A fast method to predict protein-protein interaction sites from sequence analysis.

Gallet, X, Thomas-Soumarmon, A.*, Brasseur, R.
Centre de Biophysique Moléculaire Numérique, Faculté Universitaire des Sciences Agronomiques, Passage des déportés 2, 5030 Gembloux, Belgium. * INSERM U10, Hôpital X. Bichat, 170 Bd Ney, 75018 Paris, France.

The prediction of protein interaction sites from sequences and 3-D protein structures is a very important tool to investigate the analysis of molecular recognition and in a more extensive view of biological functions. Several studies attempted to determine location of interaction domains on 3-D protein structures by analysing various parameters: hydrophobicity, accessible surface area and protrusion. Here, we present a fast procedure to predict protein-protein interaction sites from sequence. Using the Eisenberg's method[1], previous results[2,3] showed that some binding sites may be predicted by plotting the mean hydrophobicity <H> versus the mean hydrophobic moment <μ>. We analyzed the sequences of 34 known protein-protein interaction sites extracted from non-homologous proteins enclosing various biological functions[4]. Hydrophobicity profiles of sequences were analyzed using windows of 5 or 7 residues. Results show that an extensive investigation of the plot including a statistical analysis allow to detect most interaction sites, named 'Receptor-binding-domain' (RBD). Parameters and observations for RBDs prediction such as window size, average <H> and <μ> values, clusters of points and properties of amino acids, especially prolines[5] are discussed. Furthermore, RBD stretches were mapped in their corresponding 3-D structures. Molecular hydrophobic potentials (MHP) showed that interaction sites are made of both charged and hydrophobic patches. The arrangement of the latters at the protein surface is also described.

This method can be used to identify the amino acids involved in protein-protein interactions and can help to suggest ponctual mutations. Different experimental examples will be presented and discussed.

References

1. Eisenberg, D.,Weiss, R.M. and Terwilliger, T.C. (1982) Nature 299,371-374.
2. De Loof, H., Rosseneu, M., Brasseur, R., Ruysschaert, J.M. (1986) Proc. Natl. Acad. Sci. USA 83, 2295-2299.
3. Le, L., Brasseur, R., Wemers, C., Meulemans, G., Burny, A. (1988) Virus Genes 1,333-350.
4. Kini, R.M. and Evans, H.J. (1995) Biochem. Biophys. Res. Commun. 212,1115-1124.
5. Kini, R.M. and Evans, H.J. (1996) FEBS Lett 29,81-86.