Research Article

Reconstruction of Protein-Protein Interaction Network of Insulin Signaling in Homo Sapiens

Saliha Durmuş Tekir, Pelin Ümit, Aysun Eren Toku, and Kutlu Ö. Ülgen

Department of Chemical Engineering, Boğaziçi University, 34342 Bebek-İstanbul, Turkey

Correspondence should be addressed to Saliha Durmuş Tekir, salihadurmus@boun.edu.tr

Received 21 April 2010; Revised 22 July 2010; Accepted 27 October 2010

Academic Editor: Sorin Draghici

Copyright © 2010 Saliha Durmuş Tekir et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Diabetes is one of the most prevalent diseases in the world. Type 1 diabetes is characterized by the failure of synthesizing and secreting of insulin because of destroyed pancreatic $\beta$-cells. Type 2 diabetes, on the other hand, is described by the decreased synthesis and secretion of insulin because of the defect in pancreatic $\beta$-cells as well as by the failure of responding to insulin because of malfunctioning of insulin signaling. In order to understand the signaling mechanisms of responding to insulin, it is necessary to identify all components in the insulin signaling network. Here, an interaction network consisting of proteins that have statistically high probability of being biologically related to insulin signaling in Homo sapiens was reconstructed by integrating Gene Ontology (GO) annotations and interactome data. Furthermore, within this reconstructed network, interacting proteins which mediate the signal from insulin hormone to glucose transportation were identified using linear paths. The identification of key components functioning in insulin action on glucose metabolism is crucial for the efforts of preventing and treating type 2 diabetes mellitus.

1. Introduction

Signaling provides the communication of living cells by processing biological information. Mammalian cells integrate information from complex intracellular signaling pathways to make decisions in response to changes in the environment. Using systematic genome-wide and pathway specific protein-protein interaction screens, a framework of the interconnectivity of a large number of human proteins, including therapeutically relevant disease-associated proteins has been generated by these pathways. Recent developments in these protein-protein interaction networks have increased the understanding of the mechanisms of diseases with identification of drug targets and adaptation of living cells to the environment [1–5].

In mammalian cells, the balance between hepatic glucose production and glucose utilization by the tissues, such as liver, adipose, muscle, brain, and kidney provides the glucose homeostasis. In healthy individuals, the increased blood glucose levels result in secretion of insulin from $\beta$-cells of the pancreas. Insulin triggers the transportation of glucose into peripheral tissues by glucose transporter GLUT4 inhibiting hepatic glucose production [6]. By the stimulation of insulin (INS) hormone, the insulin receptor (INSR) phosphorylates insulin receptor substrate (IRS) proteins that activate two main signaling pathways. The phosphatidylinositol 3-kinase (PI3K)-AKT/protein kinase B (PKB) pathway is responsible for the metabolic actions of the insulin such as glucose uptake, glycogen synthesis, gene expression, and protein synthesis. The Ras-mitogen activated protein kinase (MAPK) pathway controls cell growth and differentiation by regulating expression of some genes and cooperating with PI3K pathway [7, 8]. Defects in insulin signaling pathways may decrease the ability of peripheral tissues to respond to insulin (insulin resistance) causing type 2 diabetes. Beside its primary role in glucose homeostasis, insulin signaling mechanism also regulates ion and amino acid transport, lipid metabolism, glycogen synthesis, gene transcription and mRNA turnover, protein synthesis and degradation, and DNA synthesis by a complex, highly integrated network activated by the insulin receptor [6, 9].
Most of research published so far reports experimental and computational work to decipher small-scale mechanisms around key proteins in insulin metabolism [6, 10–13]. However, it is very important to capture the global picture of insulin signaling in order to understand the mechanisms underlying diabetes with crosstalks between other signaling networks. This need motivated us for the reconstruction of insulin signaling network in Homo sapiens with the aim of identification of all known components together with new candidate proteins of insulin signaling. In this study, a computational framework integrating interactome data with GO annotations was used to build large scale protein interaction network which is composed of candidate proteins for insulin signaling in Homo sapiens. The reconstructed insulin signaling network was decomposed into linear paths resulting in glucose transportation to be able to identify the proteins functioning in this metabolic action of insulin. The topology of the reconstructed insulin signaling network governing glucose transportation was then analyzed to determine whether the network properties are biologically feasible or not, and to obtain detailed information about the signaling mechanisms. Moreover, graph theoretic analysis gives the proteins that are well or poorly connected in the interaction network. This study provides a comprehensive insulin signaling network with indication of key components which will facilitate a deeper understanding of underlying mechanisms of insulin-resistant states and pathophysiology of insulin deficiency.

