

Ready to start planning your care? Call us at [800-525-2225](tel:800-525-2225) to make an appointment.

X



Memorial Sloan Kettering
Cancer Center

[Make an Appointment](#)
[Back](#)

[Find a Physician](#)
[Cancer Treatment](#)
[In the News](#)
[Learn About Cancer & Treatment](#)

ABOUT US

[Our mission, vision & core values](#)

[Leadership](#)

[History](#)

[Inclusion & belonging](#)

[Annual report](#)

[Give to MSK](#)

...and lead to new strategies for drug design, say investigators ...
report published in the journal *PLoS One* in December 2011 .

Our cells contain thousands of uniquely shaped proteins. To understand how a protein functions or how it might contribute to disease, scientists often need to investigate its structure in detail.

Until now, information about protein structure has mainly been obtained through complex and time-consuming lab experiments — for example, using a method known as x-ray crystallography. For most proteins, such experiments have not yet been made and the shapes remain unknown.

It has long been known that, in theory, computer algorithms could be used to predict a protein's shape. "In the 1960s, pioneering work on protein folding revealed that for most proteins information for the shape is contained in the DNA sequence," says [Sloan Kettering Institute Computational Biology Program](#) Chair Chris Sander, who led of the study.

But until now, creating computational tools that can effectively translate genetic sequences into protein shapes has proved to be challenging. In the field of molecular biology, the challenge has been known as

the computational protein folding problem.

"The main difficulty lies in the enormous complexity of the search," Dr. Sander explains. In order to predict what shape a particular protein is likely to assume — the way the protein naturally curls and folds upon itself — a computer needs to try out an astronomically large number of possible shapes. Even for a supercomputer, such an analysis is "much worse than looking for the proverbial needle in a haystack," he adds.

Collaborating with investigators at Harvard Medical School and at the Human Genetics Foundation in Torino, Italy, Dr. Sander and his colleagues found a new way of solving the problem. Their report describes the development of a computer algorithm that can predict protein shapes with unprecedented speed and accuracy using genetic information alone.

Based on concepts from computational biology, statistical physics, and structural biology, the algorithm uses evolutionary information to shortcut the massive analysis. Evolution can provide essential clues about a protein's structure, which is inferred by analyzing the genetic sequences of thousands of similarly looking proteins — grouped in so-called families — from hundreds of different organisms, as determined using advanced genomic sequencing technologies.

The new method was tested on a number of proteins for which the shape could be verified with findings from x-ray crystallography. The investigators report that they were able to computationally predict the structure of these proteins within a few hours. Without this method, even supercomputers running for years would not be able to complete the calculation.

"We envision that the method will be a powerful complement to experimental structure prediction tools," Dr. Sander says. "Over time, it is likely to speed up basic research on cancer and other diseases, offering researchers everywhere a way to rapidly gain insight about how individual proteins are shaped during evolution and re-shaped in disease, how they function, and how their activity could be manipulated therapeutically."

© 2026 Memorial Sloan Kettering Cancer Center