

## **PRESS RELEASE**

**Source:** ELSI, Tokyo Institute of Technology, Japan

**For immediate release:** September 10, 2019

**Title:** Scientists Find Biology's Optimal "Molecular Alphabet" May Be Preordained

### **Subtitle:**

The amino acids, a fundamental set of life's building blocks, may have been adaptive throughout their evolution, suggesting a possible universal biological language.

### **Release summary:**

Life uses 20 coded amino acids (CAAs) to construct proteins. This set was likely evolutionarily "standardized" from smaller sets as organisms discovered how to make and encode them. Scientists modeled how the adaptive properties of the CAAs evolved over time. They found that sets containing even only a few CAAs were better than an enormous choice of alternatives, suggesting each time a modern CAA was discovered, it bootstrapped the set to include still more CAAs.

### **Full-text release:**

An international and interdisciplinary team working at the Earth-Life Science Institute (ELSI) at the Tokyo Institute of Technology has modeled the evolution of one of biology's most fundamental sets of building blocks and found that it may have special properties that helped bootstrap itself into its modern form.

All life, from bacteria to blue whales to human beings, uses an almost universal set of 20 coded amino acids (CAAs) to construct proteins. This set was likely "canonicalized" or standardized during early evolution; before this, smaller amino acid sets were gradually expanded as organisms developed new synthetic proofreading and coding abilities. The new study, led by Melissa Ilardo, now at the University of Utah, explored how this set evolution might have occurred.

There are millions of possible types of amino acids that could be found on Earth or elsewhere in the Universe, each with its own distinctive chemical properties. Indeed, scientists have found these unique chemical properties are what give biological proteins, the large molecules that do much of life's catalysis, their own unique capabilities. The team had previously measured how the CAA set compares to random sets of amino acids and found that only about 1 in a billion random sets had chemical properties as unusually distributed as those of the CAAs.

The team thus set out to ask the question of what earlier, smaller coded sets might have been like in terms of their chemical properties. There are many possible subsets of the modern CAAs or other presently uncoded amino acids that could have comprised the earlier sets. The team calculated the possible ways of making a set of 3-20 amino acids using a special library of 1913 structurally diverse "virtual" amino acids they computed and found there are  $10^{48}$  ways of making sets of 20 amino acids. In contrast, there are only  $\sim 10^{19}$  grains of sand on Earth, and only  $\sim 10^{24}$  stars in the entire Universe. "There are just so many possible amino acids, and so many ways to make combinations of them, a computational approach was the only comprehensive way to address this question," says team member Jim Cleaves of ELSI. "Efficient implementations of algorithms based

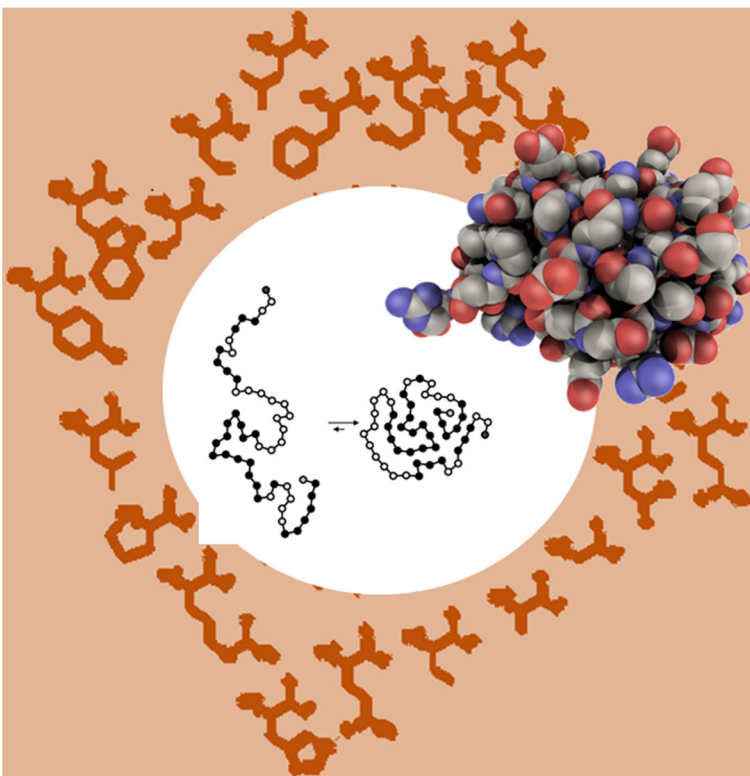
on appropriate mathematical models allow us to handle even astronomically huge combinatorial spaces," adds co-author Markus Meringer of the Deutsches Zentrum für Luft- und Raumfahrt.

As this number is so large, they used statistical methods to compare the adaptive value of the combined physicochemical properties of the modern CAA set with those of billions of random sets of 3-20 amino acids. What they found was that the CAAs may have been selectively kept during evolution due to their unique adaptive chemical properties, which help them to make optimal proteins, in turn helping organisms that could produce those proteins become more fit.