2. Materials and Methods

2.1. Network Reconstruction by GO Annotations. Figure 1 represents an overview of the computational approach integrating Gene Ontology (GO) annotations and interactome data for the reconstruction of a protein interaction network which was used to predict candidate proteins in insulin signaling in human. All known interacting human proteins obtained from BioGRID version 2.0.61 release were used as inputs to the algorithm. BioGRID (The Biological General Repository for Interaction Datasets) uses the results of high-throughput experiments and conventional studies [14]. The GO annotations (in terms of cellular component, molecular function and biological process) of the core proteins that are known to have certain functions in the insulin signaling were collected (http://www.ebi.ac.uk/QuickGO/) to form an annotation collection table (see Supplementary material available on line at doi:10.1155/2010/690925). The relevance of the human proteins to the insulin signaling was tested by employing this annotation collection table. The proteins with all three GO terms matching to those in the annotation collection table were added to the network. Thus, a high probability of having role in the insulin-interaction network is ensured for these proteins. In the second step, the interaction data among these proteins were obtained from BioGRID version 2.0.61 release, and the network architecture was constructed.

2.2. Network Decomposition Analysis. Network decomposition analysis is based on the decomposition of a protein interaction network into linear paths starting from inputs (ligands) and extending to outputs (cellular responses). In the reconstructed insulin signaling network, the linear paths from the insulin receptor to the glucose transporter GLUT4 were found by the NetSearch algorithm [15], and the specific part of the protein network governing insulin action on glucose metabolism was identified. The participation of the proteins in linear paths can be considered as an indication of their importance in the signal transduction, since any state of the signaling network is a combination of the linear paths [16]. Therefore, the participation percentages of each protein in linear paths of the reconstructed insulin signaling network were calculated to get an insight on the roles of the proteins in the signal transduction from INS to GLUT4.

2.3. Graph Theoretic Analysis. The topology of the reconstructed protein-protein interaction network functioning in glucose transportation was determined by graph theoretic analysis [17–19] based on the properties, such as the degree (connectivity) of nodes, the number of hubs (highly connected nodes), and the shortest path lengths between indirectly connected nodes, network diameter and mean path length. The graph properties of the network were found using Network Analyzer plugin (ver. 2.6.1) of Cytoscape (ver. 2.6.3). The input to the calculation is the list of binary interacting proteins. Observing the connectivity distribution of the proteins allows us to identify highly connected proteins which participate in significant numbers of interactions and play critical roles in the organization of the cellular protein interaction network. Mean path length and network diameter are calculated as the average and the maximum of the shortest path lengths, respectively.

3. Results and Discussion

In the present study, the protein-protein interaction network of insulin signaling was reconstructed in Homo sapiens with special emphasis on glucose transportation mechanism. During the reconstruction of a protein interaction network, the main problem is the existence of false positives and false negatives in the available interaction data obtained mostly by high-throughput screens [20–22]. Several approaches have been performed to improve the quality of the data by integrating different biological features, including GO annotations [23–27]. Compared to metabolic and regulatory networks, the reconstruction and analysis of signaling networks are very limited. The previous signaling network reconstruction methods are focused on integration of protein-protein interaction data with microarray gene expression profiles [15, 28–30] or a detailed literature survey on published knowledge [2, 3]. Here, we used a computational framework integrating interactome data with GO annotations.

3.1. Reconstruction of Insulin Signaling Network in H. Sapiens. 30 proteins related to human insulin signaling were identified by the literature information [6, 7, 12, 31–35] and GO annotations (Table 1). Through the literature search only experimental cases were investigated, and the proteins that are
Table 1: Core proteins of the insulin signaling network.

