COMPUTATIONAL ANALYSIS OF PROTEIN-PROTEIN INTERACTIONS IN METABOLIC NETWORKS OF ESCHERICHIA COLI AND YEAST

CAROLA HUTHMACHER
carola.huthmacher@charite.de

CHRISTOPH GILLE
christoph.gille@charite.de

HERMANN-GEORG HOLZHÜTTER
hergo@charite.de

Medical Faculty of the Humboldt University, Charite, Institute of Biochemistry, Monbijoustr. 2, 10117 Berlin, Germany

Protein-protein interactions are operative at almost every level of cell function. In the recent years high-throughput methods have been increasingly used to uncover protein-protein interactions at genome scale resulting in interaction maps for entire organisms. However, biochemical implications of high-throughput interactions are not always obvious. The question arises whether all interactions detected by *in vitro* experiments also play a functional role in the living cell. In this work we systematically analyze high-throughput protein-protein interactions stored in public databases in the context of metabolic networks. Classifying reaction pairs according to their topological distance revealed a significantly higher frequency of enzyme-enzyme interactions for directly neighbored reactions (distance = 1). To determine possible functional implications for these interactions we examined randomized networks using original enzyme interactions as well as randomly generated interaction data. A functional relevance of enzyme-enzyme interactions could be demonstrated for those reactions that exhibit low connectivity. As this is a characteristic of enzyme pairs in metabolic channeling we systematically searched the literature and indeed recovered a certain fraction of enzyme pairs that has already been implicated in metabolic channeling. However, a substantial number of enzyme pairs uncovered by our large-scale analysis remains that up to now has neither been functionally nor structurally classified and therefore present novel candidates of the metabolic channeling concept.

Keywords: enzyme interactions; network randomization; randomly generated enzyme interactions; metabolic channeling.

1. Introduction

Protein-protein interactions play critical roles for almost all cellular processes. Permanent interactions are found in multimeric proteins such as the core and the regulators of proteasomes, the small and large subunits of ribosomes, or metabolic enzymes like fatty acid synthase. Transient interactions are important among others for signal transduction, splicing, cell motility, and cell cycle. For metabolic enzymes transient interactions have been described in the context of metabolic channeling, which is assumed to result in multiple catalytic advantages [16].

Recently, systematic searches for protein interactions have been conducted at large scale using high-throughput techniques such as yeast two-hybrid methods [9, 20] and affinity purification combined with mass spectroscopy [5, 7]. These experimental techniques yield enormous amounts of binary interaction data, which are valuable resources for protein function annotation. High-throughput data are thought to be free of bias from selection processes by the experimentator or the scientific community [15] since all data obtained from one experiment are measured under identical conditions and are recorded irrespectively of existing knowledge. This opens the possibility to relate binary interaction data to other biological properties by statistical analyses. Protein interaction data has been linked to a variety of other features like gene-fusion events [4, 19], gene expression [6, 21], and interactions of homologous proteins from different organisms [12]. The observed correlations provide evidence that a large percentage of interactions are biologically meaningful. On the other hand, there are increasing arguments for a considerable percentage of false positive and false negative rates inherent in high-throughput methods since the results of different techniques vary considerably [14]. Therefore the question arises whether high-throughput interactions are partially due to artefacts owing to limitations of the experiment or whether the biological meaning is simply not yet understood.
In this work we mapped high-throughput protein-protein interaction data onto metabolic networks of the model organisms *Escherichia coli* and *Saccharomyces cerevisiae*. We analyzed interaction probabilities for neighboring enzymes in original and randomized networks using real enzyme interactions as well as randomly generated data. Significantly higher interaction rates were observed for neighboring enzymes in original networks at least for sparsely connected reactions, suggesting functional relevance for such enzyme complexes. Additionally, we identified several putative novel examples of the metabolic channeling concept among the complexes involving neighboring enzymes.

