Computational Biology and Chemistry 58 (2015) 1–8

Contents lists available at ScienceDirect

Computational Biology and Chemistry

journal homepage: www.elsevier.com/locate/compbiolchem

Systems biology approach reveals possible evolutionarily conserved moonlighting functions for enolase

Gabriela Prado Paludo [a], Karina Rodrigues Lorenzatto [a,c], Diego Bonatto [b,c,d], Henrique Bunselmeyer Ferreira [a,c,d,*]

[a] Laboratório de Genômica Estrutural e Funcional, Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil
[b] Laboratório de Radiobiologia Molecular, Centro de Biotecnologia, UFRGS, Porto Alegre, RS, Brazil
[c] Programa de Pós-Graduação em Biologia Celular e Molecular, Centro de Biotecnologia, UFRGS, Porto Alegre, RS, Brazil
[d] Departamento de Biologia Molecular e Biotecnologia, Instituto de Biociências, UFRGS, Porto Alegre, RS, Brazil

ARTICLE INFO

Article history:
Received 23 September 2014
Received in revised form 13 April 2015
Accepted 19 April 2015
Available online 21 April 2015

Keywords:
Enolase
Glycolytic enzyme
Moonlighting functions
Systems biology
Protein–protein interaction network

ABSTRACT

Glycolytic enzymes, such as enolase, have been described as multifunctional complex proteins that also display non-glycolytic activities, termed moonlighting functions. Although enolase multifunctionality has been described for several organisms, the conservation of enolase alternative functions through different phyla has not been explored with more details. A useful strategy to investigate moonlighting functions is the use of systems biology tools, which allow the prediction of protein functions/interactions by graph design and analysis. In this work, available information from protein–protein interaction (PPI) databases were used to design enolase PPI networks for four eukaryotic organisms, namely *Homo sapiens, Drosophila melanogaster,Caenorhabditis elegans*, and *Saccharomyces cerevisiae*, covering a wide spectrum of this domain of life. PPI networks with number of nodes ranging from 140 to 411 and up to 15,855 connections were generated, and modularity and centrality analyses, and functional enrichment were performed for all of them. The performed analyses showed that enolase is a central node within the networks, and that, in addition to its canonical interactions with proteins related to glycolysis and energetic metabolism, it is also part of protein clusters related to different biological processes, like transcription, development, and apoptosis, among others. Some of these non-glycolytic clusters, are partially conserved between networks, in terms of overall sharing of orthologs, overall cluster structure, and/or at the levels of key regulatory proteins within clusters. Overall, our results provided evidences of enolase multifunctionality and evolutionary conservation of enolase PPIs at all these levels.

1. Introduction

One of the major challenges of post-genomic biology is to understand how genes, proteins and other molecules interact to compose cellular systems (Zhang et al., 2007). Nowadays, the idea of "one gene, one protein, one function" has been replaced by the knowledge that many proteins display multiple functions (Jeffery, 2009). For example, the so called moonlighting proteins could display two or more different functions within a single polypeptide chain (Copley, 2012). The functions of a moonlighting protein may depend on cellular localization, cell type, oligomeric state, and/or the cellular concentration of ligands, substrates, cofactors or products (Jeffery, 1999). These different factors are not mutually exclusive and, in many cases, the functions of a protein depend on the overall physiological status of the cell. Therefore, the identification of one or more non-canonical, moonlighting functions of a given protein can be a difficult task, and, in this context, protein–protein interactions (PPI) databases can be useful tools. Systems biology approaches using PPI databases have been successfully applied for inferences on moonlighting functions for different proteins (Campanaro et al., 2007; Delprato, 2012; Gómez et al., 2011; Zhang et al., 2013a).

The development of high-throughput methods for the study of biological molecules interactions, such as protein microarrays (Jones et al., 2006), two-hybrid assays (Yu et al., 2008), co-immunoprecipitation (Hegele et al., 2012) and peptide phage display (Carducci et al., 2012), allowed the generation of robust PPI databases for four eukaryotic organisms, namely the yeast *Saccharomyces cerevisiae*, the free living nematode *Caenorhabditis elegans*, the fruit fly *Drosophila melanogaster*, and man

* Corresponding author at: Laboratório de Genômica Estrutural e Funcional, Centro de Biotecnologia, UFRGS, Av. Bento Gonçalves, 9500 – Prédio 43-421, Sala 210, Caixa Postal 15005, 91501-970 Porto Alegre, RS, Brazil. Tel.: +55 51 3308 7768.
E-mail address: henrique@cbiot.ufrgs.br (H.B. Ferreira).

http://dx.doi.org/10.1016/j.compbiolchem.2015.04.010
1476-9271/

Homo sapiens. To understand how cellular events are coordinated at the molecular level, the information contained in these databases can be arranged in a graph context to create PPI networks (Milenković and Przulj, 2008; Simonis et al., 2009). The properties of PPI networks can be explored to identify interaction possibilities whose effectiveness depend on physical binding events (Deeds et al., 2012). The cumulative effect of such events results in a distribution of protein complexes that ultimately determines cellular behavior.