| Protein Symbol | Protein Name                                                                 |
|----------------|-----------------------------------------------------------------------------|
| AHSG           | Alpha-2-HS-glycoprotein                                                     |
| AKT1           | RAC-alpha serine/threonine-protein kinase                                   |
| AKT2           | RAC-beta serine/threonine-protein kinase                                    |
| AP3S1          | AP-3 complex subunit sigma-1                                                |
| BAIAP2         | Brain-specific angiogenesis inhibitor 1-associated protein 2                |
| BCAR1          | Breast cancer antiestrogen resistance protein 1                             |
| CILP           | Cartilage intermediate layer protein 1                                     |
| ENPP1          | Ectonucleotide pyrophosphatase/phosphodiester family member 1              |
| FOXO4          | Forkhead box protein O4                                                     |
| GAB1           | GRB2-associated-binding protein 1                                           |
| GRB2           | Growth factor receptor-bound protein 2                                      |
| GRB10          | Growth factor receptor-bound protein 10                                     |
| IGFBP5         | Insulin-like growth factor-binding protein 5                                |
| IGF1           | Insulin-like growth factor 1A                                               |
| IGF1R          | Insulin-like growth factor 1 receptor                                       |
| IGF2           | Insulin-like growth factor II                                               |
| IGF2R          | Insulin-like growth factor type 2 receptor                                  |
| INS            | Insulin                                                                     |
| INSR           | Insulin receptor                                                            |
| IRS1           | Insulin-receptor substrate 1                                                |
| IRS2           | Insulin-receptor substrate 2                                                |
| PDPK1          | 3-phosphoinositide-dependent protein kinase 1                               |
| PHIP           | PH-interacting protein                                                      |
| PIK3R1         | Phosphatidylinositol 3-kinase regulatory subunit alpha                       |
| PIK3R3         | Phosphatidylinositol 3-kinase regulatory subunit gamma                       |
| SIK2           | Salt inducible protein kinase 2                                              |
| SLC2A4 (GLUT4) | Solute carrier family 2, facilitated glucose transporter member 4          |
| SOCS1          | Suppressor of cytokine signaling 1                                          |
| SORBS1         | Sorbin and SH3 domain-containing protein 1                                  |
| TSC2           | Tuberin                                                                     |

reported as functioning in insulin signaling mechanisms in human were considered as the core proteins. In addition to that, some core proteins were collected via their GO function and process terms which indicate insulin signaling explicitly. Therefore, each of these proteins is known to be essential for insulin actions. For instance, the binding of insulin receptor substrate-1 (IRS1) to the phosphorylated insulin receptor (INSR) leads to the activation of phosphatidylinositol 3-kinase (PI3K) whose regulatory subunits (PIK3R1 and PIK3R3) play pivotal roles in the metabolic and mitogenic actions of insulin. AKT1 plays an important role in GLUT4 translocation via phosphorylating and regulating components of GLUT4 complex [7, 36]. By reconstructing the protein-protein interaction network this study unravels the mechanisms around these insulin signaling proteins.

8211 interacting human proteins obtained from BioGrid 2.0.61 were tested through the GO annotations of the core proteins. If there is at least one annotation for each of the GO terms (component, function, and process) that are included in the annotation collection table, the corresponding protein was added to the network. Consequently, 6248 proteins passed this selection criterion increasing their probability to have function in insulin signal transduction. However, only 3588 of these proteins have interactome data, and of these, 365 proteins cannot be included into the network as the GO terms of their interacting partners do not coincide with those in the annotation collection table. Eventually, an interaction network of 3223 nodes and 10537 edges is obtained for insulin signaling. When the isolated smaller parts are removed the resulting protein-protein interaction network consists of 3056 proteins and 10401 interactions among them (see Supplementary material). Two of the core proteins CILP and PHIP are not included in the reconstructed network, since CILP has no interaction data, and PHIP’s interacting partners does not fulfill the selection criterion based on GO annotations.

3.2. Network Decomposition Analysis. In a protein interaction network, a signaling pathway for a specific signaling output can be identified using linear paths starting from
membrane-bound receptors and ending at that particular cellular response [23]. The linear paths of the reconstructed network (3056 proteins and 10401 interactions) were found using NetSearch algorithm of Steffen and coworkers [15]. INSR (insulin receptor) and GLUT4 (glucose transporter 4) proteins were used as the input and the output of the signaling network, respectively, for the identification of the proteins that have roles in the insulin signal transduction triggered by binding of insulin to its receptor and ending with metabolic action of glucose transportation.

