State-of-the-art challenges in the natural sciences

Synthetic Biology

ANDRZEJ B. LEGOCKI
Institute of Bioorganic Chemistry, Poznań
Polish Academy of Sciences
legocki@ibch.poznan.pl

Prof. Andrzej B. Legocki, chairman of the Academy's Division II – Biological Sciences and creator of one of Poland's leading centers for structural and molecular biology, works on resolving the conformation of proteins and nucleic acids and on identifying their functions

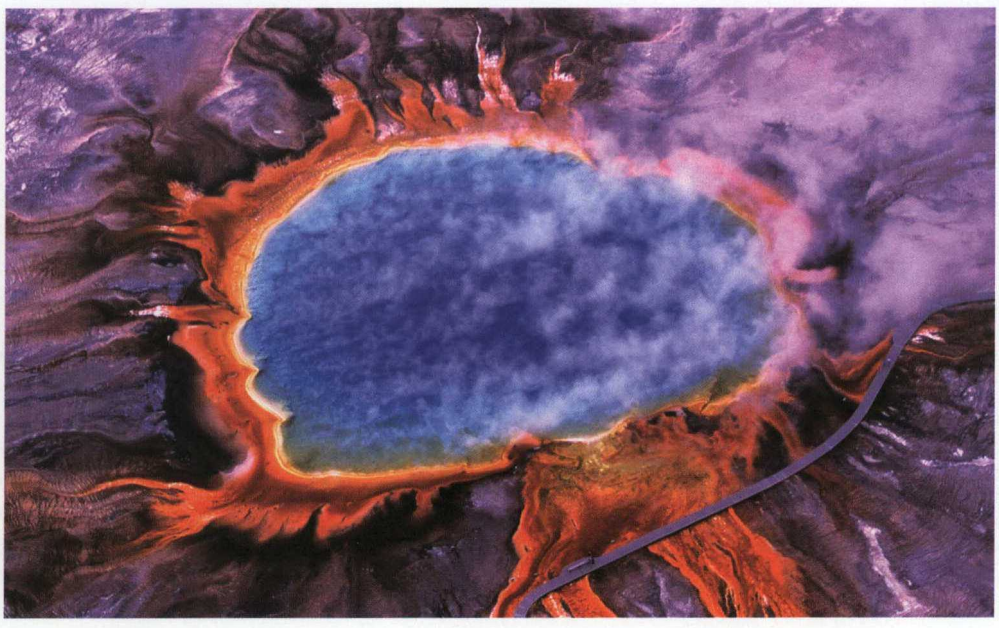
By combining the latest advances made in two major fields of science (molecular biology and applied organic chemistry), we can make fascinating new inroads in healthcare and environmental protection

In the latter half of the 20th century, the natural sciences were dominated by a reductionist approach. This culminated in the emergence of genomics, a new field of molecular biology studying the architecture of the genomes present in living organisms. The spectacular advances of genomics, with the successful completion of the human genome in 2003, have laid the basis for the development of new, state-of-the-art technologies for analyzing whole populations of proteins (proteomics), cellular metabolites (metabolomics), and the three-dimensional structure of proteins and nucleic acids.

Public databases now contain 840 fully sequenced microbial genomes with a further 1500 on the way, and DNA sequences of 5 million genes are now available from 6000 different species. Yet this knowledge of the molecular record itself is not sufficient to explain the appearance of life on Earth. Many key dilemmas (such as the chance-vs.-necessity debate) encroach into the domain of philosophical deliberation.

From systemic to synthetic

As more and more molecular data has been gathered, integrated research approaches have begun to develop, i.e. attempts have been made to generalize natural phenomena more broadly. The field of systems biology seeks ways of describing the mechanisms of natural biological systems – ranging from the cell-internal networks of individual organisms up to whole groups of genetically distant yet behaviorally intertwined species. And here, the further application of knowledge gained from organic chemistry has opened up a new field of research called *synthetic biology*. Its objective is to artificially construct molecules and biological systems that do not occur in



Yellowstone National Park

The hot springs of Yellowstone, where conditions are similar to those which prevailed on the Earth back when the first life was formed, have been found to contain many astounding organisms. They may prove very useful to scientists for technological and research purposes

