

Web-based open-source program determines protein structures

May 9 2017



Stefan Arold (right) and his team play a vital role in the local scientific community through collaboration with other research groups in the Kingdom. Credit: KAUST

ContaMiner is a web-based, open-source program developed by a unique interdisciplinary team in King Abdullah University of Science and Technology (KAUST), Saudi Arabia. This program is already saving

time for international researchers.

"How much can you understand and repair a car if you don't have a detailed picture of what is going on under the hood?" said KAUST Associate Professor Stefan Arold. "Proteins are life's workhorses: their function and dysfunction both create life and end it. Each [protein's amino acid sequence](#) folds into a particular 3-D structure that is required to support its function. If you want to understand, affect or engineer a protein's function, you need to know its 3-D structure," he explained.

The process to determine that structure begins by purifying and crystallizing the protein under investigation. The protein crystal is then bombarded by extremely powerful X-rays, which diffract in various directions, giving an indication of its structure. Researchers then apply "molecular replacement," which compares the target protein crystal to the 3-D structure of other known similar proteins.

But for the researchers to compare their protein with similar ones, first they need to know how its amino acids are arranged. ContaMiner can help researchers determine if they are even looking at the right protein to begin with.

"The protein we crystallized might not be the protein we thought it was but instead an unknown contaminant," explained Arold. Protein-based contaminants can, more often than previously thought, get crystallized instead of, or in addition to, the protein under study. These might come from the organism that originally produced the protein or occur during the purification or crystallization process.

"Scientists often waste months of work before they identify the error and the identity of the protein contaminant that they had unintentionally crystallized," said Arold.

Arold's team has worked tirelessly to compile a preliminary database, called ContaBase, of 62 known contaminants. "Contaminants were a known but under-appreciated problem because many cases went undetected," said Arold. Even in cases where contaminants are finally identified, this information often goes unpublished since the experiment was considered a failure. Often researchers report issues of contamination in online forums rather than peer-reviewed publications. "Because of this, nobody had a good idea of how many and which contaminants might occur and crystalize," he continued.

ContaMiner changes this. Now, researchers can submit their X-ray diffraction data to the program, which compares it with an updateable database of known contaminants. "If a contaminant is present, ContaMiner can typically detect it in only 5-15 minutes," said Arold.

Several hundred researchers have used the program since it was first described in late 2016 in the *Journal of Applied Crystallography*. Many from the crystallography community have helped update ContaBase to now include 71 contaminants. "It's an ongoing community effort," said Arold. ContaMiner has also been selected to be included in an online server, called CCP4, which is a highly selective collection of software related to structural biology and is the most widely used resource worldwide, explained Arold.

While having international impact, Arold and his team also play a vital role in the local scientific community through their unique set of skills. The KAUST team combines expertise in molecular biology, biochemistry, biophysics, bioinformatics and computation to investigate protein structures and function. It is the only group involved in structural biology in the kingdom and, under the supervision of structural biologist Stefan Arold, it is also setting up important local collaborations.

Currently they are collaborating with researchers at King Fahd Specialist

Hospital and Research Centre to identify gene mutations that cause diseases in the Saudi population. Arold uses his expertise, together with computational modeling, to understand why specific mutations cause protein malfunctions. "It is intriguing how much harm can be caused by a single mutation" he said. "It also gives us a glimpse of the unimaginable sophistication and complexity of our bodies.

Arold and his team also recently worked with KAUST plant scientist Mark Tester and a diverse international team to understand the molecular basis for the production of toxic compounds, called saponins, in some but not all quinoa strains.

"Does the region need more structural biologists?" asked Arold. "In my biased opinion, of course yes. In particular, experts in nuclear magnetic resonance spectroscopy, which is highly complementary to X-ray crystallography." He also advocates that biologists develop more awareness of the importance of structural biology for their research.

"Wisely, KAUST has invested heavily in [structural biology](#); and biological imaging in general is clearly an area of priority. We have outstanding resources, such as 700 and 950 megahertz [nuclear magnetic resonance](#) imaging spectrometers and the TITAN KRIOS electron microscope.

And although I am currently the only structural biologist, I might not be alone for much longer," Arold said. Negotiations are already underway to bring more talent to the institution.

More information: Arnaud Hungler et al. and ContaBase: a webserver and database for early identification of unwantedly crystallized protein contaminants, *Journal of Applied Crystallography* (2016). [DOI: 10.1107/S1600576716014965](https://doi.org/10.1107/S1600576716014965)

Provided by King Abdullah University of Science and Technology

Citation: Web-based open-source program determines protein structures (2017, May 9) retrieved 25 April 2024 from <https://phys.org/news/2017-05-web-based-open-source-protein.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.