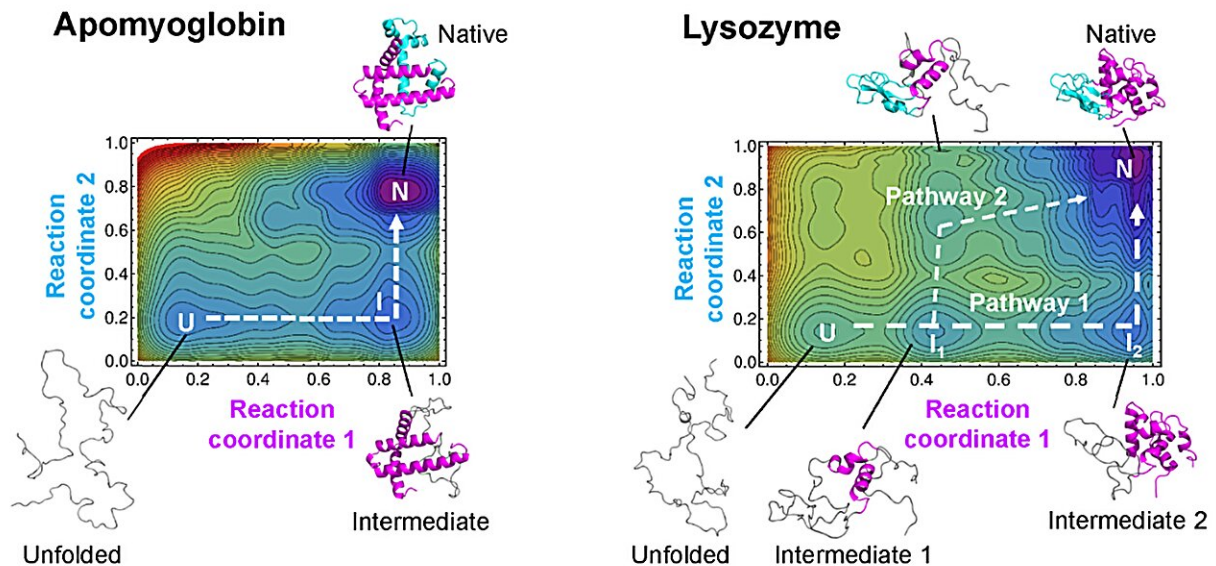


# Physical theory improves protein folding prediction

October 19 2023



Example maps with protein folding pathways. Credit: ©2023 Ooka & Arai CC-BY

Proteins are important molecules that perform a variety of functions essential to life. To function properly, many proteins must fold into specific structures. However, the way proteins fold into specific structures is still largely unknown. Researchers from the University of Tokyo have developed a novel physical theory that can accurately predict how proteins fold. Their model can predict things previous models cannot. Improved knowledge of protein folding could offer huge

benefits to medical research, as well as to various industrial processes.

The research, titled "Accurate prediction of [protein](#) folding mechanisms by simple structure-based statistical mechanical models," is published in *Nature Communications* .

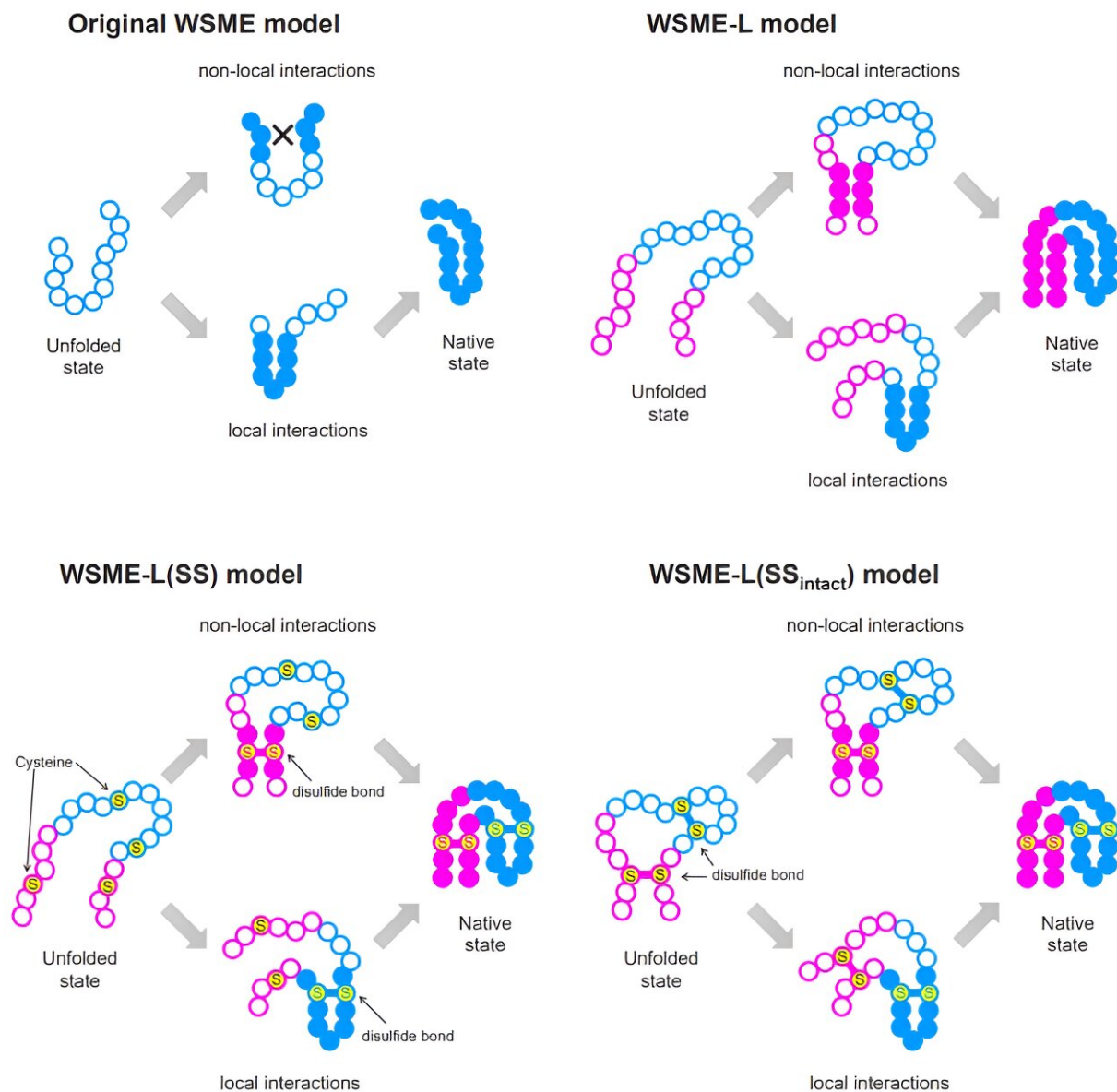
You are literally made of proteins. These chainlike molecules, made from tens to thousands of smaller molecules called amino acids, form things like hair, bones, muscles, enzymes for digestion, antibodies to fight diseases, and more. Proteins make these things by folding into various structures that in turn build up these larger tissues and biological components.