2. Methods

2.1. *Database protein-protein interactions (PPI)*

For our analyses we used experimentally detected protein-protein interactions from the four public databases DIP (version of Jan 7th 2007), IntAct (version of Dec 4th 2006), MINT (version of Feb 6th 2007), and BioGRID (version 2.0.24). Only interacting proteins, which are both linked 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 7,575 were observed in yeast and 3,288 in *E. coli*.

2.2. *Metabolic networks*

The obtained enzyme interactions were mapped onto the metabolic networks of the single-compartment organism *Escherichia coli* and the multi-compartment organism *Saccharomyces cerevisiae*. We used the metabolic network models iJR904 [18] (*E. coli*) and iND750 [3] (yeast) by Palsson and his colleagues. An important feature of these models are gene to protein to reaction associations, which allow the incorporation of further information such as protein-protein interaction data. Flux directions were assumed as proposed by Palsson. If reactions are defined as reversible we assumed a flux from left to right, meaning that metabolites on the left side of the reaction arrow are consumed. The *E. coli* set contains 931 reactions while the yeast set comprises 1,149 reactions. Genetic loci are assigned to 873 and 810 reactions, respectively. SwissProt annotation was mapped via locus information.

2.3. *Neighboring reactions*

Graphs representing metabolic networks were constructed where nodes correspond to reactions. Reactions having at least one metabolite in common were linked by edges regardless whether the metabolite is consumed or produced. Candidates were not defined as neighbors if the common metabolite is either abundant or unspecific like water and phosphate or acts as a cofactor like NAD+ and ATP. Otherwise this might result in linking reactions where a functional relevance is questionable in view of metabolic channeling. Reactions were coupled nevertheless if cofactor partners (NADH, ADP, etc.) do not occur in at least one of the reactions. By this means cofactor biosynthesis or degradation steps remain linked.

2.4. *Randomizing metabolic networks*

Networks were randomized in two different ways. In a first trial the skeleton of the original network was used and the reaction names assigned to nodes were permuted. In other words, node attributes were arbitrarily exchanged while edges remained unchanged. Since in this approach the original connectivity of reactions is ignored, we next generated networks preserving this characteristic. Reactions were grouped according to their overall connectivity and only reactions of similar connectivity were exchanged. To obtain few connectivity classes of similar size each class was defined to contain about one tenth of the reaction set. Only node labels corresponding to reactions of the same connectivity class were permuted, which roughly retains the number of neighbors for each reaction. For both types of randomization 1,000 networks were generated. The number of enzymes that form a complex and are located in direct neighborhood are counted in the original metabolic network and in the randomized ones. Homomeric interactions were not considered.

2.5. *Randomizing enzyme interactions*

We used two different approaches to generate random enzyme interactions. In the first approach interacting enzymes are drawn applying the same probability for all pairs. This probability was set to 0.5% in accordance to the observed interaction probability of *E. coli* enzymes. For the second approach we adjusted this probability to take into account the higher interaction probability of enzymes located at metabolic branchpoints [8]. If the connectivity of reactions was greater than 1 the base two logarithm of this value was multiplied to the interaction probability of reaction pairs resulting in higher probabilities for enzymes that are highly linked. 1,000 times random enzyme complexes were computed and mapped onto the original network as well as onto randomized networks as described above.

3. Results

For our analysis experimentally determined physical protein-protein interactions were obtained from the online databases DIP, IntAct, MINT, and BioGRID and mapped onto the metabolic networks of the model organisms *E. coli* and yeast. For the *E. coli* metabolic network model iJR904 enzymes corresponding to 1,759 reaction pairs were found in the interaction databases excluding those of overlapping genetic loci. In the case of the multi-compartment organism yeast enzymes corresponding to 2,508 reaction pairs were reported as interaction partners. Enzyme pairs were not included into the analysis if they belong to different cellular compartments. The focus of this work was set on enzymes that are direct neighbors within metabolic networks since such interactions have been found to play a functional role in metabolic channeling. We detected 104 pairs of reactions that are direct neighbors in the *E. coli* metabolic network and are catalyzed by enzymes recorded in protein interaction databases as partners. For yeast 103 such reaction pairs were found.

