IOWA STATE UNIVERSITY
Digital Repository

2010

Network modeling in systems biology

Tian Xia
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Part of the Electrical and Computer Engineering Commons

Recommended Citation

www.manaraa.com

Network modeling in systems biology

by

Tian Xia

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Bioinformatics and Computational Biology

Program of Study Committee:
Julie A. Dickerson, Co-major Professor
Xun Gu, Co-major Professor
Diane C. Bassham
Carl K. Chang
Zhijun Wu

Iowa State University

Ames, Iowa

2010

TABLE OF CONTENTS

ABSTRACT

A key aim of current systems biology research is to understand biology at the system level, to systematically catalogue all molecules and their interactions within a living cell, rather than the characteristics of isolated parts of a cell or organism. Network modeling is characterized by viewing cells in terms of their underlying network structure at many different levels of detail is a cornerstone of systems biology. Two emerging methodologies in network modeling provide invaluable insights into biological systems: static large-scale biological network modeling and dynamic quantitative modeling. Static large-scale biological network modeling focuses on integrating, visualizing and topologically modeling all kinds of omics data sets which are produced by innovative high throughput screening biotechnologies. Dynamic quantitative modeling focuses on exploring dynamics of biological systems by applying computational simulation and mathematical modeling.

To facilitate application of these methods in biological research and improve existing network modeling software, this work presents: i) OmicsViz and OmicsAnalyzer, software tools, dedicated to integrating and analyzing omics data sets in network context. ii) CytoModeler, software tool, dedicated to providing a bridge between static large-scale biological network modeling and dynamic quantitative modeling methods. It not only facilitates network design, model creation, and computational simulation but provides advanced visualization for simulation results. iii) Comparative network modeling application in the systems biology of the SM-SNARE protein regulation in exocytotic membrane fusion. This work presents applications of biological network modeling methods to understand regulation mechanisms in complex biological systems.

CHAPTER 1. INTRODUCTION

1.1 Background

Over a whole last century, classical reduction-based methodologies have dominated biological research, providing a wealth of knowledge about individual cellular components and their functions. However, it is increasingly evident that an isolated biological function can only rarely be attributed to an individual molecule. Instead, most biological characteristics arise from complex interactions between the cell's numerous constituents, such as proteins, DNA and RNA [1]. Therefore, a key challenge for biology in the century is to understand the structure and the dynamics of the complex intercellular web of at system level. Systems biology is a biology-based interdisciplinary study field that concentrates on the systematic studying of the complex interactions in biological systems, using a new perspective (holism instead of reduction). Network biology as one of core constituents in systems biology provides the way to understand biological processes by defining the network structure of the processes based on integration of a variety of experimental data, which allows the expertise from large and well-mapped non-biological systems to be used to characterize the intricate relationships that govern biological functions.

1.2 Network Diagram

The behavior of most complex systems, from the cell to the Internet, emerges from the orchestrated activity of many components that interact with each other through pairwise interactions. At a highly abstract level, the components can be reduced to a series of nodes that are connected to each other by links, with each link representing the interactions between two components. The nodes and links together form a network, or, in more formal

mathematical language, a graph. Biological interactions at many different levels of detail, from the atomic interactions in a folded protein structure to the relationship of organisms in a population or ecosystem, can be modeled as networks. Physical interactions between molecules, such as protein–protein, protein–nucleic-acid and protein–metabolite interactions, can easily be conceptualized using the node-link nomenclature and the union of the interactions form a network which represent the system (Figure 1.1).
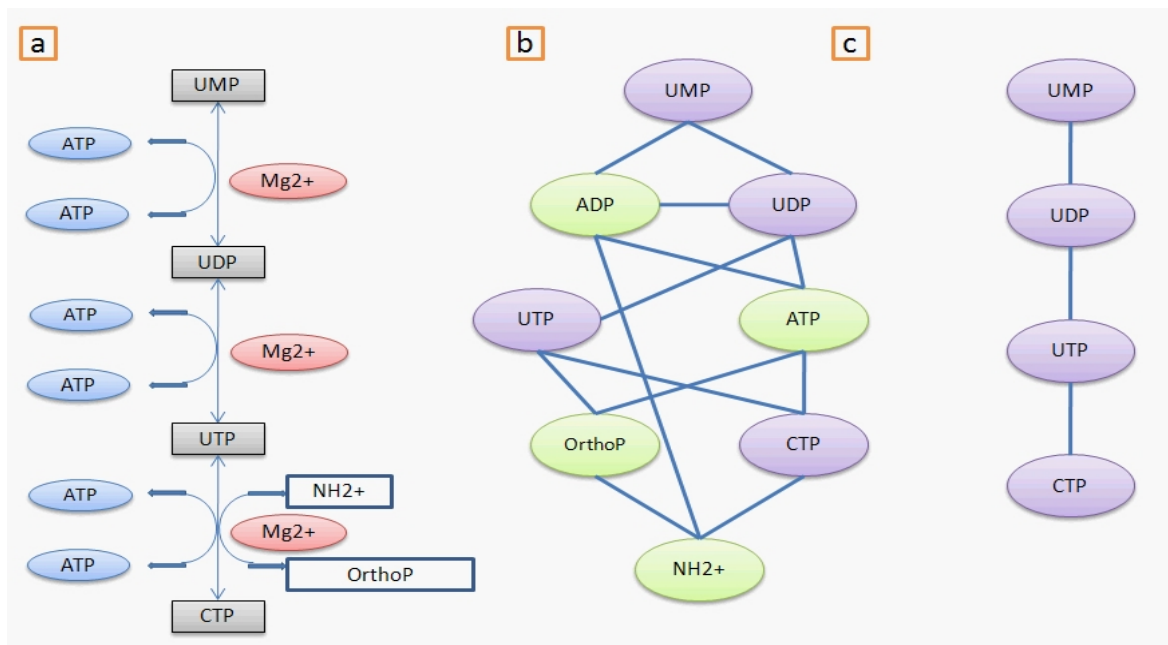
Figure 1.1 An example of biological network. To study the network characteristics of the metabolism a graph theoretic description needs to be established. Here, the graph theoretic description for a simple pathway (catalysed by Mg2+-dependant enzymes) is illustrated (a). In the most abstract approach (b) all interacting metabolites are considered equally. The links between nodes represent reactions that interconvert one substrate into another. For many biological applications it is useful to ignore co-factors, such as the high energy- phosphate donor ATP, which results in a second type of mapping (c) that connects only the main source metabolites to the main products [1].

Nevertheless, more complex functional interactions can also be considered within this representation. For example, researchers use elastic network to provide insights on conformational dynamics and the longer-scale functional behaviors of macromolecules such as protein where nodes corresponding to alpha carbons of the amino acid residue and edges are the spatial interactions between nodes (amino acids) [2].

Depending on the nature of the interactions, networks can be directed or undirected. In directed networks, the interaction between any two nodes has a well-defined direction, which represents, for example, the direction of material flow from a substrate to a product in a metabolic reaction, or the direction of information flow from a transcription factor to the gene that it regulates. In undirected networks, the links do not have an assigned direction. For example, in protein interaction networks, a link represents a mutual binding relationship: if protein A binds to protein B, then protein B also binds to protein A.

1.3   Two major methodologies to model biological systems in network

Having identified network representations of biological systems, researchers apply various approaches to link them with existing knowledge and seek better understanding of biological processes:

1.3.1   Static large-scale biological network modeling

Ever since 40 years ago, biologists have contemplated the idea of modeling biological systems by gene regulatory circuits, and even of genome-scale genetic networks. In network modeling, nodes represent biochemical species and edges represent interactions between the species. Recent progress in high-throughput screening technology in molecular biology [3, 4] enables us to focus on cell's macromolecular components and their interactions such as protein–protein, protein–DNA and protein–RNA at system scale and make insights into the fundamental properties of the collective behavior of a large number of interacting components. These kinds of biotech are able to test interactive relations of tens of thousands of biological compounds simultaneously and therefore to offer much more complete view for biological systems than traditional biological experiment techniques.

Consequently, as a preliminary step for modeling very large scale data in biological networks, new methods are needed. Static large-scale biological network modeling analysis is emerging for this need. It mainly includes the following areas:

(1) Construct/ infer functional biological network/pathway from omics and experimental data. Many studies construct or infer various biological networks, such as gene regulatory network, metabolic network, and protein- protein interaction network which describe certain biological process or species-scale system by using innovative high throughput screening technologies such as microarray and proteomics [3-7]. These constructed networks provide a whole view of target biological systems.

(2) Integrate hetero-omics data across species and data type by network model. In this step, researches try to use constructed network models to integrate all kinds of type experimental data together and then present complete view for target systems at different levels of details [8].

(3) Topologically analyze biological network. In this step, researchers try to connect the topological features of biological networks with biological function, design principles of regulation mechanism and evolution of the systems. For example, studies suggest cellular networks are scale-free and small world [1]. In a scale free network, a few highly connected nodes, dominate the network. The number of nodes with a given degree follows a power law (e.g. in protein-protein interaction network). From another perspective, studies of Hartwell et al. present a biologically relevant definition of modules as discrete units of function separable from the rest of the system [9]. These methods shed light on topologies of the cellular networks and their functional organizations. However, they either limit the analyses in single

organism interaction networks or they have subjectively and simply comparative interpretations solely based on their own topological perspectives.

(4) Comparatively analyze biological network. As with the success of BLAST in comparative genomics, in biology, a strategy of comparison of biological network across species or systems may offer a valuable framework for addressing many biological challenges, such as inference of unknown biological function or elements by comparing network model of well-studied systems with the network model of newly-investigated systems [10]. Recent articles have discussed various approaches to align and score the biological networks based on the graph theory. Restricting the alignment to mutually homologous nodes and constructing an ortholog graph, Koyuturk et al. [11] formulate the alignment of PPIN(protein protein interaction network) into a problem of searching for heavy subgraphs (highly-interconnected subgraph) in an ortholog graph. Pinter et al. [12] analyze metabolic pathways by using subtree homeomorphism, which requires that all input metabolic pathways have linear structures. Kelley and colleagues [13] use sequence alignment algorithm to evaluate similarity between pairwise networks along paths of connected nodes. Further, Sharan et al. [14] explore how to have multiple network comparisons across three species rather than pairwise.

(5) Visualize biological network and analytical results. Visualization of biological data is crucial in helping to understanding complex biological systems [15]. However, with such huge size of data sets with high heterogeneities, visualization of large scale biological networks is challenging. In this area, researchers have proposed a wide range of visualization methods such as 2D, 2.5D and 3D and develop many software tools such as Cytoscape [16]

to facilitate effective visualization of large scale biological networks and their associated analytical results [17].

1.3.2 Dynamic Quantitative modeling

Quantitative modeling in biology can be traced back to early last century when scientists applied mathematical and physical methods to study fundamental mechanisms of biological systems even without concrete biological data [18]. In particular, in the last decade, as the advances in bio-experimental, computational technologies and network modeling, quantitatively modeling biological systems has been boosted: many crucial parameters which were immeasurable before can be read out by innovative wet-lab biotechnologies and the increasing computer performances allow *in silico* simulation experiments in much wider scale and less time cost than before.

Therefore, many formalisms are developed to quantitative model biological systems, such as Boolean networks , Bayesian networks, and their generalizations, partial and ordinary differential equations, stochastic equations, algebraic equations, Monte-Carlo simulation, large-scale agent-based computer simulation and so on [19, 20].

Researchers in this methodology build dynamical model according to existing network structure and investigate system behaviors over time under various conditions and predict complex behaviors in response to complex stimuli, coupling computational and mathematical simulation and various system analysis strategies such as steady, bifurcation, sensitivity and robustness analysis. These rapid and inexpensive *in silico* experiments via dynamical modeling can be used to gain first insights, form hypotheses, and conceive and carry out meaningful tests. Utilizing dynamical modeling for understanding critical parameters,

biologists can technically and statistically design physical experiments for maximum efficacy. Resultant data from all experiments will be compared against simulations in various ways to test assumptions and hypotheses, identify new phenomena, and spark new theories.

1.4   Proposed methods and contributions of this work

This work concentrates on developing effective software tools for network modeling and an application of the network modeling methodology for the exocytosis processes.

First, we present two software tools: OmicsViz [21] and OmicsAnalyzer [22]. These tools are dedicated to providing an effective omics data integration, visualization and analysis across different species and data sets in network environment.

Second, we focus on offering biologists a user-friendly software tool to facilitate network designing, network analysis and quantitative modeling in biology. We present a software tool, CytoModeler, an open-source java application based on Cytoscape platform [16]. It provides an efficient network design interface, integrates many network analysis methods provided by Cytoscape and it plug-ins, offers a convenient built-in simulator based on ISBJava library [23] and the ability to connect to third-party sophisticated quantitative modeling software, and advanced visualization for simulation results in network context.

Third, the work combines comparative network analysis and quantitative modeling to analyze regulation mechanism of SNARE proteins and Sec1/Munc18 (SM) proteins in exocytosis.

SNARE proteins and Sec1/Munc18 (SM) proteins are two essential protein families in vesicular fusion. SNARE proteins execute the fusion process by forming a four-helical bundle complex (SNAREpin). However, the general understanding of the interactions

between SM proteins remain uncertain partly due to disparate SM-SNARE binding modes that have been observed in different organisms and secretory pathways

We built two SM-SNARE protein interaction networks (SSN) in exocytotic pathways, using substantial experimental data about the SM and SNARE protein binding mode for a non-regulated cascade-like network in yeast and a regulated feedback-loop-like neuronal synaptic network. Comparative analysis of the networks showed that different combinations of SM-SNARE binding modes may lead to divergence of regulatory function in the SM-SNARE network across species or cell types. Dynamical analysis and simulation demonstrates different system behaviors in the two SM-SNARE networks, which may explain contradictory experimental observations for the adaptation to physiological variations across the secretory pathways.

These findings are consistent with current experimental observations. This suggests that the divergence of SM-SNARE protein binding modes may lead to reconfiguration of the regulation mechanism of SM-SNARE system in different organisms or tissues. In particular, these results provide a useful approach to understand the differences among SM-SNARE experimental observations and generalize SM proteins function to two basic control mechanisms.

1.5   Organization

In Chapter 2, we introduce two software tools: OmicsViz and OmicsAnalyzer. In Chapter 3, we introduce the CytoModeler software tool, including: Architecture of CytoModeler, CytoModeler model editor, CytoModeler SBML exchanger, CytoModeler simulation interface, CytoModeler visualizer. In Chapter 4, we introduce work using

comparative network design in different secretory pathways to explain the SM-SNARE regulation mechanism, including: 1. Construction of two SM-SNARE networks. 2. Network comparison analysis to extract the crucial distinct between two SM-SNARE networks. 3. *In silico* experiments reveal differential system behaviors of SM regulation. In Chapter 5, we made conclusion for this work.

CHAPTER 2. OMICSVIZ AND OMICSANALYZER FOR INTEGRATING AND ANALYZING OMICS DATA IN NETWORK CONTEXT

2.1   Network provides graceful and concrete modeling platform to integrate complex large-scale biological datasets.

Nowadays, biological projects and experiments become much more complex and bigger in scope and data that are produced are magnitudes higher than in the past. The increasing use of innovative high-throughput technologies multiplies the amount of data generated per experiments, such as genomics, proteomics and transcriptomics [24]. The underlying data sets show a growing complexity and dynamics and are produced by numerous heterogeneous application areas. The integration of heterogeneous types of data is therefore gaining in importance: it should help to reduce the problems caused by false positives and false negatives obtained from single approaches, lead to better functional annotations for gene products and functional relationships between them, and allow the formulation of increasingly relevant biological hypotheses. Computational methods can then be used to model biological processes based on integrated data.

Networks, as specified by graphs in mathematics, can represent biological interactions and integration of biological information in the form of extensive networks consisting of vertices, denoting nodes of individual bio-entities, and edges, describing connections between vertices. In the simplest example two vertices are linked by only one relationship, but connections can express different types of relationships between two elements, such as an evolutionary relationship, the existence of a shared protein domain, the fact that they belong to the same protein family or that two genes that are co-expressed in an experiment.

An interactive visual representation of information together with data analysis techniques on top of network platform equip biologists with powerful tools to simplify the interpretation and integration of dataset and analyze the dataset, extract meaningful information and use them to answer some of the fundamental biological questions.

The development of high-level languages and platform neutral binaries, such as Java (Sun Microsystems Inc., 1994), plus the Internet as medium for dissemination has dramatically accelerated the transition time from inception to prototype to working application. Other new advances in computational technology include software to implement distributed object technology which provides easier access to remote databases, platform-neutral programming languages allowing "write once, use anywhere" applications and the development of libraries of reusable object-oriented software components, or widgets, from which developers may create customized applications.

Based upon these advances, a wide variety of tools was developed over the past years that map data on 2D networks to visualize biological interactions or relationships between bio-entities, such as Medusa [25], Cytoscape [16], Osprey [26], BioLayout Express3D [27], ProViz [28],PATIKA [29], PIVOT [30] and so on.

These techniques of visualization have roles not only in analysis, but also in building more user-friendly interfaces, implementing method to navigate large information spaces intuitively and powerful techniques to browse and query data. However, the challenges still remain due to complexity of biological dataset and interaction between the dataset and biologists. For example, integrating multiple independent datasets and mapping such datasets across species and gaining biological insight are challenging. And there is a big gap between analysis and visualization. Most existing visualization tools only incorporate a limited

number of data analysis functionalities, making it necessary to constantly switch between different applications.

Facing these challenges, we developed two plug-in software based on Cytoscape network visualization and analysis platform: OmicsViz [21] and OmicsAnalyzer [22] to improve the data integration and visualization and statistics functionalities of Cytoscape software platform.

2.2   Introduction to Cytoscape

Cytoscape is a standalone Java application. It is an open source project under LGPL license. Cytoscape mainly provides 2D representations and is suitable for large-scale network analysis with hundredth thousands of nodes and edges. It can support directed, undirected and weighted graphs and comes with powerful visual styles that allow the user to change the properties of nodes or edges. The tool provides a variety of layout algorithms including cyclic and spring-embedded layouts. Furthermore, expression data can be mapped as node color, label, border thickness, or border color.

Cytoscape comes with various data parsers or filters that make it compatible with other tools. The file formats that are supported to save or load the graphs are SIF, GML, XGMML, BioPAX, PSI-MI, SBML, and OBO. Cytoscape also allows the user to import mRNA expression profiles, gene functional annotations from the Gene Ontology (GO) and KEGG database. Users can also directly import GO Terms and annotations from OBO and Gene Association files.

It is highly interactive and the user can zoom in or out and browse the network (Figure 2.1). The status of the network as well as the edge or node properties can be saved and

reloaded. In addition, Cytoscape comes with a network manager to easily organize multiple networks. The user can have many different panels that hold the status of the network at different time points which makes it an efficient tool to compare networks between each other. It also comes with efficient network filtering capabilities. Users can select subsets of nodes and/or interactions and search for active sub networks or pathway modules. It incorporates statistical analysis of the network and it makes it easy to cluster or detect highly interconnected regions.
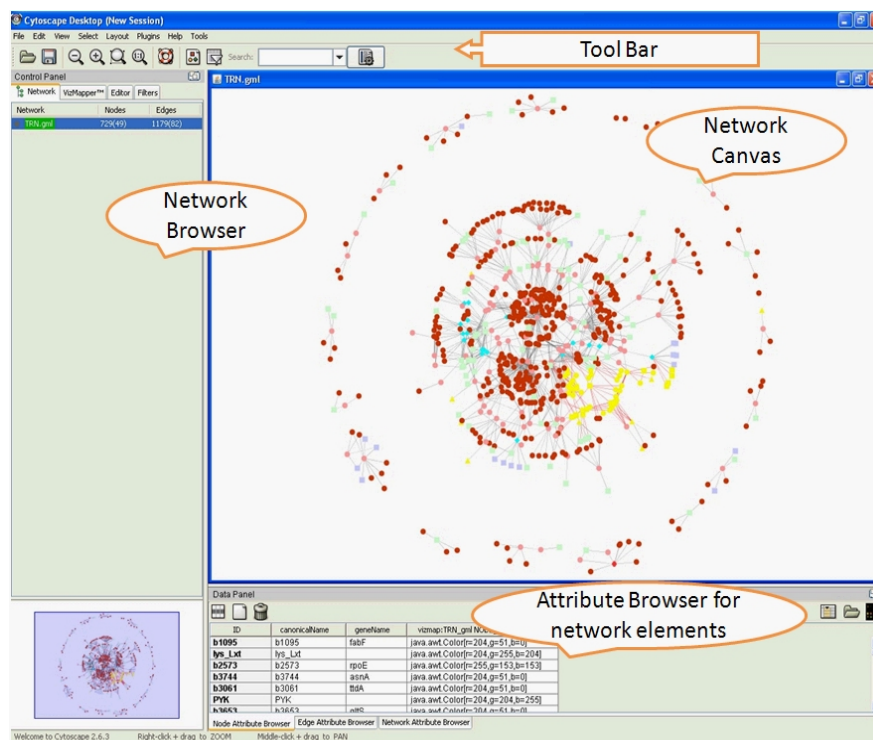
Figure 2.1 The Cytoscape Desktop. The Cytoscape canvas displays network data. The toolbar (top) contains the command buttons. The name of each command button is shown when the mouse pointer hovers over it. The Control Panel (left) displays the Network tree viewer, which lists the available networks by name and size. The Network Overview Pane (lower left) shows the current network in white and the displayed portion in blue. Finally, the Data Panel (lower right) can be used to display node, edge and network attribute data. The Cytoscape Desktop shows the sample network and expression data, with nodes colored by expression values from lowest (green) to highest (red) [31].

Cytoscape also provides a friendly extensible interface which helps the user to easily develop plug-in software to allow more specialized analysis of networks and molecular profiles.

2.3 OmicsViz: Cytoscape plug-in for visualizing omics data across species

OmicsViz is a Cytoscape plug-in for mapping and visualizing large-scale omics datasets across species, including those with many-to-many mappings between homologs. This allows users to map their data onto pathways of related model organisms. Mapping schemas across species or different experimental protocols allow users to comparatively analyze the omics data. The data can also be visualized in parallel-coordinate plots.

Node attributes in Cytoscape characterize elements with specific name IDs and diverse biological information, such as subcellular location, node type, etc. Edges reflect the interactions or other relations between the nodes. OmicsViz associates experimental datasets from transcriptomic, proteomic and metabolomic datasets with biological network models in Cytoscape according to node attribute information.

2.3.1 Omics data import and mapping

OmicsViz maps labeled measurements in the input experimental data to target nodes if the names are identical. However, due to inconsistencies in nomenclature, different names may be used to describe experimental quantities and network nodes, which actually refer to the same biological component in one species. One common example is that microarray datasets use gene probe IDs but metabolic pathways use gene locus names. OmicsViz provides a mapping file which translates between locus IDs and probe IDs to load transcriptomic data into the network model.

A unique feature of OmicsViz is that it deals with problems caused by many-to-many mappings between nodes due to homologs or uncertainty as to which node was matched or which compartment it occurs in. The matched homologs are represented by newly created nodes and connected to target nodes. Functionally, they are grouped into HomologGroups (implemented using CyGroups) with the target nodes. The cross-species transcriptomic data, therefore, can be loaded into target species network. Combining further visualization and comparative analysis, one can easily explore biological insights for both species based on prior knowledge and omics experimental observations. The imported data can be stored as a Cytoscape pvals file.

