Unraveling what it means to be alive

BDR TIMES

vol. 10
2022 SPRING

RIKEN Center for
Biosystems Dynamics Research

Dive into BDR's intriguing research

Approaching the Origin of Life with Structural Biology

It has been a while since I last talked to a structural biology researcher. I know that they use instruments like nuclear magnetic resonance (NMR) spectrometers to determine the structure of proteins, but I was very excited to be able to hear the story of what happens after the structure is determined as it sounded quite interesting.

Yakushiji (HY) ▶ You are part of Dr. Shunsuke Tagami's lab, right? Which means that your research is related to the structural analysis of DNA and proteins?
Yagi (SY) ▶ I am doing structural biology, but I'm more focused on the origin of life.

My target is proteins. Proteins are the elements that are the driving force of all kinds of biological phenomena, and I am trying to uncover how modern proteins emerged and evolved.

HY ▶ Proteins evolved…?
SY ▶ Modern proteins are made up of about 300 or 400 amino acid residues that are strung together like beads. Some of the larger proteins are made up of over 1,000 amino acid residues. But ancient proteins were likely much smaller. The smallest proteins today are about 100 amino acids in length. But even if they are only 100 residues long, there are 20 types of amino acids, so the number of possible variations is 20^{100}, or roughly 10^{130}.
HY ▶ That's an astronomical number.
SY ▶ Incidentally, the number of all atoms in the universe is considered to be 10^{80}.
HY ▶ More than the number of atoms in the universe…
SY ▶ Yes, there are so many possible variations. However, in some proteins, only a few mutations can render them useless. This means that while there are a lot of variations, there are also many useless sequences. It's almost a miracle that living organisms have been able to select only useful proteins from 10^{130} different variations.
HY ▶ Protein information is encoded in a DNA sequence, which is then transcribed into RNA, and RNA is further translated into a protein. So, I think that DNA and proteins need to evolve at the same time. It's a bit like a chicken and egg situation, isn't it?
SY ▶ That's right. One of the important molecules in this chicken and egg situation is RNA polymerase, which is responsible for transcription. My research theme is to unravel the evolution of this enzyme.
HY ▶ What about the molecule that translates the RNA sequences into proteins?
SY ▶ The main player in translation is the ribosome. The ribosome is basically made of RNA with some proteins. When pondering about the origin of life, we have to consider the co-evolution of ribosomes and RNA polymerases. By the way, have you ever heard of the "RNA world" hypothesis?
HY ▶ I have. It is one hypothesis for the origin of life which proposes that RNA was used to carry out a range of different functions, right?
SY ▶ Yes, that's right. So, there are a lot of people working on ribosomes, which are mainly composed of RNA.
HY ▶ But you are working on the RNA polymerase.

Targeting the core of RNA polymerase

SY ▶ We are trying to uncover the primitive form of RNA polymerase. RNA polymerase today is huge, but its catalytic core is composed of a small protein domain called double-psi beta-barrel (DPBB), which is about 80 amino acids in length. This DPBB structure is not only found in RNA polymerase, but in many other proteins as well.
HY ▶ Wow. They appear in so many proteins.
SY ▶ We can speculate that DPBB acquired several different functions as it evolved. So, we are trying to create an ancestral DPBB protein state.
HY ▶ I see.
SY ▶ Looking at the structure of DPBB, you can see that the structure of the N- and C-terminal halves are similar. So, it was thought that the ancestral DPBB was a dimer of these half peptides, and the current DPBB resulted from a duplication and fusion of the half gene.

I first decided to try making the sequence of the half-sized DPBB to check if it could fold into this structure. When I did, the half-sized peptide really did form the DPBB structure through homo-dimerization.

Additionally, from a structural biology perspective, it was thankfully easy to crystallize for structural analysis. When we tried to crystallize some mutants under 100 conditions, we could obtain crystals under about five or six conditions.
HY ▶ That's a high probability. Usually, you have to consider more than 1,000 conditions before you finally get one to crystallize. That's quite an efficient experimental system.
SY ▶ That's right (laughing). I think this protein has a relatively "rigid" structure making it easy to crystallize. It's also stable at high temperatures, like 80℃.
HY ▶ In the discussions on the origin of life, there are mentions of the hydrothermal vents in the deep ocean, so this seems plausible.