3.1. *Randomizing metabolic networks reduces the number of enzyme interactions involving neighboring reactions*

We compared the number of detected interactions in the metabolic networks of yeast and *E. coli* to that in randomized networks. In a first analysis step, reactions were randomly exchanged. On average 45 reaction pairs corresponding to interacting enzymes were located in direct neighborhood in the case of *E. coli* and 49 in the case of yeast, accounting for less than 50% of those in the original network (compare Table 1 run 1 with run 2). For both organisms all 1,000 randomized networks exhibited less interactions among neighboring enzymes as the original network (see Fig. 1 run 2). In a subsequent analysis we exchanged only reactions that are topologically equivalent with respect to connectivity. *E. coli* reactions were grouped into 10 classes of similar connectivity varying in size from 41 to 111 reactions and yeast reactions into 11 classes comprising minimal 66 reactions and maximal 130 reactions.

Enzymes catalyzing neighboring reactions in these randomized networks interact on average 92 times in *E. coli* and 88 times in yeast reaching almost the interaction probability of neighboring enzymes in the original metabolic networks (compare Table 1 run 1 with run 3). The number of interactions in the original networks rank among the 82.9^{th} (*E. coli*) and 87.9^{th} (yeast) percentile of the distribution obtained by randomizing the original networks (see Fig. 1 run 3). Considering only low connected reactions contained in classes 1 to 8 yields higher numbers of interactions in the original networks (see Table 2 run 4) than in both types of randomized networks (see Table 2 run 5 and 6). For *E. coli* 24 enzyme interactions were detected among direct neighbors in the original network while on average in arbitrarily randomized networks 12 interactions and in randomized networks preserving reaction connectivity 11 interactions were found. Only 4 and 2 randomized networks, respectively, exhibited as many interactions as the original networks (see Fig. 1 run 5 and 6). In the case of yeast 14 interactions among neighboring enzymes were detected in the original network and 8 and 6 interactions, respectively, in randomized networks (corresponding percentiles: 94.9 and 99.8).

Table 1. Overview on randomization runs 1-3 and results for *E. coli*. Run 3 preserves reaction connectivity.

	Run 1	**Run 2**	**Run 3**
Network	original network	permuted network	permuted network (preserving connectivity)
Enzyme interactions	original data	original data	original data
PPI involving neighboring enzymes (rel. amount)	1	0.43	0.89

3.2. *Randomly generated binary interaction data also yields higher frequency of protein interactions in original network than in permuted networks.*

Relating network topology to the ability of enzymes to interact with each other reveals that enzymes are more likely to have interaction partners if they share metabolites with many other enzymes (data not shown). This tendency is much more pronounced for yeast than for *E. coli*. To assess the bias due to increased interaction probabilities of highly connected enzymes we repeated our analysis with randomly generated enzyme interactions. In a first approach enzyme pairs were randomly selected as interaction partners with equal probability. Mapping these

Table 2. Overview on randomization runs 4-6 and results for *E. coli*. Run 6 preserves reaction connectivity. Only those reactions were considered that are assigned to connectivity classes 1 to 8.

	Run 4	**Run 5**	**Run 6**
Network	original network	permuted network	permuted network (preserving connectivity)
Enzyme interactions	original data	original data	original data
PPI involving neighboring enzymes (rel. amount)	1	0.48	0.47

Table 3. Overview on randomization runs 7-9 for *E. coli* using randomly generated enzyme interactions and interaction probabilities correlated to connectivity of reactions. Run 9 preserves reaction connectivity.

	Run 7	**Run 8**	**Run 9**
Network	original network	permuted network	permuted network (preserving connectivity)
Enzyme interactions	artificial data (interact. probs. correl. to connectivity)	artificial data (interact. probs. correl. to connectivity)	artificial data (interact. probs. correl. to connectivity)
PPI involving neighboring enzymes (rel. amount)	1	0.45	0.96

Table 4. Overview on randomization runs 10-12 for *E. coli* using randomly generated enzyme interactions and interaction probabilities correlated to connectivity of reactions. Only those reactions were considered that are assigned to connectivity classes 1 to 8. Run 12 preserves reaction connectivity.

