STATISTICAL ANALYSIS OF PROTEIN SEQUENCES TO INFER 3D STRUCTURE AND FUNCTION

Results in Brief

Shining a light on the relationship between protein sequence, structure and function
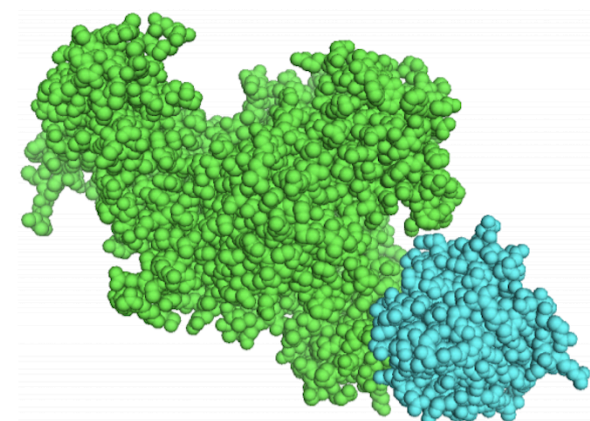
A team of mathematicians and scientists teamed up to make sense of the proteins inside all living organisms. Their groundbreaking analysis on EU project EVO-COUPLINGS could make developing medicines easier.

DIGITAL ECONOMY

FUNDAMENTAL RESEARCH

© Lucy Colwell

After water, proteins are the most abundant molecules in our bodies, making up muscles and other body tissues such as our hair. They are an essential component of all living organisms and scientists will be able to better understand them and develop medicines, thanks to new computational analysis lead by Marie Curie fellow Dr Lucy Colwell.

Her team of mathematicians, chemists and biochemists at Cambridge University have discovered relationships between the sequences of proteins, their 3D structures and functions. "These findings will focus research in the future. The relationship between protein sequence and protein structure and function is one of the great problems of our time," Dr Colwell, who

worked on the Evo-Couplings project, says.

New drug breakthroughs could be made more easily as a result of EVO-COUPLINGS as the team's work improves scientists' ability to predict the tertiary structure and the [interaction partners ⧉](#) of a protein.

The research, undertaken with the support of the Marie Curie programme, was a novel approach that was initially received skeptically by peers. But that has since changed. "I'm excited to see how mainstream these ideas have become," she said. "Before our work it was much more difficult to predict protein structure and interactions from sequences alone. This approach is becoming standard in the field and has formed a key component of a number of recent important advances."

Cracking the matrix

The researchers included mathematicians who developed methods that use random matrix theory – a probabilistic approach developed by physicists – to help the chemists and biochemists analyse protein sequence data.

Experimental techniques in recent years have enabled natural scientists to gather large amounts of data for research but sifting through that information to find what is useful can be a headache. "My job is to cut through the noise," said Dr Colwell. "To use protein sequences to predict structure, we had to first identify and remove the 'noise' in the data caused by the fact that different proteins are related to each other. This signal has to be 'normalised' out of the data before mathematical models can be built that make useful predictions."

Analysts use data visualisation methods to help identify structure in a range of fields: these methods can be applied to any type of data. Dr Colwell is now working with Google to use advances in machine learning to spot patterns that are difficult for humans to identify.

Keywords

EVO-COUPLINGS   Proteins   protein structure   sequence   random matrix
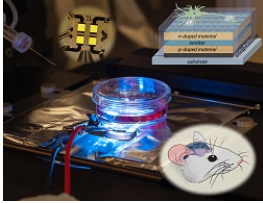
machine learning

Discover other articles in the same domain of application
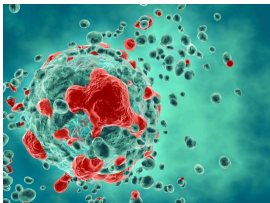
Shedding more light on how the body controls our immune systems
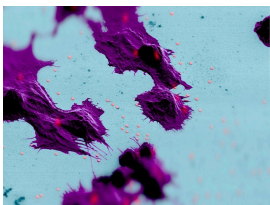
4 January 2022

OLEDs for better control of neurons

27 November 2017

Antibody-recruiting sugar-based molecules: The 'sweet' approach to targeted cancer therapy

26 June 2020

Miniscule lasers within cells advance imaging in biomedical research

24 February 2020

Project Information

EVO-COUPLINGS

Funded under

Grant agreement ID: 631609

Project closed

Start date
1 January 2015

End date
31 December 2018

Specific programme "People" implementing the Seventh Framework Programme of the European Community for research, technological development and demonstration activities (2007 to 2013)

Total cost
€ 100 000,00

EU contribution
€ 100 000,00

Coordinated by
THE CHANCELLOR MASTERS AND SCHOLARS OF THE UNIVERSITY OF CAMBRIDGE
🇬🇧 United Kingdom

Last update: 8 July 2019

Permalink: https://cordis.europa.eu/article/id/379505-shining-a-light-on-the-relationship-between-protein-sequence-structure-and-function