The mapping rules between the labels in the data file, *X*, and the node names in the active graph, *Y*, $F(X) \to Y$ where $X = \{x_1, ..., x_N\}$ and $Y = \{y_1, ..., y_M\}$ can be summarized as:

If $F(x_i)$ maps to a unique value of y_j (i.e., a one-to-one mapping), then the data associated with x_i is mapped as attributes of node, y_j.

If $F(x_i)$ maps to multiple values of *Y*, then the data is mapped to the set of nodes, $\{y \mid y = F(x_i)\}$.

If multiple values of *X,* defined as the set, X_i, map to a set of nodes in the graph, i.e., $\bigcup_{x \in X_i} \{y \mid y = F(x)\}$, then additional "homolog" nodes and edges connected to these nodes are added to the graph. The attributes of the newly-created nodes are the same as the parent node.

OmicsViz uses data files in formats ranging from comma separated value files to Excel spreadsheets to take advantage of data from PLEXdb [32, 33], GEO, and other microarray databases. OmicsViz can map multiple datasets to networks at the same time. This helps

users to compare results across datasets. The plug-in dynamically updates the CyNode

attribute panel and lists all selected datasets.

2.3.2 Data Visualization

OmicsViz provides line chart visualization component for loaded omics datasets using

the JFreeChart library [34] (Figure 2.2). The experimental data associated with nodes in the

network can be displayed on a parallel coordinate plot. The parallel coordinates of line chart

delineate the variability of omics data with respect to different treatments and time points.

The plot window can be controlled by users by adjusting parameters, such as line colors, axis

ranges and labels
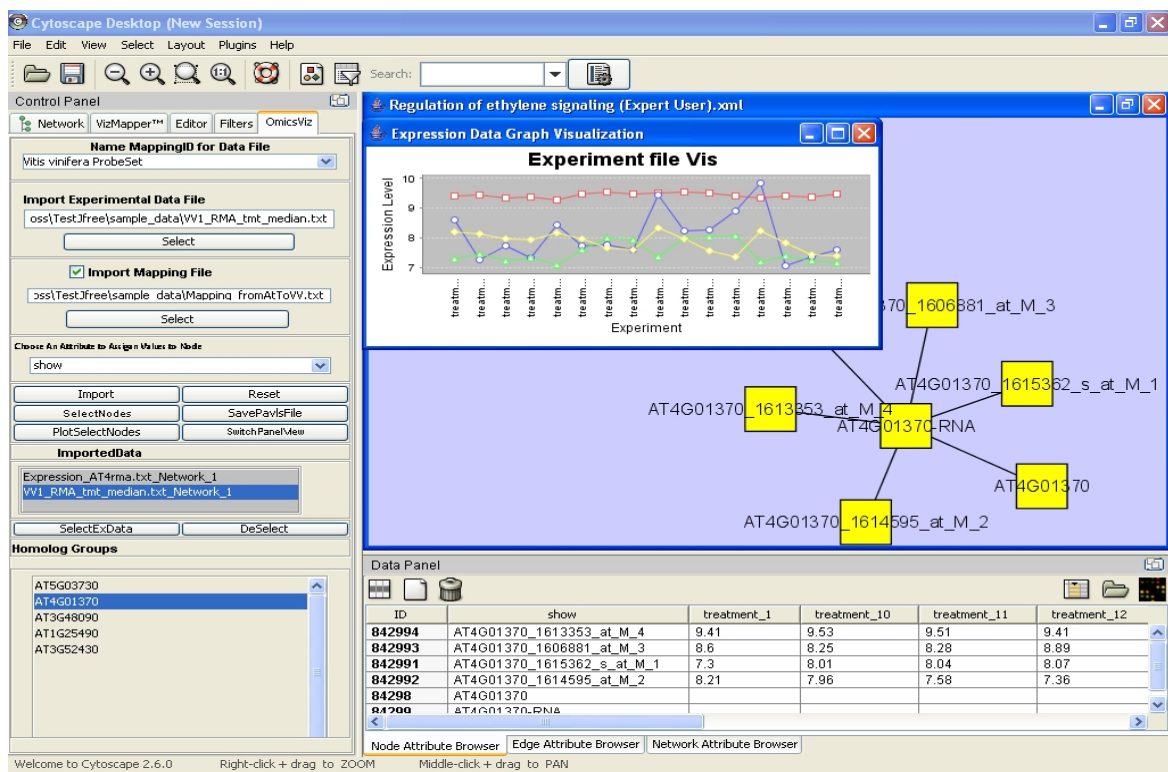
Figure 2.2 OmicsViz plugin control panel screenshot. The highlighted yellow nodes show mapped Vitis homologs onto an Arabidopsis gene locus. The parallel coordinate plot of the expression data is shown in the upper right corner [21].

2.3.3   Usage example

OmicsViz is also able to import transcriptomic data into Cytoscape network platform across species to facilitate further comparative survey. For example, the Arabidopsis genome has been completely sequenced and the annotation is well developed. Grape (Vitis vinifera), has only recently been sequenced and most genes are electronically inferred. The reciprocal better BLAST hits method [35] can produce whole-genome scale homolog mappings between species. This forms a basis for the mapping of Omics data from a less-studied organism to the network model developed for the plant model organism, Arabidopsis. The better BLAST hit method gives a set of possible homologs between species. This example uses TBLASTN for each Arabidopsis protein sequence against the Vitis consensus sequences on the Affymetrix Vitis GeneChip to find the top five grape hits and BLASTX to find the top five hits from Arabidopsis for each Vitis sequence. If Arabidopsis gene A and Vitis probeset B are reciprocal hits, A and B then are grouped as a possible homolog pair. This gives a many-to-many mapping relationship. Mappings for many plant species can be found in the sample data files available with the OmicsViz plugin. The microarray data is from experiment VV2 in PLEXdb (www.plexdb.org). OmicsViz also manages multiple datasets which are associated with specific networks. It enables users to compare datasets of interest and dynamically update the CyNode attribute panel and switch between corresponding datasets.

2.4   OmicsAnalyzer: a Cytoscape plug-in suite for modeling omics data

The widespread use of omics-related biotechnologies has led to many heterogeneous omics datasets. Software packages facilitate the interpretation of omics data by using graph models where nodes represent bio-elements and edges are bio-interactions between these

elements [36]. However, the complexity of biological systems and omics datasets require an integrative analytical environment, which allows [24] network manipulation at different levels of detail, flexible mappings between multiple omics datasets coupled with data visualization and providing statistical analysis and system simulation and visualization of resulting analysis.

Cytoscape [16], an open-source platform soft-ware dedicated to network modeling in biology, provides a highly flexible environment for combining these elements. To meet the above requirements, OmicsAnalyzer provides (i) integrates hetero-omics data within species or across species, (ii) performs statistical analysis correlation for exploring relations between the datasets, (iii) and visualizes the omics information and resulting analysis in a network context.

2.4.1   Architecture of OmicsAnalyzer

The key features of OmicsAnalyzer are presented in three levels (Figure 2.3): data collection and pre-processing level, statistical level and visualization level which function both independently as stand-alone tools and also communicate each other to facilitate the systems level analysis.
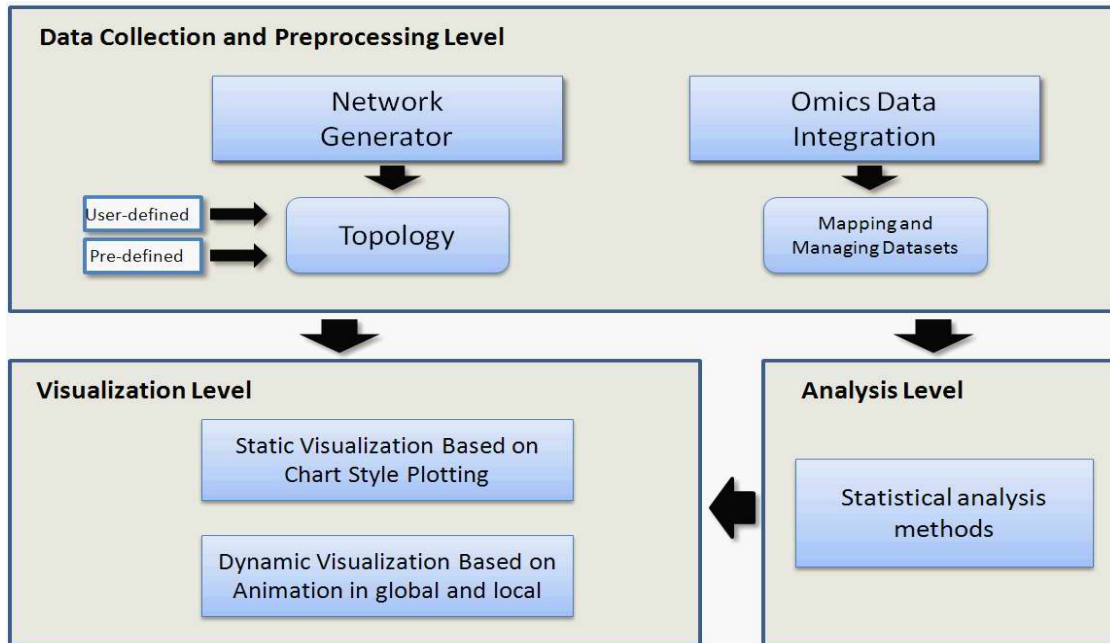
Figure 2.3 Architecture of OmicsAnalyzer

2.4.2 Data collection and pre-processing

In the first level, data collection and pre-processing, there are two modules: network generator and omics data integration. The network generator enables the user to define subnetworks of interest with respect to topology at different levels of detail such as cycle, dense module or network motif, instead of limiting user to the entire network which is imported from outside databases. OmicsAnalyzer manipulates the network by communicating with the Cytoscape platform or third-part plug-ins. Specific functionality includes (i) selection of arbitrary subnetworks using a graphical editor provided by Cytoscape, (ii) defining special-interest networks based on network motifs, cycles, modules, shortest paths determined by existing Cytoscape plug-ins, such as SubgraphCreator. The Omics data integration module is based on our previously developed plug-in, OmicsViz[21]. This module integrates homo or hetero omics datasets with networks, maps omics datasets across species and also manages the imported datasets. The integration allows users to

compare omics data across related networks to gain insights into biological dynamics and function.

2.4.3   OmicsAnalyzer Statistical Analysis

OmicsAnalyzer offers a variety of statistical functionalities for data analysis. At first, relationships or common patterns in the data can be found by calculation of correlation. OmicsAnalyzer provides three approaches to calculate correlation including Pearson, Spearman's rank and Kendall correlation coefficient. The Pearson correlation coefficient is visualized with a color coded diagram frame (Figure 2.4). Alternatively, Spearman's rank order correlation coefficient may be used, which is more robust against outliers. For the interactive analysis when a reference node is selected and the correlation with all remaining nodes is visualized by different node colors.
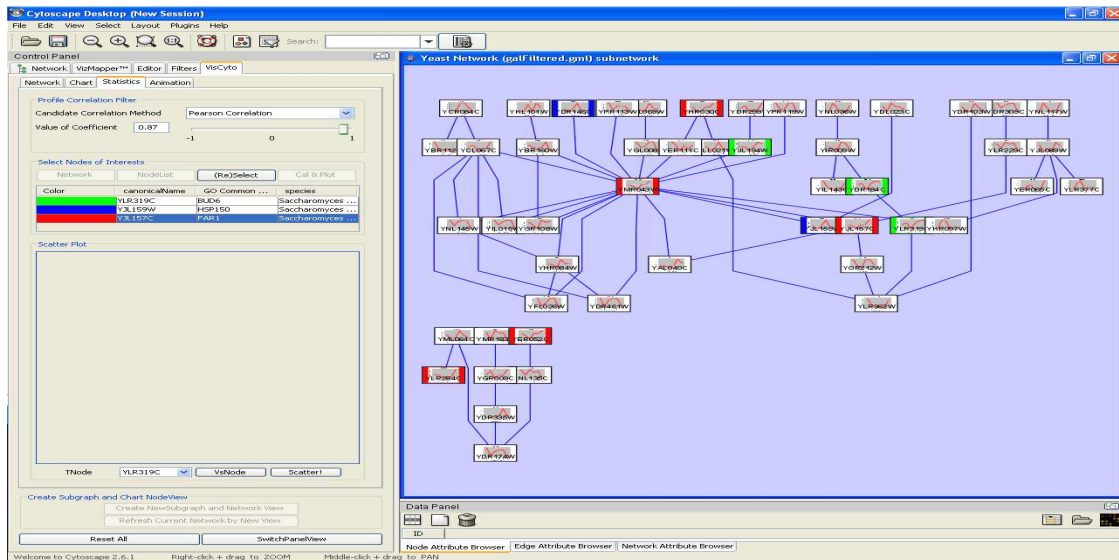
Figure 2.4 OmicsAnalyzer statistical panel provides the correlation calculation and cluster functions. The three nodes (highlighted by green, blue, red) are user-selecting key nodes. OmicsAnalyzer then calculates correlation relationship between the three key nodes and other nodes in the network. If the resulting correlation coefficient between two nodes is beyond the threshold value, OmicsAnalyzer categorizes them in the same cluster and highlighted with corresponding color of key node. The picture shows three resulting clusters.

To create a correlation network from a number of selected substance nodes, the correlation between all possible pairs of nodes can be calculated, Subsequently all nodes are assigned to clusters, based on the best matching SOM node. As a result the substances are grouped according to similar patterns. These clusters can be color-coded (Figure 2.4). Similar patterns are easily discovered visually even if they are widely spread over the picture.

Relationships or common patterns in the data can also be found by plotting the measurement values for a defined set of substances inside a scatter plot matrix (Figure 2.5). This matrix displays the measurement values for all combinations of the selected substance nodes.
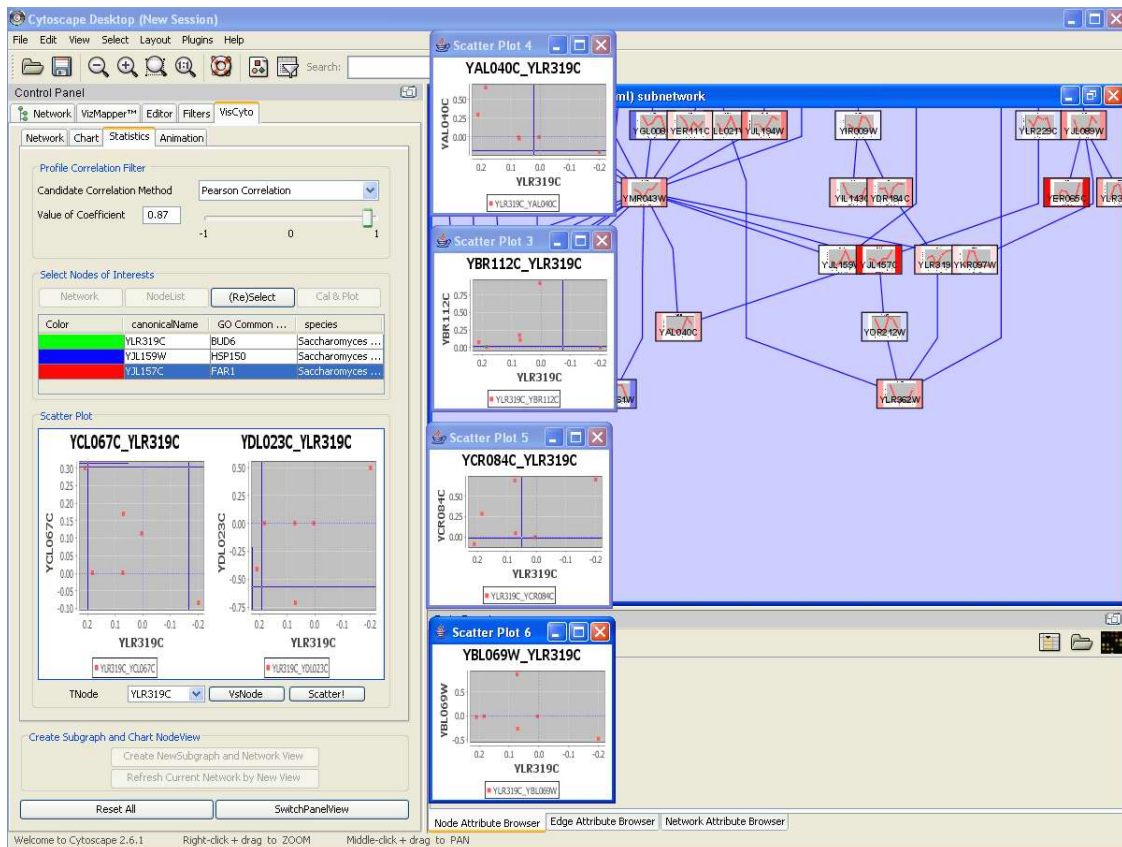
Figure 2.5 OmicsAnalyzer offers scatter plotting function to show the correlation relationship between pairwise experimental data.

2.4.4   OmicsAnalyzer Visualization

OmicsAnalyzer provides an innovative approach to visualizing experimental data in network context. Experimental data associated to specific biological compounds such as gene and protein are firstly plotted in chart-style pictures (line, bar and pie charts provided) and the pictures are then rendered onto nodes in network view which correspond these biological compound (Figure 2.6). Experimental data of different genotypes or plants may be shown within a single diagram inside each substance node, or shown in separate diagrams. The drawing style of the diagrams may be modified with a number of parameters such as series colors, the display of range or category labels, and line widths. As the system supports replicate measurement values in the data input form, the standard deviation (SD) or the standard error of the mean (SEM) may be shown as an error bar in both kinds of diagrams.

Figure 2.6 OmicsAnalyzer shows experimental data onto to the node in network context with multiple chart styles.

2.4.4.1 OmicsAnalyzers Animation

Animation supports efficient and timely analysis of large amounts of multi-dimensional data. OmicsAnalyzer Animation highlights dynamic changes over experimental conditions (time, treatment, dosage, tissue, disease) and allows the user to observe important sub pathways and key genes that influence other parts of the pathway. There two animation modes in OmicsAnalyzer: global and local. In global mode, OmicsAnalyzer plays the animation in the context of the whole network for each node and each edge by creating color-mapping for the network components according to their associated experimental data. In local mode, user can select the specific targets from the whole network such as gene clusters, protein complex and sub pathways of interests, and show the animation within individual panel with chart style (Figure 2.7-Figure 2.9). In the panel, the user can easily zoom in and out to examine the dynamics of experimental data in details.
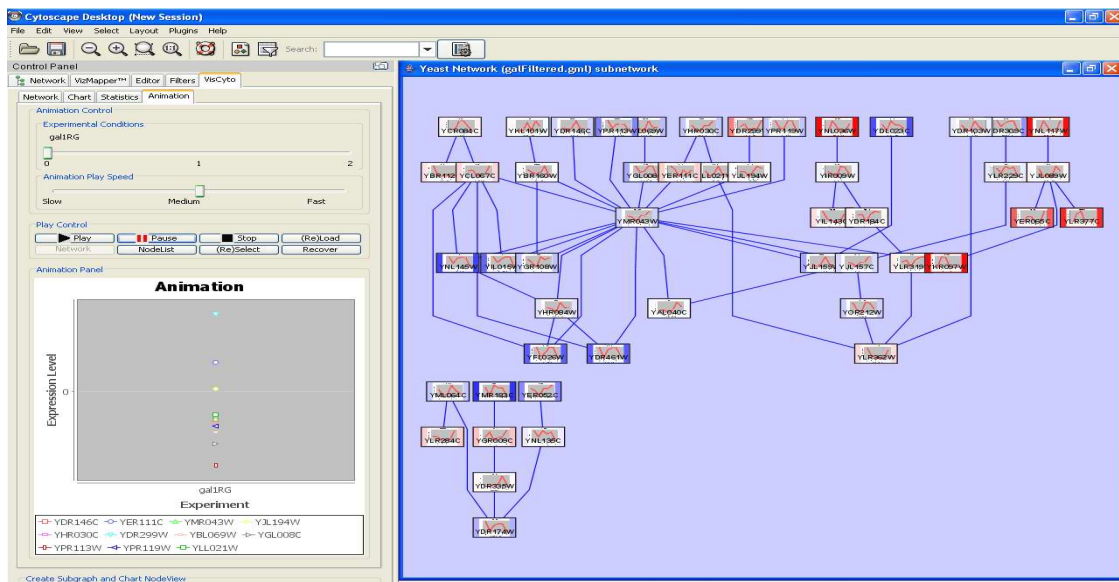
Figure 2.7 OmicsAnalyzer provides animation functions in two ways: global and local. In Cytoscape main panel, OmicsAnalyzer animate the whole network according experimental value at each condition. A color mapping is created to map experimental value to specific color. In this picture, blue corresponds MIN experimental value, -1 and red corresponds MAX experimental value, 1. The colors representing other experimental values continuously transfer between blue and red. In left corner panel, OmicsAnalyzer offers a local animation view for specific nodes which mostly attract user's interest, which are picked by user. This the first frame of the animation.
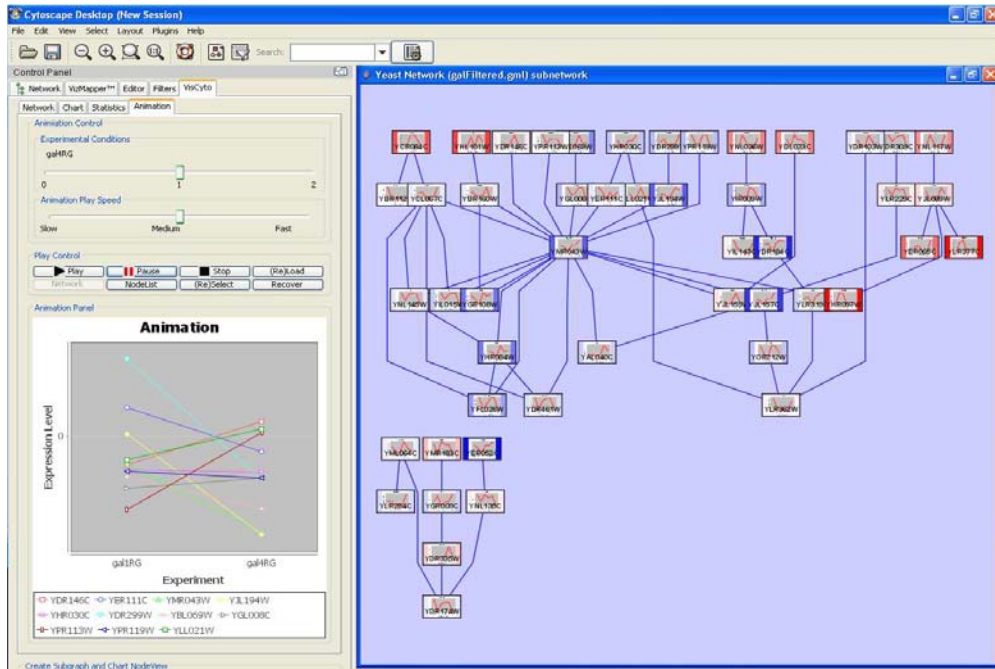
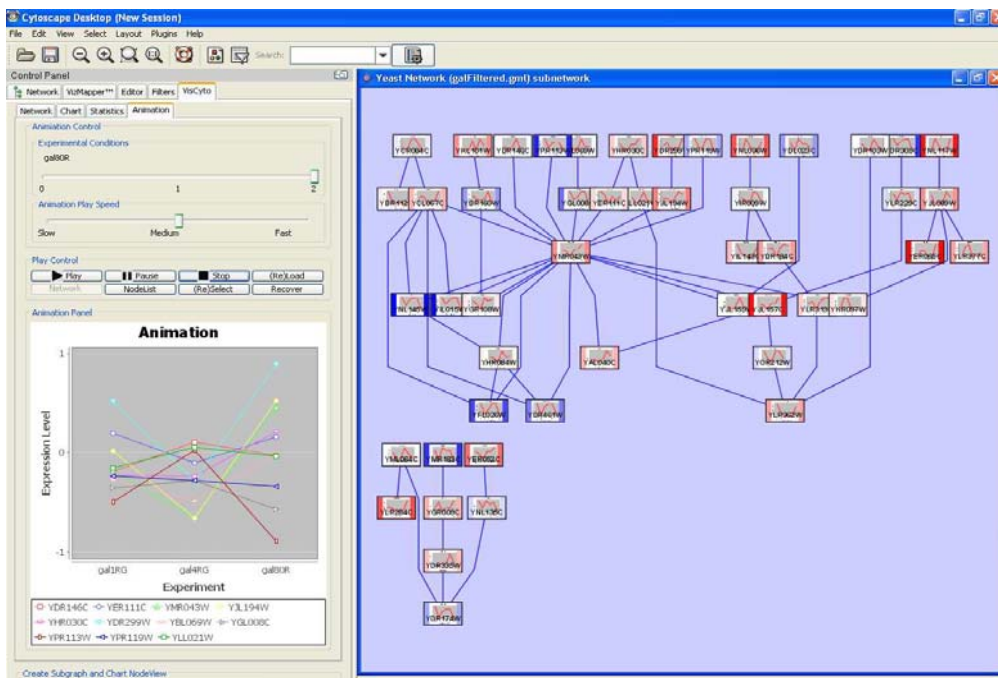**Figure 2.8 The second frame of the animation.**

Figure 2.9 The third frame of the animation.

