Interactive Visualization of Metabolic Pathways

Marc Streit Michael Kalkusch Dieter Schmalstieg

ICG - Graz University of Technology, Austria

ABSTRACT

This work enriches the visualization of complex graphs from the biomedical domain called metabolic pathways by means of interactive linked views in 2D and 3D. Due to the high degree of complexity of pathways biologists are restrained by currently available systems in the process of hypothesis generation and validation. We propose an interactive visual data mining process that is supported by the integration of methods as detail-on-demand, focus and context, and semantic zoom.

CR Categories and Subject Descriptors: I.3.6 [Computer Graphics]: Methodology and Techniques – Interaction Techniques.

Additional Keywords: Metabolic pathways, visualization, hierarchical graph, complex graph, non-planar graph.

1 INTRODUCTION

Metabolic pathways represent cellular functions and are therefore important for the daily work of biologists and pathologists. Metabolic pathways are directed non-planar graphs consisting of two types of nodes namely chemical substances and enzymes. The chemical substances are called compounds or metabolites. One compound is part of a chemical reaction. A compound is denoted as substrate, if it is a source substance and product if it is an output substance.

Inside regular metabolic pathways enzymes are linked to compounds only and may have multiple fan-in and multiple fan-out. Metabolic pathways can be grouped hierarchically in 4 levels as shown in figure 1. The first 3 levels represent different bio-molecular properties while the fourth bottom level consists of chemical reactions only. Since pathways describe biological properties one pathway can be linked or embedded inside another pathway. This is a side effect of flattening the non-planar graph.

Chemical reactions are not visualized at the moment, but they are accessible via the integrated web browser and KEGG [1], a public accessible pathway data base. Enzymes and compounds can appear multiple times inside different sub-groups as well as in one sub-graph. We propose the usage of multiple views showing different portions and scales of the metabolic graphs [2]. The graphs are also available in multiple layout variants, essentially semantic level of detail. Linked views allow showing all instances simultaneously. We address this issue by visually linking multiple nodes in 2.5D views [3].

2 OVERVIEW

We focused on enriching metabolic pathways by applying information visualization paradigms such as multiple views,

Inffeldgasse 16, 8010 Graz, Austria
[streit | kalkusch | schmalstieg]@icg.tugraz.at

Figure 1. 4 levels of abstraction of metabolic pathways. Root level contains 10 groups, 2nd level 9 groups, 3rd level shows a specific pathway, and 4th level shows a chemical reactions.

linking and brushing, focus and context, and detail-on-demand [4]. The contribution is not a new layout algorithm for non-planar graphs, but a new way to interact with huge graphs using several linked views. We relay on existing layout algorithms and on hand-crafted graphs as used in KEGG [1]. The KEGG layout is used as a background texture and only the nodes are superimposed on top of this texture [5]. This approach does not require to layout the edges while allowing interactive manipulation based on the familiar KEGG representation.

Each of the nodes can be selected and detailed information on the compound or enzyme is presented in the linked web browser (see fig. 2). If a link to an enzyme or a compound is clicked in the web browser the corresponding node in the 2D pathway is highlighted. If another pathway is selected inside the web browser the new pathway is loaded inside the 2D view. Figure 2 shows the CO2–fixation (kegg:map00720) with the compound “L-Malate” selected.

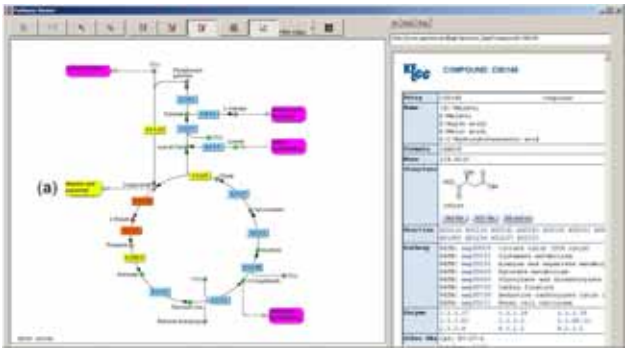

Figure 2. Starting at the selected compound “L-Malate” (shown in red) adjacencies are color coded from orange to yellow. Details on “L-Malate” are presented in the linked web browser.

A neighbor sub-graph - consisting of several nodes with a defined distance from a selected node - gives the viewer important clues on the biological function of the sub-graph. Especially if the breadth search reaches links to other pathways (denoted as (a) in fig. 2), neighborhood provides important meta-information that is otherwise difficult to derive in particular if topologically adjacent

nodes are not in immediate spatial proximity in screen space. The referred pathway can be displayed in other views.

3 Multiple Views Using 2D and 3D

Once the user starts switching pathways it is difficult not to lose the context. Adding two more views of the root level and the intermediate level of the hierarchy and linking them to the current selected pathway helps the user to keep the context [4] (see fig. 3). The current pathway is highlighted in the two other views. We also propose to show several pathways as planes in 3D to be able to present [3] and interact with them concurrently as proposed by [5]. This also enables us to depict multiple occurrences of one enzyme or compound in several pathways. The pathways can either be selected by the user due to special interests of the user or switched and loaded depending on one enzyme or one compound.