Enolase, also known as phosphopyruvate hydratase, is a glycolytic enzyme that catalyzes the dehydratation of 2-phosphoglycerate to phosphoenolpyruvate. This protein is present in all three domains of life and is one of the most abundantly expressed cytosolic proteins in many organisms (Piast et al., 2005). In addition to its classical involvement in glycolysis, other functional roles have been described for enolase. For example, enolase plays an important role in mouse mast cell differentiation (Ryu et al., 2012). Knockdown of enolase expression in different tumor cell lines causes a dramatic increase in their sensitivity to microtubule targeting drugs, suggesting that enolase expression levels can affect the sensitivity of tumor cell lines to anti-tubulin drugs (Georges et al., 2011). Moreover, enolase (and other glycolytic enzymes) has been involved in a specific apoptosis mechanism, in which it is externalized and participate in immune modulation, holding promise for understanding and addressing causes of autoimmune and inflammatory pathology (Ucker et al., 2012). Although the multifunctionality of enolase has been increasingly evident, it is not yet clear how many different functions this protein can play and whether its non-glycolytic functions are evolutionarily conserved or not.

In this work, a systems biology approach was used to investigate the moonlighting functions of enolase in different eukaryotic organisms. This approach allowed the comparison of possible enolase interactions and functions in *S. cerevisiae D. melanogaster*, *C. elegans* and man. Our results showed that enolase is a central node in all designed networks and is involved in conserved interactions with proteins related to different cellular functions. Besides its canonical interactions with proteins related to glycolysis and energetic metabolism, enolase was also found as part of conserved proteins clusters associated with several other biological processes. The importance of enolase moonlighting functions in such evolutionarily conserved networks is discussed.

2. Material and methods

2.1. Enolase ortholog analysis

Multiple sequence alignments were performed with the Clustal Omega software (http://www.ebi.ac.uk/Tools/msa/clustalo/) (Sievers et al., 2011) to define enolase orthologs used for network design. Human enolase α isoform (gi|1167843|emb|CAA34360.1| alpha-enolase [*H. sapiens*]) was established as reference, along with the single *S. cerevisiae* enolase (gi|171457|gb|AAA88713.1| enolase [*S. cerevisiae*]). Enolase A (gi|22945470|gb|AAN10458.1| enolase, isoform A [*D. melanogaster*]), and enolase 1 (gi|3879986| emb|CAA92692.1| Protein ENOL-1, isoform a [*C. elegans*]) were defined by sequence identity and similarity criteria as the orthologs in *D. melanogaster* and *C. elegans*, respectively.

2.2. Protein–protein interaction network design and global topological analysis

Enolase PPI networks of model organisms were designed using the metasearch tool STRING 8.2 (http://www.string-db.org). In this sense, the following parameters were used: active prediction methods all enabled except text mining; no more than 50 interactions; and high confidence score (0.700). To calculate the confidence score of a specific connection, various major sources of interaction/association data were benchmarked independently by STRING (Szklarczyk et al., 2011), and combined scores were computed to indicate higher confidences when more than one type of information supported any association. Taking these parameters into account, networks were generated for *H. sapiens*, *D. melanogaster*, *C. elegans* and *S. cerevisiae*. STRING searches were subsequently analyzed using Cytoscape 2.6.3 (http://www.cytoscape.org) (Shannon et al., 2003). The networks were analyzed in terms of global topology, to define the number of protein clusters (subnetworks), using Cytoscape AllegroMCODE plugin (http://www.allegroviva.com/allegromcode) (Gudbjartsson et al., 2000). The parameters used in AllegroMCODE to generate subnetworks were: loops included; degree cutoff 2, node score cutoff 0.2, *K*-Core 2, and maximum depth of network 100. Each subnetwork obtained from this analysis was evaluated in terms of gene ontologies (GOs).

2.3. Network centrality analyses

Networks centralities (Scardoni et al., 2009), based on node degree (the number of edges connected to each node) and betweenness (the number of shortest paths from all vertices to all others that pass through each node), were calculated by Cytoscape CentiScaPe 1.21 plugin (http://www.cbmc.it/~scardonig/centiscape/centiscape.php). Highly connected nodes, considered hubs, were defined as those with node degree value higher than the node degree value threshold for the total network. Bottleneck nodes were defined as those with betweenness values higher than the network betweenness threshold.

2.4. Gene ontology analyses

The generated protein clusters were analyzed by the Biological Networks Gene Ontology (BiNGO) (version 2.44) Cytoscape plugin http://www.cytoscape.org (Maere et al., 2005) to identify major associated biological processes. The degree of functional enrichment for a given cluster and category (*P*-value) was calculated using hypergeometric distribution, and multiple test correction was also valued by the false discovery rate (FDR) algorithm, fully implemented with a significance level of $P < 0.05$.

2.5. Protein orthology evaluation

The degree of evolutionary network conservation among different organisms was evaluated by analyzing the type and number of ortholog proteins found in each network. For this purpose, we compared ortholog proteins between each two networks, considering all possible pairs, and, finally, among all networks. Amino acid sequences from probable ortholog proteins were submitted to multiple alignments using the Clustal Omega software. Only proteins with similarity higher than 50% were considered, according to Pearson (2013), in order to define orthologs. The functional categories of the orthologs were identified using the KOG tool (http://genome.jgi.doe.gov/Tutorial/tutorial/kog.html) (Götz et al., 2011). To compare the proteins present in the networks, the identity of proteins were verified in specific databases for each organism (*H. sapiens*: http://www.genecards.org; *D. melanogaster*: http://flybase.org; *C. elegans*: http://www.wormbase.org; *S. cerevisiae*: http://www.yeastgenome.org).