Disponible en ligne sur ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com

PATHOLOGIE BIOLOGIE

Pathologie Biologie 57 (2009) 324–333

Review

Interaction networks: From protein functions to drug discovery. A review

Les réseaux d'interactions : de la fonction des protéines à la conception de médicaments. Une revue

E. Chautard [a], N. Thierry-Mieg [b], S. Ricard-Blum [a,*]

[a] *UMR 5086 CNRS, institut de biologie et chimie des protéines, université Lyon 1, IFR, 128 biosciences Lyon-Gerland, 7, passage du Vercors, 69367 Lyon cedex 07, France*

[b] *TIMC-IMAG Laboratory, TIMB Team, faculté de médecine, 38706 La Tronche cedex, France*

Received 22 September 2008; accepted 17 October 2008
Available online 13 December 2008

Abstract

Most genes, proteins and other components carry out their functions within a complex network of interactions and a single molecule can affect a wide range of other cell components. A global, integrative, approach has been developed for several years, including protein–protein interaction networks (interactomes). In this review, we describe the high-throughput methods used to identify new interactions and to build large interaction datasets. The minimum information required for reporting a molecular interaction experiment (MIMIx) has been defined as a standard for storing data in publicly available interaction databases. Several examples of interaction networks from molecular machines (proteasome) or organelles (phagosome, mitochondrion) to whole organisms (viruses, bacteria, yeast, fly, and worm) are given and attempts to cover the entire human interaction network are discussed. The methods used to perform the topological analysis of interaction networks and to extract biological information from them are presented. These investigations have provided clues on protein functions, signalling and metabolic pathways, and physiological processes, unraveled the molecular basis of some diseases (cancer, infectious diseases), and will be very useful to identify new therapeutic targets and for drug discovery. A major challenge is now to integrate data from different sources (interactome, transcriptome, phenome, localization) to switch from static to dynamic interaction networks. The merging of a viral interactome and the human interactome has been used to simulate viral infection, paving the way for future studies aiming at providing molecular basis of human diseases.

Résumé

La plupart des gènes, des protéines et des autres constituants cellulaires exercent leurs fonctions au sein d'un réseau complexe d'interactions et une seule molécule peut affecter un ensemble d'autres molécules. Une approche globale, intégrative, a été développée depuis plusieurs années pour construire des réseaux d'interactions protéine–protéine (interactomes). Dans cette revue sont décrites les méthodes dites « haut débit » qui permettent d'identifier plusieurs milliers d'interactions en parallèle. L'information minimale pour décrire une expérience d'interaction moléculaire (MIMIx) a été définie pour le stockage des données dans des bases de données d'interactions publiquement disponibles. Plusieurs exemples de réseaux d'interactions sont donnés, des machines moléculaires (protéasome) ou des organelles (phagosome, mitochondrie) jusqu'aux organismes entiers (virus, bactéries, levure, mouche et vers). Les tentatives de construction de l'interactome humain entier ainsi que les méthodes d'analyse globale des réseaux d'interactions sont brièvement présentées. Ces études ont permis de prédire les fonctions de protéines, d'étudier des voies de signalisation ou métaboliques, de déterminer les mécanismes moléculaires de certaines pathologies (cancer, infections virales) et seront très utiles pour identifier de nouvelles cibles thérapeutiques et pour la mise au point de nouveaux médicaments. Un des défis majeurs est maintenant d'intégrer des données de différentes sources (interactome, transcriptome, phénome, localisome) pour passer des réseaux statiques aux réseaux

* Corresponding author.
E-mail address: s.ricard-blum@ibcp.fr (S. Ricard-Blum).

0369-8114/$ – see front matter
doi:10.1016/j.patbio.2008.10.004

dynamiques. La fusion d'un interactome viral à l'interactome humain a ainsi permis de simuler une infection virale. Cette approche ouvre la voie à des études visant à élucider les bases moléculaires des pathologies humaines.