Sota Yagi

Special Postdoctoral Researcher in the Laboratory for Advanced Biomolecular Engineering. He is a father of two children and devotes his weekends to housework and looking after his children. Since he often goes to the park, he knows the particulars of all the parks in his neighborhood, such as which one has a lot of swings, or the one with the sandbox with cat poo.

Hideki Yakushiji

Business developer based in Kobe. He has a broad background in areas such as analytical chemistry, optics, biotechnology and IT. He is involved in a wide range of activities to assist in commercializing technologies and ideas born from academia, including setting up opportunities for idea sharing, finding investors, and strategic planning.

Simplification of amino acid components in sequence

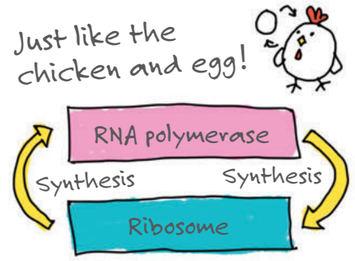
SY ▶ This protein is made up of a sequence of 43 amino acids, but seven of the 20 amino acid types are not used. In other words, it is made of 13 different amino acids. So, my next question was, "Is it possible to further reduce the number of amino acid types?"
HY ▶ How can you do that?
SY ▶ First, we reduce the number of amino acid types one by one, such as making a variant without methionine and another variant without leucine. I made six different variants, each without one type of amino acid, and to my surprise, I found that they all formed DPBB structures.
HY ▶ Wow. It must be a very strong structure.
SY ▶ Not only that, but I also tested variants without two or three types, and those also formed similar structures.
HY ▶ That's amazing.
SY ▶ And…
HY ▶ (There's more!)
SY ▶ In the end, I made a variant without all six types of amino acids I was testing, but this variant did not form any structure in solution.
HY ▶ (Too bad…)
SY ▶ But…
HY ▶ But…?
SY ▶ The crystallization process was a part of my routine, so I tried crystallizing this variant as well. Surprisingly, it did crystallize. I was able to confirm that the structure was almost identical to the others.
HY ▶ So, you can make this structure with only seven amino acids? That's amazing.

Primitive amino acids and primitive protein

HY ▶ Actually, it's even more interesting to look at these seven amino acids in a codon table.
HY ▶ By codon table, you are referring to the table showing how the DNA sequences correspond to amino acids.
SY ▶ Yes. Looking at this table, you can see that these seven amino acids are clustered near the bottom right side.
HY ▶ I see that.
SY ▶ In my recent work, I was able to reproduce the DPBB structure using only seven of the 20 amino acid types. The amino acids used in this experiment also have relatively simple molecular structures, so it's easy to think that these amino acids also existed in the ancient world. If so, it is conceivable that DPBB could have emerged even in early Earth.
HY ▶ I understand how the structure could be made, but I'm curious about its function.
SY ▶ That is a frequently asked question.
To be honest, I don't have the full answer yet, but we have found that it binds to DNA. β-barrel structures similar to DPBB are found in transcription factors and ribosomal proteins. So, now, I'm picturing that the ancient DPBB structure may have also bound to RNA and other nucleotides to do something.
HY ▶ In any case, nothing will start if they don't bind together, and when they were bound together, by chance, it might have turned useful. This cycle was probably repeated over and over. If you think about it in terms of the Earth's timescale, there is a hundred million years to experiment as much as you want.
SY ▶ It's very difficult to figure out how proteins evolved in the early world. It's kind of like science fiction thinking. So, I think one solution to investigate protein evolution is to experimentally reproduce the evolutionary process. Now that I have been able to simplify RNA polymerase to this level, I would like to look at the evolution of RNA polymerase even further in the past.