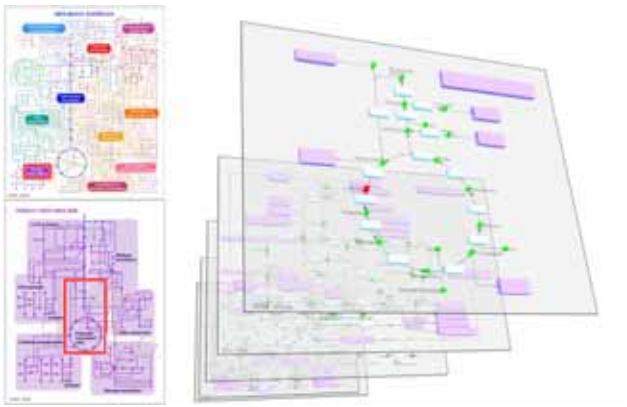

Figure 3. 3D view of 5 pathways (fig.1). Top left image shows root level of the hierarchy. Bottom left image shows middle level of the hierarchy with the currently selected pathway highlighted.

3.1 Interaction with Multiple Views

When working with large 2D pathways as shown in figure 4 we also present an overview map of the 2D pathway. The enzyme MDH (EC1.1.1.37) is selected inside the 2D view. The 3D view shows all other pathways containing this enzyme by connecting the enzymes with yellow lines. If an enzyme is represented several times inside the pathway several lines point to them. For example in figure 4 the second plane in the 3D view contains the enzyme MDH three times.

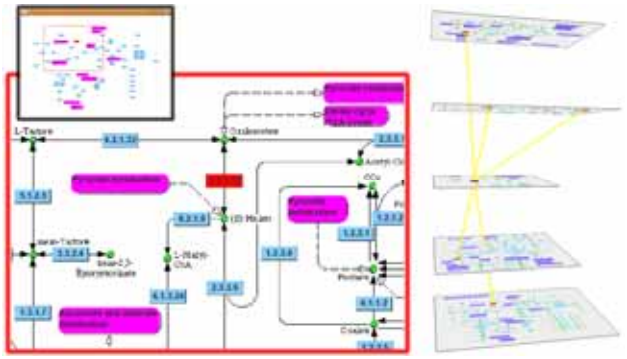

Figure 4. Enlarged 2D graph with overview map on top combined with 3D view showing all 5 pathways containing the selected enzyme “1.1.1.37” connected by yellow lines.

3.2 Multi User Setup

Our approach targets on either a multiple desktop setup or a multi-user setup including a projector for collaborative data analysis. (see fig. 5). This setup consists of a projection wall and 2 LCD screens. The projection wall shows the magnified 2D pathway “Glycine” (kegg:map00260) with the highlighted enzyme EC1.5.99.1, an overview map of the 2D pathway (bottom left), and a 3D view of all 5 pathways containing EC1.5.99.1 (see also fig. 4). The left LCD screen displays the CO2-fixation pathway as a transparent 3D view linked to the intermediate level and to the root level on the right side of the screen (compare to fig. 3). The right LCD screen shows a third pathway (kegg: map00630) connected to the web browser showing EC1.5.3.1.

4 Conclusion

Multiple views enable the user to trace several links of an enzyme or compound visually instead of opening many websites in a row and keeping the threads together in the mind of the user. Our proposed 3D representation in combination with the traditional 2D pathway combined with the web browser enables the user to perceive hidden properties of the complex graphs. Feedback from our medical and biomedical partners was remarkable. A few users with a high degree of functional fixation had problems with the distortion inside the 3D perspective. Orthographic projections were more appealing to them. A flexible combination of several 2D pathway views and 3D pathway views enriched with focus + context overview maps are the most successful layouts. The next step is to integrate the fourth abstraction layer namely the chemical reactions.

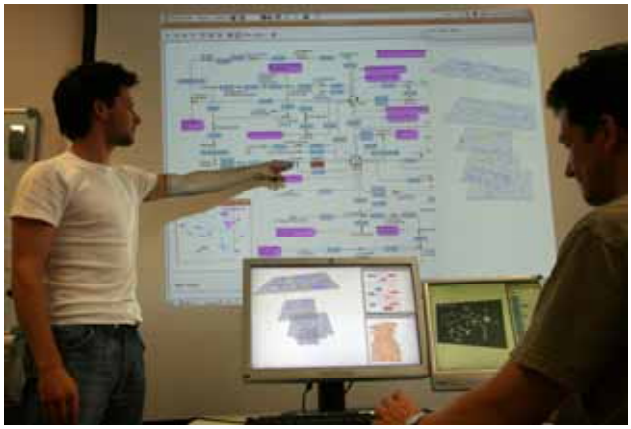

Figure 5. Multi-user setup combining all proposed techniques.

References

[1] Hiroyuki Ogata, Susumu Goto, Kazushige Sato, W. Fujibuchi, H. Bono, and M. Kanehisa. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 28(1): pages 27–30, 2000, http://www.genome.jp/kegg/

[2] GeneView Visualization Framework, http://www.GeneView.org

[3] Ulrik Brandes, Tim Dwyer, and Falk Schreiber. Visualizing related metabolic pathways in two and a half dimensions. In Giuseppe Liotta (editor), *Graph Drawing*, Perugia, pages 111–122, 2004

[4] Robert Kosara, Helwig Hauser, and Donna L. Gresh. An interaction view on information visualization. *Proceedings of EuroGraphics* 2003, pages 123–137, 2003

[5] Marc Streit. Metabolic Pathway Visualization Using Gene-Expression Data. *Master thesis*, Graz University of Technology, Austria, June 2007