The shortest path length between INSR and GLUT4 was found as 4, since the shortest 7 linear paths include 5 proteins connected linearly by 4 interactions. In order to determine the optimum path length for the identification of the linear paths functioning specifically in glucose transportation, the paths were searched by increasing the maximum path length by one each time (Table 2). The number of core proteins and the interacting proteins included in the linear paths were investigated to determine the critical path length and participating proteins that have roles in glucose transport response of the signaling network. Between INSR and GLUT4, a path length of 6 resulting in 7176 linear paths was chosen to be optimum, as it provides a balance between smaller path length and participating core proteins. The criterion of small path length is reasonable, since signaling mechanisms are known to give such responses very quickly [37]. Increasing the maximum path length from 6 to 7 causes the number of the core proteins that participate in the linear paths to increase only by one, from 17 to 18, despite a nearly two fold increase in the number of interacting proteins. Increasing the path length more than 6 would result in nearly same signaling mechanisms around these 17 core proteins with longer paths covering more proteins in the insulin signaling network. Therefore, these 498 proteins and 2887 interactions (see Supplementary material) that function in the linear paths at a path length of 6 constitute the insulin signaling pathway having roles in glucose translocation.

Bottleneck proteins are known as the key connectors that are central to many shortest paths in an interaction network [38]. To identify the bottlenecks in the signal
transduction from INSR to GLUT4, the percentage of each protein contributing to 7176 linear paths were calculated. INSR, GLUT4 and DAXX (death domain-associated protein 6) participate in all the linear paths since they are the input, the output, and the unique protein that connects GLUT4 to the network as its interacting partner, respectively. 10 following proteins with the highest participation in these linear paths (Table 3) should be investigated with special care owing to their critical roles in transducing the signal from INSR to GLUT4. These proteins are highly encountered in linear paths as many are bound to the input or output proteins. SMAD2 (mothers against decapentaplegic homolog 2), MAPK1 (mitogen activated protein kinase 1), and JAK2 (tyrosine protein kinase JAK2) interact with INSR in the reconstructed network. AR (androgen receptor), MDM2 (E3 ubiquitin protein ligase Mdm2), NR3C1 (glucocorticoid receptor), UBE2I (SUMO-conjugating enzyme UBC9), HDAC1 (histone deacetylase 1), and PML (probable transcription factor PML) interact with DAXX which is the only protein having an interaction with GLUT4 in the network. On the other hand, TP53 (tumor protein p53) was also found to be a bottleneck in the PPI network, since it has interactions with other bottleneck proteins such as MAPK1, MDM2, UBE2I, HDAC1, PML, and NR3C1. Since these bottleneck proteins control most of the signal transduction from insulin to the glucose transporter protein, their mutations may cause glucose transportation system to fail resulting in insulin resistance.

The most promising result of architecture of the reconstructed network is about DAXX protein, since it connects GLUT4 to the network. Although its physical interaction with GLUT4 was reported [41], its functional roles in insulin signaling mechanism remain elusive except very few studies [42]. Therefore, the ending part of the interaction network functioning through glucose transportation should be investigated thoroughly to discover the effects of DAXX on GLUT4.

3.3. Graph Theoretic Analysis. The reconstructed insulin signaling network was represented by an undirected interaction graph with 498 nodes and 2887 edges. The topological analysis was performed using Network analyzer plugin of Cytoscape. The network diameter and the mean path length were found as 5 and 2.9, respectively, indicating the small-world topology. A comparative analysis of the graph theoretic properties of several protein interaction networks (Table 4) similarly reveals the small-world architecture. The comparison of the number of nodes and edges of the present network with those of the other PPI networks indicates that the reconstructed insulin signaling network is highly connected, that is, its average connectivity is 11.6. The small network diameter and the low mean path length result from this architecture since any two nodes in the network are connected by shorter paths through high number of neighbouring proteins. The connectivity (k, the number of links per node) distribution of the nodes in the reconstructed graph was found as scale-free (Figure 2) following nearly a power law model \( P(k) \approx k^{-\gamma} \) \( \gamma = 1.53 \) \( R^2 = 0.83 \). Having small-world properties with scale-free topology is a general characteristic of complex biological networks [43–47]. The node of GLUT4 with only one edge was excluded in the inner diagram of Figure 2 since it is an outlier point.