CHAPTER 3. CYTOMODELER, A TOOL FOR BRIDGING LARGE SCALE NETWORK ANALYSIS AND DYNAMIC QUANTITATIVE MODELING

Recent emergences of innovative biotechnologies support the large scale biological network analysis and quantitative modeling. However, connection of the two methodologies is still difficult. Here, we present CytoModeler, an open-source Java application based on the Cytoscape platform. It creates a bridge between large scale network analysis and quantitative modeling by fusing the functionality of Omics analysis of Cytoscape platform, various computational simulators and network context visualizations.

3.1   Introduction

In last decade, emergent biotechnologies, such as high-throughput screening have produced a huge amount of large-scale biological datasets. These data potentially provide a description of many biological processes and functions at a system level. Therefore, dedicated software tools that enable biologists to explore these new data and connect them with biology systems are needed [24]. In systems biology, two kinds of software tools are extensively used to model and analyze biological processes at system level: large scale network analysis and quantitative differential equation modeling.

Large scale network analysis applies network to model large scale omics data, where nodes represent biological species and edges between nodes stand for reaction between the species [1, 37, 38]. First, diverse types of omics data can be reconstructed as networks or integrated into existing networks, where the nodes and edges own a variety of attributes, such as species name, cellular location, molecular type, organisms and so on, which describe the network components. Furthermore, based on the networks, numerous analytical methods, ranging from statistics to machine learning, are developed to extract hidden information and

to translate the omics data to enhance our understanding of biological systems as whole. To facilitate such complicated analyses of large scale networks, it is necessary to develop efficient and user-friendly tools. Significant progress has been made in development of software which focus on large scale network and omics data analysis [16, 36, 39-42] over last ten years. Cytoscape, being a popular open source software platform among these tools, serves as a community-based framework for systems biology modeling. It provides not only a rich functionality for large scale network visualization and for viewing and editing data, but also a highly flexible and extensible plug-in mechanism [43]. To date, over a hundred Cytoscape plug-ins are developed, which can be categorized as: analysis, network and attribute I/O, network inference, functional enrichment, communication/scripting plug-ins [44]. Through these functional plug-ins, users can do much more than viewing a network but can compare different networks [45], cluster networks into subnetworks or other statistical analyses according to the contained omics data [46-48], examine Gene Ontology (GO) terms for a set of genes [49], align protein sequences [50], and view structures of protein [51] .

The development of quantitative modeling has a much longer history than large-scale network analysis. It can be traced back to the last century when theoretical models were developed to study complex biological system such as metabolic analysis [52-54]. In this method, researchers build mathematical equations for biological systems and simulate, predict and test behaviors of the systems *in silico*. Furthermore, through analysis of the steady states, robustness/sensitivity of the equations, researchers try to interpret how the systems are regulated and how they operate. In last decade, advances in bio-experimental and computational technologies have boosted the quantitative modeling of biological systems. Many crucial parameters which were immeasurable before can be read out by innovative

wet-lab biotechnologies. Therefore, many software tools have been developed for quantitatively creating and analyzing biological models [55]. For example, the Systems Biology Toolbox [56] is a commercial software tool based on MATLAB; COPASI [57] provides predefined rate laws functions and also stochastic and deterministic simulation algorithms; Systems Biology Workbench (SBW) [58] offers a tool suite which include many functional modules ranging from JDesigner for graphical model editing, Jarnac for command line based simulation [59], to a Bifurcation Discovery Tool. Several standard formats are also proposed for facilitating exchange of modeling information such as systems biology markup language (SBML) [60].

While a number of active software tools of these two methods grow fast, the communication between them is missing. It is necessary that information of biological systems inferred from large scale network analysis are further quantitatively modeled for examining dynamics of the systems with respect to various physiological conditions or time course, which is crucial to understand function of the systems and guiding design of next step wet-lab experiments. Therefore, this work developed CytoModeler, a software tool dedicated to bridging large scale network analysis and quantitative modeling by providing the following features: i) Access to a wide range of functionality for large scale network analysis by being a plug-in tool on Cytoscape platform, ii) A built-in compact quantitative model editor which enables model creating, importing and exporting with several formats including SBML level1 and 2, iii) A modeling exchanger through which users can transfer modeling information between Cytoscape and quantitative modeling tools such as Systems Biology Toolbox (SBT) on MATLAB [56], COPASI [57] and Systems Biology Workbench (SBW) [58], iv) Advanced visualization for quantitative simulation results in a network context with

both static and dynamic point views. Therefore, CytoModeler enables biologists to integrate large scale network analysis and quantitative modeling into an analytical workflow of systems biology.

The rest of the paper is organized as follows: we first introduce the architecture of CytoModeler. Second, we demonstrate functionality of the built-in model editor. Third, we present a built-in CytoModeler model simulator, communication with third part simulators, and visualization of simulation results. Finally, we present a case study about how to use CytoModeler to model exocytosis pathway.

3.1.1 Architecture of CytoModeler

To serving a bridge between large scale analysis and quantitative modeling, CytoModeler is built on Java programming language and Cytoscape platform and connects to software tools dedicated to quantitative modeling in systems biology. Figure 3.1 shows an architecture diagram of CytoModeler.

The architecture of CytoModeler is categorized into four functional components (Figure 3.1). The first component, CytoModeler Model Editor, is a built-in equation editor. It offers concise model editing function by using CyNodes, CyNetworks and HyperEdges (CyNode: a basic element in Cytoscape to represent node in biological network; CyNetwork, a basic element in Cytoscape to represent biological network; HyperEdge, a special type of edge which is able to represent biochemical reaction.) of the Cytoscape platform. It also provides a model transformer which enables communication with other Cytoscape function plug-ins. Furthermore, it allows users to import and export the edited model as several formats such as SBML [60]. The CytoModeler Simulation Interface gives users multiple choices for

simulating the model. It is based on an improved version of ISBJava [23]. This simulator

offers many numerical solvers including stochastic algorithm (detailed in Simulation section).

If a user needs more sophisticated systems analysis such as bifurcation analysis,

CytoModeler Simulation Interface can connect with the modeling information with other

dedicated analytical tools such as SBT on MATLAB, SBW and COPASI to facilitate the

further system analysis using Java interface and SBML [60] standard file format. The

architecture is extensible so it can be integrated with more third part systems biology tool.

The third component, CytoModeler SBML Exchanger, serves as communication center for

exchanging modeling information between Cytoscape, SBT on MATLAB, SBW and

COPASI using SBML standard format. The fourth component, CytoModeler Visualizer

provides user advanced visualization for simulation in network context. There are a various

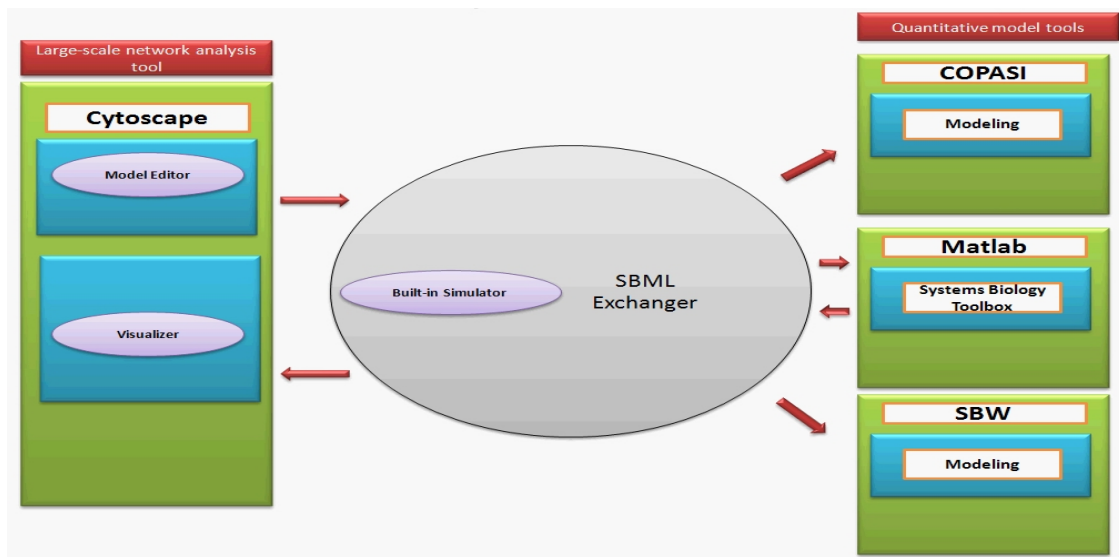visualization modes for user including both static and dynamic point of views.

Figure 3.1 Architecture of CytoModeler

3.1.2 CytoModeler Model Editor

The CytoModeler Model Editor (CMME) is a plug-in component. CMME has a graphical editing function for kinetic equation that uses the Cytoscape graphical interface framework and network element classes which allows access to all the functionality for large scale network analysis provided by Cytoscape itself. CMME also uses the Cytoscape node attribute panel for a graphical representation of kinetic model and also a form-based editor.

The graphical style kinetic model editing interface in CMME is based on basic elements of Cytoscape network (Figure 3.2 left). The editor uses CyNodes (Figure 3.2 left) to represent biochemical species such as protein, gene, and ligands in all kinds of biological kinetic models. To represent reactions between the species, the editor uses HyperEdges, a particular Cytoscape edge which consists of a group of CyEdges and a unique CyNode -- Connector node (Figure 3.2 left) and is able to represent a variety of reaction types. With a set of CyNodes and HyperEdges, a CyNetwork shows a biological kinetic model in the Cytoscape canvas where user can freely manipulate the model elements using a mouse.
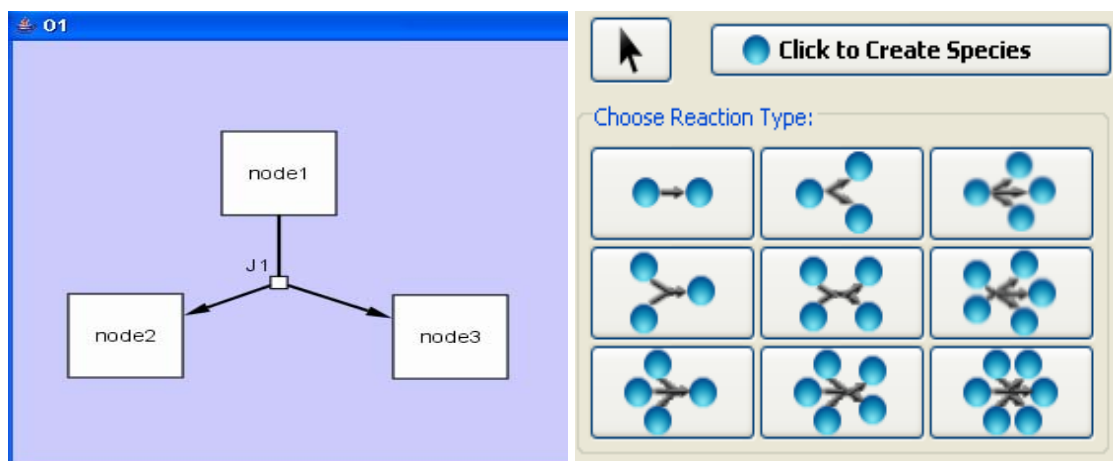

Figure 3.2 CyNode, HyperEdge, CyNetwork and Reaction palette. Three CyNodes and one HyperEdge in a CyNetwork 01 construct a biochemical reaction J1, which has an identifier connector labeled by J1(left) and Reaction palette. The palette provides 9 different type reactions (right) .

To help user easily create different reactions, CytoModeler Model editor provides a reaction palette (Figure 3.2 right) where there are 9 different reaction types such as one to one reaction, and triple to triple reaction. User can add the reactions to a set of species nodes by predefined order.

Once a reaction is fully created, user can input relate model parameters, such as the initial population of species, rate constants, reaction stoichiometry and so on for the reaction and its reactants and products through reaction and species property editors (Figure 3.3) which can be invoked by double clicking the edges, connector nodes or species nodes.

User can also edit kinetic models through the Cytoscape data panel (Figure 3.4). Cytoscape use node, edge, and network attribute to integrate biological information into network model. The Cytoscape data panel permits browsing the attributes in form-based styles. A row stands for an entity such as a node, edge or network. Every column of the row presents all attribute values associated with the entity. CytoModeler uses CyNode (species node) to represent biological species and a HyperEdge to represent a biochemical reaction. In a HyperEdge, there is a unique CyNode, called Connector which identifies the HyperEdge and defines the reaction. Therefore, all kinetic model information can be stored in attributes of all the CyNodes including nodes for species and connector nodes for reactions. Once a user opens the Cytoscape data panel in node attribute tab, they can view all the information and also input information to the current model such as initial concentration and reaction rate parameters by selecting the form cells of corresponding model elements. At the same time, the action will accordingly bring highlights of these edited model elements in a graphical network. Therefore, it provides an interactive feeling between model editing and the network context.
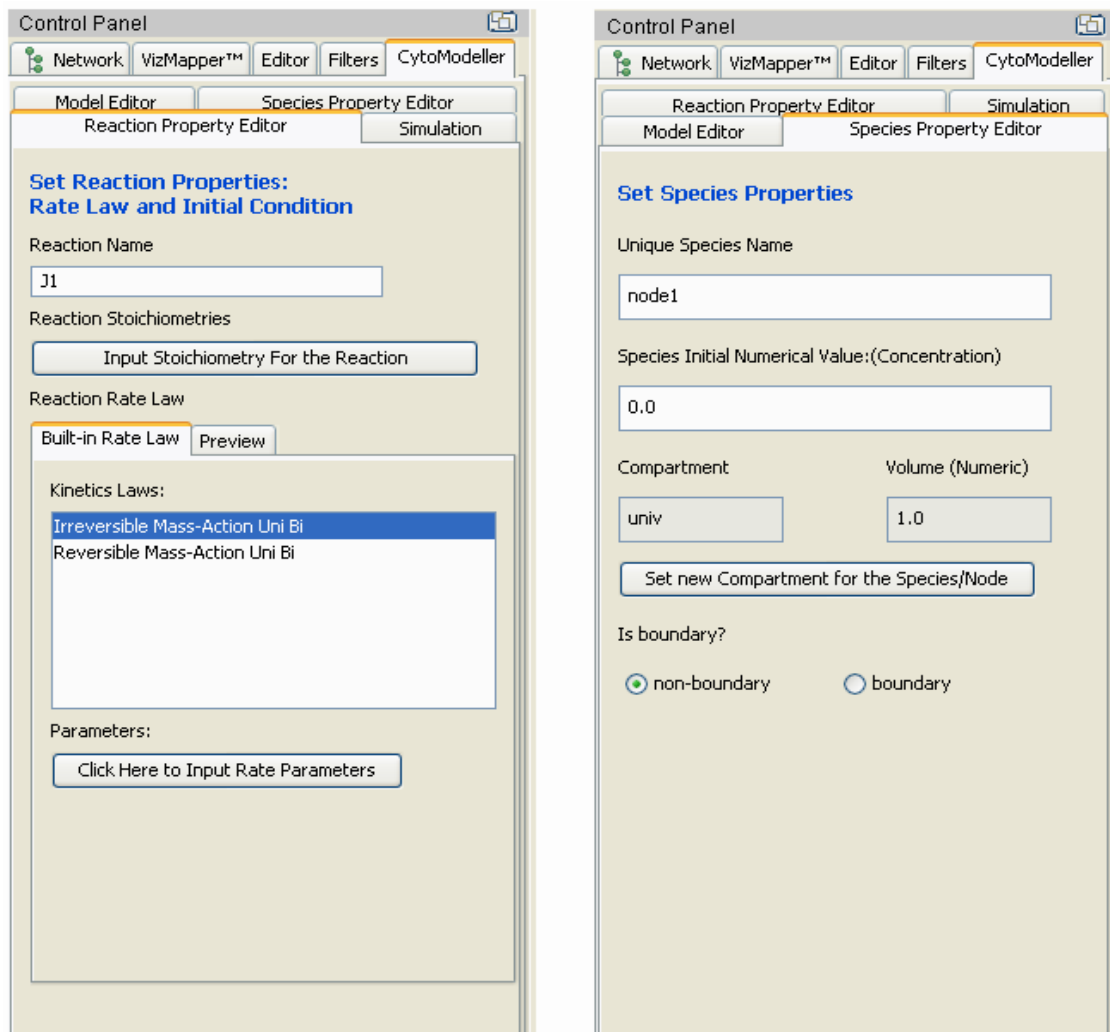
Figure 3.3 Reaction and species property editors. The left is reaction property editor and the right is species property editors. These editors provides various interface to allow user input model associated parameters.

Figure 3.4 CytoModeler form-based editor

In addition to the direct editing function, The CytoModeler Model Editor also provides a model transformer which can transform normal CyNetwork into CMME model. It has a mechanism simple and easy to use: Through click the "Create Model from Current Network" button, user can transform nodes and edges from other CyNetworks into CytoModeler model network's species and reactions. By this, it provide a way to make CyModeler use of functions of Cytoscape large scale network analysis. For example, by using Cytoscape, a user can construct a gene regulatory network from a microarray dataset and further group the network into many smaller network clusters according to gene expression profiles. When a user is interested in certain network cluster and therefore wants to further build kinetic model for the clusters to explore the dynamics of these gene interactions, the model transformer at this point gives user an easy and efficient way to transfer all the genes in the network cluster into a kinetic model by just clicking the button. In another case, a user can apply a comparative network analysis for a large biological network on Cytoscape through functional plug-ins netMatch [45] and NetworkAnalyzer [61]. The result of the analysis reveals an interesting pathway hidden in the large network. The user wants to further simulate the pathway to its behavior under different experimental conditions. The CytoModeler Model editor now can fulfill this task easily by running the model transformer.

3.1.3   CytoModeler SBML Exchanger

CytoModeler SMBL Exchanger is a functional component responsible for exchanging kinetic modeling information such as model import, output and exchange between CytoModeler functional components, Cytoscape platform and other quantitative modeling software tools by using systems biology markup language (SBML) format. Due to the

diverse architecture of different software tools, a common standard file format is crucial. We used SBML format.

SBML is an eXtensible Markup Language (XML) which records mathematical expression information using of MathML [62]. SBML represents a model using hierarchical set of components which describe all kinds of biological and mathematical information at different level of details for the model such as species, reactions, compartments where reactions occur, function definitions for a mathematical functions which are used in defining reaction rate laws, parameters in the rate laws, initial assignments for model initial conditions and so on (refer to www.sbml.org for details). There are currently three levels of SBML. Compared to level 1, SBML level 2 integrate more components to describe model. Since SBML is widely accepted by systems biology community, it enables reuse of model across different software platforms and easy publication of model in the community.

CytoModeler Exchanger is based on the SBML level 1 and 2 formats. It enables export and import of a model from SBML format and transforms current Cytoscape network into SBML format and allows data exchange across platforms.

3.1.4   CytoModeler Simulation Interface

To facilitate the analysis of dynamical behaviors of biological systems, the CytoModeler simulation interface provides two user-friendly methods for system simulation or further sophisticated analyses.
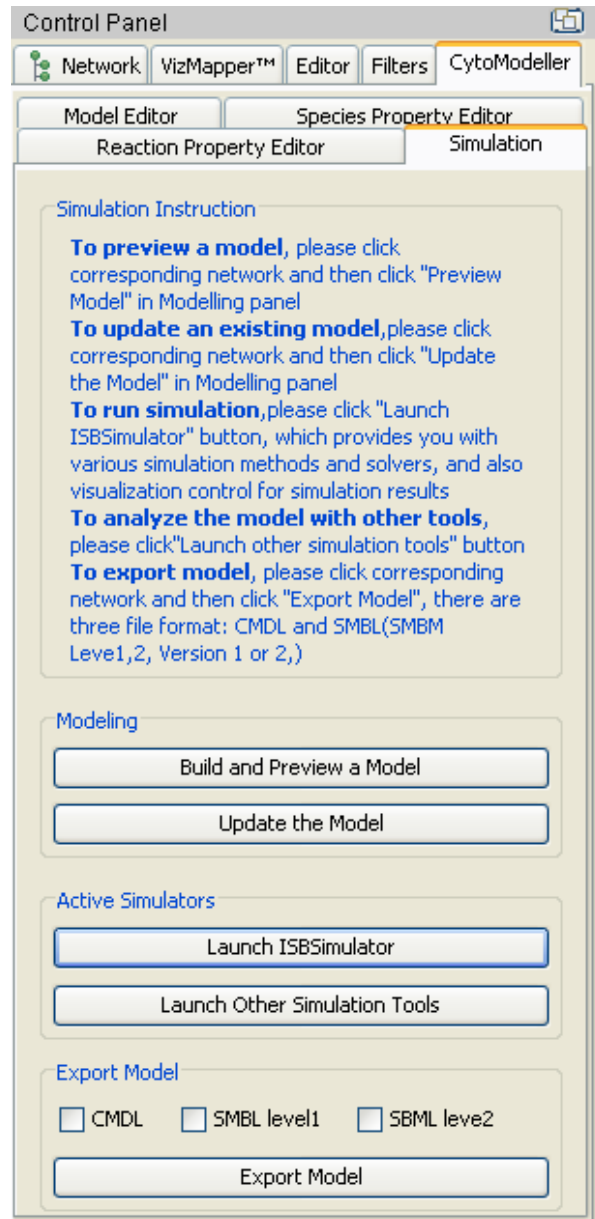
Figure 3.5 Simulation component interface. It provides options for users to apply simulation or further systems analysis.

The first option is a simulation component (Figure 3.5) which provides model preview

and update and various simulators which are based on ISBJava library [23] with both

deterministic and stochastic algorithms, including stochastic simulators: Gibson-Bruck,

Gillespie, Tauleap-complex and Tauleap-simple and deterministic simulators: ODE- 5th –

order-Runge-Kutta-adaptive, ODEtoJava-dopr54-adaptive and ODEtoJava-imex443-stiff.

