Researcher unraveling SARS-CoV-2 spike protein through music (Update)

April 10 2020, by Kim Martineau

MIT Professor Markus Buehler designs new proteins with the help of artificial intelligence. He recently translated the spike protein of the novel coronavirus (SARS-Cov-2) into sound to visualize its vibrational properties, as seen here, which could help in finding ways to stop the virus. Primary colors represent the spike's three protein chains. Credit: Markus Buehler

The proteins that make up all living things are alive with music. Just ask Markus Buehler: The musician and MIT professor develops artificial intelligence models to design new proteins, sometimes by translating them into sound. His goal is to create new biological materials for sustainable, non-toxic applications. In a project with the MIT-IBM Watson AI Lab, Buehler is searching for a protein to extend the shelf-life of perishable food. In a new study in *Extreme Mechanics Letters*, he and his colleagues offer a promising candidate: a silk protein made by honeybees for use in hive building.

In [another recent study](#), in APL Bioengineering, he went a step further and used AI discover an entirely new [protein](#). As both studies went to print, the Covid-19 outbreak was surging in the United States, and Buehler turned his attention to the spike protein of SARS-CoV-2, the appendage that makes the novel coronavirus so contagious. He and his colleagues are trying to unpack its vibrational properties through molecular-based sound spectra, which could hold one key to stopping the virus. Buehler recently sat down to discuss the art and science of his work.

Q: Your work focuses on the alpha helix proteins found in skin and hair. Why makes this protein so intriguing?

A: Proteins are the bricks and mortar that make up our cells, organs, and body. Alpha helix proteins are especially important. Their spring-like structure gives them elasticity and resilience, which is why skin, hair, feathers, hooves, and even cell membranes are so durable. But they're not just tough mechanically, they have built-in antimicrobial properties. With IBM, we're trying to harness this biochemical trait to create a protein coating that can slow the spoilage of quick-to-rot foods like strawberries.

Q: How did you enlist AI to produce this silk protein?

A: We trained a deep learning model on the Protein Data Bank, which contains the amino acid sequences and three-dimensional shapes of about 120,000 proteins. We then fed the model a snippet of an amino acid chain for honeybee silk and asked it to predict the protein's shape, atom-by-atom. We validated our work by synthesizing the protein for the first time in a lab—a first step toward developing a thin antimicrobial, structurally-durable coating that can be applied to food. My colleague, Benedetto Marelli, specializes in this part of the process. We also used the platform to predict the structure of proteins that don't yet exist in nature. That's how we designed our entirely new protein in the APL Bioengineering study.

Q: How does your model improve on other protein prediction methods?

A: We use end-to-end prediction. The model builds the protein's structure directly from its sequence, translating amino acid patterns into three-dimensional geometries. It's like translating a set of IKEA instructions into a built bookshelf, minus the frustration. Through this approach, the model effectively learns how to build a protein from the protein itself, via the language of its amino acids. Remarkably, our method can accurately predict protein structure without a template. It outperforms other folding methods and is significantly faster than physics-based modeling. Because the Protein Data Bank is limited to proteins found in nature, we needed a way to visualize new structures to make new proteins from scratch.

Q: How could the model be used to design an actual protein?

A: We can build atom-by-atom models for sequences found in nature that haven't yet been studied, as we did in the APL Bioengineering study using a different method. We can visualize the protein's structure and use other computational methods to assess its function by analyzing its stablity and the other proteins it binds to in cells. Our model could be used in drug design or to interfere with protein-mediated biochemical pathways in infectious disease.
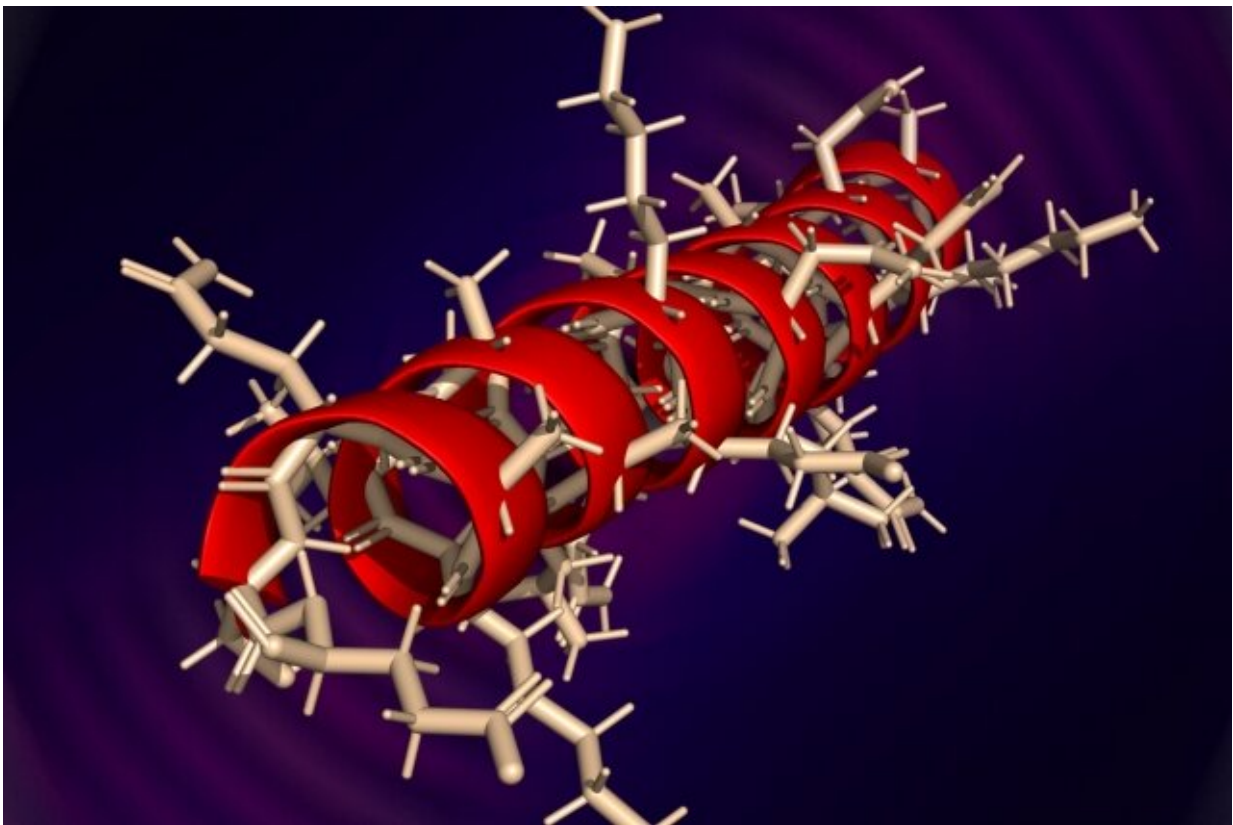
With the MIT-IBM Watson AI Lab, Markus Buehler and his colleagues used artificial intelligence to reproduce the alpha-helical protein that honeybees produce in making silk to build their hives. Credit: Markus Buehler