The hubs of the insulin signaling network were determined as GRB2 (growth factor receptor bound protein 2), HDAC1, AR, and TP53 having connectivity values of 88, 84, 83 and 74, respectively. GRB2 has a vital role in signaling by receptor protein tyrosine kinases, where its SH2 and SH3 domains bind to the receptors and effectors and it functions in the insulin signaling through lots of proteins including IRS1 [34, 35]. It was reported that HDAC inhibition in human primary myotubes increases endogenous GLUT4 gene expression [48]. Investigating all HDAC proteins in the reconstructed network (HDAC1, 2, 3, 4, 5, 9) may provide potential drug targets for the treatment and management of insulin resistance and type 2 diabetes. Similar to HDAC1, TP53 has a repressive effect on transcriptional activity of the GLUT4 gene promoters. Mutations within its DNA-binding domain were found to impair this repressive effect resulting in increased glucose metabolism and cell energy supply facilitating tumor growth [49]. AR functions mainly as a ligand-activated transcription factor. Besides, it was reported to induce the rapid activation of kinase signaling cascades [50]. In addition to having a high degree in the protein interaction network of insulin signaling, AR was also found to have the highest participation in the linear paths from INSR to GLUT4 (Table 3). This is one of the promising results of this study indicating critical nodes in the insulin signaling governing glucose transportation.

4. Conclusions

There is a growing need for a comprehensive protein-protein interaction network of insulin signaling, especially covering its part on glucose metabolism with the aim to solve the type 2 diabetes problem. Here, we integrated GO annotations and interactome data for the reconstruction of a protein interaction network of insulin signalling, considering the relevance of the proteins as well as their interactions. Starting with 30 insulin signalling-related proteins, the proposed method resulted in an interaction network of 3056 proteins and 10,401 protein-protein interactions for human insulin signaling. The linear paths transducing the signal from the insulin receptor to the glucose transporter protein include 498 proteins with 2887 physical interactions and constitute the network of signaling for glucose transportation. The key components of the reconstructed network were identified as bottlenecks and hubs since they are crucial for the signal processing being central to many signaling paths and having many neighboring proteins, respectively. The mechanisms around these components, for example, directed interactions, activation, or inhibition effects in the reconstructed insulin signaling network, are potential targets for further analyses to gain insight on causes and results of type 2 diabetes. Additionally, DAXX protein requires special care being the unique protein that connects the flowing information to the GLUT4 protein. Finally, other putative insulin signaling
proteins having interaction with GLUT4 should be searched to obtain a robust network. This large-scale protein-protein interaction network allows us to consider any signaling node within its global working mechanism which is required by the holism perspective of systems biology approach.

Acknowledgments

The financial support for this research was provided by the Research Funds of Boğaziçi University and TÜBİTAK through Projects nos. 109A501P and 108M649, respectively. The doctoral scholarship for S. D. Tekir is sponsored by TÜBİTAK is gratefully acknowledged.