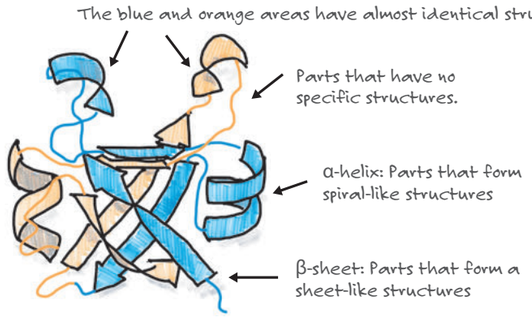
Clues for researchers focusing on evolution of translation

HY ▶ The tools used are so advanced nowadays that it seems like a world apart from 20 years ago.
SY ▶ That's true. This time, we used computational calculations to design the sequences and to predict their structures, but I couldn't do that myself. So, we collaborated with Kam Zhang and his team in the Laboratory for Structural Bioinformatics. I have also recently been using an AI system called AlphaFold.
HY ▶ We talked earlier about the chicken and egg problem, but because RNA polymerase is also made of proteins, I guess research on ribosomes, which make proteins from RNA, is also important.
SY ▶ This time, we have discovered that the core of the RNA polymerase protein seems to be composed of only seven amino acids types.
HY ▶ Oh, I see. So, the next step would be to make a ribosome that can polymerize these seven amino acids.
SY ▶ That's right. Modern proteins consist of 20 amino acid types, so we need to know what kind of function is needed to polymerize the 20 different amino acid types in the translation process. But we discovered that the core structure of RNA polymerase can be made from just seven amino acid types, so it should also be possible to reduce the number of amino acids needed to seven when considering the evolution of the translation system.

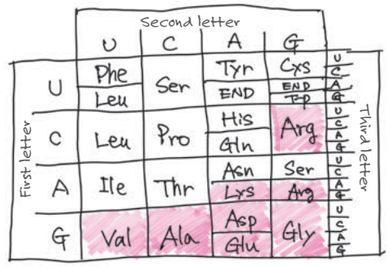


HY ▶ It sounds like the bar has been lowered a bit.
SY ▶ There are a lot of people around the world doing research on ribosomes, and I think our results provide them with some clues.
HY ▶ Your contribution to this topic was made possible because you went against the grain and chose to study RNA polymerase, which no one was really working on.
SY ▶ So, I hope we can collaborate with teams working on translation in the future.

DPBB = Catalytic core found in many proteins



Codon table
✓ Three nucleotides encode one amino acid
✓ The combinations are shown in the table
E.g. If the nucleotide sequence is UGG, then it encodes the amino acid Trp (tryptophan)



→ The seven amino acids used this time are found positioned near the bottom right of the table...!?

Postscript

This time, we started with the surprise fact that the "origin of life" is actually one of the fundamental themes of biology. It was a very interesting interview because of my own personal interest in the topic. It also made me realize that technology has advanced to a point where we can experimentally verify different ideas.

Read other interviews ▶

BDR Research Highlights

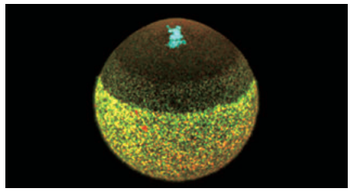
Research highlights articles and press releases between November 2021 to February 2022.
Read these and other articles on the website.

01

Successful fertilization requires careful coordination of chromosomes

For a fertilized egg to develop normally, it must inherit one set of genomic DNA from the mother and one set from the father and maintain diploidy, and errors in the process can lead to miscarriage or birth defects. In mammals, it is known that immediately after fertilization the fertilized egg transiently becomes a triploid (one paternal and two maternal genomes). Masashi Mori in the Lab for Chromosome Segregation and his colleagues have developed a microscopy strategy that allowed them to directly visualize chromosomal dynamics in a fertilized mouse egg. They discovered that the unfertilized egg organizes its internal protein infrastructure in a way that biases sperm fusion at sites far from the maternal chromosomes. When fertilization occurs, the same arrangement of proteins helps to sequester paternally and maternally contributed DNA so that the sperm chromosome is protected from being released along with the extra maternal genome copy, thus ensuring that the fertilized egg has a complete set of maternal and paternal DNA. This suggests that the intracytoplasmic sperm injection used in fertility treatments requires careful control of sperm injection location.