After completion of model construction , users can go to the "Simulation" tab, preview the

whole model input by them in human-readable format, update the model by just clicking on

buttons in the tab, and then directly launch the built-in simulation interface (Figure 3.6) (This

interface is implanted and improved from ISBJava library [23] by providing more functions

such as advanced visualization control for simulation results). The interface enables the user

to access many control elements to manipulate simulation such as different simulators listed

above, time starting point and ending point, and step size. After setting all the control

parameters, user can run the simulations and the result will be shown in a traditional X-Y

plot other advanced visualization options in the simulation interface menu you can select

(detailed in CytoModeler Visualizer section).
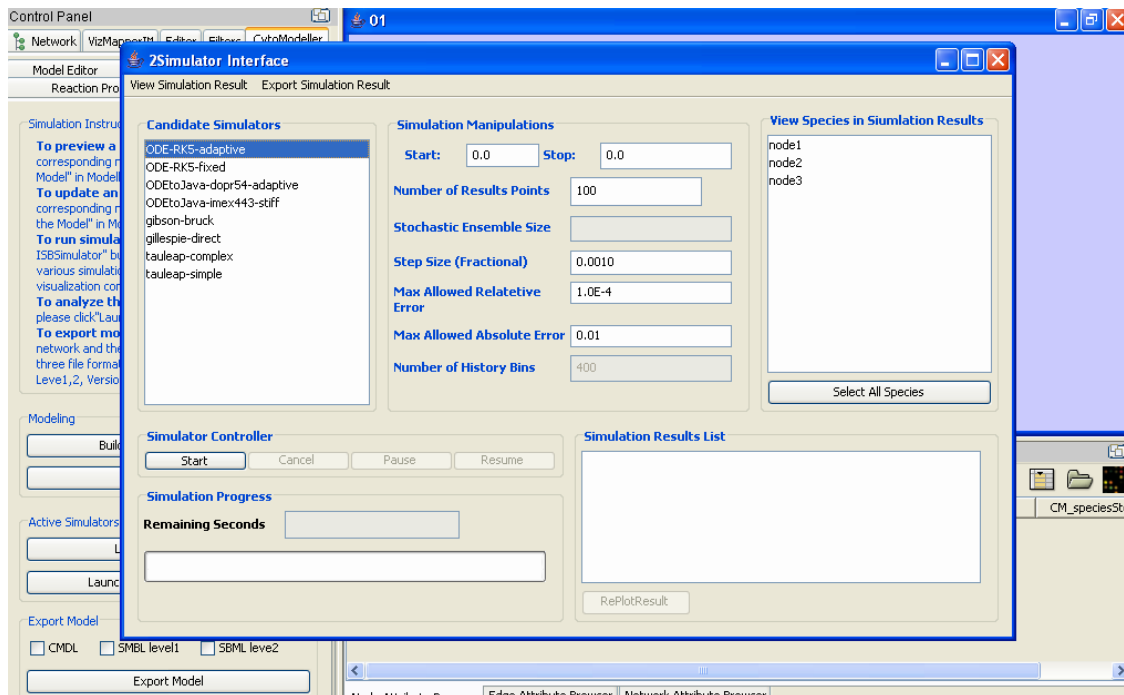
Figure 3.6 A built-in simulator interface. The simulator is based on ISBJava library[23] and improved with many new features.

The second option for simulation is to connect the third-party software tools with built model information, which may provide not only simulation but also more sophisticate functionality for systems analysis. These tools might include parameter sensitivity, systems robustness or bifurcation which is valuable for providing insights into regulation mechanism of biological system. CytoModeler uses SBML Exchanger to seamlessly communicate with three quantitative modeling software tools: Systems Biology Toolbox (SBT) on MATLAB, Systems Biology Workbench (SBW) and COPASI.

Systems Biology Toolbox (SBT) is a tool built on the MATLAB platform. CytoModeler only work with Systems Biology Toolbox version2. It provides a wide range of functionality for quantitative analysis. These functions are represented in MATLAB script command and function programming style. For example, "cm = SBmodel()" command will create a new blank model object with name cm. If user is familiar with MATLAB, they can easy use SBT2. In particular, SBT v2 allows high speed simulation using CVODEs integrator package from SUNDIALS, which is stated the 30-150 times faster than ODE15s simulation solver and therefore quite useful for large-scale network model simulation. The CytoModeler simulation interface connects SBT with SBML format modeling information edited by CytoModeler editor when user request launch SBT tool. The SBML-based modeling information is automatically packaged into a SBT model object and then user can use the object to perform any analysis provide by SBT2.

Systems Biology Workbench (SBW), compared to SBT, is more interactive quantitative model tool. It is actually a software suite which consists of many functional modules: JDesigner for graphical model creation, Network Object Model (NOM) for SBML model manipulation, Jarnac [59] for simulation, Oscill8 for bifurcation analysis and so on. These

modules are interconnected by peer to peer architecture and through any of them user can invoke other modules (SBW also can interconnect with CellDesigner [63], another popular modeling tool supporting Systems Biology Graphical Notation (SBGN) [64] format). CytoModeler simulation interface enables users to send created model to SBW modules and therefore provides a functional fusion between Cytoscape, a large scale network analysis tool and these quantitative modeling modules.

Similarly, CytoModeler simulation interface interacts with COPASI [57] with a tree-based representation for model elements, which is quite useful for systems with large size. Because CytoModeler simulation interface is designed in an extensible JAVA API, it easily implements further integration for more quantitative modeling software with Cytoscape for functional enrichment.

3.1.5   CytoModeler Visualizer

The CytoModeler visualizer enables advanced visualization for simulation result. Traditional visualization of simulation result usually consists of static X-Yplots of time versus concentrate or phase plane plots. As model sizes increase, the traditional X-Y plot gets cluttered and confusing and also lacks interaction with the model itself. To improve the traditional visualization, CytoModeler visualizer connects plotting of simulation result with network elements of model in Cytoscape interface. It enables an interactive visualization environment for user and simulation results and facilitates understanding of systems dynamical behaviors.

There are two options for advanced visualization in the network context: static and dynamic. In static visualization mode, user can individually plot X-Y chart for simulation

result of each species in its corresponding CyNode view (it is similar with VANTED style

[36]) (Figure 3.7) or plot the X-Y chart in the Cytoscape data panel (Figure 3.8). In this

choice, the X-Y chart for visualization of simulation result and network model canvas are

shown in the same window. When user clicks any of the series for one biochemical species in

the X-Y chart, it is highlighted and the node representing the species will be also highlighted

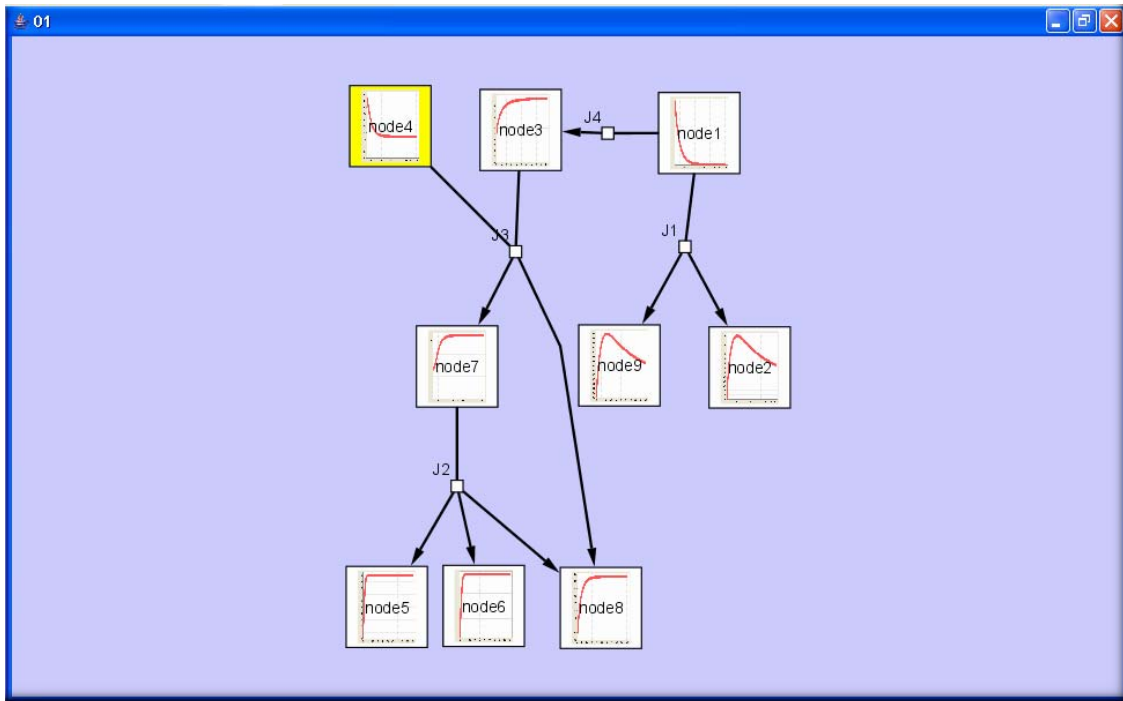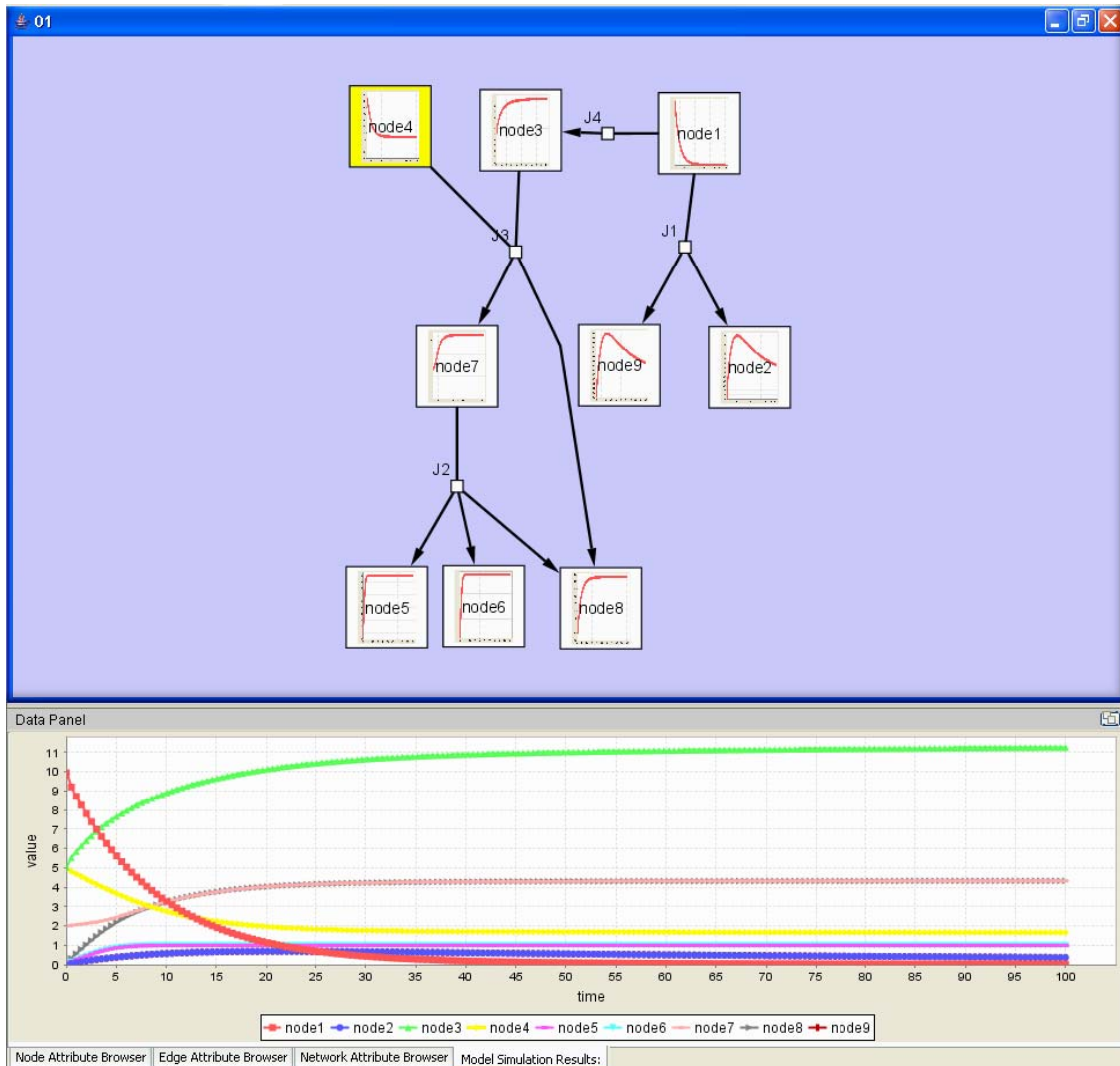(Figure 3.8). It therefore enables an interactive visualization mode for user.

Figure 3.7 Static visualization of simulation result in network context

Figure 3.8 Static X-Y plotting visualization of simulation result in Cytoscape data panel. It provides an interactive way between simulation result and the network context. In the figure, one series in the X-Y plot in Data Panel is selected, the corresponding node in network canvas is also highlighted.

In the dynamic visualization mode, CytoModeler visualizer uses animation to dynamically visualize simulation results in network element views. It offers user multiple choices: i) dynamically plotting X-Y chart for simulation result value of each species at each time point in its corresponding CyNode view (Figure 3.8), ii) dynamically cascading view (Figure 3.9). In this choice, numeric values of simulation result of a biochemical species are represented as areas with specific color. Different values are represented by different sizes of

areas. The whole network is shown as dynamical cascade and it enables clear views for how the concentration/population of every biochemical species changes, iii) shaking view. This choice provides more features for dynamical visualizing simulation result. It first calculates derivative of every series and then horizontally shakes CyNodes whose derivative of corresponding series is positive and vertically shakes CyNodes whose derivative of corresponding series is negative. Amplitude of the shaking is measured by numeric values of the derivatives. And the CytoModeler visualizer also uses two colors to identify different changing trends for each biochemical species. One color stands for increase of concentration/population of a biochemical species, another for decrease. For all the dynamic visualization methods, CytoModeler visualizer implements them in extensible Java and Cytoscape API and it can be easily extend for any new dynamic visualization methods supported by Cytoscape visual mechanisms.
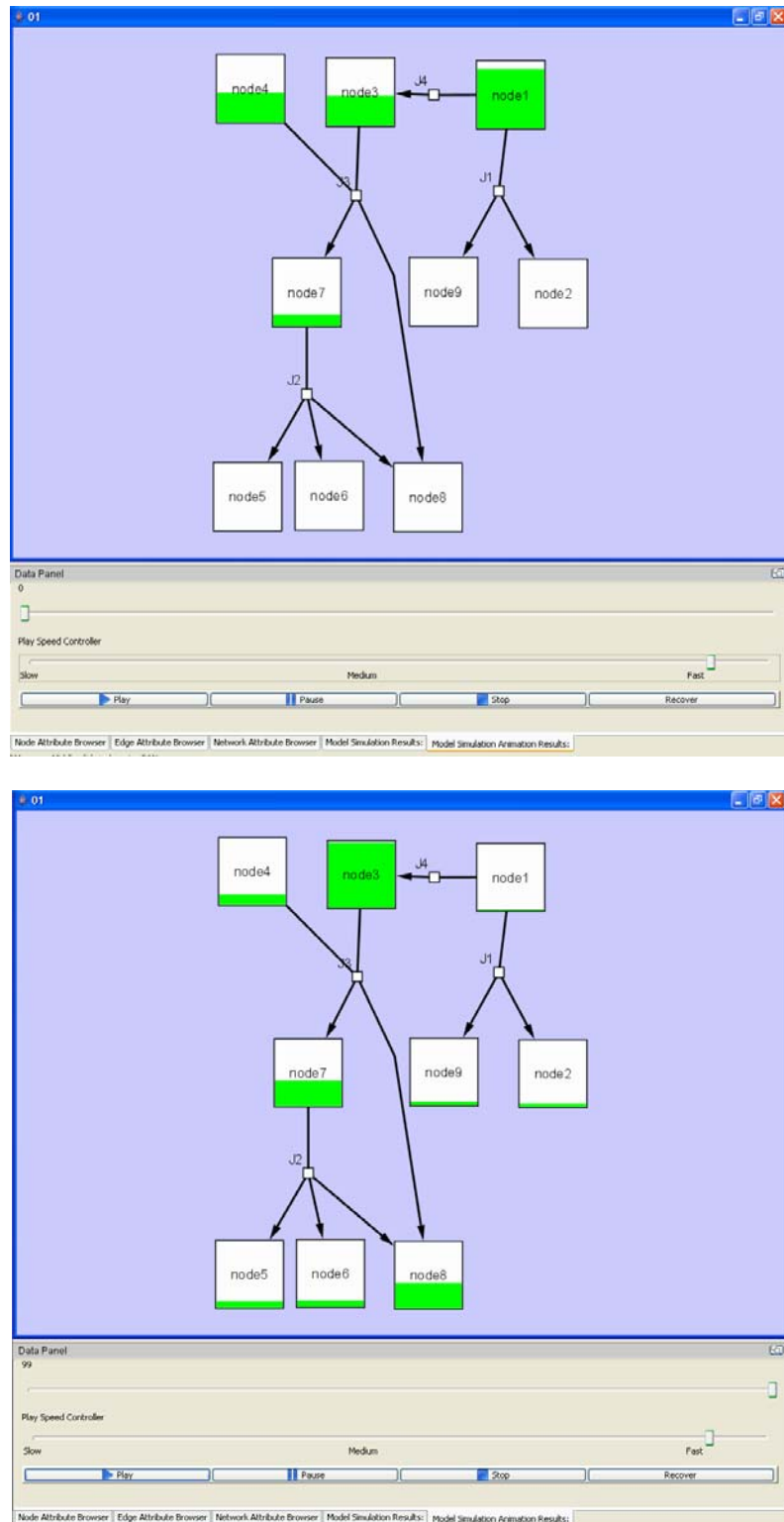
42

Figure 3.9 Dynamic visualization of simulation results in cascading mode. There shows two time frames: 0, 99 for the whole model.

3.2 An Usage Example: Build SM-SNARE Model Using CytoModeler

Next, we show a usage example for CytoModeler by applying our previous SM-SNARE regulation model. We first used CytoModeler editor to create two network models for yeast and neuronal SM-SNARE regulation pathways (Figure 3.10 and Figure 3.11). All the proteins are presented in CyNode and all reactions between them are represented in HyperEdges. After inputting all parameter, the two models were simulated in the built-in simulator. The results are shown in static X-Y plot in Cytoscape Data Panel (Figure 3.8) and network context (Figure 3.13) and dynamically shown in cascading animation with comparison for two models' simulation results (Figure 3.14).
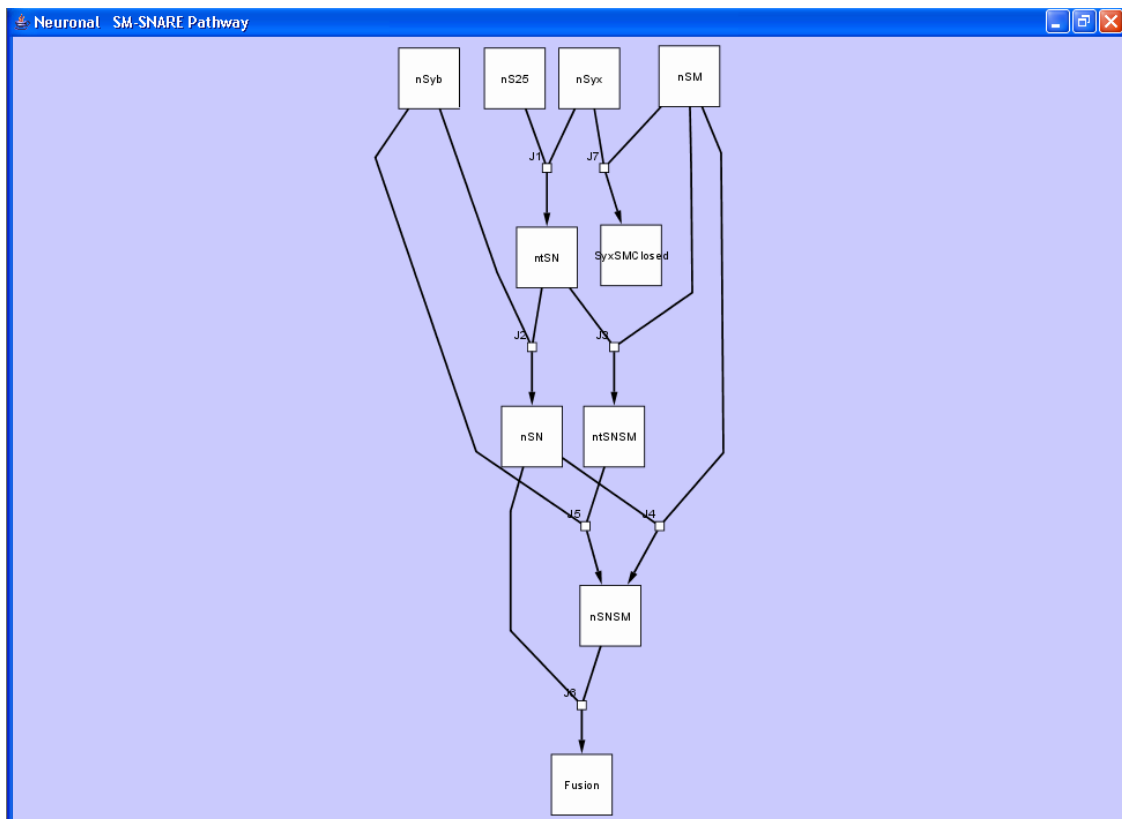
Figure 3.10 Neuronal SM-SNARE regulation pathway in CytoModeler
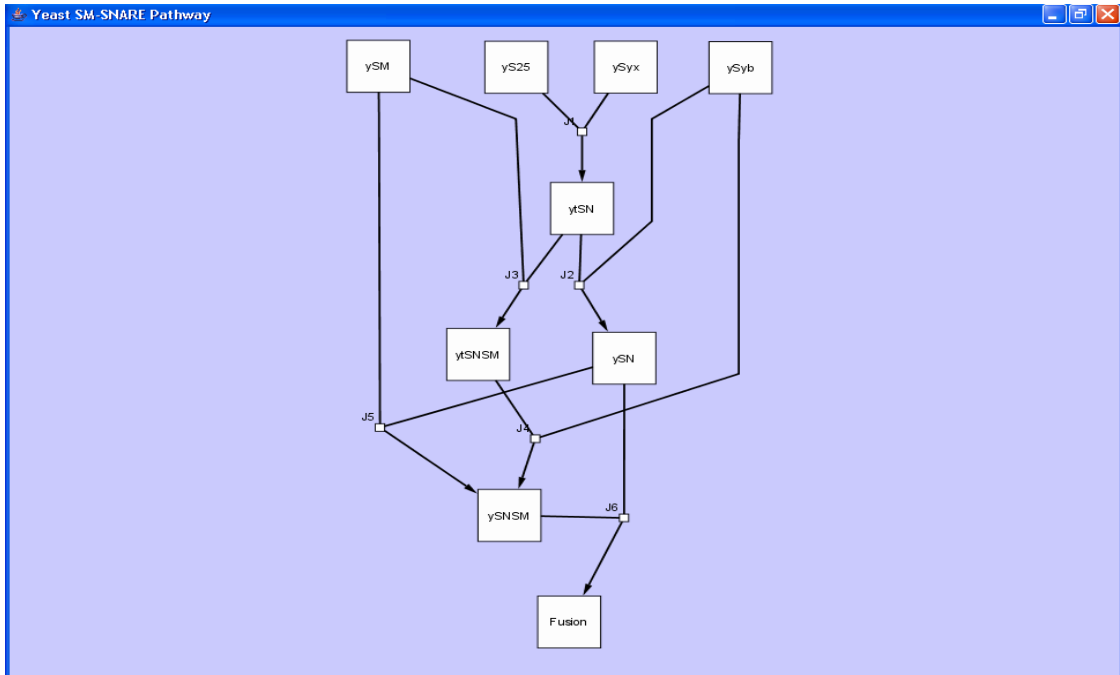
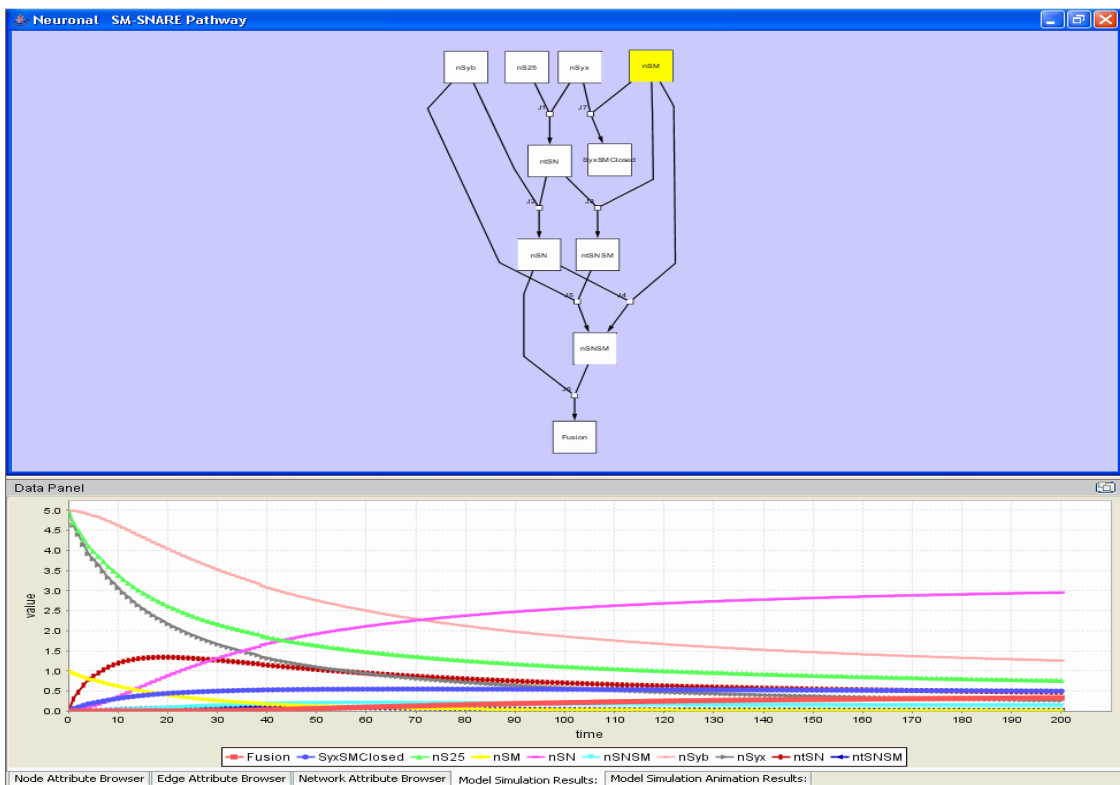**Figure 3.11 Yeast SM-SNARE pathway in CytoModeler**