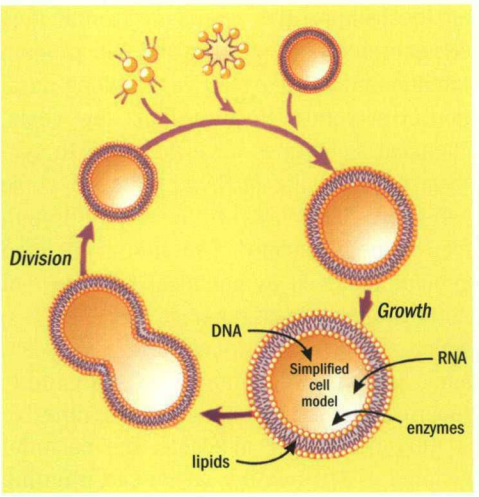
nature, ones which exhibit traits specifically chosen to be useful. This means not so much copying and improving upon existing natural systems, as creating new simple biological systems that do not exist in nature, to be used for technological and research purposes. Over the coming two decades, synthetic biology can be expected to play a key role in the development of biomedicine, the pharmaceutical industry, environmental protection, and even reducing our dependence on conventional energy sources.

Minimal cell

Living cells are the most perfect "reactors" for manufacturing all sorts of organic products. Plant cells, for instance, are capable of producing 200,000 different types of metabolites! This is the result of plants' evolutionary adaptation to being fixed to a single location, which means often being exposed to difficult environmental conditions. Plants have sizeable genomes, coding various defensive and adaptive mechanisms that are activated in reaction to threats or stress stimuli. Microorganisms (bacteria and viruses), on the other hand, have significantly smaller genomes. One of the first practical problems facing synthetic biology is how to create simple, synthetic or semi-synthetic cellular systems. The point of departure here is a "minimal cell" concept. Defining such a cell, or more precisely its genome, posed an intriguing challenge for both fundamental research and developing technological applications. How large a set of genes would be needed to support the simplest life functions - including the self-replicating capability, independent metabolism, and maintaining stability? The organisms with the smallest known genomes are the human parasites *Mycoplasma genitalium* (540,000 nucleotides coding 517 genes) and aphid symbiotic bacteria *Buchnera aphidicola* (450,000 nucleotides, 450 genes). By comparison, the genome of the model bacteria *Escherichia coli* consists of 4.6 million nucleotides and codes 4300 genes, whereas the human genome runs 3.2 billion nucleotides long and codes 25,000 genes.

This "minimal cell" concept touched off an extremely interesting line of inquiry of much wider import, pertaining to the earliest stages of the evolution of life on Earth. We

Andrzej B. Legocki



Simplified diagram of the growth and division of liposome bubble structures as models of primeval cellular membranes (from Szostak, J.W. et al., 2001)

can assume that prebiotic events followed significantly simpler scenarios than the modern cell cycle, at the very least because there was a significantly smaller number of genes and molecular components involved. In this sense the simple artificial cell systems now being developed can be viewed as possible models of the early stages of life on our planet. They are being obtained by two different approaches. The reductionist approach works by reducing the number of components of a present-day cell using the methods of genetic engineering, whereas the alternative evolutionary approach starts with the simplest cellular structures and then gradually builds up from there, increasing their complexity. Most of the monomers that make up biological macromolecules can be produced from simple inorganic compounds, through chemical reactions. This may indeed have been how the chain of events that led to the emergence of life was triggered shortly after our Earth was formed. It is estimated that no more than 250-350 genes are required to maintain the life functions of bacteria.

