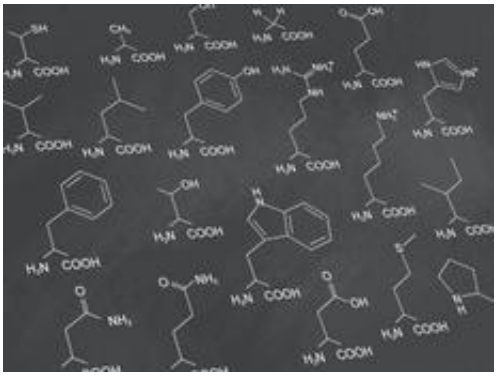
A glimpse into the evolution of proteins

March 6 2013, by Fabio Bergamin



Not all of the twenty or so amino acids of today's proteins (depicted, the structural formulae of a selection) were already part of the living world four billion years ago. Credit: Christine Khammash / ETH Zurich

(Phys.org) —ETH-Zurich researchers look several billion years back in time, when life on Earth was just beginning. In a laboratory experiment, they examined how a primitive protein was able to evolve. However, the journey into the past also afforded the scientists a glimpse into the future of synthetic biology.

Life did not emerge overnight. The first primitive life forms that developed on Earth around four billion years ago had little in common with today's organisms. They probably managed without proteins. And the first proteins that formed in the following few hundred million years to become essential elements of the living world also differed from those of today: scientists assume that the first proteins were composed of a

reduced repertoire of only seven or eight different amino acids. The repertoire of today's proteins, however, typically contains twenty amino acids.

A team of researchers headed by Donald Hilvert, a professor at the Laboratory of Organic Chemistry, has now simulated how the minimal starting amino acid repertoire might have expanded in the course of evolution. Not only do the researchers draw conclusions regarding the past from their work, they also gain important insights into future directions for synthetic biology.

Evolution of a simplified enzyme

The ETH-Zurich researchers conducted evolution experiments with a synthetic protein they developed some years ago in their lab. It is based on an enzyme that actually exists, a so-called chorismate mutase. While the real enzyme is composed of all twenty amino acids, however, the synthetic protein contains a reduced repertoire of only nine. "Our synthetic chorismate mutase has the same function as its natural counterpart, but is less stable and less active," says Hilvert.

The scientists have now examined how far the simplified protein can be optimised and how exactly it behaves if it is allowed to expand the amino-acid repertoire – just as it may have occurred during millennia of evolution. To this end, they introduced the corresponding gene for the protein into bacteria that do not contain a natural chorismate mutase and allowed the bacteria to grow in a bioreactor for one and a half months. In the course of this experiment, mutations naturally arose in the genetic material. The researchers then studied the extent to which the synthetic protein and its gene had changed.

Surprisingly only minor changes

Sure enough, the gene and the protein were altered in the experiment – only marginally but with major effects: the enzymatic activity of the protein and its stability increased. The fact that the changes in the sequence of amino acids were comparatively minor surprised the scientists. While the protein consists of roughly 100 individual amino acids, only two of them changed. At one site, the amino acid isoleucine was replaced with the amino acid threonine, and at another site leucine with valine. "Structurally speaking, these are only very minor changes. The new amino acids aren't all that different from the ones they replace," says Hilvert. The scientists would have expected a greater number of more significant changes. "Even if the two mutations are very conservative, we were able to show that they give the protein a clear selective advantage," explains Hilvert.

Insights for synthetic biology

The work of the ETH-Zurich scientists not only affords a glimpse into the history of evolution as it might have occurred almost four billion years ago; the researchers were also able to draw conclusions from it of relevance to synthetic biology. Part of this research field involves expanding the functionality of proteins and organisms through molecules that do not exist naturally.

These days, synthetically produced enzymes are already used in medicine and industry. In the future, not only might these contain the typical twenty natural amino acids, but also additional synthetic ones. The hope is that a larger repertoire will provide more possibilities in the design of artificial enzymes, such as with regard to their catalytic functionality. Moreover, enzymes made of synthetic amino acids are expected to be broken down more slowly in the body or a manufacturing process, which could be very beneficial for certain medical or industrial applications.

"One lesson from this project is that you shouldn't consider only spectacular synthetic amino acids in developing the amino-acid repertoire further in synthetic biology," says Peter Kast, a professor at the Laboratory of Organic Chemistry, who was also involved in the project. "Amino acids with minor differences from natural ones also proved to be very useful."

Why does biology "only" use twenty standard amino acids?

In fact, the genetic code upon which all present life on Earth is based would permit more than the twenty universal amino acids it contains. After all, the individual amino acids of proteins are encoded at the genetic level by a sequence of three so-called nucleotide bases (a base triplet) each. The alphabet of the genetic substance DNA contains four different bases and the number of possible triplets from these is sixty-four (four to the power of three). With the genetic code, up to sixty-four different amino acids could therefore be described. Instead, nature "only" programs a limited set of twenty amino acids and a redundant code: typically, individual amino acids are each described by several triplets. But why?

One reason is probably that this redundancy cushions the impact of mutations in the DNA, which was evidently an advantage during evolution. "With the repertoire of twenty universal amino acids, evolution achieved a momentary optimum," says ETH-Zurich Professor Donald Hilvert.

Apart from the twenty universal amino acids, a more complex incorporation mechanism allows the use of a twenty-first amino acid called selenocysteine. It is found in many living organisms, including humans. In certain microorganisms, the archaea, there is even a twenty-

second amino acid: pyrrolysine. This leads scientists to conclude that the genetic code has continued to evolve over the last billion years. Hilvert: "Evolution obviously isn't over. It could well be that additional amino acids will be incorporated into the code in due course."

More information: Muller, M. et al. Directed Evolution of a Model Primordial Enzyme Provides Insights into the Development of the Genetic Code. *PLOS Genetics*, 2013, 9: e1003187, doi: 10.1371/journal.pgen.1003187

Provided by ETH Zurich