomical number of sequences, reaching 20^{100}, far surpassing the estimated number of atoms in the universe (10^{80}). However, it is believed that the vast majority of such enormous sequences do not result in functional proteins. Therefore, the emergence of the first proteins on early Earth and their subsequent evolution into the diverse and complex proteins of life today seems almost miraculous.

In project, we aim to: 1) reproduce the processes of protein emergence and diversification on early Earth, and 2) analyze ongoing events of novel protein emergence." This will help us verify whether the Protein Big Bang is a miracle or an event that could easily occur. Furthermore, by utilizing the analyzed mechanism of novel protein emergence, we aim to 3) efficiently design proteins with completely new structures and functions for applications such as addressing environmental issues.

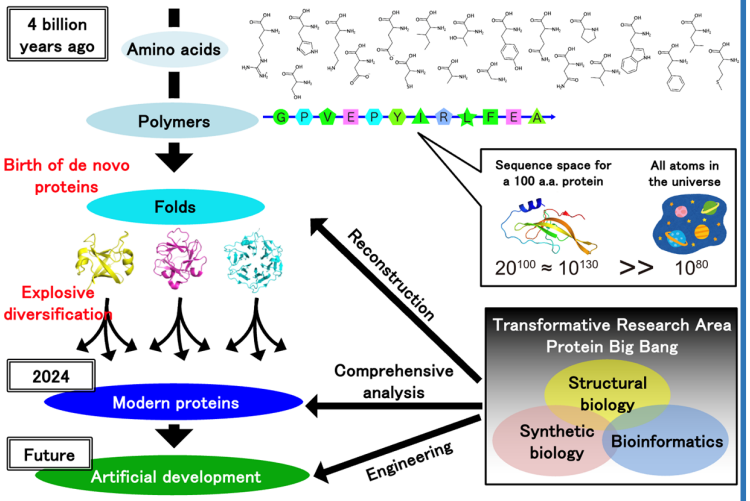


Figure 1. Protein Big Bang

●Ancient Novel Protein Birth and Evolution

We will experimentally reproduce ancient protein evolution processes to elucidate the mechanism of novel fold emergence from simple sequences. Furthermore, we will analyze the evolutionary pathways from simple proteins to giant protein complexes.

●Current Novel Protein Birth and Evolution

We will verify what novel protein folds are emerging within the vast biological resources present on Earth today.

●Future Applications of Novel Proteins

Design and artificial evolution of proteins with novel folds and functions, aiming for future practical applications.

Expected Research Achievements

●Research Objectives

・Ancient Novel Protein Birth and Evolution

We will attempt mutual conversions for various fundamental protein folds and estimate how many of them have independently emerged. Furthermore, we will create intermediate states of structural evolution of giant protein.
Additionally, to reproduce the process where entirely novel folds emerge through the linkage of short peptides, we will create libraries by linking secondary structure peptide fragments. Then, we will select sequences with superior stability and conduct detailed property and structure analysis to confirm fold emergence.

・Current Novel Protein Birth and Evolution

From vast uncultivated microbial genome and metagenome data, we will estimate de novo proteins with novel folds and their functions. Furthermore, for sequences whose structures cannot be predicted even with the latest AI, experimental structure analysis will be conducted.

・Future Applications of Novel Proteins

We will conduct experimental evolution to add functionality to various novel fold and verify the evolutionary potential of each fold. To expedite such research, we will develop high-speed fold screening by foldability and functional evaluation methods.

●Ripple Effects of the Research Field

・Elucidation of the Mechanisms of Life's Emergence in the Universe

As the next step after the Protein Big Bang project, we aim to advance co-evolution with RNA and reconstruct life as a system where numerous molecules work together in harmony. Furthermore, advancing such research could lead to the reconstruction of life systems using molecular species and side chains different from those on Earth (xenobiology). By integrating with astrochemistry and planetary chemistry, large-scale fusion research is expected to elucidate the entire mechanism of life's emergence in the universe.

・Achieving a Sustainable Society through the Use of Artificial Proteins

The design of proteins with functions surpassing those of natural proteins can become a powerful tool for solving important issues such as healthcare and environmental problems. The development of peptides and proteins (such as antibodies) with pharmaceutical activity is currently being vigorously pursued, and it is expected that creating artificially engineered cells to express them will enable continuous on-site synthesis of pharmaceuticals and more. Additionally, enzymes for the decomposition of harmful compounds can contribute to the removal of pollutants from factories.

Homepage Address, etc.	Website of PI: https://www.bdr.riken.jp/en/research/labs/tagami-s/index.html Website of the project: under preparation