Paired mutations: a new approach to discovering the shape of proteins

A new study published in PNAS describes an innovative approach to identify the modular structure of proteins from their sole sequence information

29 November 2017

It is a bit like business partners: if one of the two parties changes strategy – voluntarily or otherwise – to keep the business going, the other has to adapt in turn. The leap from business ventures to the structure of proteins might seem a little bold. Yet, this concept of “balanced changes” is precisely the guiding principle of an important new study that just appeared in PNAS, the journal of the National Academy of Sciences of the United States.
The study represents a significant advancement in the fascinating and complex problem of how the sequence, structure and function of proteins are tied together.

The team of scientists, comprising researchers from SISSA and Philadelphia’s Temple University, began their research from an established fact: within the three-dimensional structure of protein, certain amino acids interact so closely with each other that the mutation of one of them must be counteracted by compensatory mutations of the others to keep the protein functional.

As the scientists explain: “By analyzing the repertoire of mutations across thousands of members of a protein family, we identified the so-called ‘sequence covariations’, that is the specific positions that exhibit a high frequency of paired mutations. We know from previous studies that when a recurring ‘co-mutation’ is observed, the two mutated sites are usually close to each other or interact in some way. Our new study has shown that we can go even further: from these co-mutations it is possible to uncover the protein’s macrostructure, its fundamental structural and functional units.”

The approach, which was validated in well-known contexts, can now be used for more reliable protein structure predictions as well as to shed light on the functional implications of structural domains. These, in fact, usually underpin the capability of proteins to change conformation, e.g. in response to the binding of other molecules.

As the authors of the new research explain: “There are three main levels of analysis in the study of proteins: the first is the sequence of amino acids, the second is the three-dimensional structure that these filaments take on very short time after they are synthesized, while the third regards their function. In recent years, research has been very much focused on the connection of the latter two aspects. With our new study, we have taken a step back towards the first level which, as we have shown, can provide us with much more information than expected about the other two aspects.” They continue: “What emerges is that local changes in proteins can have structural repercussions at a much larger scale, even the entire molecule. The biological functionality of proteins in fact, is increasingly recognised to depend on their collective structural and dynamical properties. We can harness this principle to improve our predictions of protein structure and function.” Thus by studying thousands of protein variants belonging to the same family and identifying the pairs of compensatory mutations, the scientists were able to infer which groups of amino acids were likely in the same structural domain, even if they were far away along the
sequence. This analysis hence provided extremely valuable information about the so-called macro-domains, i.e. the modular structural blocks constituting proteins. Yet how did the scientists prove that the method works? “Very simple,” they explain. “By using this approach, we first tried identify the structural modules of well-studied proteins by using as input solely their sequence information and using statistical inference methods. We did not use at all their structure as input, so that we could use it a posteriori to compare it with our predictions. And the agreement was excellent. Now we have a valuable tool that, thanks to computer-based statistical analysis, can be used to broaden our predictive capabilities in a very wide range of applicative avenues, from biophysics to biomedicine.”

---

**Full paper:**
Patterns of coevolving amino acids unveil structural and dynamical domains
by D. Granata, L. Ponzoni, C. Micheletti and E. Carnevale
Proceedings of the National Academy of Sciences, USA, 2017

Link to the article: [http://www.pnas.org/content/early/2017/11/27/1712021114.full](http://www.pnas.org/content/early/2017/11/27/1712021114.full)

**CREDITS IMAGE:** Luca Ponzoni

**CONTACTS:**
Nico Pitrelli
pitrelli@sissa.it
Tel. +39 0403787462 Cell. +39 3391337950

Donato Ramani
ramani@sissa.it
Tel. +39 0403787230 Cell. +39 3428022237

[https://www.facebook.com/sissa.school/](https://www.facebook.com/sissa.school/)
[@Sissaschool](https://www.facebook.com/sissa.school/)

Maggiori informazioni sulla SISSA: [www.sissa.it](http://www.sissa.it)