	Run 10	**Run 11**	**Run 12**
Network	original network	permuted network	permuted network (preserving connectivity)
Enzyme interactions	artificial data (interact. probs. correl. to connectivity)	artificial data (interact. probs. correl. to connectivity)	artificial data (interact. probs. correl. to connectivity)
PPI involving neighboring enzymes (rel. amount)	1	~1	~1

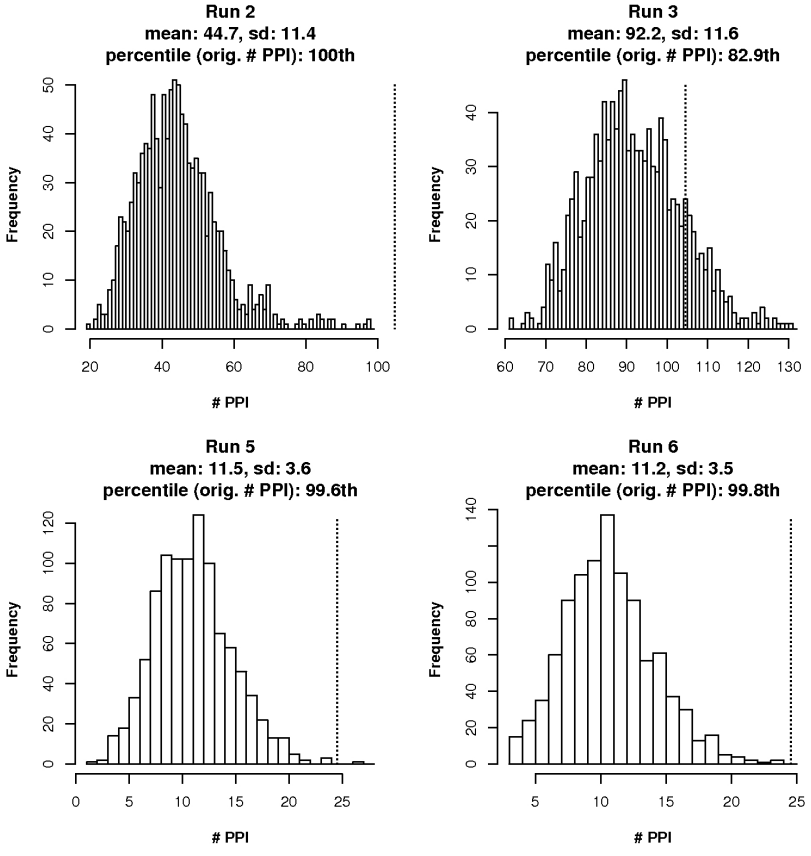

Fig. 1. PPI number distribution obtained by randomizing the *E. coli* network (see Table 1 and 2). Dotted lines represent number of interactions reported for neighboring enzymes in original network.

interaction data onto the original network and onto the two types of permuted networks resulted for all networks in similar numbers of interactions among neighboring enzymes (data not shown). In contrast, using in accordance to the real data higher interaction probabilities for those enzymes located at metabolic junctions yielded similar results as for experimentally observed interactions (compare Table 1 run 1-3 with Table 3 run 7-9). Significantly less interactions were found for neighboring enzymes in arbitrarily randomized networks than in the original network (compare Table 3 run 7 with run 8) but almost as high amounts in randomized networks where exchanges were restricted to reactions of similar connectivity (compare Table 3 run 7 with run 9). Considering only reaction pairs of low connectivity unveiled that interaction probabilities for neighboring enzymes in the original network and in the different types of randomized networks resemble (see Table 4 run 10-12).