Figure 3.12 Simulation results are shown in Cytoscape Data Panel. One series is selected and the corresponding node/species nSM is simultaneously highlighted.
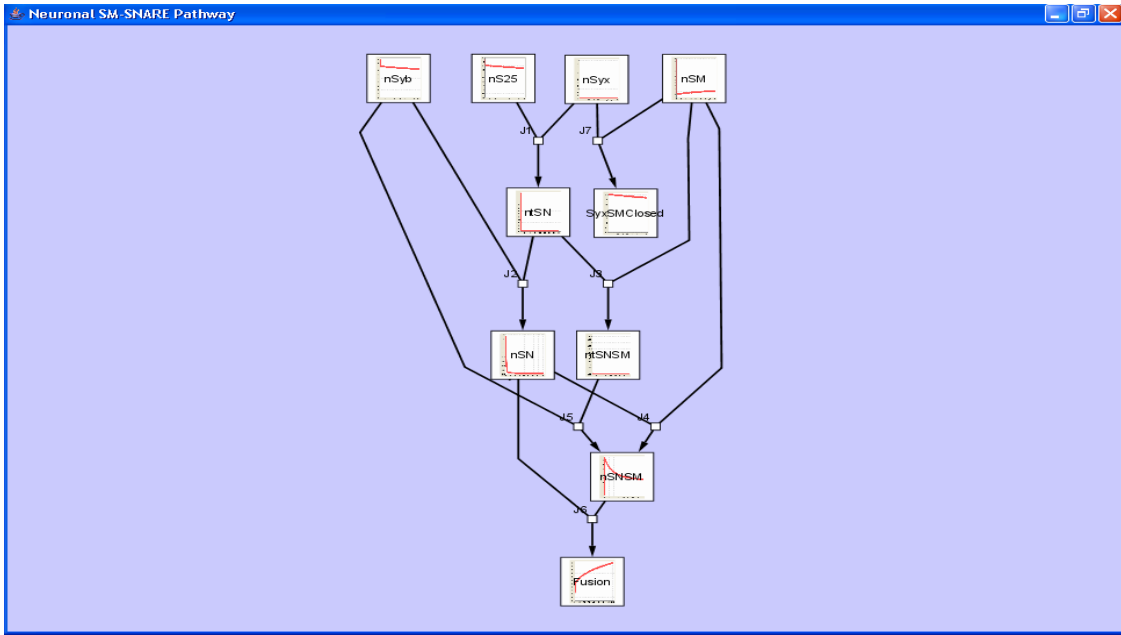
Figure 3.13 X-Y plotting simulation results in network context
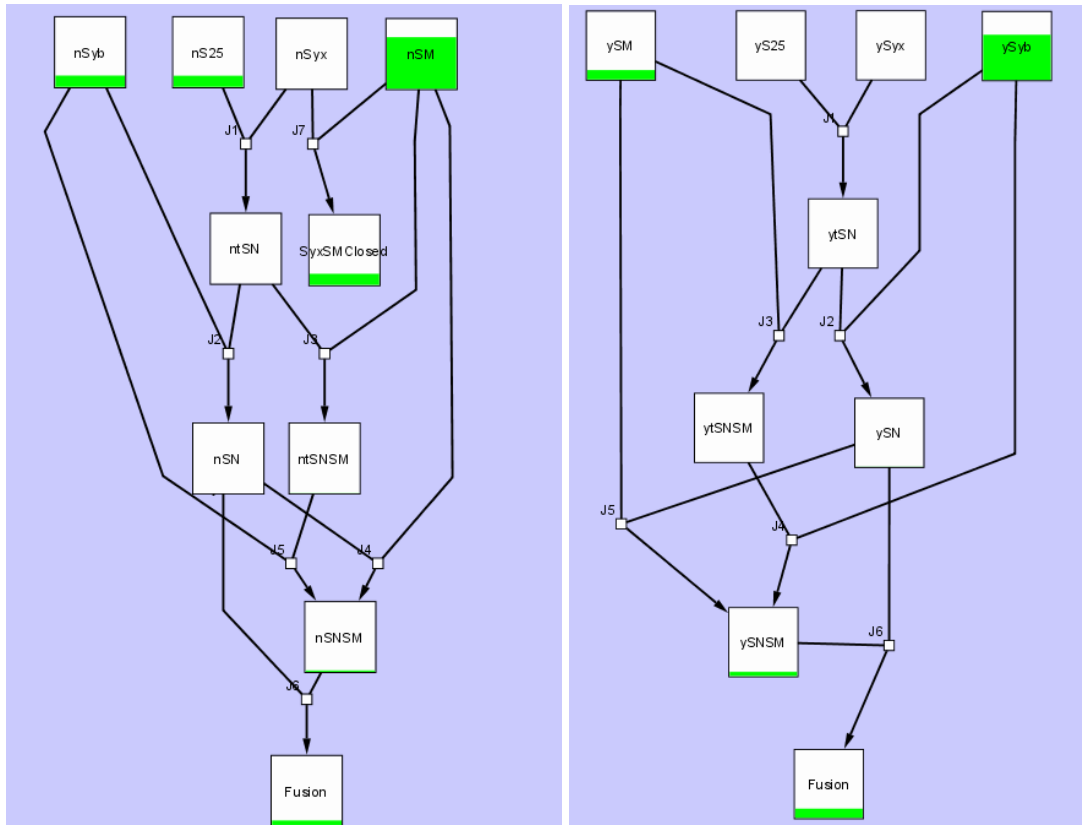
Figure 3.14 A comparison for two models at final time frame of simulation results

CHAPTER 4. QUANTITATIVE MODELING SNARE-SM REGULATION

4.1 Background

4.1.1 Exocytosis background

Exocytosis forms the underpinning mechanism of cell to cell communication in multicellular organisms by means of leading the traffic of vesicles to bind to and fuse with plasma membrane, and thereby releasing its vesicle contents -- a wide range of extracellular signals into targeting cells (Figure 4.1). Great efforts have been made to decipher the nature of the process. It is thought that the central players in trafficking events is an array of evolutionarily conserved proteins which work as a team and orchestrate the process of vesicle docking and priming as well as the process of vesicle fusion. This cooperation necessitates a wealthy amount of interconnections and the process is best described as a network which governs the progression of exocytosis and emerges distinct structures and dynamics at different transition states.
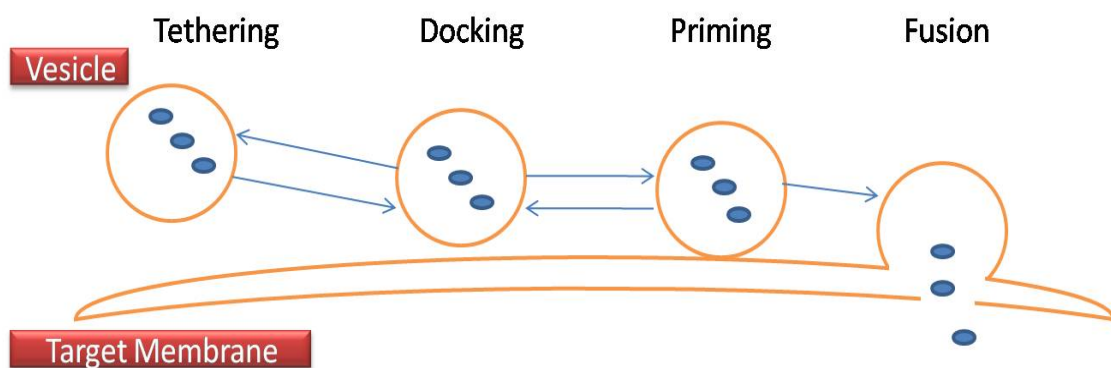
Figure 4.1 Exocytotic pathway. Exocytosis consists of multiple intermediate steps such as vesicle tethering, docking to target membrane, and priming at the membrane and fusing with the membrane to release biochemical contents the vesicle carries.

SNARE proteins since they were first identified in late 1980s have been viewed as a
interaction hub in the interactome functioning as a engine to drive membrane fusion [65].
Three or four of them coming from four subfamilies Qa-, Qb, Qc, and R- SNAREs [2, 3]
(originally categorized as v-SNARE and t-SNARE according to their localization on vesicle
or acceptor membrane [66]) initially assemble into a parallel four helical bundle trans-
complex (FHC/SNAREpin) to link two membranes together and then proceed to a cis-
complex to drive the membranes so close that the readily fuse with each other [67] (Figure
4.2 and Figure 4.3). More recent studies suggest that SNARE proteins collaborate with a
series of interaction partners and form a SNARE cycle to conduct the process of fusion [68].
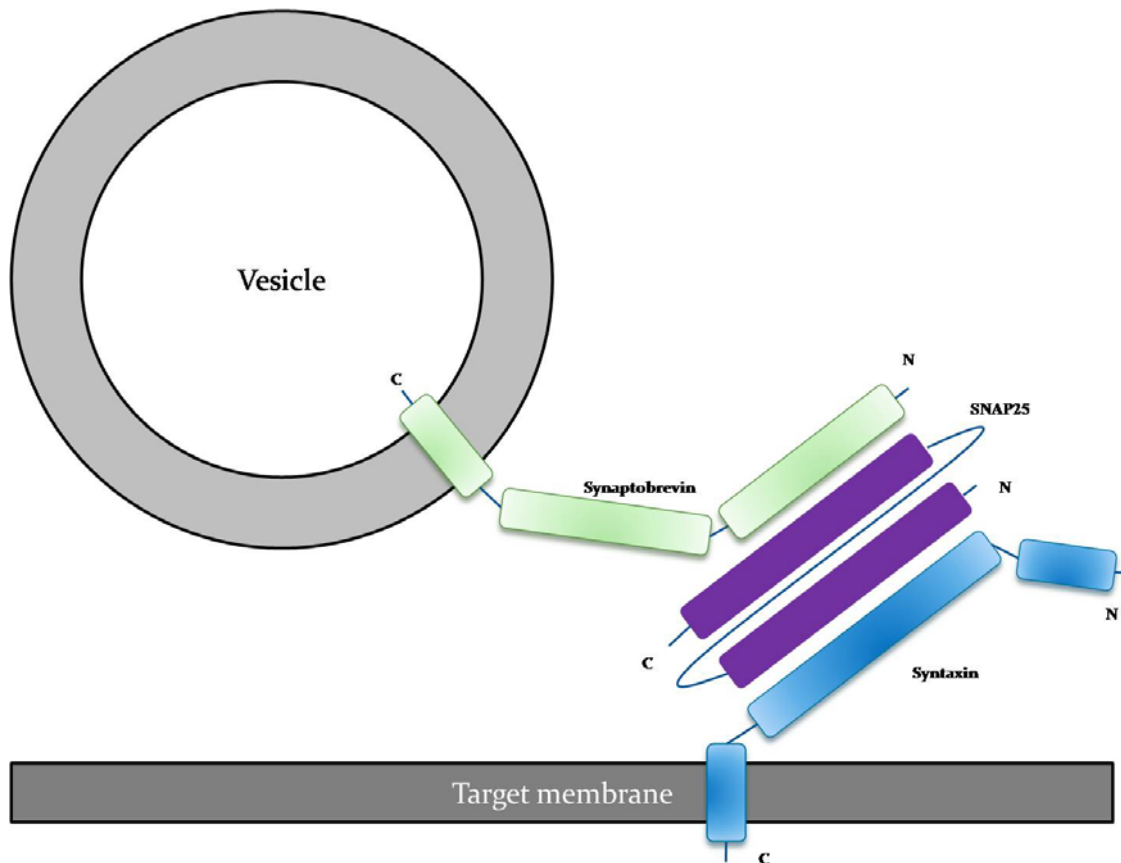
Figure 4.2 Molecular machinery of SNAREpin complex. SNAREpin is four-helical bundle complex, formed by synaptobrevin (green), SNAP25 (purple), syntaxin (blue) protein interactions.
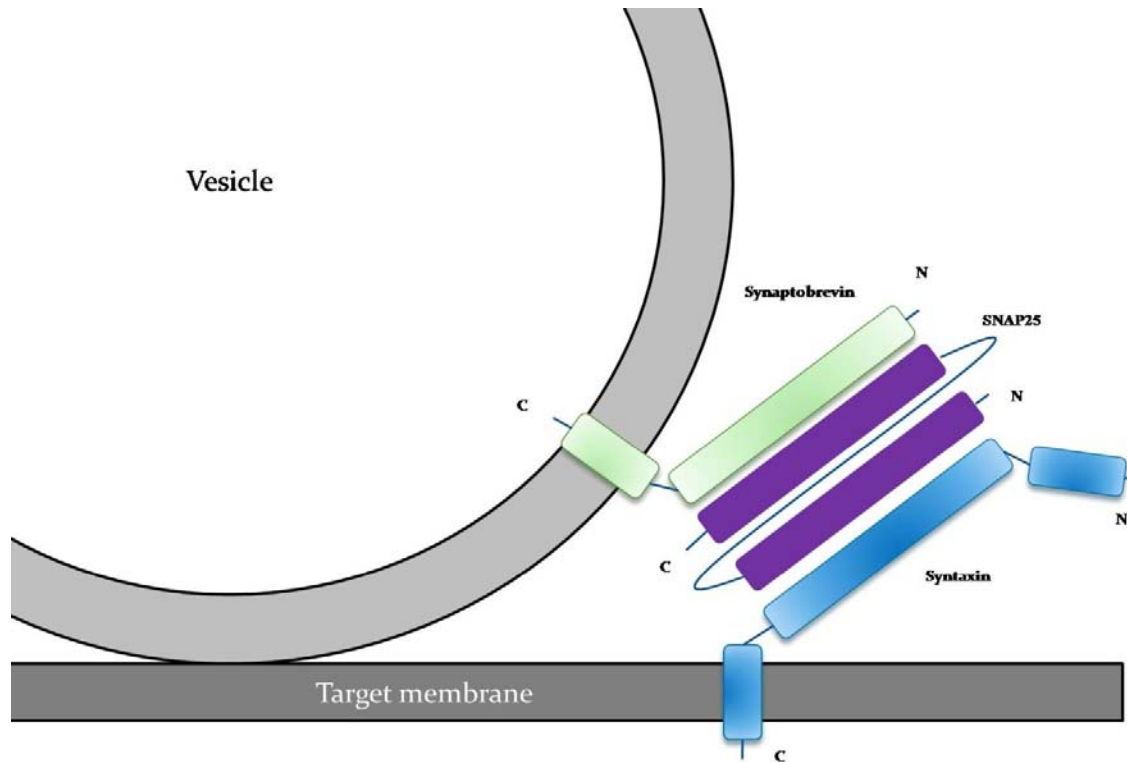
Figure 4.3 SNAREpin serves a fusion engine to execute vesicle fusion with target membrane. It drives two membranes together and then stimulates membrane fusion between them.

In addition to SNAREs, there are many regulatory protein involved in exocytosis to control progression of exocytosis. SM proteins (Sec/Munc-18 like proteins) , an essential regulatory protein, function as organizers to control SNAREpin conformation through diverse modes [69]. In regulated exocytosis, several late regulatory proteins including complexin and synaptotagmin [70] mediate crucial transition of SNAREpin states from loose trans to cis complexes [65, 71]. Complexins whereby freezing the fusion machinery at a late stage of SNARE cycle accumulate and prime SNAREpins to fuse in response to subsequent calcium influx detected by synaptotagmin, which is thought to serve as a universal sensor of calcium signaling [72-75]. Recent high-throughput proteomics studies revealed many other proteins involved in exocytotic membrane fusion and the whole exocytosis process [76-78].

Figure 4.4shows an exocytotic interactome describing a protein-protein interaction network which control exocytosis process).
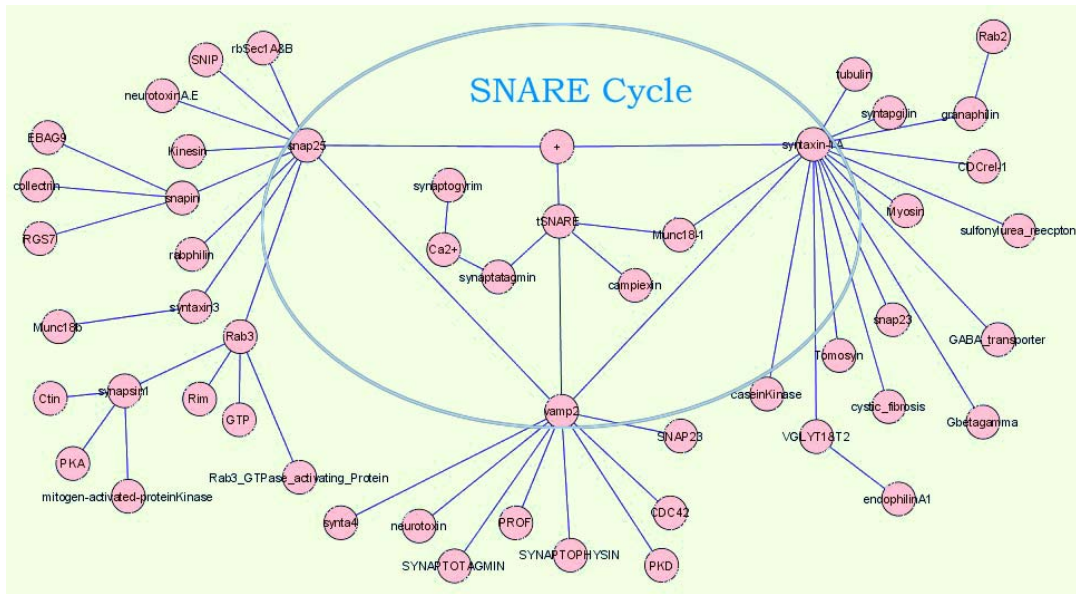
Figure 4.4 Exocytotic interactome

The integration of data with large scale and existing observations remains challenging and new methods which enable to reveal regulation mechanism of complex exocytotic system which consists of interactions between SNARE and many regulatory proteins are needed [79].

4.1.2 SM-SNARE regulation background

Cellular processes are regulated through intricate interactions of biological molecules, represented in network interactome models as 'nodes' and 'edges' [80]. Systems biology network modeling can be extensively applied to interpret experimental observations, explore regulation mechanisms of biological processes, and study the functional roles of specific molecular species. However, the complexity of cellular networks increases when there are

uncertainties or diversities in network topological structures, especially across organism or cell types.

Sec/Munc-18 proteins (SM proteins) which regulate SNARE (soluble N-ethylmaleimide-sensitive factor attachment protein receptor) proteins in exocytotic vesicular fusion, exemplify this complexity. Vesicular fusion during exocytosis, merging two separate phospholipid membranes into one, is one of the most fundamental cellular processes in all eukaryotic cells underlying many crucial biological processes. Like many biological processes, vesicular fusion is mediated by the interactions of several evolutionarily conserved protein families such as SM and SNARE. The SNARE protein family play a central role [79] in vesicular fusion, where t-SNAREs (target membrane associated) and v-SNAREs (vesicle-associated) assemble into a parallel four helical bundle complex (SNAREpin) to drive the two membranes close enough so that they readily fuse with each other [65].

It is also evident that the SM protein family is essential in vesicle fusion, playing a crucial role in regulating SNARE proteins. In contrast to the consensus about the functional mechanisms of the SNARE family, contradictory experimental observations have been reported about the SM-SNARE regulation mechanism [81, 82]. SM proteins exhibit diverse binding modes to SNAREs. These binding modes can be generally categorized according to the binding partners of SM: the t-SNARE complex (pattern 1), the SNAREpin complex (pattern 2), or the mono-SNARE (pattern 3) [69, 81-83]. However, even within a certain binding pattern, SM proteins that come from different organisms or trafficking pathways exhibit hetero binding modes to their cognate SNAREs. A classic example is the neuronal SM protein Munc18-1 (pattern 3) which embraces monomeric syntaxin-1(t-SNARE) in closed conformation and prevents the formation of the SNAREpin complex. This suggests a

negative functional role for Munc18-1 (closed mode of binding of Munc18-syntaxin) [81, 84]. Other SM proteins such as Sly1 and Vps45 bind monomeric syntaxin-1 by interacting with its N-terminal peptide, which allows syntaxin-1 assembled into SNAREpin to facilitate the fusion [82]. This heterogeneity of binding modes between SM and SNARE proteins introduces uncertainties and complexity into the interaction network of regulation in vesicular fusion, and thus greatly complicates the understanding of key functional roles of the SM protein family in exocytosis. SM proteins have been proposed to be both positive and negative regulators in yeast and neuron [33, 85-90]. Rothman and Melia's study for SM regulation in SNARE system provided predictive insights at the system level: the dynamics of the dual roles of SM may determine which outcome dominates in observed overall exocytosis [33].