Mori M, Yao T, Mishina T, et al. *J Cell Biol* 220, e202012001 (2021)

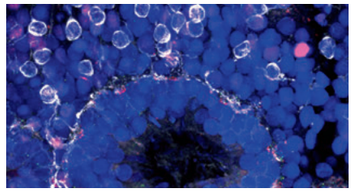


02

Improved retinal transplant technique almost ready for clinical trials

Retinitis pigmentosa is a hereditary disease in which photoreceptors in the retina die, leaving people with complete loss of vision or progressive loss in certain spots. One promising therapy is to replace the part of the retina at the back of the eye with a new retinal sheet, including the photoreceptors, grown from stem cells. For this regenerative cell therapy to work, the new light receptors in the graft must connect to neurons in the host retina, allowing light from the outside world to be relayed to the brain, which is how we see. Researchers led by Michiko Mandai (Lab for Retinal Regeneration) have now used a genetic modification to improve human-derived retina transplants grown in the lab. After transplant into damaged rat retinas, timed removal of certain cells from the grafts allowed better connections to host retinas, which resulted in more responsiveness to light in the damaged eyes. Because the retinal sheets were generated from stem cells of human origin, this represents one of the final steps necessary before this technique can be tested in human clinical trials for repairing retinal degeneration.

Yamasaki S, Tu HY, Matsuyama T, et al. *iScience* 25, 103657 (2022)

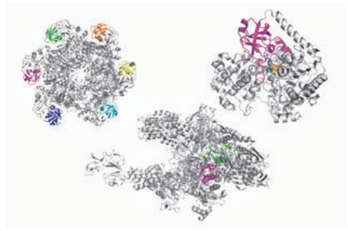


03

Plausible steps toward the evolution of a key protein fold of RNA polymerases

Proteins are the most important functional molecules of living systems, but their origins are one of the great mysteries in the life sciences. RNA polymerase is an essential enzyme that is involved in the transcription of DNA into messenger RNA. The key central region of RNA polymerase known as a double-ψ β-barrel (DPBB) is found in several other key proteins, suggesting that it evolved early in the origin of life. Sota Yagi and Shunsuke Tagami (Lab for Advanced Biomolecular Engineering) have discovered that the DPBB fold can be built from just seven of the 20 amino acids found in modern proteins. They also found that these amino acids can be coded for by only a small subset of the molecular coding features of the modern genetic code. Their experiments demonstrated that DPBBs can arise from the combination of two identical small sections of protein, which suggests that the ancient folded structure originated as a homodimer of a short unstructured peptide. This work indicates that folded proteins could have emerged much more easily than protein scientists had previously imagined.

Yagi S, Padhi AK, Vucinic J, et al. *J Am Chem Soc* 143, 15998-16006 (2021)



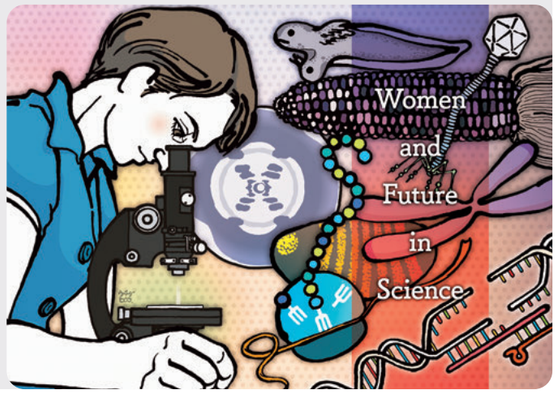
Shining a light on women in science

News

The RIKEN BDR launched a new seminar series called, "Women and Future in Science", in November 2021 that features women researchers working in life science-related fields and their achievements, with the aim to enhance the visibility of women researchers in the scientific community and to encourage and empower other budding women scientists, including postdoctoral researchers and students, at BDR and at other institutions to strive for the scientific accomplishments achieved by the speakers. As of February 2022, the seminar has been held three times with talks from six talented women based in Japan and the United States, and more are in the process of being scheduled.

The idea for the Women and Future in Science Seminar was proposed by BDR's Diversity Working Group, a group that was initially formed through discussions of several BDR team leaders who felt a strong need to raise awareness and address issues related to diversity and equality at the Center. While there will be an emphasis on featuring talks by early-career women scientists working on cutting-edge research, the organizers are also planning to include talks from more established women scientists to share their work, their experiences and their wisdom.

