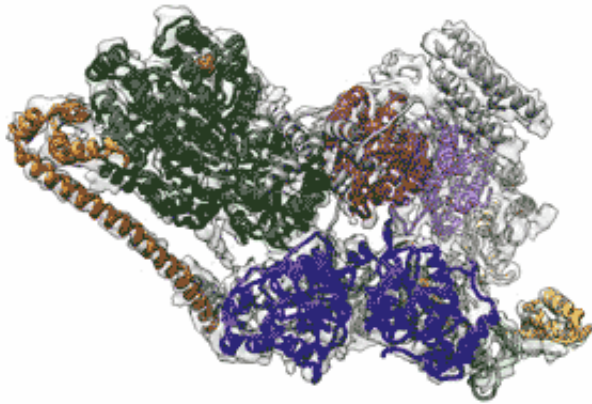


Scientists map key DNA protein complex at near-atomic resolution

September 13 2017



The cryo-EM structure of Human Transcription Factor II (TFIIH). The atomic coordinate model, colored according to the different TFIIH subunits, is shown inside the semi-transparent cryo-EM map. Credit: Basil Greber/Berkeley Lab and UC Berkeley

Chalking up another success for a new imaging technology that has energized the field of structural biology, researchers at the Department of Energy's Lawrence Berkeley National Laboratory (Berkeley Lab) obtained the highest resolution map yet of a large assembly of human proteins that is critical to DNA function.

The scientists, who are reporting their achievement Sept. 13 in an advanced online publication of the journal *Nature*, used cryo-electron microscopy (cryo-EM) to resolve the 3-D structure of a [protein complex](#) called transcription factor IIH (TFIIH) at 4.4 angstroms, or near-atomic resolution. This [protein](#) complex is used to unzip the DNA double helix so that genes can be accessed and read during transcription or repair.

"When TFIIH goes wrong, DNA repair can't occur, and that malfunction is associated with severe cancer propensity, premature aging, and a variety of other defects," said study principal investigator Eva Nogales, faculty scientist at Berkeley Lab's Molecular Biophysics and Integrated Bioimaging Division. "Using this structure, we can now begin to place mutations in context to better understand why they give rise to misbehavior in cells."

TFIIH's critical role in DNA function has made it a prime target for research, but it is considered a difficult protein complex to study, especially in humans.

Advanced technology needed to map complex protein

"As organisms get more complex, these proteins do too, taking on extra bits and pieces needed for regulatory functions at many different levels," said Nogales, who is also a UC Berkeley professor of molecular and cell biology and a Howard Hughes Medical Institute investigator. "The fact that we resolved this protein structure from human cells makes this even more relevant to disease research. There's no need to extrapolate the protein's function based upon how it works in other organisms."

Biomolecules like proteins are typically imaged using X-ray crystallography, but that method requires a large amount of stable sample for the crystallization process to work. The challenge with TFIIH is that it is hard to produce and purify in large enough quantities, and

once obtained, it may not form crystals suitable for X-ray diffraction.

Enter cryo-EM, which can work even when sample amounts are very small. Electrons are sent through purified samples that have been flash-frozen at ultracold temperatures to prevent crystalline ice from forming.

Cryo-EM has been around for decades, but major advances over the past five years have led to a quantum leap in the quality of high-resolution images achievable with this technique.

"When your goal is to get resolutions down to a few angstroms, the problem is that any motion gets magnified," said study lead author Basil Greber, a UC Berkeley postdoctoral fellow at the California Institute for Quantitative Biosciences (QB3). "At high magnifications, the slight movement of the specimen as electrons move through leads to a blurred image."

From single pictures to movies

The researchers credit the explosive growth in cryo-EM to advanced detector technology that Berkeley Lab engineer Peter Denes helped develop. Instead of a single picture taken for each sample, the direct detector camera shoots multiple frames in a process akin to recording a movie. The frames are then put together to create a high-resolution image. This approach resolves the blur from sample movement. The improved images contain higher quality data, and they allow researchers to study the sample in multiple states, as they exist in the cell.

Since shooting a movie generates far more data than a single frame, and thousands of movies are being collected during a microscopy session, the researchers needed the processing punch of supercomputers at the National Energy Research Scientific Computing Center (NERSC) at Berkeley Lab. The output from these computations was a 3-D map that

required further interpretation.

"When we began the data processing, we had 1.5 million images of individual molecules to sort through," said Greber. "We needed to select particles that are representative of an intact complex. After 300,000 CPU hours at NERSC, we ended up with 120,000 images of individual particles that were used to compute the 3-D map of the protein."

To obtain an atomic model of the protein complex based on this 3-D map, the researchers used PHENIX (Python-based Hierarchical ENvironment for Integrated Xtallography), a software program whose development is led by Paul Adams, director of Berkeley Lab's Molecular Biophysics and Integrated Bioimaging Division and a co-author of this study.

Not only does this structure help with basic understanding of DNA repair, the information could be used to help visualize how specific molecules are binding to target proteins in drug development.

"In studying the physics and chemistry of these biological molecules, we're often able to determine what they do, but how they do it is unclear," said Nogales. "This work is a prime example of what structural biologists do. We establish the framework for understanding how the molecules function. And with that information, researchers can develop finely targeted therapies with more predictive power."

More information: The cryo-electron microscopy structure of human transcription factor IIH, *Nature* (2017).

[nature.com/articles/doi:10.1038/nature23903](https://doi.org/10.1038/nature23903)

Provided by Lawrence Berkeley National Laboratory

Citation: Scientists map key DNA protein complex at near-atomic resolution (2017, September 13) retrieved 27 April 2024 from <https://phys.org/news/2017-09-scientists-key-dna-protein-complex.html>

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.