By knowing more about this folding process, researchers can better understand more about the processes that constitute life itself. Such knowledge is also essential to medicine, not only for the development of new treatments and industrial processes to produce medicines, but also for knowledge of how certain diseases work, as some are examples of protein folding gone wrong. So, to say proteins are important is putting it mildly; proteins are the stuff of life.

Encouraged by the importance of protein folding, Project Assistant Professor Koji Ooka from the College of Arts and Sciences and Professor Munehito Arai from the Department of Life Sciences and Department of Physics embarked on the task of improving upon the prediction methods of protein folding. This task is formidable for many reasons. In particular, the computational requirements to simulate the dynamics of molecules necessitate a powerful supercomputer.

Recently, the artificial intelligence-based program AlphaFold 2 has accurately predicted structures resulting from a given amino acid sequence; but it cannot give details of the way proteins fold, making it a black box. This is problematic, as the forms and behaviors of proteins

vary such that two similar ones may fold in radically different ways. So, instead of AI, the research duo needed a different approach: statistical mechanics, a branch of physical theory.



Four iterations of WSME, from the original to the new, and two specialized versions for more specific circumstances. Credit: ©2023 Ooka & Arai CC-BY

"For over 20 years, a theory called the Wako-Saitô-Muñoz-Eaton (WSME) model has successfully predicted the folding processes for proteins comprising around 100 [amino acids](#) or fewer, based on the native protein structures," said Arai.

"WSME can only evaluate small sections of proteins at a time, missing potential connections between sections farther apart. To overcome this issue, we produced a new model, WSME-L, where the L stands for 'linker.' Our linkers correspond to these nonlocal interactions and allow WSME-L to elucidate the folding process without the limitations of protein size and shape, which AlphaFold 2 cannot," Arai added.

But it doesn't end there. There are other limitations of existing protein folding models on which Ooka and Arai set their sights. Proteins can exist inside or outside of living cells; those within are in some ways protected by the cell, but those outside cells, such as antibodies, require additional bonds during folding, called [disulfide bonds](#), which help to stabilize them. Conventional models cannot factor in these bonds, but an extension to WSME-L called WSME-L(SS), where each S stands for sulfide, can do so.

To further complicate things, some proteins have disulfide bonds before folding starts, so the researchers made a further enhancement called WSME-L(SS<sub>intact</sub>), which factors in that situation at the expense of extra computation time.

"Our theory allows us to draw a kind of map of protein folding pathways in a relatively short time; mere seconds on a desktop computer for short proteins, and about an hour on a supercomputer for large proteins, assuming the native protein structure is available by experiments or AlphaFold 2 prediction," said Arai.

"The resulting landscape allows a comprehensive understanding of

multiple potential folding pathways a long protein might take. And crucially, we can scrutinize structures of transient states. This might be helpful for those researching diseases like Alzheimer's and Parkinson's—both are caused by proteins which fail to fold correctly. Also, our method may be useful for designing novel proteins and enzymes which can efficiently fold into stable functional structures, for medical and industrial use," Arai continued.

While the models produced here accurately reflect experimental observations, Ooka and Arai hope they can be used to elucidate the folding processes of many proteins that have not yet been studied experimentally. Humans have about 20,000 different proteins, but folding processes have been thoroughly studied in only around 100 of these.

**More information:** Accurate prediction of protein folding mechanisms by simple structure-based statistical mechanical models, *Nature Communications* (2023). [DOI: 10.1038/s41467-023-41664-1](https://doi.org/10.1038/s41467-023-41664-1)

Provided by University of Tokyo

Citation: Physical theory improves protein folding prediction (2023, October 19) retrieved 21 May 2024 from <https://phys.org/news/2023-10-physical-theory-protein.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--