For this complex system, a single mathematical or computational representation of the network does not sufficiently capture the whole range of system behaviors. So this work integrates comparative network modeling to solve the discrepancies in existing experimental observations and to elucidate the regulation mechanism of SM-SNARE system. Comparative methods enable us to infer protein functions and explore the regulation mechanism of cellular systems by comparing targets with well characterized and evolutionary related proteins and systems across species [10]. In addition, network modeling enables us to investigate complex biological functions without a priori mechanistic knowledge, by decomposing them into relatively independent and less-complicated sub-functions which are carried out by combinations of a set of biological network modules or motifs [9, 91, 92].

This work compares two ensemble protein interaction maps for SM-SNARE network in exocytotic pathways based on the binding mode information: the cascade-like non-regulated

SSN in yeast and the feedback-loop-like regulated SSN neuronal synaptic pathways. Comparative topological analysis revealed that the closed mode of binding of Munc18-syntaxin may be the critical factor that brings the complexity to highly regulated synaptic exocytosis in terms of network topology and system behaviors, compared to constitutive exocytosis. Next, computational simulations reveal bifurcation behavior in the neuronal system that might explain the discordance in studies of over-expressed SM protein. Furthermore, *in silico* mutation experiments confirm that the complex behaviors result from the closed mode of binding of Munc18-syntaxin. This analysis should help address the following important questions: What kind of functional role does SM play in exocytosis? Why are conflicting results observed in overexpression experiments of neuronal SM protein Munc18-1? What underlying mechanisms of regulation contribute to the complexity with respect to the design principles of the exocytotic system?

4.2 Comparative construction of two SM-SNARE networks (SSN)

This study began by constructing two SM-SNARE networks for yeast and neuronal synaptic exocytosis respectively to capture a wide range of SM-SNARE regulation. During the last decades, extensive genetic, physiological, biophysical and biochemical studies have focused on these two model systems because they represent two fundamental types of exocytosis: non-regulated (or constitutive) and regulated. In yeast, the exocytotic pathway is non-regulated.

The yeast pathway operates continually and supplies a continuous stream of vesicles containing lipids and proteins for the plasma membrane. Yeast exocytotic SNAREs Sso1p (yeast syntaxin/t-SNARE), Sec9p (yeast SNAP25/t-SNARE) and Snc1/2p (yeast

synaptobrevin/ t-SNARE) mediate the vesicular fusion process. Sso1p and Sec9p preassemble into the acceptor t-SNARE complex. Then, Snc1/2p associates with the acceptor complex to form the SNAREpin complex, which acts as an engine to release biochemical energy to drive the vesicular and plasma membranes together. The yeast SM protein, Sec1p, regulates the SNARE complexes and the fusion rate by directly binding the t-SNARE complex (pattern 1) and the assembled SNAREpin (pattern 2) [86, 93]. In Figure 4.5, the nodes represent the SNARE, SM, and reactant complex. The edges describe the interactions between them. The process of exocytotic fusion is formalized as a feedforward set of interactions between the SNARE proteins and the SM protein. (To simplify the notation and emphasize homology between the networks, Table 1 gives the naming conventions for all models).

In neurons, the synaptic exocytosis pathway is highly regulated in time and space, and it controls specialized neuron communication and the release of neurotransmitters contained by synaptic vesicles in response to calcium signals. Despite the regulation, the core molecular machinery of the synaptic exocytosis pathway is evolutionarily related with that of yeast; for example, neuronal t-SNAREs, syntaxin-1 and SNAP25 pre-assemble into an acceptor t-SNARE complex. The complex later reacts with VAMP (synaptobrevin/vesicle associated membrane protein) to form an assembled SNARE complex/SNAREpin. The neuronal SM protein Munc18-1 also binds to the intermediate acceptor t-SNARE complex (pattern 1) and assembled SNAREpin (pattern 2) to facilitate the membrane fusion. Munc18-1 has an extra binding mode (closed mode of binding of Munc18-syntaxin) with syntaxin-1, which stabilizes syntaxin-1 in closed conformation, blocking the formation of the SNAREpin complex [94].

These dynamical models for each network enable examination for the behavior of the system (**Methods**). The models consist of differential equations with a set of initial conditions and reaction rates which are expressed by first- and second-order rate constants, based on the interaction diagrams in Figure 4.5 (see **Methods** for details). Following previous modeling studies in chromaffin and P12 cells [95, 96], the concentrations of the SNAREpin and SNAREpin/SM complex quantitatively describe the progression of fusion (detailed in **Methods**).

Table 4.1 Term abbreviations for yeast and neuronal SM-SNARE system.

	Syntaxin	SNAP25	Synaptobrevin	SM	t-SNARE complex	t-SNARE/SM complex	SNAREpin	SNAREpin/SM complex
Yeast	ySyx Sso1p	yS25 Sec9p	ySyb Snc1/2p	ySM Sec1p	ytSN	ytSNSM	ySN	ySNSM
Neuron	nSyx syntaxin-1	nS25 SNAP25	nSyb VAMP2	nSM Munc18-1	ntSN	ntSNSM	nSN	nSNSM
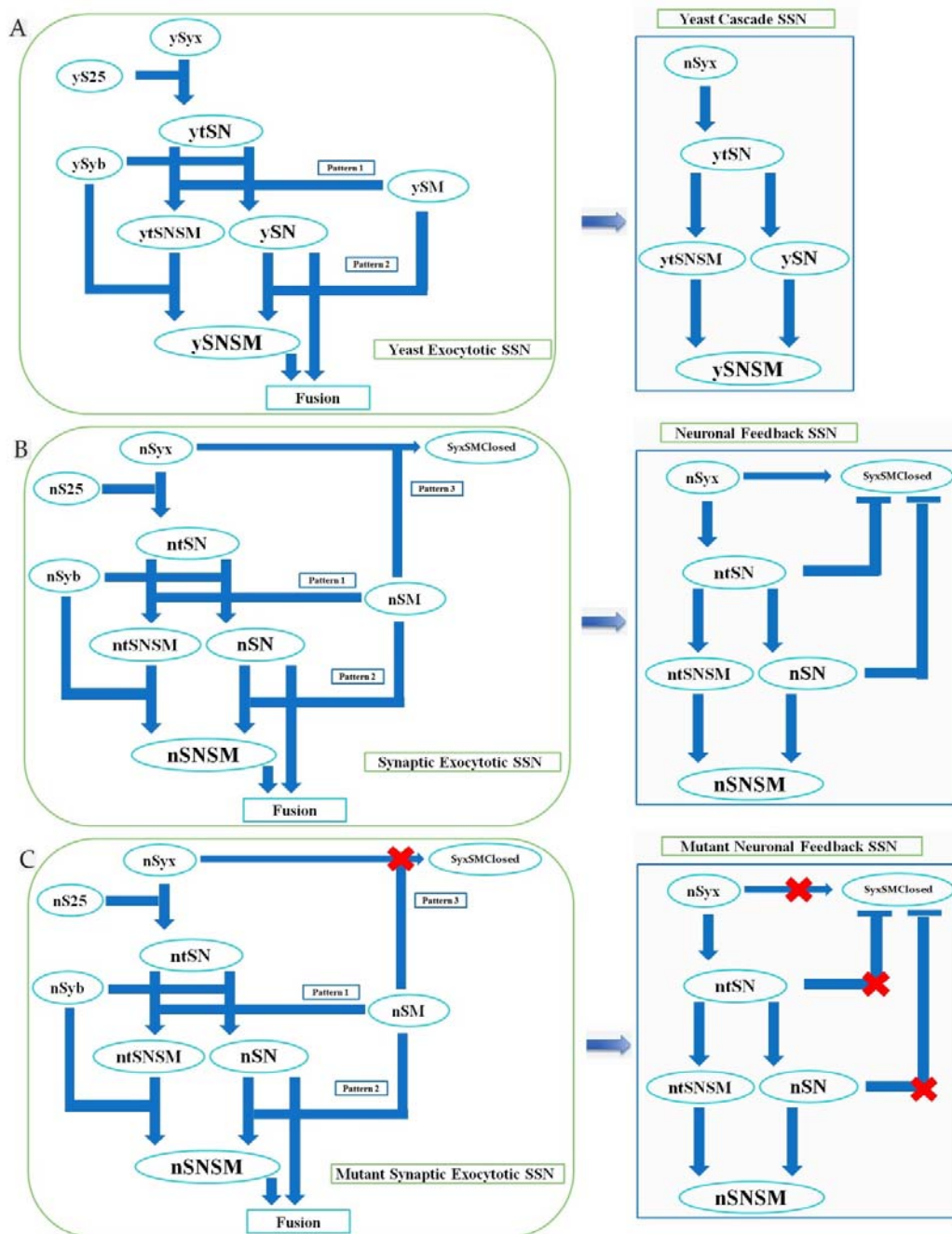
Figure 4.5 SM-SNARE network diversity (A) Formulated diagram of the yeast SSN in exocytosis. ySyx: yS25, ySyb and ySM describe Sso1p(yeast syntaxin), Sec9p(yeast SNAP25), Snc1/2p(yeast synaptobrevin) and Sec1p(SM), respectively. The logic network diagram on the left shows the cascade-like yeast SSN. (B) Formulated diagram of neuronal network in synaptic exocytosis, nSyx, nS25,nSyb and nSM to describe syntaxin-1, SNAP25, VAMP(neuronal synaptobrevin) and Munc18-1(neuronal SM), respectively. The logic network diagram on the left shows the feedback structure with modulation in the neuronal synaptic system. (C) Formulated diagram of mutant neuronal network in synaptic exocytosis, nSyx, nS25, nSyb and nSM to describe syntaxin-1, SNAP25, VAMP (neuronal synaptobrevin) and Munc18-1(neuronal SM), respectively. The logic network diagram on the left shows that the feedback structure is blocked due to mutant nSyx.

4.3 Network comparison analysis extracts the critical distinction between two SM-SNARE networks

The two SM-SNARE networks from yeast and neuron illustrate the design principles that underlie the divergence and complexity in the yeast and neuronal exocytotic systems. Network comparison analysis [10] explores the differences with respect to network structure, since the topological diversity of biological networks usually reflects the diversities of function, evolutionary selection, and regulation mechanism of cellular processes [91, 97]. The analysis showed that the unique neuronal SM-SNARE binding mode (closed mode of binding of Munc18-syntaxin) might be a critical factor that led to the structural divergence between the two SM-SNARE networks in yeast and neuron. In the yeast SSN, every component piece of SNAREpin/SM is sequentially assembled to an intermediate protein complex through a series of discrete levels. Therefore the network is cascade-like (Figure 4.5**A**). In the neuronal SSN, there is a similar cascade branch to the one in yeast. However, there is another branch which is introduced by the neuronal closed mode of binding, Munc18-syntaxin. Due to this extra branch, nSyx (syntaxin-1) is inhibited by nSM (Munc18-1) or it plays another functional roles by interacting nSM (Munc18-1), such as helping vesicle docking [90]. These two branches actually form a feedback loop because the acceptor t-SNARE complex and SNAREpin which formed through the cascade branch can also interact with nSM forming the SNAREpin/SM complex. This sequesters nSM (Munc18-1) away from the nSyx (syntaxin-1) and keeps nSyx from being inhibited in the closed mode (Figure 4.5**B**).

The neuronal SM-SNARE binding mode (closed mode of binding of Munc18-syntaxin) radically changes the topology of the SM-SNARE network in neurons from that in yeast,

even as it conserves the cascade-like branch. This suggests that the binding mode drives the divergence of the SM-SNARE network regulation in the secretory pathways in the different systems, and introduces the complexity into the neuronal system.

4.4 Comparative *in silico* experiments reveal differential system behaviors of SM regulation

To further examine the effects of divergent SM-SNARE network structure on system behaviors, we investigated system behavior in response to SM regulation both in yeast and neurons, using the system models described above. We also compared our simulated results with experimental observations to validate the models and connect the model prediction with functional role of SM regulation.

4.4.1 Yeast SSN Model predicts SM stimulates fusion in yeast

The first simulation models system behaviors of the cascade-like yeast SM-SNARE network with respect to the ySM protein concentration. The ySM positively regulates the fusion process as the amount of fusion shows a hyperbolic response to the ySM protein concentration. Figure 4.6**A** presents fusion curves of five simulated experiments with different ySM concentrations. Figure 4.6**B** depicts the steady-state fusion level of the system with respect to varying the initial concentration of ySM protein. This demonstrates that ySM plays a positive role that stimulates membrane fusion in both rate and amount.

The simulation analysis agrees with experimental observations [86] in lipid mixing assays. This kind of assay utilizes the FRET pair fluorophores: nitrobenzoxadiazole (NBD) and rhodamine. The emission of NBD is muted by fluorescence resonance energy transfer. When fusion occurs between the fluorescent donor and unlabeled acceptor vesicles, the fluorescent probes (NBD and rhodamine) labeled on donor vesicles are diluted in the new

membrane and the emission of NBD is enhanced. The fluorescent intensity can measure the fusion level. More importantly, the reconstitution lipid mixing assay allows biologists to investigate the fusion event *in vitro* by precisely controlling the concentration ratio of SNARE proteins or other regulatory proteins.

In [86], recombinant yeast SM protein Sec1p was added to the yeast SNARE reconstitution liposome system at different concentrations. The effects of the stimulation on fusion show a monotonous dependent on Sec1p concentration (Figure 6D in [86]). When the levels of Secp1 are increased in the assay, an monotonic increase in fusion is observed [86]. Therefore, the experimental observations verify the yeast SSN model prediction.

Figure 4.6 Yeast SM-SNARE system analysis *in silico* **(A) Fusion curves of five** *in silico* **experiments for different initial concentrations of the ySM protein using the network from figure 1A. (B) Final fusion levels of the yeast SSN at steady state with respect to concentration of the ySM protein.**

4.4.2   Neuronal SM-SNARE model predicts SM stimulates fusion in neuron but in more complex way

The neuronal SSN model allows computational exploration of the system behaviors in the feedback-loop-like neuronal SM-SNARE network with respect to the nSM protein concentration. The results show a bifurcation. At reasonable physiological levels (the concentration of nSM is less than nSyx [82, 89]) nSM effectively stimulated the fusion.

However, under extreme conditions where concentration of nSM is larger than nSyx nSM concentration shows a negative relationship with fusion efficiency. It is hard to achieve this response under normal physiological conditions *in vivo*, as it requires the level of nSM protein concentration to be much larger than that of t-SNARE. This contradicts the fact that syntaxin-1(tSNARE) outnumbers Munc18-1(nSM) [82, 89].

Specifically, we found when the concentration of nSM is below the bifurcation point, nSM stimulates the fusion process; once the concentration goes beyond the point, then nSM starts to inhibit fusion. Therefore, by analyzing the bistable system response to nSM protein amount, we can explain the conflicting experimental data. Figure 4.7A shows fusion curves of five simulated experiments with different nSM concentration. Figure 4.7A shows that nSM plays either a positive or negative role depending on the dose. Overall, nSM stimulates the fusion level compared to the basal SNARE system without nSM regulation. However, when the level of nSM increases beyond the bifurcation point (abnormal physiological condition), the stimulation efficiency decreases and nSM may also play a negative factor at the extreme physiological condition. This system behavior differs from that of the cascade SM-SNARE system in which the fusion level monotonically increases with the level of ySM.

Figure 4.7**B** describe fusion amount of the neuronal SSN with respect to the initial concentration of the nSM protein at steady state. When the concentration of nSM is less than or equal to SNAREs, it functions as a promoter to stimulate fusion and the stimulation reaches maximum when the concentration ratio of SM and t-SNARE is 1. When the concentration of nSM increases beyond the concentration of t-SNARE, its stimulation effect decreases. Beyond the bifurcation point, nSM functions as a negative regulator which inhibits fusion.
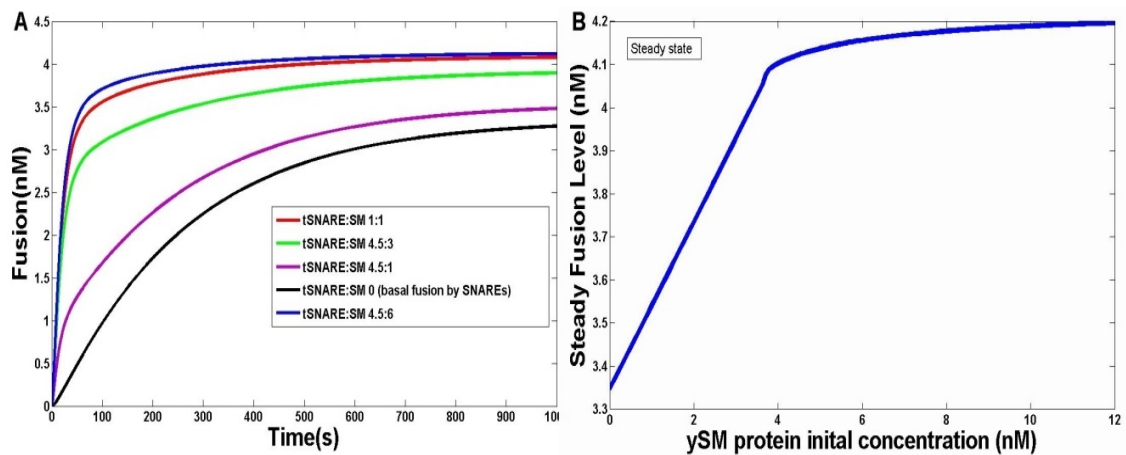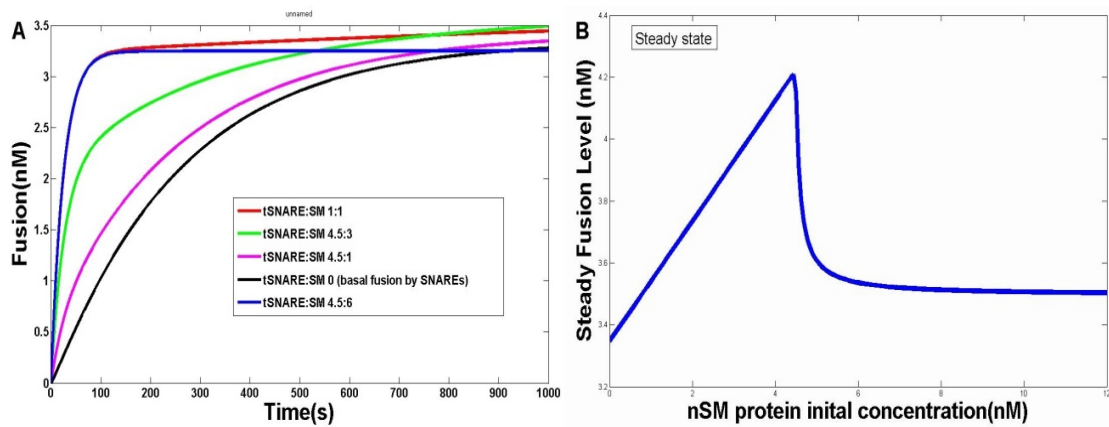
Figure 4.7 Neuronal SM-SNARE system analysis *in silico* (A) Fusion curves of five *in silico* experiments for different initial concentrations of the nSM protein using the network from figure 1B. (B) Final fusion levels of the neuronal SSN at steady state with respect to the concentration of the nSM protein.

This finding reconciles current conflicting experimental observations for functions of neuronal SM Munc18-1. In addition, it reveals the regulation mechanism between neuronal SNARE and SM proteins. Shen's [33] *in vitro* lipid mixing assay study shows that Munc18-1 stimulates liposome fusion and the stimulation effect achieve maximum when ratio of syntaxin-1/Munc18-1 is around 1, which exactly confirms the neuronal SSN model prediction. An *in vivo* study by Wu et al. [85], showed that overexpression of the *Drosophila* SM protein strongly inhibits evoked and spontaneous neurotransmitter release which is mostly based on synaptic vesicle fusion with the target membrane. Simulations of SM overexpression in the neuronal SSN (Figure 4.8) predict the inhibiting effects [85, 98]. Furthermore, increasing the concentrations of both nSM and nSxy prevented the inhibition of fusion. This reproduces observations in studies [85, 99] and confirms the hypothesis that the relative levels of nSM with respect to nSxy are critical for the efficiency of exocytotic fusion [81].
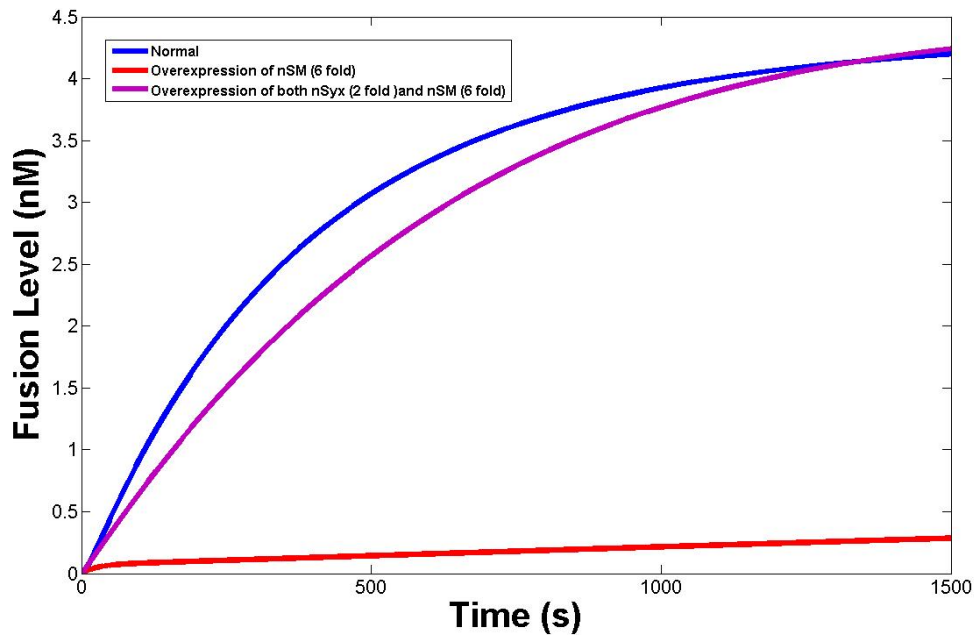
Figure 4.8 nSM Overexpression analysis *in silico* **The relative levels of nSM and nSyx are crucial for fusion outcome. When nSM is overexpressed, fusion is inhibited (red line); when both nSM and nSyx are overexpressed, fusion proceeds (purple line).**

4.4.3   Simulated mutations confirm in neuronal SM-SNARE network

The above *in silico* analyses show different behaviors of yeast and neuronal SM-SNARE systems. According to the comparative network analysis, the neuronal closed mode of binding of Munc18-syntaxin of SM-SNARE is a key factor in the divergence of the two systems in terms of network topology. We performed a perturbation analysis *in silico* to eliminate the neuronal nSM binding to closed nSyx (closed mode of binding of Munc18-syntaxin)(see Figure 4.5**C**). To comparatively analyze the effect of the *in silico* mutation, the mutant nSM system was simulated with five different nSM concentrations. Figure 4.9**A** depicts the fusion curves of the five simulations for the mutant nSM system. Figure 4.9**B** compares the amount of fusion at steady-state for the three systems in Figure 4.5with respect to the initial concentration of nSM protein. The simulation results show that the mutant

neuronal SSN shows similar behavior with the yeast SSN. The fusion is amplified by increasing the concentration of mutant nSM protein and the stimulation effect monotonically increases with the concentration of the mutant nSM protein. This indicates that neuronal nSM binding to closed nSyx (closed mode of binding of Munc18-syntaxin) may be the key factor in inducing the structural divergence of SM-SNARE network in yeast and neuron.
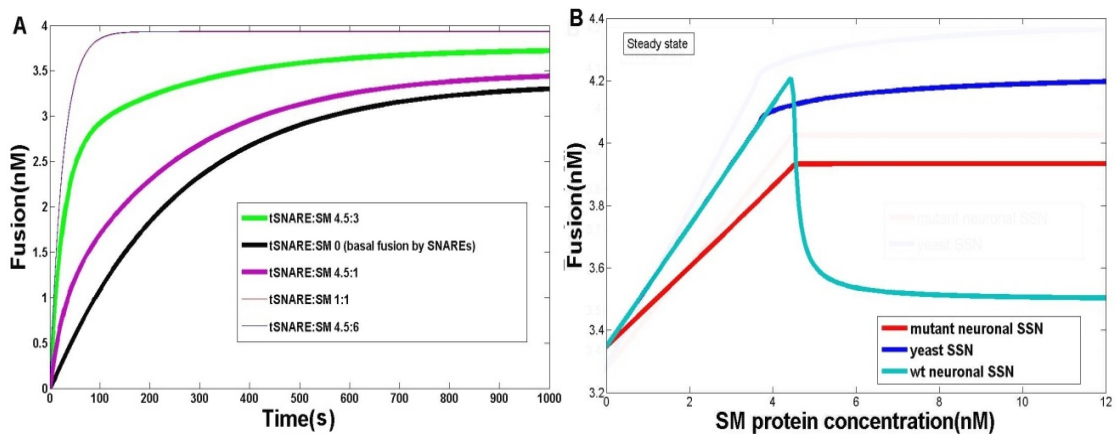
Figure 4.9 Mutant neuronal SM-SNARE system analysis *in silico* (A) Fusion curves of five *in silico* experiments with different initial concentrations of the nSM protein in the mutant neuronal SSN system which eliminates the nSM(Munc18-1) binding to closed nSyx(syntaxin-1) (Figure 1C). (B) A comparison of fusion levels between the yeast SSN, neuronal SSN, mutant SSN at steady state with respect to the SM protein concentration.

