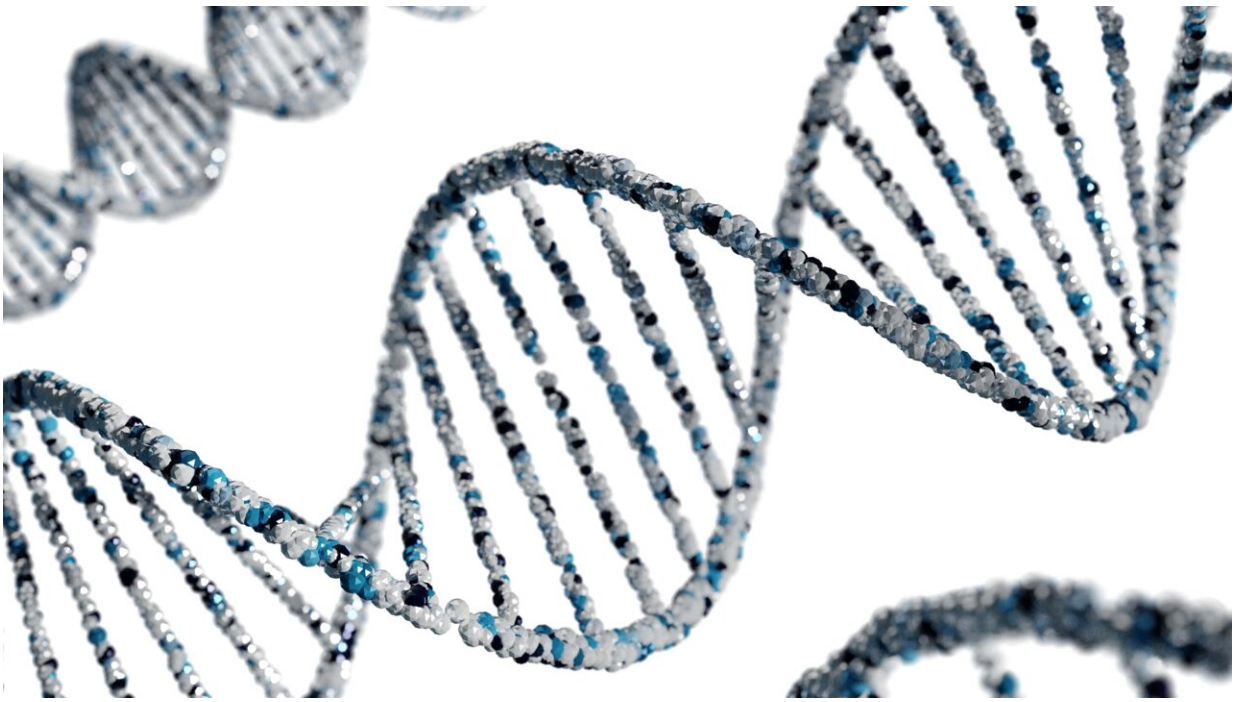
# Chemists find new pharmaceutically active substances from billions of newly combined molecules

September 2 2024, by Daniel Meierhans



Credit: Unsplash/CC0 Public Domain

Nowadays, there's lots of buzz about spectacular new medical treatments, such as personalized cancer therapy with modified immune cells or antibodies. Such treatments, however, are very complex and expensive and so find only limited application. Most medical therapies

are still based on small chemical compounds that can be produced in large quantities and thus at low cost.

The bottleneck in the development of new molecular therapies is the limited number of new active substances that can be found using current techniques. A method developed in the 2000s at Harvard and ETH Zurich promises to provide a remedy: DNA-encoded chemical libraries (DEL).

To date, DEL technology could be used to produce millions of chemical compounds and test their effectiveness in one go. However, the drawback with this was that the researchers could build only small molecules from a few chemical building blocks. Chemists at ETH Zurich have now refined and significantly improved this process.

With the help of the new method, published in the journal *Science*, researchers can now automatically synthesize and test not just a few million, but billions of different substances within a few weeks. The method can also be applied to produce much larger drug molecules, such as ring-shaped peptides, which can be used to target additional pharmacological targets.

## Creating and testing all combinations

"The first active substances developed with the help of early DEL technology are currently in advanced clinical trials. This new DEL method once again massively expands the possibilities," Jörg Scheuermann explains.