This seminar series is held online via Zoom, and anyone outside BDR can also participate by registering in advance. Go to the Women and Future in Science Seminar website for upcoming dates and to register: https://www2.bdr.riken.jp/bdrdiversity/seminars/wfs/



Peek-a-LABoo

Team Leader Fumiaki Obata joined RIKEN BDR in April 2021 from his former post at The University of Tokyo to lead the new Laboratory for Nutritional Biology. He first began working with fruit flies, *Drosophila melanogaster*, during his doctoral research in the laboratory of Dr. Masayuki Miura at The University of Tokyo, a renowned scientist in the field of cell death. While analyzing a cell-death defect mutant fly line, he uncovered a phenotype linked to metabolism, leading him to shift to metabolism-related research. He began investigating how food (i.e. nutrients), which are the basic elements of metabolic pathways, influences lifespan. After obtaining his doctorate, he joined the laboratory of Dr. Alex Gould in the U.K., who is well known for metabolism research using fruit flies, where he made another surprising finding that a certain dietary condition changed the gut microbiota, and this change led to a longer lifespan. These successive serendipitous discoveries have been a guiding force in expanding the scope of his research interests. We talked to Dr. Obata to learn more about the Laboratory for Nutritional Biology.

Laboratory for Nutritional Biology

Q How many members are in the laboratory?

A There are currently (as of February) five people in the laboratory–three students who came with me from The University of Tokyo, a part-time administrative staff, and me. From April, one of the students will become a RIKEN Special Postdoctoral Researcher (SPDR) and we will also have one research scientist, one technical staff, and one master's program student joining the laboratory bringing the total number to eight people.

Q What are the main research themes of the laboratory?

A The overarching theme of the laboratory is investigating how diet, particularly the nutrients in food, influences health and lifespan using the fruit fly as a model. The advantage of using *Drosophila* is firstly their lifespan is short–two to three months–which allows us to evaluate healthspan quickly, cheaply, and without needing a lot of space. The fact that they are easy to genetically manipulate is also another advantage. We also need to consider how gut microbiota is affected in response to diet conditions. Understanding how each nutrient or bacteria affects lifespan requires actually genetically manipulating the host (*Drosophila*), thus with methods already established for manipulating these three properties, the *Drosophila* is an excellent model for use in our research.

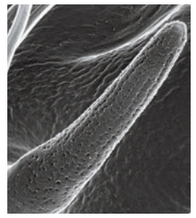
Q What are some of the strengths of the laboratory?

A It is often overlooked, but many studies using model organisms are carried out without considering diet (food) or gut microbiota, which presents a possibility that the data collected has been unknowingly affected by any of these factors. One of the main strengths of our laboratory is that, in addition to being able to manipulate *Drosophila* genes, we are equipped to precisely modify the dietary nutrients as well as the gut microbiota. When conducting experiments, we need to clearly know the conditions of the nutrients and gut microbiota as well as the genome sequences of *Drosophila* and form conclusions after taking these aspects into consideration.

Q Is there a long-term goal you are working toward in your research?

A We humans often eat three times a day, consuming large amounts of nutrients through our meals. The booming health food industry suggests that there are many people who are careful about certain nutrients they intake. It's said that there are over 26,000 types of nutrients, and we go about our daily lives without being aware of the amount of each nutrient contained in the food we eat. Our healthspan is likely greatly affected by whether there are large or small amounts of certain nutrients (e.g. amino acids) in our food. I hope to be able to explain how these nutrients affect our healthspan from a molecular perspective. Translating our work into real-world settings is also something I am interested in. Understanding the fundamental relationship between nutrients, gut microbiota, metabolism and healthspan is of course our underlying goal, but I think our research is easily applicable to humans as well. I would be very pleased if we can translate our findings to benefit society. For example, many people nowadays are taking nutritional supplements, but it is still not clear whether they are good for our health or how and when to take them for them to be effective. We may be able to propose possible interventions that are backed by scientific evidence.

On the cover!



A growing baby corn?

This is an electron microscope image of an olfactory sensillum found on the head of a fruit fly. Nanometer-sized (one-millionth of a millimeter) odorant molecules can pass through tiny pores (nanopores) on the sensillum surface to stimulate the nerves inside for detecting odors. In contrast, micrometer-sized (one-thousandth of a millimeter) dust and virus particles are prevented from entering the body because they cannot pass through the pores.

◎Credit:
Lab for Morphogenetic Signaling

BDR TIMES vol.10
Issued April 1, 2022
Published by RIKEN
Center for Biosystems
Dynamics Research (BDR)
E-mail: bdr-riken@ml.riken.jp
https://www.bdr.riken.jp/en/

RIKEN 2022-023