3.3. *Known and yet uncharacterized enzyme complexes*

Our analyses of experimentally observed enzyme interactions as well as randomly generated interaction data mapped onto original metabolic networks and permuted networks suggest a functional role of interactions involving neighboring enzymes at least for those catalyzing sparsely connected reactions.

Reaction pairs corresponding to neighboring enzymes that are reported to interact can be grouped into three categories: both reactions consume the same metabolite (A), both reactions produce the same metabolite (B), or one reaction produces a metabolite that is consumed by the other reaction (C). Some reaction pairs have both common substrates and products and thus simultaneously belong to more than one category. Enzyme complexes of category C present candidates for metabolic channeling. Of the *E. coli* reaction set 37 reaction pairs were assigned to type A, 23 pairs to type B, and 48 reaction pairs to type C. In the case of yeast group A comprises 27 reaction pairs, group B 13 reaction pairs, and group C 73 pairs.

Scanning PubMed revealed that several previously known and characterized enzyme interactions were recovered by our large-scale analysis. This includes the tryptophan biosynthetic enzyme complex of anthranilate synthase and anthranilate phosphoribosyltransferase [1] as well as interactions involving the TCA enzymes malate dehydrogenase and citrate synthase [11]. In addition, enzyme complexes corresponding to 39 pairs of neighboring reactions in the *E. coli* network and 32 yeast reaction pairs were unveiled as potential candidates for metabolic channeling but have not been further experimentally characterized yet (for details see [8]). An example is given by the *E. coli* enzymes NAD synthase and the glycine cleavage system. Both enzymes produce metabolites (NAD and ammonia, respectively) that are consumed by the interaction partner. Furthermore, an interaction between the yeast enzymes fatty-acyl-CoA synthase and 1-acyl-glycerol-phosphate acyltransferase was observed. The first enzyme activates fatty acids of varying length with Co-A, which are substrates of the phosphatidate producing second enzyme thereby linking the fatty acid biosynthesis pathway to the phospholipid biosynthesis pathway.

Additionally, our analysis detected interactions that are not conform with the metabolic channeling concept (category A and B). This includes complexes comprising glutamine-fructose-6-phosphate transaminase (GFPT) from yeast, which is found to associate with several enzymes that use glutamine and produce glutamate like GFPT itself. The same holds for asparagine synthase, which uses the amido group of glutamine to generate asparagine from aspartate.

4. Discussion

Meanwhile, systematic and exhaustive analysis of protein-protein interactions in cells and tissues has become an important field of experimental cell biology and biochemistry. The protein-protein interaction networks resulting from these studies are on one hand impressive when depicted as complex network graphs but on the other hand they help only little to understand regulation and dynamics on the molecular

level unless the physiological relevance of a protein-protein interaction established under *in vitro* conditions is substantiated in the context of the underlying reaction network. One useful validation may consist in mapping proteins with reported interactions onto the topology of the corresponding metabolic network. In this work we have carried out a systematic statistical analysis of enzyme-enzyme interactions in metabolic networks based on currently available data in protein-protein interaction databases. We focused on interactions involving direct neighbors of a network as such interactions are best understood in the context of metabolic channeling.