He and his research group at the Institute of Pharmaceutical Sciences are among the pioneers of DEL technology, which is considered to be the key to utilizing the combinatorial possibilities in the chemical production of molecules in practice.

The aim of combinatorial chemistry is to produce as many molecular variants as possible from individual building blocks. From all these combinations, the researchers fish out those that demonstrate the desired activity. The number of different molecules grows exponentially with the number of synthesis cycles and with the number of different building blocks that are combined in each synthesis cycle.

## Using DNA code to identify the active molecules

For researchers to be able to identify the individual active compounds in the rapidly growing "molecular soup" in efficacy tests, the DEL method attaches a defined short fragment of DNA to the molecule in parallel with each active-ingredient building block. This creates a unique DNA sequence as a readable barcode for each combination of building blocks.

For example, the entire soup of molecules can be tested for its ability to bind to a specific protein, and individual DNA segments can be amplified and clearly identified using the PCR (polymerase chain reaction) technique familiar from COVID tests.

## Preventing exponential growth of contamination

Chemical reality, however, has thus far severely limited the possibilities of DEL technology. The process of linking the DNA fragments with the chemical building blocks is invariably reliable, but the effectiveness with which those building blocks link together chemically varies depending on the combination. As a result, the DNA code loses its uniqueness.

The same code can refer not just to the complete molecule with all building blocks, but also to truncated variants containing only some of the building blocks. These impurities also increase exponentially with each round of synthesis. In practice, this has limited the manageable size

of DEL libraries to combinations of three to four connected blocks and thus to several million different compounds.

## Self-purification built in

Scheuermann's research team have now found a way to prevent the increasing contamination of the molecular library: to purify the DEL that has been synthesized down to the very last building block. The ETH researchers' method is based on two main parts.

First, synthesis of the molecules is coupled to magnetic particles that can be handled easily and automatically. This enables washing cycles, among other things. Second, the team introduced a second chemical coupling component on the particles that can bind only to the last of the planned building blocks.

All truncated molecules that are missing, say, the last building block, can be removed in a single washing step. In the end, the library has only those molecules that contain all the building blocks specified in the DNA code.

## Conflict with combinatorial chemistry

As elegant as the method looks on paper, it was difficult to implement, as Scheuermann says, "It was particularly challenging to find magnetic particles that don't interfere with the enzymatic coupling of DNA fragments. In the course of their doctoral projects, Michelle Keller and Dimitar Petrov from my group invested a lot of time and energy to make sure the method works reliably."

The idea of performing such combinatorial chemistry on particles emerged back in the 1990s, but only now have the ETH researchers been

able to put this into practice for library synthesis.

## More diverse and larger molecules

The self-purifying DEL technology goes beyond allowing the handling of much larger libraries of several billion molecules; it also lets researchers synthesize bigger molecules consisting of five or more building blocks.

"Before, we could search for small active substances that fit like a key into the lock of the active site of therapeutically relevant proteins, but now we can search for larger ones as well. These larger active substances can dock not only in a protein's active centers, but also to other specific areas of a protein's surface, for example, in order to prevent it from binding to a receptor," Scheuermann says.

Fundamental biological research also benefits from the possibility of finding molecules that bind to certain protein surfaces, as this makes it possible to label and examine proteins in their cellular context. Moreover, the ETH method could be a boon for major international research initiatives such as Target 2035.

This initiative addresses the ca. 20,000 human proteins and aims to find, by 2035, a molecule for each of them that binds specifically to that one protein and can therefore influence its function.

## Spin-off service for industry and science

To make the technology available to the pharmaceutical industry and for basic research as efficiently as possible, Scheuermann and his team will establish a spin-off company. This company will offer the entire process: from the development of DEL collections and automated synthesis to

automated efficacy testing and DNA-based identification of the molecules.

"We're seeing immense interest from industry and research, especially in cyclic molecules, which to date haven't been accessible in large numbers," Scheuermann says.

**More information:** Michelle Keller et al, Highly pure DNA-encoded chemical libraries by dual-linker solid-phase synthesis, *Science* (2024). DOI: 10.1126/science.adn3412

Provided by ETH Zurich