References

[1] H. Ruffner, A. Bauer, and T. Bouwmeester, “Human protein-protein interaction networks and the value for drug discovery,” Drug Discovery Today, vol. 12, no. 17–18, pp. 709–716, 2007.
[2] K. Oda and H. Kitano, “A comprehensive map of the toll-like receptor signaling network,” Molecular Systems Biology, vol. 2, Article ID 2006.0015, 2006.
[3] K. Oda, Y. Matsuoka, A. Funahashi, and H. Kitano, “A comprehensive pathway map of epidermal growth factor receptor signaling,” Molecular Systems Biology, vol. 1, Article ID 2005.0010, 2005.
[4] J. A. Papin and B. O. Palsson, “The JAK-STAT signaling network in the human B-cell: an extreme signaling pathway analysis,” Biophysical Journal, vol. 87, no. 1, pp. 37–46, 2004.
[5] D. E. Moller, “New drug targets for type 2 diabetes and the metabolic syndrome,” Nature, vol. 414, no. 6865, pp. 821–827, 2001.
[6] B. Cheatham and C. R. Kahn, “Insulin action and the insulin signaling network,” Endocrine Reviews, vol. 16, no. 2, pp. 117–142, 1995.
[7] C. M. Taniguchi, B. Emanuelli, and C. R. Kahn, “Critical nodes in signalling pathways: insights into insulin action,” Nature Reviews, vol. 7, no. 2, pp. 85–96, 2006.
[8] J. Avruch, “Insulin signal transduction through protein kinase cascades,” Molecular and Cellular Biochemistry, vol. 182, no. 1–2, pp. 31–48, 1998.
[9] A. Avogaro, S. V. de Kreutzenberg, and G. P. Fadini, “Oxidative stress and vascular disease in diabetes: is the dichotomization of insulin signaling still valid?” Free Radical Biology and Medicine, vol. 44, no. 6, pp. 1209–1215, 2008.
[10] Y. H. Chew, Y. L. Shia, C. T. Lee et al., “Modeling of glucose regulation and insulin signaling pathways,” Molecular and Cellular Endocrinology, vol. 303, no. 1–2, pp. 13–24, 2009.
[11] K. Schmelzle, S. Kane, S. Gridley, G. E. Lienhard, and F. M. White, “Temporal dynamics of tyrosine phosphorylation in insulin signaling,” Diabetes, vol. 55, no. 8, pp. 2171–2179, 2006.
[12] A. R. Sedaghat, A. Sherman, and M. J. Quon, “A mathematical model of metabolic insulin signaling pathways,” American Journal of Physiology, vol. 283, no. 5, pp. E1084–E1101, 2002.
[13] B. Cheatham, C. J. Vlahos, L. Cheatham, L. Wang, J. Blenis, and C. R. Kahn, “Phosphatidylinositol 3-kinase activation is required for insulin stimulation of pp70 S6 kinase, DNA synthesis, and glucose transporter translocation,” Molecular and Cellular Biology, vol. 14, no. 7, pp. 4902–4911, 1994.
[14] C. Stark, B. J. Breitkreutz, T. Reguly, L. Boucher, A. Breitkreutz, and M. Tyers, “BioGRID: a general repository for interaction datasets,” Nucleic Acids Research, vol. 34, pp. D535–539, 2006.
[15] M. Steffen, A. Petti, J. Aach, P. D’haeseleer, and G. Church, “Automated modelling of signal transduction networks,” BMC Bioinformatics, vol. 3, article 34, 2002.
[16] S. Durmuş Tekir, K. Yağcı Arga, and K. O. Ülgen, “Drug targets for tumorigenesis: insights from structural analysis of EGFR signaling network,” Journal of Biomedical Informatics, vol. 42, no. 2, pp. 228–236, 2009.
[17] N. Pržulj, D. A. Wigle, and I. Jurisica, “Functional topology in a network of protein interactions,” Bioinformatics, vol. 20, no. 3, pp. 340–348, 2004.
[18] A.-L. Barabási and Z. N. Oltvai, “Network biology: understanding the cell’s functional organization,” Nature Reviews Genetics, vol. 5, no. 2, pp. 101–113, 2004.
[19] R. Albert, “Scale-free networks in cell biology,” Journal of Cell Science, vol. 16, no. 2, pp. 117–142, 1995.
[20] C. M. Deane, L. Salwinski, I. Xenarios, and D. Eisenberg, “Protein interactions: two methods for assessment of the reliability of high throughput observations,” Molecular & Cellular Proteomics, vol. 1, no. 5, pp. 349–356, 2002.
[21] E. Sprinzak, S. Sattath, and H. Margalit, “How reliable are experimental protein-protein interaction data?” Journal of Molecular Biology, vol. 327, no. 5, pp. 919–923, 2003.
[22] K. Y. Arga, Z. I. Onsan, B. Kirdar, K. O. Ülgen, and J. Nielsen, “Understanding signaling in yeast: insights from network analysis,” Biotechnology and Bioengineering, vol. 97, no. 5, pp. 1246–1258, 2007.

Table 4: Graph properties of protein interaction networks.

| Model                  | Number of nodes | Number of interactions | Mean path length | Diameter |
|------------------------|-----------------|------------------------|------------------|----------|
| Insulin signaling      | 498             | 2887                   | 2.9              | 5        |
| EGFR signalling (Oda et al. 2005) | 329             | 1795                   | 4.7              | 11       |
| DIP (S. cerevisiae)    | 4773            | 15444                  | 5.0              | 13       |
| DIP (S. cerevisiae)    | 5798            | 20098                  | 4.9              | 12       |
| DIP (D. melanogaster)  | 7451            | 22819                  | 4.4              | 11       |
| DIP (C. elegans)       | 2638            | 4030                   | 4.8              | 14       |
| DIP (H. pylori)        | 710             | 1420                   | 4.1              | 9        |
| DIP (H. sapiens)       | 1065            | 1369                   | 6.8              | 21       |
| DIP (E. coli)          | 553             | 761                    | 5.5              | 16       |
| DIP (M. musculus)      | 329             | 286                    | 3.6              | 9        |