They found that even hypothetical sets containing only one or a few modern CAAs were especially adaptive. It was difficult to find sets even among a multitude of alternatives that have the unique chemical properties of the modern CAA set. These results suggest that each time a modern CAA was discovered and embedded in biology's toolkit during evolution, it provided an adaptive value unusual among a huge number of alternatives, and each selective step may have helped bootstrap the developing set to include still more CAAs, ultimately leading to the modern set.

If true, the researchers speculate, it might mean that even given a large variety of starting points for developing coded amino acid sets, biology might end up converging on a similar set. As this model was based on the invariant physical and chemical properties of the amino acids themselves, this could mean that even Life beyond Earth might be very similar to modern Earth life. Co-author Rudrarup Bose, now of the Max Planck Institute of Molecular Cell Biology and Genetics in Dresden, further hypothesizes that "Life may not be just a set of accidental events. Rather, there may be some universal laws governing the evolution of life."

#### Images:



Title: How amino acid properties affect protein folding

Caption: The coded amino acids each have unique properties that help biological proteins fold optimally. The set of amino acids biology uses appears to have been evolutionarily optimized.

Credit: H.J. Cleaves, ELSI/ Wikimedia Commons

### Reference:

Melissa Ilardo<sup>1</sup>, Rudrarup Bose<sup>2</sup>, Markus Meringer<sup>3</sup>, Bakhtiyor Rasulev<sup>4</sup>, Natalie Grefenstette<sup>5</sup>, James Stephenson<sup>6,7</sup>, Stephen Freeland<sup>8</sup>, Richard J. Gillams<sup>9,10</sup>, Christopher J. Butch<sup>9,11,12</sup> & H. James Cleaves II<sup>9,12,13\*</sup>

Adaptive Properties of the Genetically Encoded Amino Acid Alphabet Are Inherited from Its Subsets. *Scientific Reports*, DOI: 10.1038/s41598-019-47574-x

1. University of Utah Hematology, UC Berkeley Integrative Biology, George and Dolores Eccles Institute of Human Genetics, 15 N 2030 E, Room: 3240, Salt Lake City, UT, 84112, USA.
2. National Institute of Science Education and Research Bhubaneswar, P.O. Jatni, Khurda, 752050, Odisha, India.
3. German Aerospace Center (DLR), EarthObservation Center (EOC), Münchner Straße 20, 82234, Oberpfaffenhofen-Wessling, Germany.
4. Department of Coatings and Polymeric Materials, North Dakota State University, Fargo, ND, 58108, USA.
5. Department of Chemistry, University College London, 20 Gordon Street, London, WC1H 0AJ, UK.
6. European Molecular Biology Laboratory–European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, UK.
7. Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK.
8. University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, MD, 21250, USA.
9. Earth-Life Science Institute, Tokyo Institute of Technology, 2-12-1-IE-1 Ookayama, Meguro-ku, Tokyo, 152-8550, Japan.
10. Structural Genomics Consortium, Nuffield Department of Medicine, University of Oxford, Old Road Campus Research Building, Oxford, OX3 7DQ, UK.
11. Department of Chemistry, Emory University, 1515 Dickey Drive, Atlanta, GA, USA.
12. Blue Marble Space Institute for Science, 1001 4th Ave, Suite 3201, Seattle, WA, 98154, USA.
13. Institute of Advanced Study, 1 Einstein Drive, Princeton, NJ, 08540, USA.

\*Corresponding authors email: henderson.cleaves@gmail.com

### Contacts:

Thilina Heenatigala  
Director of Communications  
Earth-Life Science Institute (ELSI),  
Tokyo Institute of Technology  
E-mail: thilinah@elsi.jp  
Tel: +81-3-5734-3163

Henderson James Cleaves II  
Associate Professor  
Earth-Life Science Institute (ELSI),

Tokyo Institute of Technology  
E-mail: henderson.cleaves@gmail.com  
Tel: +1-85-8366-3049

**More information:**

**Tokyo Institute of Technology (Tokyo Tech)** stands at the forefront of research and higher education as the leading university for science and technology in Japan. Tokyo Tech researchers excel in fields ranging from materials science to biology, computer science, and physics. Founded in 1881, Tokyo Tech hosts over 10,000 undergraduate and graduate students per year, who develop into scientific leaders and some of the most sought-after engineers in industry. Embodying the Japanese philosophy of “monotsukuri,” meaning “technical ingenuity and innovation,” the Tokyo Tech community strives to contribute to society through high-impact research.

**The Earth-Life Science Institute (ELSI)** is one of Japan’s ambitious World Premiere International research centers, whose aim is to achieve progress in broadly inter-disciplinary scientific areas by inspiring the world’s greatest minds to come to Japan and collaborate on the most challenging scientific problems. ELSI’s primary aim is to address the origin and co-evolution of the Earth and life.

**The World Premier International Research Center Initiative (WPI)** was launched in 2007 by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) to help build globally visible research centers in Japan. These institutes promote high research standards and outstanding research environments that attract frontline researchers from around the world. These centers are highly autonomous, allowing them to revolutionize conventional modes of research operation and administration in Japan.