4.5   Discussion

In summary, we applied comparative systems biology methodologies to analyze the dynamics of the SM protein family that regulates the SNARE protein in exocytotic fusion in both yeast and neuron, and explored the connection between the dynamics and the network structure. Cellular functions are mediated by the networks of interacting molecules that operate at different levels and different structure of organizations [9, 91]. Through comparative network construction and topological analysis, this approach investigates system level differences between the regulation mechanisms of SM-SNARE interactions in exocytotic fusion. The results suggest that the yeast SM-SNARE network has a distinct regulation architecture from the neuronal SM-SNARE network. The unique neuronal SM

protein binding to closed syntaxin-1 is a critical factor in the divergence of network topology, which transforms the neuronal SM-SNARE network from a cascade into a feedback structure by introducing an interaction-specific alteration to the SM-SNARE network. The distinct binding mode introduces complexity into system regulation: *in silico* experiments indicate that feedback-loop-like neuronal SM-SNARE presents bifurcation behaviors with respect to the neuronal SM protein concentration, compared to the simple monotonic system response in yeast SM-SNARE system. It suggests that when the concentration of neuronal SM protein goes beyond a certain value, its functional role reverses and it inhibits exocytotic fusion. This explains the conflicting experimental observations for neuronal SM protein and Munc18-1.

4.5.1 The SM proteins

SM is confirmed as the positive regulator on exocytotic membrane fusion both in the yeast and neuronal system. In yeast, SM monotonically stimulates the fusion progression as its concentration increases, according to both *in silico* and *in vitro* experiments. In neurons, SM affects the fusion rate depending on the ratio of the concentrations of nSM and nSyx. However, going beyond the bifurcation point is unlikely to occur under normal physiological conditions as studies have shown that the t-SNARE protein nSyx outnumbers nSM [82, 89]. Going beyond the bifurcation point requires that the level of nSM exceeds the amount of t-SNARE protein.

This work brings a new perspective on why many studies which explore the functional role of SM protein by applying overexpression of the SM proteins appear contradictory. However, it should be noted that the mechanism of neuronal exocytosis *in vivo* is more

complex and further experiments which can monitor the dynamics of SM protein concentration in temporal and spatial extents are needed to validate the bifurcation behaviors.

4.5.2   The unique binding mode of nSM (Munc18-1)

Another intriguing observation is that the neuronal SM protein Munc18-1 binding to closed syntaxin-1 appears to be the critical difference between the yeast and neuronal SSNs. The classic binding mode of neuronal SM protein was identified as a negative factor for exocytotic membrane fusion because it stabilizes the closed syntaxin-1 and therefore prevents it from formation of SNAREpin. New insights from Sudhof's lab [90] demonstrate that the neuronal SM protein binding mode functions facilitate the vesicle docking to target membrane. Our study demonstrates both functions at the level of regulatory mechanism. This binding mode combined with other modes of Munc18-1 to t-SNARE and SNAREpin complexes creates a feedback SM-SNARE regulation system, which precisely controls exocytosis by cooperatively working with other regulatory factors, complexin and synaptotagmin proteins [79] to meet the requirements of fast neuronal release.

4.5.3   The SM-SNARE network

This study connects topology of SM-SNARE network with the regulation mechanism of systems and biological functions in exocytosis. We identified two structures for the SM-SNARE network: cascade in yeast and feedback in neurons. Both models have demonstrated robustness in a large working parameter space (**Methods**) and reproduce a variety of functional experimental observations.

Mounting evidence suggests that complex biological functions can be decomposed into relatively independent subfunctions. The modules/motifs responding to input signal and

surrounding perturbations depending on their topology [91]. These combinations of modules/motif not only provide functional robustness but also flexibility that may be critical for system to evolve and adapt in a great range of evolutionary and functional constraints. More complex function can be hierarchically built via the combinations of simple and robust functional modules/motif [100].

This SSN analysis presents the modular design principle in the exocytosis process. The SSN subfunction executes the exocytosis function by combining with other subfunctions, such as signaling calcium influx. In simple constitutive exocytosis of yeast, the cascade SM-SNARE network robustly functions in exocytotic membrane fusion. In contrast, neuron exocytosis must be tightly regulated to temporal and spatial factors. To meet the variances in physiological requirements, the new module/motif introduced by neuronal-specific Munc18-1 binding to closed syntaxin-1 combines with the conserved cascade-like module produce the new feedback-loop-like module/motif. The new combination allows more complex functions in neuronal exocytotic fusion than in yeast. Working with other subfunctions operated by other modular combinations such as protein complexin and synaptotagmin, the SM-SNARE system forms an exquisite regulation mechanism to control complex neuronal exocytosis.

4.6   Methods

4.6.1   Mathematical formulation of SM-SNARE regulations

For the two SM-SNARE networks, we constructed dynamical models for them to examine effects of the different network topologies of regulation, on the behavior of the system.

The models consist of differential equations with a set of parameters and non-zero initial conditions, based on previously characterized interaction diagrams. We in our models

integrated multiple simultaneous input signals and examine quantitative effects of them on

the systems by considering supply stimulus and degradations of SM and SNARE proteins

because numerous studies *in vivo* exam the physiological behaviors by controlling supply

input signals such as overexpression or deletion of SM genes. For each model, the state of the

system is described by the concentrations of all relevant protein ($C_1(t), C_2(t)..., C_n(t)$). The

rates of reactions are dependent on these concentrations and on biochemical rate constant

parameters ($k_1, k_2, ..., k_n$) and input supply rates ($p_1, p_2, ..., p_n$) and protein degradation rates

($u_1, u_2, ..., u_n$). To describe the temporal behavior, systems of ODEs are provided in general

form:

$$\frac{dC_i}{dt} = f_i(\boldsymbol{C}, \boldsymbol{K}, \boldsymbol{S}, \boldsymbol{U}) \tag{1}$$

$$i = 1, ..., n$$

which describes the evolution of a set of state variables $\boldsymbol{C} = C_1(t), C_2(t)..., C_n(t)$; $\boldsymbol{K}, \boldsymbol{S}, \boldsymbol{U}$

represent vectors of system parameters. It is assumed that the state C evolves in a subset $\boldsymbol{C}$ of

Euclidean space, in which it is positive or non-negative that is always satisfied in

biochemical application. In our model, we used protein family names: syntaxin, SNAP25,

VAMP and SM to present model variable names for specific SNARE proteins and SM

protein in different organisms, respectively. The whole systems are shown in Table 4.2 and

Table 4.3.

Table 4.2 The yeast and neuronal SM-SNARE systems

| Yeast SSN Reactions	Neuronal SSN Reactions
$ySyx + yS25 \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ytSN$	$nSyx + nS25 \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ntSN$
$ytSN + ySyb \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} ySN$	$ntSN + nSyb \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} nSN$
$ySN + ySM \underset{k_{-4}}{\overset{k_4}{\rightleftarrows}} ySNSM$	$nSyx + nSM \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} nSyxSMClose$
$ytSN + ySM \underset{k_{-5}}{\overset{k_5}{\rightleftarrows}} ytSNSM$	$nSN + nSM \underset{k_{-4}}{\overset{k_4}{\rightleftarrows}} nSNSM$
$ytSNSM + ySyb \underset{k_{-6}}{\overset{k_6}{\rightleftarrows}} ySNSl$	$ntSN + nSM \underset{k_{-5}}{\overset{k_5}{\rightleftarrows}} ntSNSM$
$ySNSM \xrightarrow{k_{fsnsm}} fusion_1$	$ntSNSM + nSyb \underset{k_{-6}}{\overset{k_6}{\rightleftarrows}} nSNSM$
$ySN \xrightarrow{k_{fsn}} fusion_2$	$nSNSM \xrightarrow{k_{fsnsm}} fusion_1$
$fusion = fusion_2 + fusion_1$	$nSN \xrightarrow{k_{fsn}} fusion_2$
	$fusion = fusion_2 + fusion_1$

Table 4.3 The equation sets of modeling yeast and neuronal SM-SNARE systems

Yeast SSN Equation	Neuronal SSN Equation

Yeast SSN Equation

$$\frac{d(yS25)}{dt} = -v_1 + j_1$$

$$\frac{d(ySyx)}{dt} = -v_1 + j_2$$

$$\frac{d(ySyb)}{dt} = -v_2 - v_6 + j_3$$

$$\frac{d(ySM)}{dt} = -v_4 - v_5 + j_4$$

$$\frac{d(ytSN)}{dt} = v_1 - v_2 - v_5 + j_5$$

$$\frac{d(ytSNSM)}{dt} = v_5 - v_6 + j_6$$

$$\frac{d(ySN)}{dt} = v_2 - v_4 - v_{fsn} + j_7$$

$$\frac{d(ySNSM)}{dt} = v_6 + v_4 - v_{fsnsm} + j_8$$

$$\frac{d(fusion)}{dt} = v_{fsn} + v_{fsnsm}$$

$$j_1 = p_1 - u_1 \cdot yS25$$
$$j_2 = p_2 - u_2 \cdot ySyx$$
$$j_3 = p_3 - u_3 \cdot ySyb$$
$$j_3 = p_4 - u_4 \cdot ySM$$
$$j_5 = -u_5 \cdot ytSN$$
$$j_6 = -u_6 \cdot ytSNSM$$
$$j_7 = -u_7 \cdot ySN$$
$$j_8 = -u_8 \cdot ySNSM$$
$$v_1 = k_1 \cdot ySyx \cdot yS25 - k_{-1} ytSN$$
$$v_2 = k_2 \cdot ytSN \cdot ySyb - k_{-2} \cdot ySN$$
$$v_4 = k_4 \cdot ySN \cdot ySM - k_{-4} \cdot ySNSM$$
$$v_5 = k_5 \cdot ytSN \cdot ySM - k_{-5} \cdot ytSNSM$$
$$v_6 = k_6 \cdot ytSNSM \cdot ySyb - k_{-6} \cdot ySNS$$
$$v_{fsn} = k_{fsn} \cdot ySN$$
$$v_{fsnsm} = k_{fsnsm} \cdot ySNSM$$

Neuronal SSN Equation

$$\frac{d(nS25)}{dt} = -v_1 + j_1$$

$$\frac{d(nSyx)}{dt} = -v_1 - v_3 + j_2$$

$$\frac{d(nSyb)}{dt} = -v_2 - v_6 + j_3$$

$$\frac{d(nSM)}{dt} = -v_4 - v_5 - v_3 + j_4$$

$$\frac{d(ntSN)}{dt} = v_1 - v_2 - v_5 + j_5$$

$$\frac{d(ntSNSM)}{dt} = v_5 - v_6 + j_6$$

$$\frac{d(nSN)}{dt} = v_2 - v_4 - v_{fsn} + j_7$$

$$\frac{d(nSNSM)}{dt} = v_6 + v_4 - v_{fsnsm} + j_8$$

$$\frac{d(nSyxSMClosed)}{dt} = v_3 + j_9$$

$$\frac{d(fusion)}{dt} = v_{fsn} + v_{fsnsm}$$

$$j_1 = p_1 - u_1 \cdot nS25$$
$$j_2 = p_2 - u_2 \cdot nSyx$$
$$j_3 = p_3 - u_3 \cdot nSyb$$
$$j_3 = p_4 - u_4 \cdot nSM$$
$$j_5 = -u_5 \cdot ntSN$$
$$j_6 = -u_6 \cdot ntSNSM$$
$$j_7 = -u_7 \cdot nSN$$
$$j_8 = -u_8 \cdot nSNSM$$
$$j_9 = -u_9 \cdot nSyxSMClosed$$
$$v_1 = k_1 \cdot nSyx \cdot nS25 - k_{-1} ntSN$$
$$v_2 = k_2 \cdot ntSN \cdot nSyb - k_{-2} \cdot nSN$$
$$v_3 = k_3 \cdot nSyx \cdot nSM$$
$$\quad - k_{-3} \cdot nSyxSMClosed$$
$$v_4 = k_4 \cdot nSN \cdot nSM - k_{-4} \cdot nSNSM$$
$$v_5 = k_5 \cdot ntSN \cdot nSM - k_{-5} \cdot ntSNSM$$
$$v_6 = k_6 \cdot ntSNSM \cdot nSyb - k_{-6} \cdot nSNS$$
$$v_{fsn} = k_{fsn} \cdot nSN$$
$$v_{fsnsm} = k_{fsnsm} \cdot nSNSM$$

4.6.2 Bifurcation and Robustness Analysis of Parameters

The kinetics simulation and search of the existence of a bistability were implemented in Matlab7.0R. Differential equations were solved using the ODE23s routine. The bifurcation analysis was completed in Systems Biology Toolbox [56].

For testing the robustness of parameters, we generated 2000 random parameter sets using Latin Hypercube Sampling when all parameters are varied +/-30% relative to their original values. Here, we used a uniform probability distribution for each parameter (Table 4.4).

Initial concentration of SM protein (µM)
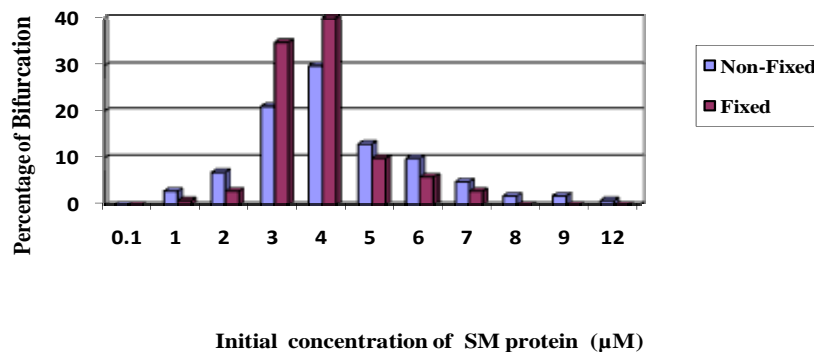
Figure 4.10 Robustness of bifurcation in feedback neuronal SM-SNARE system. For any given initial concentration of SM protein (0.1µM, 12 µM), the percentage of parameter sets that exhibits bifurcation are plotted. No Fix: All of the parameters are varied +/-30% from its default value. Fix: The supply input rate of SNARE and SM family proteins (p_1-p_4) are fixed while all other parameters are varied +/-30% from its default value for each sample run.

4.6.3 Initial Conditions, Parameters and Units

Initial Conditions and Units. The concentrations of reactant proteins are given in molar units. For non-soluble proteins such as v-SNARE, Syb, we followed the [96] work and based the protein concentration estimation on the concentration of secretory vesicles in molar. During the exocytosis process, the vesicles cluster in active zone area in pool form. The size of vesicle pool varies with respect to different cell types or even species from 200 to

3000[101]. Therefore, the molar concentration of vesicles was estimated in the range of 0.2-30 nm. Accordingly, the concentration of Syb is considered to be in an identical range of vesicle concentration (0.2-30 nm). The t-SNARE proteins such as SNAP25 and syntaxin are thought to be vastly expressed *in vivo* and the studies [96, 102] evaluated the concentration of these protein in a range of 0.1-100 µm. The essential regulatory proteins, SM/Munc18 and displacement factor proteins such as Munc13 are known to be expressed at much lower levels, compared to SNARE proteins. The concentrations of the proteins were taken in range of 1-30 nm [78, 96, 101, 103].

Rate constants. In our models, where available, we have relied on *in vivo* and *in vitro* biochemical experiments for parameter values [33, 95, 102, 104-107]. In cases where the values of biochemical parameter were not known yet, we estimated physically reasonable values based on a previous modeling study [96] which provided invaluable information on mining biochemical experiments for parameter values *in vivo/in vitro* and also approaches to estimating unknown parameters. It should be stressed that these available rate constants are measured independently and under different secretion systems which may be different quantitatively. However, because exocytosis process is highly conserved throughout different cell types, we integrated these rate constants into our kinetic equations which aim at providing insights into fundamental regulation mechanisms of protein interaction among two most essential protein families (SM and SNARE) during almost every type of exocytosis process [82, 94] and therefore our models can served as a framework for integration refinement from different systems through adding system-specific regulatory steps or fitting newly characterized kinetic features.

Table 4.4 System Parameters of the Models   Est. Estimated; Units: The total amounts of different species are expressed in units of molar (M). The first and second rate constants presented in units of s^{-1} and $M^{-1} \cdot s^{-1}$, respectively. The supply input rate constants (p_i, i=1…4) are presented in $M \cdot s^{-1}$. The degradation rate constants (u_i, i=1…4) are presented in s^{-1}.

Parameter	Value	Reference
k_1	$1.6 \times 10^4 \, M^{-1} s^{-1} \sim 7 \times 10^4 \, M^{-1} s^{-1}$	[108-111]
k_{-1}	$5 \times 10^{-3} \, s^{-1} \sim 7 \times 10^{-2} \, s^{-1}$	[108-110]
k_2	$4 \times 10^6 \, M^{-1} s^{-1} \sim 6 \times 10^6 \, M^{-1} s^{-1}$	[108-110]
k_{-2}	$2 \times 10^{-4} \, s^{-1} \sim 7 \times 10^{-4} \, s^{-1}$	[108-110]
k_3	$6 \times 10^6 \, M^{-1} s^{-1} \sim 7 \times 10^6 \, M^{-1} s^{-1}$	[107]
k_{-3}	$9 \times 10^{-4} \, s^{-1} \sim 7 \times 10^{-3} \, s^{-1}$	[106, 107]
k_4	$6 \times 10^6 \, M^{-1} s^{-1} \sim 9 \times 10^6 \, M^{-1} s^{-1}$	[33, 107]
k_{-4}	$6 \times 10^{-5} \, s^{-1} \sim 5 \times 10^{-4} \, s^{-1}$	[33, 107]
k_5	$3 \times 10^6 \, M^{-1} s^{-1} \sim 7 \times 10^6 \, M^{-1} s^{-1}$	[107]
k_{-5}	$5 \times 10^{-4} \, s^{-1} \sim 7 \times 10^{-3} \, s^{-1}$	[107]
k_6	$3 \times 10^6 \, M^{-1} s^{-1} \sim 7 \times 10^6 \, M^{-1} s^{-1}$	[107]
k_{-6}	$5 \times 10^{-4} \, s^{-1} \sim 7 \times 10^{-3} \, s^{-1}$	[107]
k_{fsn}	$1.5 \times 10^{-3} \, s^{-1} \sim 5 \times 10^{-3} \, s^{-1}$	[88, 112]
k_{fsnsm}	$1.5 \times 10^{-2} \, s^{-1} \sim 1 \times 10^{-1} \, s^{-1}$	[33, 88]
p_7	$3 \times 10^{-8} \, M \, s^{-1}$	Est. from [96,
p_2	$3 \times 10^{-8} \, M \, s^{-1}$	Est. from [96,
p_3	$3 \times 10^{-8} \, M \, s^{-1}$	[96, 102]
p_4	$8 \times 10^{-9} \, M \, s^{-1}$	Est. from [96,
u_1	$1 \times 10^{-6} \, s^{-1} \sim 3 \times 10^{-6} \, s^{-1}$	Est. from [96,
u_2	$1 \times 10^{-6} \, s^{-1} \sim 3 \times 10^{-6} \, s^{-1}$	Est. from [96,
u_3	$1 \times 10^{-6} \, s^{-1} \sim 3 \times 10^{-6} \, s^{-1}$	Est. from [96,
u_4	$1 \times 10^{-6} \, s^{-1} \sim 3 \times 10^{-6} \, s^{-1}$	Est. from [96,
u_5	$3 \times 10^{-7} \, s^{-1} \sim 6 \times 10^{-7} \, s^{-1}$	Est. from [96,
u_6	$3 \times 10^{-7} \, s^{-1} \sim 6 \times 10^{-7} \, s^{-1}$	Est. from [96,
u_7	$3 \times 10^{-7} \, s^{-1} \sim 6 \times 10^{-7} \, s^{-1}$	Est. from [96,
u_8	$3 \times 10^{-7} \, s^{-1} \sim 6 \times 10^{-7} \, s^{-1}$	Est. from [96,
u_9	$3 \times 10^{-7} \, s^{-1} \sim 6 \times 10^{-7} \, s^{-1}$	Est. from [96,

CHAPTER 5. CONCLUSION

5.1   Summary

This work proposes a framework for exploring the regulation mechanism of exocytotic membrane fusion and facilitating application of systems biology methodologies in exocytosis process.

Exocytosis is a key process in numerous biological systems. As innovative biotechnologies develop, many proteins are revealed in exocytosis process which is therefore shown in complex systems behaviors. It requires new methods to analyze such complex systems and explain experimental observations. We first use comparative network designing combining quantitative modeling to analyze the regulation of two key proteins families: SM and SNARE in exocytotic membrane fusion. It demonstrates that the topological structure of the SM-SNARE interaction network motif is an omnipresent factor that small change in structure of networks can dramatically modify the exocytotic system behaviors. It successfully reproduced current experimental observations and provided a bifurcation mechanism to explain conflicts observed in current studies. Therefore it provides an alternative way to bridge the consensus among SM-SNARE protein experimental observations and further analyze regulatory relationship between exocytotic proteins.