We previously found that adjacent enzymes are more likely to be recorded in protein interaction databases than distant enzymes [8]. This can be interpreted in the sense that a large proportion of such enzyme interactions plays indeed a functional role. However, one has to ascertain that this is definitively attributed to biochemical function and not due to a bias from network topology and the way pairs of reactions are iterated. Permutation of networks and randomization of interaction data are strategies to discriminate between both. We counted protein interactions in neighboring reactions and performed the same analysis for randomized networks. Arbitrary permutation resulted in significantly less interactions among neighboring enzymes than in the original network (compare Table 1 run 1 with run 2) while permutation restricted to exchanges of reactions that have similar connectivity, in order to preserve topological properties of the network, yielded only slightly less interactions (compare Table 1 run 1 with run 3). The observation could be misinterpreted in the sense that neighboring enzymes interact more frequently than expected due to functional relevance of these interactions, which is related to network topology. Repetition of this analysis using randomly generated enzyme interactions resulted in similar findings (compare Table 1 run 1-3 with Table 3 run 7-9) proving this predication in its general form wrong. In this approach higher interaction probabilities were chosen for highly connected reactions in accordance to the real interaction data. However, when using randomly generated enzyme interactions and including only sparsely linked reaction pairs into the statistical analysis there we found practically no difference in the amount of next-neighbor interactions in the permuted networks and the original network (see Table 4 run 10-12) whereas for actually reported enzyme interactions the same permutation procedures resulted in a drop of more than 50% (compare Table 2 run 4 with run 5 and 6). This points to functional relevance of enzyme interactions for sparsely connected neighboring reactions. Although this statement is based on low numbers of observed interactions it is supported by the fact that the trend could be demonstrated for two distinct organisms. For enzyme interactions involving reactions with high connectivity functional importance cannot be concluded from our analysis. Anyhow, direct neighborhood of enzymes does not seem to play a role for these interactions. Since enzymes at metabolic branching points are often key enzymes controling whole pathways a regulatory function of interactions involving such enzymes is conceivable due to induced conformational changes, which alter enzyme activity.

Our result is conform with the concept of metabolic channeling, which often involves

consecutive reaction steps where intermediates are not or only little involved in other pathways. Examples of channeling where intermediates do not play important roles elsewhere are found in the synthesis pathways of pyrimidine [2], lumazine [10] and tryptophan [13, 17].

The number of well-studied cases is still small as biochemical evidences for a direct interaction of enzymes are difficult to obtain. Therefore we aimed to uncover putative novel examples of the metabolic channeling concept by analyzing high-throughput interaction data. Less than 3% of metabolic enzyme interactions reported for *E. coli* and yeast come into question for channeling of reaction intermediates, suggesting that metabolic channeling is not an ubiquitous phenomenon in cells. However, more than half of these candidate enzyme pairs have not been further experimentally characterized, in the case of *E. coli* even more than 80%, indicating that not all examples are known yet. Low connectivities of involved reactions might be a usefull discriminator for functional relevance of enzyme interactions as our results have shown. A putative metabolic channeling enzyme complex detected by our genome-wide analysis is that of fatty-acyl-CoA synthase and 1-acyl-glycerol-phosphate acyltransferase, which might have importance during cell devision as fatty acids are directly transfered to an enzyme of the phospholipid biosynthesis, which is needed for the synthesis of cellular membranes. By this means fatty acids are rather converted to phospholipids than being processed by competing enzymes. Intriguingly, our analysis revealed a substantial number of neighboring enzymes forming a complex for which the metabolic channeling concept does not apply as the two coupled reactions utilize a common substrate or generate the same product. It has to be noted that this finding has to be considered with care as it is based on *a priori* assumptions concerning the directionality of fluxes in the network. Nevertheless, it is worthwhile to ask for a possible functional relevance of such constellations deviating from the classical channeling concept. This includes several interactions of glutamine-fructose-6-phosphate transaminase (GFPT) with other enzymes utilizing amido nitrogen from glutamine as well. The functional relevance of these interactions might be the bundling of enzymes that either produce or consume the same metabolite. TAP pull-down assays revealed that GFPT furthermore interacts with glutamate synthase (precursor), which may form glutamine and 2-oxoglutarate from glutamate. Consequently, this interaction gives supply of glutamine for GFPT from which associated enzymes could participate.

References

[1] Browne, B. A., Itzel Ramos, A., and Downs, D. M., PurF-independent phosphoribosyl amine formation in yjgF mutants of Salmonella enterica utilizes the tryptophan biosynthetic enzyme complex anthranilate synthase-phosphoribosyltransferase, *J. Bacteriol*, 188(19):6786–6792, 2006.