My name is LUCA

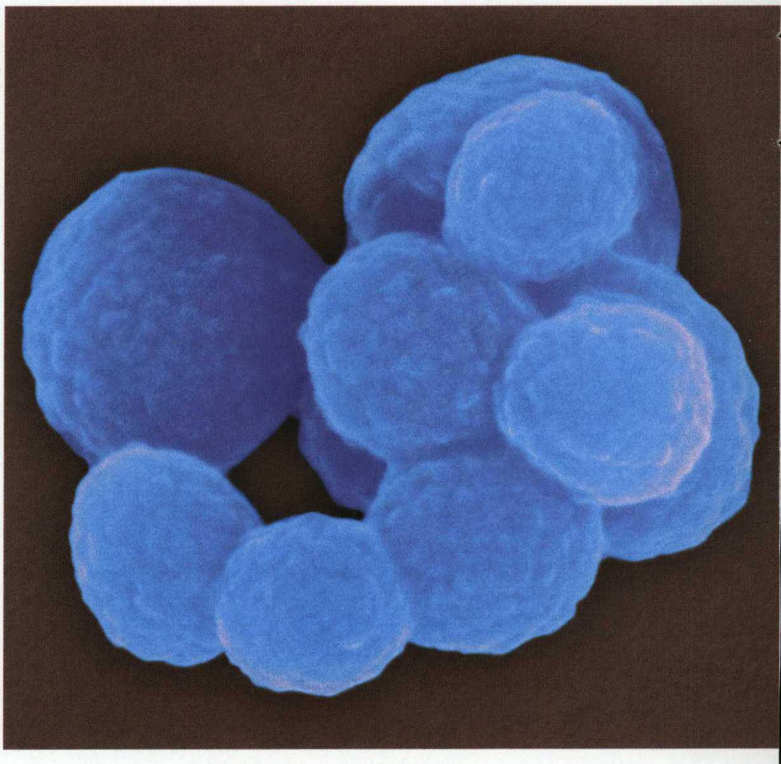
The study of how archaic genomes were organized 3.5 billion years ago could bring us nearer to comprehending the amazing complexity and diversity exhibited by living organisms. The idea of a *Last Universal Common Ancestor* (or LUCA) is derived from the Darwinian concept that all living organisms could be traced back to a single ancestor. However, attempting to recreate the LUCA genome is a purely theoretical task, because it is now impossible to retrace

all the complexities of the horizontal flow of genes between species and the random mutations that occurred over the long duration of evolution. Over the past few years, chemical techniques have managed to synthesize long chains of DNA, and even whole genomes. This marks a huge breakthrough for synthetic biology. The first completed genome obtained by chemical synthesis, in 2004, was the bacteriophage ΦX174 DNA, which is 5386 nucleotides long. Five years later, the same J. Craig Venter Institute, reported that it had managed to synthesize a 100-times longer molecule: the bacterial genome of *Mycoplasma genitaliom*, running 582,970 nucleotides long. Its synthesis involved obtaining and then linking together 101 fragments, each five to seven thousand nucleotides in length. A special technology was developed to introduce such long DNA chains into an artificial *S. cerevisiae* yeast chromosome. This synthesis has been recognized as a promising milestone towards the goal of constructing a cell machinery in the laboratory that is comprised of de novo created elements and circuits.

Out of the depths

The search for organisms that may have retained useful ancient properties that cannot be found anywhere else has turned its attention to the world's oceans. Many surprising microorganisms were discovered living in the depths, under conditions similar to those that prevailed on the Earth back when the first life was forming. For example, *Methanococcus jannaschii*, a kind of autotrophic bacteria (which utilize simple inorganic compounds to sustain their life functions) with a genome containing 1.7 million nucleotides, coding 1800 genes, was discovered at a depth of 2600 m (where the pressure is 200 atmospheres), in a volcanic area of the Eastern Pacific. These bacteria are able to transform hydrogen and carbon dioxide into methane used as feedstock for other fuels. The discovery of this bacterial property has encouraged scientists to try to harness it to generate "next-generation" fuels from CO_2, which is the main culprit of the greenhouse effect. The initiators of this idea (from the J. Craig Venter Institute) have declared that they expect to obtain the first bacteria to be used for technological purposes by the end of 2010! If they keep their promise and it does prove to be feasible to generate methane from carbon dioxide, a bold new era of ecological biotechnology will begin.