Secondly, to facilitate application similar research for biologists, we developed a software tool, CytoModeler. It provides easy way to design biological network and execute model simulation. More important, it provides a way to bridging two kinds of systems biology software tools: large scale network analysis and quantitative modeling. CytoModeler is built on large scale network analysis platform Cytoscape. Therefore, it can integrate all

functionality provided by Cytoscape and many functional plug-in tools, which are dedicated to modeling and analyzing large-scale network and Omics data. On the other hand, CytoModeler also provides built-in quantitative model creation, simulation function and provides a seamless connection with three popular quantitative modeling software tools: Systems Biology Toolbox, Systems Biology Workbench and COPASI. It enables user to take advantage of sophisticated modeling analysis functions for biological systems provided by these tools. Therefore, CytoModeler fuses these two major software tools in systems biology.

5.2   Limitation and future work

While SM-SNARE regulation plays a key role in controlling exocytotic membrane fusion, many other proteins are suggested important during the process, such as complexin, synaptotagmin, Munc13, we introduced in Chapter 1, and so on. This work is mainly just focused on SM and SNARE regulation and therefore further work which integrate more proteins with larger system context are needed. Also, the work suggested a bifurcation mechanism for explaining conflicting experiment observations for SM-SNARE regulation. Therefore, further experiments are needed to employ to test the predictions, such as designing mutant experimental system in vivo or in vitro to test our hypothesis. For CytoModeler, it provides a new way to visualize simulation results in network context. However, robustness and sensitivity of model are also key features for insuring correctly modeling and also potentially reflect design principle of target biological systems. Therefore, the effective visualization of robustness and sensitivity of model are needed.

REFERENCES CITED

1. Barabasi AL, Oltvai ZN: **Network biology: understanding the cell's functional organization.** *Nat Rev Genet* 2004, **5:**101-113.
2. Tama F, Sanejouand YH: **Conformational change of proteins arising from normal mode calculations.** *Protein Eng* 2001, **14:**1-6.
3. Ito T, Chiba T, Ozawa R, Yoshida M, Hattori M, Sakaki Y: **A comprehensive two-hybrid analysis to explore the yeast protein interactome.** *Proc Natl Acad Sci U S A* 2001, **98:**4569-4574.
4. Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P, et al: **A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae.** *Nature* 2000, **403:**623-627.
5. Shapira SD, Gat-Viks I, Shum BO, Dricot A, de Grace MM, Wu L, Gupta PB, Hao T, Silver SJ, Root DE, et al: **A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection.** *Cell* 2009, **139:**1255-1267.
6. Rual JF, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, Li N, Berriz GF, Gibbons FD, Dreze M, Ayivi-Guedehoussou N, et al: **Towards a proteome-scale map of the human protein-protein interaction network.** *Nature* 2005, **437:**1173-1178.
7. Li S, Armstrong CM, Bertin N, Ge H, Milstein S, Boxem M, Vidalain PO, Han JD, Chesneau A, Hao T, et al: **A map of the interactome network of the metazoan C. elegans.** *Science* 2004, **303:**540-543.
8. Ge H, Walhout AJ, Vidal M: **Integrating 'omic' information: a bridge between genomics and systems biology.** *Trends Genet* 2003, **19:**551-560.
9. Hartwell LH, Hopfield JJ, Leibler S, Murray AW: **From molecular to modular cell biology.** *Nature* 1999, **402:**C47-52.
10. Sharan R, Ideker T: **Modeling cellular machinery through biological network comparison.** *Nat Biotechnol* 2006, **24:**427-433.
11. Koyuturk M, Kim Y, Topkara U, Subramaniam S, Szpankowski W, Grama A: **Pairwise alignment of protein interaction networks.** *J Comput Biol* 2006, **13:**182-199.
12. Pinter RY, Rokhlenko O, Yeger-Lotem E, Ziv-Ukelson M: **Alignment of metabolic pathways.** *Bioinformatics* 2005, **21:**3401-3408.
13. Kelley BP, Sharan R, Karp RM, Sittler T, Root DE, Stockwell BR, Ideker T: **Conserved pathways within bacteria and yeast as revealed by global protein network alignment.** *Proc Natl Acad Sci U S A* 2003, **100:**11394-11399.
14. Sharan R, Suthram S, Kelley RM, Kuhn T, McCuine S, Uetz P, Sittler T, Karp RM, Ideker T: **Conserved patterns of protein interaction in multiple species.** *Proc Natl Acad Sci U S A* 2005, **102:**1974-1979.
15. O'Donoghue SI, Gavin AC, Gehlenborg N, Goodsell DS, Heriche JK, Nielsen CB, North C, Olson AJ, Procter JB, Shattuck DW, et al: **Visualizing biological data-now and in the future.** *Nat Methods*, **7:**S2-4.

16. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: **Cytoscape: a software environment for integrated models of biomolecular interaction networks.** *Genome Res* 2003, **13:**2498-2504.

17. Gehlenborg N, O'Donoghue SI, Baliga NS, Goesmann A, Hibbs MA, Kitano H, Kohlbacher O, Neuweger H, Schneider R, Tenenbaum D, Gavin AC: **Visualization of omics data for systems biology.** *Nat Methods*, **7:**S56-68.

18. Westerhoff HV, Palsson BO: **The evolution of molecular biology into systems biology.** *Nat Biotechnol* 2004, **22:**1249-1252.

19. Mogilner A, Wollman R, Marshall WF: **Quantitative modeling in cell biology: what is it good for?** *Dev Cell* 2006, **11:**279-287.

20. de Jong H: **Modeling and simulation of genetic regulatory systems: a literature review.** *J Comput Biol* 2002, **9:**67-103.

21. Xia T, Dickerson JA: **OmicsViz: Cytoscape plug-in for visualizing omics data across species.** *Bioinformatics* 2008, **24:**2557-2558.

22. Xia T, Dickerson AJ: **OmicsAnalyzer: a Cytoscape plug-in suite for modeling omics data.** *To be submitted to Journal of Bioinformatics* 2010.

23. Ramsey S, Orrell D, Bolouri H: **ISBJava available at http://magnet.systemsbiology.net/software/ISBJava/.**

24. Joyce AR, Palsson BO: **The model organism as a system: integrating 'omics' data sets.** *Nat Rev Mol Cell Biol* 2006, **7:**198-210.

25. Hooper SD, Bork P: **Medusa: a simple tool for interaction graph analysis.** *Bioinformatics* 2005, **21:**4432-4433.

26. Breitkreutz BJ, Stark C, Tyers M: **Osprey: a network visualization system.** *Genome Biol* 2003, **4:**R22.

27. Freeman TC, Goldovsky L, Brosch M, van Dongen S, Maziere P, Grocock RJ, Freilich S, Thornton J, Enright AJ: **Construction, visualisation, and clustering of transcription networks from microarray expression data.** *PLoS Comput Biol* 2007, **3:**2032-2042.

28. Iragne F, Nikolski M, Mathieu B, Auber D, Sherman D: **ProViz: protein interaction visualization and exploration.** *Bioinformatics* 2005, **21:**272-274.

29. Demir E, Babur O, Dogrusoz U, Gursoy A, Nisanci G, Cetin-Atalay R, Ozturk M: **PATIKA: an integrated visual environment for collaborative construction and analysis of cellular pathways.** *Bioinformatics* 2002, **18:**996-1003.

30. Orlev N, Shamir R, Shiloh Y: **PIVOT: protein interacions visualizatiOn tool.** *Bioinformatics* 2004, **20:**424-425.

31. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, Christmas R, Avila-Campilo I, Creech M, Gross B, et al: **Integration of biological networks and gene expression data using Cytoscape.** *Nat Protoc* 2007, **2:**2366-2382.

32. Shen L, Gong J, Caldo RA, Nettleton D, Cook D, Wise RP, Dickerson JA: **BarleyBase--an expression profiling database for plant genomics.** *Nucleic Acids Res* 2005, **33:**D614-618.

33. Wise RP, Caldo RA, Hong L, Shen L, Cannon E, Dickerson JA: **BarleyBase/PLEXdb: A Unified Expression Profiling Database for Plants and Plant Pathogens.** *Methods Mol Biol* 2007, **406:**347-364.

34. Gilbert D, Morgner T: **JFreeChart, a free java class library for generating charts.**

35. Stuart JM, Segal E, Koller D, Kim SK: **A gene-coexpression network for global discovery of conserved genetic modules.** *Science* 2003, **302:**249-255.

36. Junker BH, Klukas C, Schreiber F: **VANTED: a system for advanced data analysis and visualization in the context of biological networks.** *BMC Bioinformatics* 2006, **7:**109.

37. Butts CT: **Revisiting the foundations of network analysis.** *Science* 2009, **325:**414-416.

38. Barabasi AL: **Scale-free networks: a decade and beyond.** *Science* 2009, **325:**412-413.

39. Jourdan F, Breitling R, Barrett MP, Gilbert D: **MetaNetter: inference and visualization of high-resolution metabolomic networks.** *Bioinformatics* 2008, **24:**143-145.

40. Ludemann A, Weicht D, Selbig J, Kopka J: **PaVESy: Pathway Visualization and Editing System.** *Bioinformatics* 2004, **20:**2841-2844.

41. Suderman M, Hallett M: **Tools for visually exploring biological networks.** *Bioinformatics* 2007, **23:**2651-2659.

42. Byrnes RW, Cotter D, Maer A, Li J, Nadeau D, Subramaniam S: **An editor for pathway drawing and data visualization in the Biopathways Workbench.** *BMC Syst Biol* 2009, **3:**99.

43. Killcoyne S, Carter GW, Smith J, Boyle J: **Cytoscape: a community-based framework for network modeling.** *Methods Mol Biol* 2009, **563:**219-239.

44. www.cytoscape.org.

45. Ferro A, Giugno R, Pigola G, Pulvirenti A, Skripin D, Bader GD, Shasha D: **NetMatch: a Cytoscape plugin for searching biological networks.** *Bioinformatics* 2007, **23:**910-912.

46. Barsky A, Gardy JL, Hancock RE, Munzner T: **Cerebral: a Cytoscape plugin for layout of and interaction with biological networks using subcellular localization annotation.** *Bioinformatics* 2007, **23:**1040-1042.

47. Vlasblom J, Wu S, Pu S, Superina M, Liu G, Orsi C, Wodak SJ: **GenePro: a Cytoscape plug-in for advanced visualization and analysis of interaction networks.** *Bioinformatics* 2006, **22:**2178-2179.

48. Yang L, Walker JR, Hogenesch JB, Thomas RS: **NetAtlas: a Cytoscape plugin to examine signaling networks based on tissue gene expression.** *In Silico Biol* 2008, **8:**47-52.

49. Maere S, Heymans K, Kuiper M: **BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks.** *Bioinformatics* 2005, **21:**3448-3449.

50. Baumbach J, Wittkop T, Emig D: **BLAST2SimilarityGraph.** 2009.

51. Morris JH, Huang CC, Babbitt PC, Ferrin TE: **structureViz: linking Cytoscape and UCSF Chimera.** *Bioinformatics* 2007, **23:**2345-2347.

52. Glass L, Kauffman SA: **The logical analysis of continuous, non-linear biochemical control networks.** *J Theor Biol* 1973, **39:**103-129.

53. Kacser H, Burns JA: **The control of flux.** *Symp Soc Exp Biol* 1973, **27:**65-104.

54. Reder C: **Metabolic control theory: a structural approach.** *J Theor Biol* 1988, **135:**175-201.

55. Endler L, Rodriguez N, Juty N, Chelliah V, Laibe C, Li C, Le Novere N: **Designing and encoding models for synthetic biology.** *J R Soc Interface* 2009, **6 Suppl 4:**S405-417.

56. Schmidt H, Jirstrand M: **Systems Biology Toolbox for MATLAB: a computational platform for research in systems biology.** *Bioinformatics* 2006, **22:**514-515.

57. Hoops S, Sahle S, Gauges R, Lee C, Pahle J, Simus N, Singhal M, Xu L, Mendes P, Kummer U: **COPASI--a COmplex PAthway SImulator.** *Bioinformatics* 2006, **22:**3067-3074.

58. Sauro HM, Hucka M, Finney A, Wellock C, Bolouri H, Doyle J, Kitano H: **Next generation simulation tools: the Systems Biology Workbench and BioSPICE integration.** *Omics* 2003, **7:**355-372.

59. Sauro HM: **Jarnac: a system for interactive metabolic analysis.** *In J-H Hofmeyr, J M*

Rohwer, and J L Snoep, editors Animating the Cellular Map: Proceedings of the 9th International Meeting on BioThermoKinetics, Stellenbosch University Press 2000.

60. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, Arkin AP, Bornstein BJ, Bray D, Cornish-Bowden A, et al: **The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models.** *Bioinformatics* 2003, **19:**524-531.

61. Assenov Y, Ramirez F, Schelhorn SE, Lengauer T, Albrecht M: **Computing topological parameters of biological networks.** *Bioinformatics* 2008, **24:**282-284.

62. Ausbrooks R, Stephen Buswell, David Carlisle, Stéphane Dalmas, Stan Devitt, Angel Diaz, Max Froumentin, Roger Hunter, Patrick Ion, Michael Kohlhase, Robert Miner, Nico Poppelier, Bruce Smith, Neil Soiffer, Robert Sutor, Stephen Watt: **Mathematical Markup Language (MathML) Version 2.0 (2nd Edition).** *World Wide Web Consortium* 2003.

63. Funahashi A, Jouraku A, Matsuoka Y, Kitano H: **Integration of CellDesigner and SABIO-RK.** *In Silico Biol* 2007, **7:**S81-90.

64. Le Novere N, Hucka M, Mi H, Moodie S, Schreiber F, Sorokin A, Demir E, Wegner K, Aladjem MI, Wimalaratne SM, et al: **The Systems Biology Graphical Notation.** *Nat Biotechnol* 2009, **27:**735-741.

65. Jahn R, Scheller RH: **SNAREs--engines for membrane fusion.** *Nat Rev Mol Cell Biol* 2006, **7:**631-643.

66. Rothman JE: **Mechanisms of intracellular protein transport.** *Nature* 1994, **372:**55-63.

67. Geppert M, Sudhof TC: **RAB3 and synaptotagmin: the yin and yang of synaptic membrane fusion.** *Annu Rev Neurosci* 1998, **21:**75-95.

68. Sudhof TC: **Membrane fusion as a team effort.** *Proc Natl Acad Sci U S A* 2007, **104:**13541-13542.

69. Burgoyne RD, Morgan A: **Membrane trafficking: three steps to fusion.** *Curr Biol* 2007, **17:**R255-258.

70. Koh TW, Bellen HJ: **Synaptotagmin I, a Ca2+ sensor for neurotransmitter release.** *Trends Neurosci* 2003, **26:**413-422.

71. Wong JL, Koppel DE, Cowan AE, Wessel GM: **Membrane hemifusion is a stable intermediate of exocytosis.** *Dev Cell* 2007, **12:**653-659.

72. Giraudo CG, Eng WS, Melia TJ, Rothman JE: **A clamping mechanism involved in SNARE-dependent exocytosis.** *Science* 2006, **313:**676-680.

73. Schaub JR, Lu X, Doneske B, Shin YK, McNew JA: **Hemifusion arrest by complexin is relieved by Ca2+-synaptotagmin I.** *Nat Struct Mol Biol* 2006, **13:**748-750.

74. Tang J, Maximov A, Shin OH, Dai H, Rizo J, Sudhof TC: **A complexin/synaptotagmin 1 switch controls fast synaptic vesicle exocytosis.** *Cell* 2006, **126:**1175-1187.

75. Jackson MB, Chapman ER: **Fusion pores and fusion machines in Ca2+-triggered exocytosis.** *Annu Rev Biophys Biomol Struct* 2006, **35:**135-160.

76. Morciano M, Burre J, Corvey C, Karas M, Zimmermann H, Volknandt W: **Immunoisolation of two synaptic vesicle pools from synaptosomes: a proteomics analysis.** *J Neurochem* 2005, **95:**1732-1745.

77. Burre J, Volknandt W: **The synaptic vesicle proteome.** *J Neurochem* 2007, **101:**1448-1462.

78. Takamori S, Holt M, Stenius K, Lemke EA, Gronborg M, Riedel D, Urlaub H, Schenck S, Brugger B, Ringler P, et al: **Molecular anatomy of a trafficking organelle.** *Cell* 2006, **127:**831-846.

79. Jahn R, Lang T, Sudhof TC: **Membrane fusion.** *Cell* 2003, **112:**519-533.

80. Cusick ME, Klitgord N, Vidal M, Hill DE: **Interactome: gateway into systems biology.** *Hum Mol Genet* 2005, **14 Spec No. 2:**R171-181.

81. Rizo J, Sudhof TC: **Snares and Munc18 in synaptic vesicle fusion.** *Nat Rev Neurosci* 2002, **3:**641-653.

82. Toonen RF, Verhage M: **Munc18-1 in secretion: lonely Munc joins SNARE team and takes control.** *Trends Neurosci* 2007, **30:**564-572.

83. Dacks JB, Field MC: **Evolution of the eukaryotic membrane-trafficking system: origin, tempo and mode.** *J Cell Sci* 2007, **120:**2977-2985.

84. Chen X, Lu J, Dulubova I, Rizo J: **NMR analysis of the closed conformation of syntaxin-1.** *J Biomol NMR* 2008, **41:**43-54.

85. Wu MN, Littleton JT, Bhat MA, Prokop A, Bellen HJ: **ROP, the Drosophila Sec1 homolog, interacts with syntaxin and regulates neurotransmitter release in a dosage-dependent manner.** *Embo J* 1998, **17:**127-139.

86. Scott BL, Van Komen JS, Irshad H, Liu S, Wilson KA, McNew JA: **Sec1p directly stimulates SNARE-mediated membrane fusion in vitro.** *J Cell Biol* 2004, **167:**75-85.

87. Wu MN, Schulze KL, Lloyd TE, Bellen HJ: **The ROP-syntaxin interaction inhibits neurotransmitter release.** *Eur J Cell Biol* 2001, **80:**196-199.

88. Tareste D, Shen J, Melia TJ, Rothman JE: **SNAREpin/Munc18 promotes adhesion and fusion of large vesicles to giant membranes.** *Proc Natl Acad Sci U S A* 2008, **105:**2380-2385.

89. Schutz D, Zilly F, Lang T, Jahn R, Bruns D: **A dual function for Munc-18 in exocytosis of PC12 cells.** *Eur J Neurosci* 2005, **21:**2419-2432.

90. Gerber SH, Rah JC, Min SW, Liu X, de Wit H, Dulubova I, Meyer AC, Rizo J, Arancillo M, Hammer RE, et al: **Conformational switch of syntaxin-1 controls synaptic vesicle fusion.** *Science* 2008, **321:**1507-1510.

91. Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U: **Network motifs: simple building blocks of complex networks.** *Science* 2002, **298:**824-827.

92. Schmid EM, McMahon HT: **Integrating molecular and network biology to decode endocytosis.** *Nature* 2007, **448:**883-888.

93. Togneri J, Cheng YS, Munson M, Hughson FM, Carr CM: **Specific SNARE complex binding mode of the Sec1/Munc-18 protein, Sec1p.** *Proc Natl Acad Sci U S A* 2006, **103:**17730-17735.

94. Sudhof TC, Rothman JE: **Membrane fusion: grappling with SNARE and SM proteins.** *Science* 2009, **323:**474-477.

95. Hua Y, Scheller RH: **Three SNARE complexes cooperate to mediate membrane fusion.** *Proc Natl Acad Sci U S A* 2001, **98:**8065-8070.

96. Mezer A, Nachliel E, Gutman M, Ashery U: **A new platform to study the molecular mechanisms of exocytosis.** *J Neurosci* 2004, **24:**8838-8846.

97. Wuchty S, Oltvai ZN, Barabasi AL: **Evolutionary conservation of motif constituents in the yeast protein interaction network.** *Nat Genet* 2003, **35:**176-179.

98. Schulze KL, Littleton JT, Salzberg A, Halachmi N, Stern M, Lev Z, Bellen HJ: **rop, a Drosophila homolog of yeast Sec1 and vertebrate n-Sec1/Munc-18 proteins, is a negative regulator of neurotransmitter release in vivo.** *Neuron* 1994, **13:**1099-1108.

99. Mitchell SJ, Ryan TA: **Munc18-dependent regulation of synaptic vesicle exocytosis by syntaxin-1A in hippocampal neurons.** *Neuropharmacology* 2005, **48:**372-380.

100. Ma W, Lai L, Ouyang Q, Tang C: **Robustness and modular design of the Drosophila segment polarity network.** *Mol Syst Biol* 2006, **2:**70.

101. Sudhof TC: **The synaptic vesicle cycle.** *Annu Rev Neurosci* 2004, **27:**509-547.

102. Lang T, Bruns D, Wenzel D, Riedel D, Holroyd P, Thiele C, Jahn R: **SNAREs are concentrated in cholesterol-dependent clusters that define docking and fusion sites for exocytosis.** *Embo J* 2001, **20:**2202-2213.

103. Ashery U, Varoqueaux F, Voets T, Betz A, Thakur P, Koch H, Neher E, Brose N, Rettig J: **Munc13-1 acts as a priming factor for large dense-core vesicles in bovine chromaffin cells.** *Embo J* 2000, **19:**3586-3596.

104. Rickman C, Meunier FA, Binz T, Davletov B: **High affinity interaction of syntaxin and SNAP-25 on the plasma membrane is abolished by botulinum toxin E.** *J Biol Chem* 2004, **279:**644-651.

105. Weninger K, Bowen ME, Chu S, Brunger AT: **Single-molecule studies of SNARE complex assembly reveal parallel and antiparallel configurations.** *Proc Natl Acad Sci U S A* 2003, **100:**14800-14805.

106. Pevsner J, Hsu SC, Braun JE, Calakos N, Ting AE, Bennett MK, Scheller RH: **Specificity and regulation of a synaptic vesicle docking complex.** *Neuron* 1994, **13:**353-361.

107. Burkhardt P, Hattendorf DA, Weis WI, Fasshauer D: **Munc18a controls SNARE assembly through its interaction with the syntaxin N-peptide.** *Embo J* 2008, **27:**923-933.

108. Nicholson KL, Munson M, Miller RB, Filip TJ, Fairman R, Hughson FM: **Regulation of SNARE complex assembly by an N-terminal domain of the t-SNARE Sso1p.** *Nat Struct Biol* 1998, **5:**793-802.

109. Pobbati AV, Stein A, Fasshauer D: **N- to C-terminal SNARE complex assembly promotes rapid membrane fusion.** *Science* 2006, **313:**673-676.

110. Margittai M, Widengren J, Schweinberger E, Schroder GF, Felekyan S, Haustein E, Konig M, Fasshauer D, Grubmuller H, Jahn R, Seidel CA: **Single-molecule fluorescence resonance energy transfer reveals a dynamic equilibrium between closed and open conformations of syntaxin 1.** *Proc Natl Acad Sci U S A* 2003, **100:**15516-15521.

111. Fasshauer D, Margittai M: **A transient N-terminal interaction of SNAP-25 and syntaxin nucleates SNARE assembly.** *J Biol Chem* 2004, **279:**7613-7621.

112. Lu X, Zhang F, McNew JA, Shin YK: **Membrane fusion induced by neuronal SNAREs transits through hemifusion.** *J Biol Chem* 2005, **280:**30538-30541.