

Predicting protein-protein interactions

Nicholas Burgoyne and Richard Jackson

Now that both the individual protein structures and certain aspects of their molecular biology can be determined at a genomic scale there is a growing knowledge gap emerging in biochemistry. Although we often know the structures of two interacting proteins, the structure of them in complex is still difficult to determine. Predictions of the interface can be very instructive for a molecular biologist, by guiding experimental analysis of the complex. Predictions of the structure of a complex can also be useful in the structure determination processes of protein complexes by NMR, X-ray crystallography and electron microscopy.

Predictions of protein-protein interfaces are very useful, but the process is difficult. This is primarily due to the small set of structural examples that are currently known. Just as it is assumed that there are a given number of protein folds, there is likely to be a defined number of interactions between them. The known protein interfaces show great diversity in terms of the size of the buried interface and the chemical composition of their binding surfaces. Despite this fact, there are sufficient similarities to predict protein interfaces by taking the mean interface-like properties from known complexes.

It may also be possible to concentrate the prediction on a certain subset of the entire protein-protein interface. It is known that most protein-interfaces contain clefts, into which are placed the sidechains of its interacting protein partner. Alanine-scanning mutagenesis and evolutionary conservation of residues around the clefts suggests that some residues are more important to the interaction than others. It is interesting to note that most interfaces have these important clefts, and that their properties are largely consistent across the known interfaces.

It is our intention to attempt to predict these interface-clefts from a protein surface by relying on the known properties that give the clefts the ability to bind small regions of their interaction partners. This would indicate which regions of the surface are most-likely to interact with the sidechains of the binding protein. Experimental evidence (mentioned previously) indicates that there can often be more than one of these functionally important sites in a given complex, it is also known that these sites often do not alter conformation on formation of the protein complex. Therefore the prediction of the relevant clefts could act as a basis for the generation of possible protein-protein interactions.

This work is ongoing.

Funding

This work is funded by the MRC.