BE&W/PHOTOTAKE



What is required to create a "minimal cell"? How many genes will it need to have to function? *Mycoplasma genitaliom* is the living organism with one of the smallest genomes – having just 540,000 nucleotides coding 517 genes

Engineering new proteins

Life processes are based on a relatively small number of protein structures, representing just a small fraction of the theoretically possible combinations of 20 amino acids. Experiments have even shown that it is possible to form active protein molecules built out of a 9-amino-acid alphabet. We do not know why certain protein structures did appear in evolution, but others did not. This may have been chance, or it may have resulted from favoured thermodynamic properties of certain structures, such as their greater durability within a specific environment. New avenues of research in computer-simulated protein engineering seek to identify protein models built out of amino acid sequences that do not occur in nature, for which the term "never born proteins" has been adopted. Several research consortia set up with this specific aim are now working on developing such proteins. Work is also underway to increase the durability of physiologically important proteins (such as serum

proteins), by adding certain non-naturally-occurring amino acids to increase resistancy of protein molecules against chemical degradation within the cell. Precise "biomolecule sensors" are also being designed to monitor metabolic pathways for cancer diagnostics.

Artificial membranes

Research efforts towards creating a minimal cell comprised of laboratory-derived biological components utilize liposomes (liquid bilayer vesicles), which can play the role of biological membranes. Lipid vesicles exhibit a natural tendency for a certain type of self-organization. Liposome "bubbles" are able to integrate protein and nucleic acid molecules in functional compartments. The first success here was the replication of the bacteriophage Qβ RNA, inside appropriately prepared liposomes. A functional membrane must allow transport of small molecular substances across the boundary and show a capability to grow. Making artificial cellular membranes functioning in a controlled manner seems to be one of the main challenges of synthetic biology.

For health and the environment

There are now many research programs underway in the field of synthetic biology. Some of them aim to produce specific biotechnological products, such as the EU program "Programming bacterial catalysis *à la carte*," which has an objective related to environmental protection. This project focuses on obtaining bacteria cells with genomes significantly reduced in size, still retaining their capacity to degrade dangerous chloroaromatic compounds.

Another group of important synthetic biology projects seeks to alter the metabolic pathways in model organisms so that they produce specifically desired biopreparations. An excellent example of this can be found in attempts to have yeast cells biosynthesize artimisinic acid - a substance that is a precursor of an important anti-malerial drug (artemisinin) but can be found in nature only in small quantities (in the sweet wormweed plant *Artemisia absinthium*).

Malaria has plagued mankind since the time of the first hominids and takes a greater toll nowadays than it ever has, with nearly half a billion people contracting it each year. It is caused by several species of protists which are now growing more resistant to quinine-based drugs, and so a new drug urgently needs to be developed. Yeast cells were chosen to host artemisinin biosynthesis, with 12 artificial genes added to reprogram the cellular processes of what is known as the mevalonate pathway.

Whether such new biological advances can successfully be put to practical use hinges upon many factors. Public acceptance is extremely important. One unsuccessful attempt is exemplified by the story of the transgenic *Golden Rice* variety, which produces β-karoten (provitamin A), an important dietary supplement especially valuable in regions affected by famine and malnutrition. Lines of this modified rice were first developed back in the year 2000. Although the achievement was widely publicized and despite the great demand for such "enhanced" rice, it has so far not been introduced into widespread cultivation.

Importance of being prudent

The pursuit of synthetic biology may hold vast benefits in store, but as is often the case with scientific progress, it does also entail certain risks. Scientists need to provide the clearest possible answers to the questions synthetic biology raises. Does the manipulation of synthetic genomes and semi-synthetic biological installations involve greater risk than working with genomes and systems of natural origin? Might introducing such genomes and self-replicating DNA chains into biotechnological use prove to be a dangerous move? Could genome chemical synthesis technology end up recreating dangerous pathogens that have been eliminated earlier from the environment?

It will not always be possible to find convincing answers to all the questions and doubts that arise, or at least not quickly. Prudence and accountability, values which should be at the forefront of every scientist's mind, are extremely important if not fundamental to the ethos of scientific inquiry. ■

Further reading:

Legocki A.B. (2009). Naukowe definicje życia [Scientific Definitions of Life]. *Nauka, 1*, 123-133.

Szostak J.W., Bartel D.P., Luisi L. (2001). Synthesizing life. *Nature, 409*, 387-390.