[2] Christopherson, R. I., Traut, T. W., and Jones, M. E., Multienzymatic proteins in mammalian pyrimidine biosynthesis: channeling of intermediates to avoid futile cycles, *Curr. Top Cell. Regul.*, 1859–1877, 1981.

[3] Duarte, N. C., Herrgard, M J., and Palsson, B., Reconstruction and validation of

Saccharomyces cerevisiae iND750, a fully compartmentalized genome-scale metabolic model, *Genome Res.*, 14(7):1298–1309, 2004.

[4] Enright, A. J., Iliopoulos, I., Kyrpides, N. C., and Ouzounis, C A., Protein interaction maps for complete genomes based on gene fusion events, *Nature*, 402(6757):86–90, 1999.

[5] Gavin, A. C., Bosche, M. *et al.*, Functional organization of the yeast proteome by systematic analysis of protein complexes, *Nature*, 415(6868):141–147, 2002.

[6] Ge, H., Liu, Z., Church, G. M., and Vidal, M., Correlation between transcriptome and interactome mapping data from Saccharomyces cerevisiae, *Nat. Genet.*, 29(4):482–486, 2001.

[7] Ho, Y., Gruhler, A. *et al.*, Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry, *Nature*, 415(6868):180–183, 2002.

[8] Huthmacher, C., Gille, C., and Holzhütter, H.-G., A comprehensive statistical analysis of protein interactions in metabolic networks reveals novel enzyme pairs potentially involved in metabolic channeling, *J. Theor. Biol.*, 2007 (submitted).

[9] Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y., A comprehensive two-hybrid analysis to explore the yeast protein interactome, *Proc. Natl. Acad. Sci. USA*, 98(8):4569–4574, 2001.

[10] Kis, K. and Bacher, A., Substrate channeling in the lumazine synthase/riboflavin synthase complex of Bacillus subtilis, *J. Biol. Chem.*, 270(28):16788–16795, 1995.

[11] Lindbladh, C., Rault, M., Hagglund, C., Small, W. C., Mosbach, K., Bulow, L., Evans, C., and Srere, P. A., Preparation and kinetic characterization of a fusion protein of yeast mitochondrial citrate synthase and malate dehydrogenase, *Biochemistry*, 33(39):11692–11698, 1994.

[12] Liu, Y., Liu, N., and Zhao, H., Inferring protein-protein interactions through high-throughput interaction data from diverse organisms, *Bioinformatics*, 21(15):3279–3285, 2005.

[13] Matchett, W. H., Indole channeling by tryptophan synthase of neurospora, *J. Biol. Chem.*, 249(13):4041–4049, 1974.

[14] von Mering, C., Krause, R. *et al.*, Comparative assessment of large-scale data sets of protein-protein interactions, *Nature*, 417(6887):399–403, 2002.

[15] Mrowka, R., Patzak, A., and Herzel, H., Is there a bias in proteome research?, *Genome Res.*, 11(12):1971–1973, 2001.

[16] Ovadi, J., Physiological significance of metabolic channelling, *J. Theor. Biol.*, 152(1):1–22, 1991.

[17] Pan, P., Woehl, E., and Dunn, M. F., Protein architecture, dynamics and allostery in tryptophan synthase channeling, *Trends. Biochem. Sci.*, 22(1):22–27, 1997.

[18] Reed, J. L., Vo, T. D., Schilling, C. H., and Palsson, B., An expanded genome-scale model of Escherichia coli K-12 (iJR904 GSM/GPR), *Genome Biol.*, 4(9):R54, 2003.

[19] Tsoka, S. and Ouzounis, C. A., Prediction of protein interactions: metabolic enzymes are frequently involved in gene fusion, *Nat. Genet.*, 26(2):141–142, 2000.

[20] Uetz, P., Giot, L. *et al.*, A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae, *Nature*, 403(6770):623–627, 2000.

[21] Walhout, A. J. M., Reboul, J. *et al.*, Integrating interactome, phenome, and transcriptome mapping data for the C. elegans germline., *Curr. Biol.*, 12(22):1952–1958, 2002.