Available online at www.sciencedirect.com
ScienceDirect
Journal of Theoretical Biology 252 (2008) 456–464
Journal of Theoretical Biology
www.elsevier.com/locate/yjtbi

A computational analysis of protein interactions in metabolic networks reveals novel enzyme pairs potentially involved in metabolic channeling ☆

Carola Huthmacher*, Christoph Gille, Hermann-Georg Holzhütter

Institute of Biochemistry, University Medicine Berlin—Charité, Monbijoustr. 2, 10117 Berlin, Germany

Received 31 May 2007; received in revised form 7 September 2007; accepted 26 September 2007
Available online 5 October 2007

Abstract

Protein–protein interactions are operative at almost every level of cell structure and function as, for example, formation of sub-cellular organelles, packaging of chromatin, muscle contraction, signal transduction, and regulation of gene expression. Public databases of reported protein–protein interactions comprise hundreds of thousands interactions, and this number is steadily growing. Elucidating the implications of protein–protein interactions for the regulation of the underlying cellular or extra-cellular reaction network remains a great challenge for computational biochemistry. In this work, we have undertaken a systematic and comprehensive computational analysis of reported enzyme–enzyme interactions in the metabolic networks of the model organisms *Escherichia coli* and *Saccharomyces cerevisiae*. We grouped all enzyme pairs according to the topological distance that the catalyzed reactions have in the metabolic network and performed a statistical analysis of reported enzyme–enzyme interactions within these groups. We found a higher frequency of reported enzyme–enzyme interactions within the group of enzymes catalyzing reactions that are adjacent in the network, i.e. sharing at least one metabolite. As some of these interacting enzymes have already been implicated in metabolic channeling our analysis may provide a useful screening for candidates of this phenomenon. To check for a possible regulatory role of interactions between enzymes catalyzing non-neighboring reactions, we determined potentially regulatory enzymes using connectivity in the network and absolute change of Gibbs free energy. Indeed a higher portion of reported interactions pertain to such potentially regulatory enzymes.

Keywords: Enzyme interaction; Mapping; Yeast; *E. coli*

1. Preface

Inventing and permanently updating the concept of metabolic control theory (MCT) is doubtless one of most important contributions of Reinhart Heinrich to computational systems biology. In its original formulation (Heinrich and Rapoport, 1974) MCT quantifies the control of enzymes within the system. Later, in view of accumulating experimental evidences for the involvement of enzyme complexes in metabolic regulation, he and his co-workers generalized the concept of MCT (Schuster and Heinrich, 1992) and based on that he noticed a dramatic change in the control strengths of enzymes when undergoing interactions with other enzymes (Kholodenko et al., 1995). Inspired by this work, we look here at enzyme–enzyme interactions in metabolic networks from a metabolome-wide perspective.

2. Introduction

Many proteins mediate their function through physical interactions with other proteins. The resulting protein complexes vary in composition (homomeric or heteromeric) and affinity strengths. For example, the cytoskeleton consists of filaments that are stable assemblies of actin, tubulin, and other cytoskeleton proteins defining the

☆This article is devoted to our memorable dear friend and colleague Prof. Reinhart Heinrich who unexpectedly died in October 2006. Our analysis was largely inspired by his famous work on the control and design of metabolic networks.

*Corresponding author. Tel.: +49 30 528466; fax: +49 30 528937.
E-mail address: carola.huthmacher@charite.de (C. Huthmacher).

0022-5193/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2007.09.042

cellular shape. Stable protein interactions are found as well in "molecular machines" like proteasomes or the large and the small subunits of ribosomes, which are composed of different subunits needed for full functioning. Rather transient protein interactions occur in signal transduction pathways where extra-cellular ligand binding to a particular receptor on the cell surface triggers a protein interaction cascade within the cell changing the cellular status.