Keywords: Bioinformatics; Databases; Disease; Drug discovery; Interaction network; Interactome; Protein–protein interactions; Protein–carbohydrate interactions

Mots clés : Bioinformatique ; Bases de données ; Découverte de médicaments ; Interactions protéines–protéines et protéines–sucres ; Rôle des protéines ; Interactome ; Pathologies ; Réseau d'interactions

1. Introduction

Biological systems are made up of very large numbers of different components interacting at various scales. Most genes, proteins and other cell components carry out their functions within a complex network of interactions and a single component can affect a wide range of other components. Interactions involved in biological processes have been first characterized individually, but this ''reductionist'' approach suffers from a lack of information about time, space, and context in which the interactions occur in vivo. A global, integrative, approach has been developed for several years, focusing on the building of protein–protein interaction maps (interactomes). These interaction networks are complex systems, where new properties arise. This is part of an emergent field, called systems biology which is ''the study of an organism, viewed as an integrated and interacting network of genes, proteins and biochemical reactions which give rise to life'' (http://www.systemsbiology.org/). This interdisciplinary approach, involving techniques from the mathematical, computational, physical and engineering sciences is required to understand complex networks. The systems biology approach has been recently applied to the study of proteases that operate in linear pathways or in regulatory circuits forming a protease web [1,2]. The overexpression or the reduced expression of a protease may perturb the protease web leading either to further connections or to a loss of interactions that may initiate and propagate pathological events [1]. Degradomics, the application of genomic and proteomic approaches to identify the protease and protease-substrate repertoires (degradomes) on an organism-wide scale, has been developed by using specific DNA microarrays to analyze the expression of proteases and inhibitors on a system-wide basis (70-mer oligonucleotide probes for all 1561 human and murine proteases, inactive homologues and inhibitors) and mass spectrometry-based proteomics. Elucidating the substrate degradomes of proteases will help to understand the function of proteases in development and disease and the identification of central proteases will identify new drug targets and will help predicting the potential for side effects due to the interconnected nature of the protease web [2,3]. Systems biology may also be helpful in medicine where treatments focused on components are currently disease-driven, aimed for normalcy and additive, whereas systems biology looks at interrelationships and dynamics. Systems approach will lead to individualized, time-sensitive, space-sensitive and synergistic treatments taking into account the multidimensional use of drugs [4].

An interactome is the whole set of molecular physical interactions between biological entities in cells and organisms and it is essential to understand how gene functions and regulations are integrated at the level of an organism. Indeed, many proteins mediate their biological function through protein interactions, which are involved in supramolecular assemblies (collagens, elastic fibers, actin filaments), in the building of molecular machines (molecular motors, ribosomes, proteasome) and in major biological processes such as immunity (antigen–antibody interaction), metabolism (enzyme–substrate interaction), signalling (interaction of messenger molecules, hormones, neurotransmitters with their cognate receptors), and gene expression (DNA–protein interactions). Furthermore, the sequencing of the genome and advances in proteomics lead to the identification of proteins of unknown functions. Interaction networks might give clues on the functions of these newly discovered proteins or on new functions of already identified proteins. The systematic identification of interactions for a given proteome has been proposed as a potentially informative functional strategy [5,6]. In this review we will:

- describe high-throughput methods used to identify interactions and to build interaction networks, including standard format of reporting and a brief description of publicly available interaction databases;
- give some examples of interaction networks from subcellular compartments to whole organisms (yeast, bacteria, fly, worm);
- explain how to analyze the interaction networks to get information on the biological functions of proteins, and to predict the behavior of the network by simulating constraints induced by physiopathological processes leading to a more rational approach in therapeutics and drug discovery.

2. Building interaction networks

Manual curation and text mining are used to extract interaction data from the literature. New interaction data are collected by high throughput experimental methods, including yeast two hybrid (Y2H), tandem-affinity mass spectrometry (TAP-MS), protein complementation assays and protein arrays. In addition, several methods aiming at predicting interactions (inference, Rosetta stone) have been developed to establish comprehensive maps of hypothetical protein–protein interactions.

2.1. Experimental methods

In yeast two-hybrid, the two proteins to be tested for interactions are expressed in yeast as hybrid fusion proteins