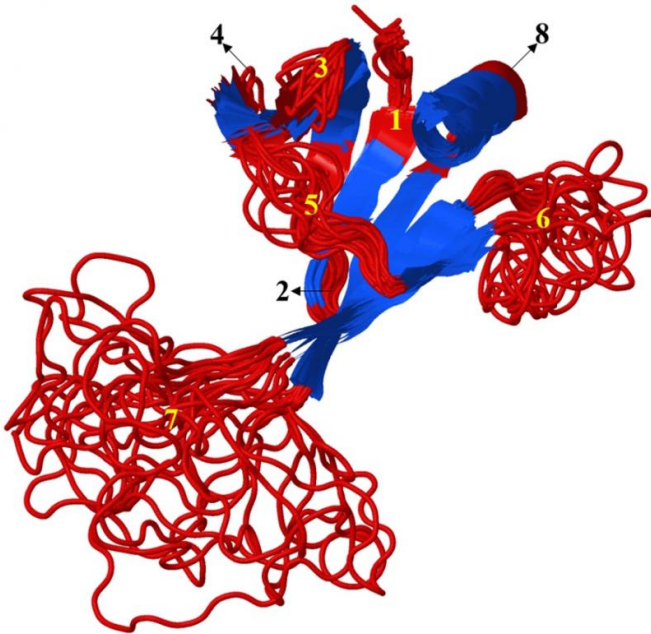


# Bioinformatics computer programs help biologists understand intrinsically disordered proteins

21 July 2017, by Leah Small



A protein with disordered regions shown in red and numbered. The disordered regions are far more flexible than the blue, structured regions. Credit: Virginia Commonwealth University

Proteins are the building blocks of life and biological agents. They are drivers of growth and development and the spread of viruses and bacteria, and have key roles in disease pathways and virtually all cellular functions. As scientists gain knowledge about proteins, the mechanisms behind biological mysteries are revealed.

To help shed light on the workings of proteins, Virginia Commonwealth University researcher Lukasz Kurgan, Ph.D., vice chair of the Computer Science Department in the School of Engineering, has developed a series of bioinformatics programs to assist biologists in developing insights into the

functions of [intrinsically disordered proteins](#). This group of proteins lacks a fixed structure, which means they are totally or partially flexible and amorphous.

Over the last several decades, scientists have sequenced 85 million unique proteins, structured and unstructured alike, but still don't know what the vast majority of these proteins do. As more proteins are discovered, more sophisticated computer programs must be developed to help determine their functions.

"We have manually curated but understand less than 1 percent of these proteins, and right now there's over 80 million to solve," said Kurgan, a Qimonda-endowed professor and data scientist. "A program can solve these proteins faster than a single human and can help researchers speed up hypothesis generation."

## Solving the puzzle

Determining a [protein's](#) function becomes even more challenging when a protein is completely or partially disordered. When a protein does have a defined structure, researchers use prior knowledge and bioinformatics programs to first decipher that structure, which then helps determine function. If the protein is disordered, biologists turn to programs built by Kurgan and other computer scientists that use [predictive models](#) to generate workable hypotheses on the protein's [function](#).

Since 2008, Kurgan has developed four programs for this purpose. This spring, his team was awarded a \$500,000 grant from the National Science Foundation to develop subsequent programs. So far, Kurgan's programs have more than 7,000 users from more than 1,300 cities in 96 countries.

Kurgan has also developed six programs that determine whether a protein is disordered or not. In 2012, his MFDp program was ranked third out of 28 participants in the biannual worldwide CASP10 experiment, which evaluates the effectiveness of computer and human predictors of intrinsic disorder. In 2014, Kurgan's lab released DisoRDPbind, the first [program](#) to predict multiple functions of intrinsically disordered proteins.

being done," Kurgan said. "It's not based on the efforts of one specific researcher or group. Collectively we help each other."

Provided by Virginia Commonwealth University

Kurgan's programs use existing collections of data on proteins whose functions have been determined to build predictive models to map the functions of unknown intrinsically disordered proteins.

"The details are not easy. Building these models takes a little bit of art, theory and experience," Kurgan said.

### **New ideas for disorder**

It is a commonly accepted fact among scientists that disordered proteins, similar to their structured counterparts, have essential functions. This assertion was at first met with disbelief, as is initially common with many scientific discoveries.

"About 30 years ago, when disordered proteins were discovered, there were a lot of deniers. Some people said that this is just noise in the protein structures," Kurgan said. "Now, disorder as a mechanism of biology is an accepted fact. Just because a protein has no defined structure, doesn't mean it's useless. It just works in a different way."

Now, deciphering disorder is a collaborative effort in the scientific community, and various programs from multiple entities come together to provide different approaches to determine the functions of disordered proteins. Kurgan has worked with researchers from the University of South Florida, Indiana University, and Tianjin and Nankai Universities in China, on a study that used his programs to discover the incidence of intrinsic disorder in close to 1,000 species from all kingdoms of life. Several other collaborative studies have focused on the functional roles of intrinsic disorder in HIV, Hepatitis C and Dengue viruses.

"Collectively we can push the boundaries of what is

APA citation: Bioinformatics computer programs help biologists understand intrinsically disordered proteins (2017, July 21) retrieved 26 May 2019 from <https://phys.org/news/2017-07-bioinformatics-biologists-intrinsically-disordered-proteins.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.*