Interaction partners for which the biological function is known allow inferences on uncharacterized proteins (Wahlout et al., 2000; Brun et al., 2003). Therefore, several attempts have been made in the last years to analyze protein–protein interactions on genome-scale providing new working hypotheses for proteins of interest. Novel high-throughput methods like yeast two-hybrid systems (Y2H) (Fields and Song, 1989) and tandem affinity purification (TAP) (Puig et al., 2001) followed by mass spectroscopy analysis (MS) have generated interaction maps for several organisms including *Saccharomyces cerevisiae* (Fromont-Racine et al., 1997; Uetz et al., 2000; Ito et al., 2001; Gavin et al., 2002; Ho et al., 2002) and *Escherichia coli* (Butland et al., 2005; Arifuzzaman et al., 2006). For a large portion of these interactions the biological relevance is not known yet, which demands further investigation.

In this study, we focus on protein interactions involving metabolic enzymes. Interactions between metabolic enzymes have been studied in multiple small-scale experiments. A well-studied example is the tryptophan synthase complex (Hyde et al., 1988). This multi-enzyme complex consists of two alpha and two beta subunits, which catalyze the conversion of indoleglycerol-phosphate to indole and subsequently the synthesis of tryptophan from indole and serine. The complex was found to function substantially more efficient than the separated individual enzymes. Analysis of the structure of the enzyme complex revealed a tunnel connecting the active sites of both subunits thereby facilitating direct transfer of the non-polar intermediate indole to the next enzyme and preventing it from getting lost by passing the outer cellular membrane.

The tryptophan complex is the prime example of the metabolic channeling hypothesis, which was proposed in the 1980s (Srere, 1987). At this time experimental findings suggested high intracellular inhomogeneity and macromolecular crowding affecting cellular fluidity and diffusion rates (Luby-Phelps, 1994). The fact that cells are able to overcome these limitations on free diffusion motivated the hypothesis of metabolic channeling. It suggests the association of enzymes that catalyze consecutive reactions of a metabolic pathway thereby enabling channeling of reaction intermediates from one enzyme to the next entailing several catalytic advantages (Ovádi, 1991).

Publicly accessible databases like DIP, IntAct, MINT, and BioGRID provide a great number of protein–protein interactions obtained from high-throughput assays as well as literature searches. These data facilitate large-scale analyses of protein interactions in metabolic networks giving further information on functional relevance of interactions among enzymes. Here we present a computational analysis of enzyme interactions in the model organisms *E. coli* and *S. cerevisiae*. Our approach relates the topological distance between two enzymes, defined by the shortest connecting path in the network, to the occurrence of reported protein–protein interactions. This analysis reveals that for neighboring reactions the fraction of interacting enzymes is higher than for reactions that are not adjacent in the network. Case-by-case inspection of the interactions between enzymes sharing a metabolite reveals a substantial number of enzyme pairs, which have not been further characterized in the literature and therefore present a set of candidates for novel metabolic channeling examples.

We also analyzed some features of those enzymes, which physically interact but catalyze topologically more distant reactions. An increased share of potentially regulatory enzymes in long-distance enzyme–enzyme interactions was found.

3. Methods

3.1. Databases of reported protein–protein interactions

For our analyses we used experimentally detected protein–protein interactions from the four public databases DIP (version of January 7, 2007), IntAct (version of December 4, 2006), MINT (version of February 6, 2007), and BioGRID (version 2.0.24). Only interactions where both proteins are referenced to SwissProt were included in our analyses in order to benefit from SwissProt annotations such as genetic locus information. The resulting set of more than 170,000 binary interactions includes high-throughput data as well as manually curated data. On the basis of SwissProt almost 18,000 interactions among proteins with assigned EC number were identified originating from more than 100 different organisms. Of these interactions 7575 were observed in yeast and 3288 in *E. coli*.

The obtained enzyme interactions were mapped onto the metabolic networks of the single-compartment organism *E. coli* and the multi-compartment organism *S. cerevisiae*, respectively. Enzyme interactions were mapped onto the networks via genetic loci annotation. We only considered interactions of proteins from distinct genetic loci and excluded homomeric interactions. We used the metabolic network models by Palsson and his colleagues for *E. coli* (iJR904, Reed et al., 2003) and yeast (iND750, Duarte et al., 2004). The *E. coli* set contains 931 reactions while the yeast set comprises 1149 reactions. Genetic loci are assigned to 873 and 810 reactions, respectively. SwissProt information was annotated via locus information.