Review: Protein Binding Dependence on Sequence

A. Dievsky

June 10, 2013

Protein-protein binding permeates every biological process, notably, but not limited to, enzymatic transformation and antibody immune response. While many interactions are already known, the effect of point mutations on the binding is often a mystery. This effect is responsible for many hereditary illnesses (*e. g.,* a mutation in the enzyme changes its affinity to the substrate), and it also can be used to improve the synthetic antibody effectiveness. Therefore it's imperative to have a technique which tells how the affinity of a certain interaction will change under a given mutation.

The binding effectiveness is often expressed in the terms of the free energy of the binding, *i.e.* the difference between the bound state energy and the unbound state one. This value is always non-positive, and the larger is its absolute value, the stronger is the binding (the more energy will be needed to disassemble the complex). It is well known that in most cases only a handful of the residues, called 'hotspots', are responsible for the majority of the binding energy.

One of the standard methods of inspecting the mutations' effects is called 'alanine scanning'. It consists of physically generating a library of mutated proteins by consequently changing each amino acid residue to alanine (alanine being the most chemically neutral amino acid) and then evaluating the free energy of the binding *via* an *in vitro* experiment. However, this method is somewhat consuming both in time and in other resources, and considers only alanine substitution, so a computational method would be of much use. Most existing computational methods also consider only alanine substitutions. They mostly consist of obscenely computationally expensive molecular dynamics simulators, and oversimplified approaches based on empirical energy functions.

The article "BeAtMuSiC: prediction of changes in protein-protein binding affinity on mutations" by Y. Dehouk *et al.* offers a new method suffering from neither of mentioned shortcomings. It employs a set of statistical potentials adapted to a coarse-grained representation of protein structures and offers fast assessment of all possible mutations in the protein complex. The method is freely available as a web service.

The method considers two models: in the first one, the proteins have independent folds, while in the other they are folded only in the complex. The actual prediction is the combination of the two models' outputs. Each model represents

the change of the binding energy through a generalized linear model, with sigmoid functions applied to the solvent accessibility of the mutated residue. The model coefficients (*i.e.*, statistical potentials) depend on the mutation being studied, pairwise inter-residue distances *etc.*

The method was validated on several public datasets. The results were generally good, with correlation between actual and predicted values noticeable by naked eye. However, the authors note that the coarse-grained approach sometimes fails on the complexes where the affinity is dependent on a singular residue, such as the bovine pancreatic trypsin inhibitor (BPTI) + bovine β-trypsin complex. In these cases the predictions can deviate grossly from the actual values.

The method also participated in CAPRI, a contest where various algorithms were offered a challenge to predict mutations' effects on two newly synthesized influenza inhibitors. The predictions were checked against the previously unreleased experimental data. The authors report that BeAtMuSiC was 'among the top performers'.

The authors hypothesize that their web service would be useful for protein engineering, where it can be used to drastically reduce the pool of candidates for further practical examination, while it can also be applied to the studies of the pathological consequences of non-synonymous SNPs.

All that said, the presented approach actually strikes one as a little *too* coarse-grained; however, the validation results suggest that the simplification is justified.

References

[1] Dehouk Y. *et al.*, *BeAtMuSiC: prediction of changes in protein-protein binding affinity on mutations,* Nucleic Acid Res., 2013, epub ahead of print

[2] Jones S. and Thornton, J. M., *Principles of protein-protein interaction*, Proc. Natl. Acad. Sci. USA, 93, 13–20

[3] Moreira, I. S., Fernandes, P. A. and Ramos, M. J., *Hot spots – a review of the protein-protein interface determinant amino-acid residues,* Proteins, 68, 803–812.

[4] Thorn, K. S. and Bogan, A. A. *ASEdb: a database of alanine mutations and their effects on the free energy of binding in protein interactions,* Bioinformatics, 17, 284–285.