Q: What's the benefit of translating proteins into sound?

A: Our brains are great at processing sound! In one sweep, our ears pick up all of its hierarchical features: pitch, timbre, volume, melody, rhythm, and chords. We would need a high-powered microscope to see the equivalent detail in an image, and we could never see it all at once. Sound is such an elegant way to access the information stored in a protein.

Typically, sound is made from vibrating a material, like a guitar string, and music is made by arranging sounds in hierarchical patterns. With AI we can combine these concepts, and use molecular vibrations and neural networks to construct new musical forms. We've been working on methods to turn protein structures into audible representations, and translate these representations into new materials.

Q: What can the sonification of SARS-CoV-2's "spike" protein tell us?

A: Its protein spike contains three protein chains folded into an intriguing pattern. These structures are too small for the eye to see, but they can be heard. We represented the physical protein structure, with its entangled chains, as interwoven melodies that form a multi-layered composition. The spike protein's amino acid sequence, its secondary structure patterns, and its intricate three-dimensional folds are all featured. The resulting piece is a form of counterpoint music, in which notes are played against notes. Like a symphony, the musical patterns reflect the protein's intersecting geometry realized by materializing its DNA code.

Q: What did you learn?

A: The virus has an uncanny ability to deceive and exploit the host for its own multiplication. Its genome hijacks the host cell's protein

manufacturing machinery, and forces it to replicate the viral genome and produce viral proteins to make new viruses. As you listen, you may be surprised by the pleasant, even relaxing, tone of the music. But it tricks our ear in the same way the virus tricks our cells. It's an invader disguised as a friendly visitor. Through music, we can see the SARS-CoV-2 spike from a new angle, and appreciate the urgent need to learn the language of proteins.

Q: Can any of this address Covid-19, and the virus that causes it?

A: In the longer term, yes. Translating proteins into sound gives scientists another tool to understand and design proteins. Even a small mutation can limit or enhance the pathogenic power of SARS-CoV-2. Through sonification, we can also compare the biochemical processes of its spike protein with previous coronaviruses, like SARS or MERS.

In the music we created, we analyzed the vibrational structure of the spike protein that infects the host. Understanding these vibrational patterns is critical for drug design and much more. Vibrations may change as temperatures warm, for example, and they may also tell us why the SARS-CoV-2 spike gravitates toward human cells more than other viruses. We're exploring these questions in current, ongoing research with my graduate students.

We might also use a compositional approach to design drugs to attack the virus. We could search for a new protein that matches the melody and rhythm of an antibody capable of binding to the spike protein, interfering with its ability to infect.

Q: How can music aid protein design?

A: You can think of music as an algorithmic reflection of structure. Bach's Goldberg Variations, for example, are a brilliant realization of counterpoint, a principle we've also found in proteins. We can now hear this concept as nature composed it, and compare it to ideas in our imagination, or use AI to speak the language of protein design and let it imagine new structures. We believe that the analysis of sound and music can help us understand the material world better. Artistic expression is, after all, just a model of the world within us and around us.

Viral Counterpoint of the Coronavirus Spike Protein (2019-nCoV): soundcloud.com/user-275864738/ … ke-protein-2019-ncov

More information: Zhao Qin et al. Artificial intelligence method to design and fold alpha-helical structural proteins from the primary amino acid sequence, *Extreme Mechanics Letters* (2020). DOI: 10.1016/j.eml.2020.100652

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology

Citation: Researcher unraveling SARS-CoV-2 spike protein through music (Update) (2020, April 10) retrieved 9 April 2024 from https://phys.org/news/2020-04-unraveling-sars-cov-spike-protein-music.html