

DMSP – Database for Modeling Signaling Pathways

Combining Biological and Mathematical Modeling Knowledge for Pathways

M. Visvanathan¹, M. Breit¹, B. Pfeifer¹, C. Baumgartner¹, R. Modre-Osprian², B. Tilg¹

¹Institute for Biomedical Engineering, University for Health Sciences, Medical Informatics and Technology, Hall in Tyrol, Austria

²ARC Seibersdorf research GmbH, Hall in Tyrol, Austria

Summary

Objectives: Presently, the protein interaction information concerning different signaling pathways is available in a qualitative manner in different online protein interaction databases. The challenge here is to derive a quantitative way of modeling signaling pathways from a qualitative way of modeling signaling pathways from a qualitative level. To address this issue we developed a database that includes mathematical modeling knowledge and biological knowledge about different signaling pathways.

Methods: The database is part of an integrative environment that includes environments for pathway design, visualization, simulation and a knowledge base that combines biological and modeling information concerning pathways. The system is designed as a client-server architecture. It contains a pathway designing environment and a simulation environment as upper layers with a relational knowledge base as the underlying layer.

Results: DMSP – Database for Modeling Signaling Pathways incorporates biological datasets from online databases like BIND, DIP, PIP, and SPiD. The modeling knowledge that has been incorporated is based on a literature study. Pathway models can be designed, visualized and simulated based on the knowledge stored in the DMSP. The user can download the whole dataset and build pathway models using the knowledge stored in our database. As an example, the TNF α pathway model was implemented and tested using this approach.

Conclusion: DMSP is an initial step towards the aim of combining modeling and biological knowledge concerning signaling pathways. It helps in understanding pathways in a qualitative manner from a qualitative level. Simulation results enable the interpretation of a biological system from a quantitative and system-theoretic point of view.

Keywords

Database, modeling, signaling and pathways

Methods Inf Med 2008; 47: 140–148

doi:10.3414/ME0461

1. Objectives

Signal transduction plays a central role in cellular activities. Cellular functions are controlled by dynamic patterns of interaction among signaling proteins that arise when a cell surface receptor detects a change in its environment [1]. The signaling proteins control the huge diversity of physiological responses in cells. Thus, improved understanding of cellular signaling has a number of potential and practical applications, from rational design of drugs to vaccines in the engineering of cells for biotechnology purpose [2].

Currently data integration is a growing field in the application of database services in systems biology. A quantitative database including mathematical modeling and biological data about different signaling pathways will significantly help in understanding these pathways. This will provide a valuable insight into the mechanisms of interactions among proteins, by identifying specific components and reactions in signaling pathways [3].

The biological pathways are split mainly into metabolic and biochemical, transcription, regulation and protein synthesis and signal transduction. These pathways have distinct attributes, to be kept and managed in several database systems. The signal transduction pathways are responsible for coordinating metabolic processes with transcription and protein synthesis.

Currently there are several databases [4] like DIP [5], BIND [6], PIP [7] and SPiD [8] that store only biological information about proteins, genes and their interactions. These

databases however do not provide mathematical modeling information. There are also databases like DOQCS [9] that maintain mathematical models of cellular signaling pathways without biological knowledge.

Software platforms allowing the modeling and execution of biomedical applications such as analysis and management of proteomics experiments [28] are also important to manage the complexity of biomedical experiments. In systems biology synergistic applications of experiments, theory and modeling towards understanding biological processes as a whole system ultimately require an integrative database in order to express and correlate queries. To address this issue we built up a database called DMSP (<https://sourceforge.net/projects/dmsp11>) that combines modeling and biological data. Here, the modeling data includes kinetic constants, kinetic equations and initial concentrations that are used to build mathematical models. The biological data includes descriptions about the proteins and their interaction and other information concerning signaling pathways. Therefore, the DMSP is part of an integrative environment where the data can be retrieved and analyzed in a visualization and simulation environment.

The paper is organized as follows: Section 2 presents an overview of the integrative environment including the database, the conceptual idea behind the data model, the data integration process and the data handling of the database. Section 3 describes the structure of the database and demonstrates results experimented on a TNF α pathway [15]. In Section 4 we discuss this approach and delineate further activities.

2. Methods Involved in Data Integration and Handling

DMSP is part of an integrative environment that was designed as a client-server system (Fig. 1). It contains a Pathway Designing-Visualization Environment (PDVE) and Pathway Simulation Environment (PSE) as upper layers coupled with a relational database as the database layer called Database for Modeling Signaling Pathways (DMSP). This integrative environment enables us to design a signaling pathway cartoon model using the pathway designing-visualization environment. The pathway model developed in this environment incorporates biological and modeling datasets from the database. Furthermore, the model can be exported to the simulation module as an XML-based model. Only the information required for simulating the signaling pathway is extracted from the XML-based model. Finally, in the simulation environment the mathematical model of the pathway can be simulated and analyzed.

The database technology that is being used is MS-SQL Server 2000 (<http://www.microsoft.com/sql>). The Vector PathBlazer [10] acts as the pathway visualization and pathway designing environment. The simulation is done in a module developed under MATLAB® version 7.0 (<http://www.mathworks.com/>).

Online protein interaction databases like BIND [6] or SPiD [8] provide information about the proteins and genes including their names, its synonyms, or others. These databases also provide describing information of different reactions that occur between different proteins. The DMSP database schema now incorporates biological datasets about the components, the reactions, and the pathways from the above mentioned online repositories. The components can further be divided into proteins, DNA, sub-molecules and others. Those components are involved in different reactions building different protein complexes. The protein complexes based on protein-protein interactions within a pathway are also maintained in DMSP.

In general, mathematical models of signaling pathways use differential equations

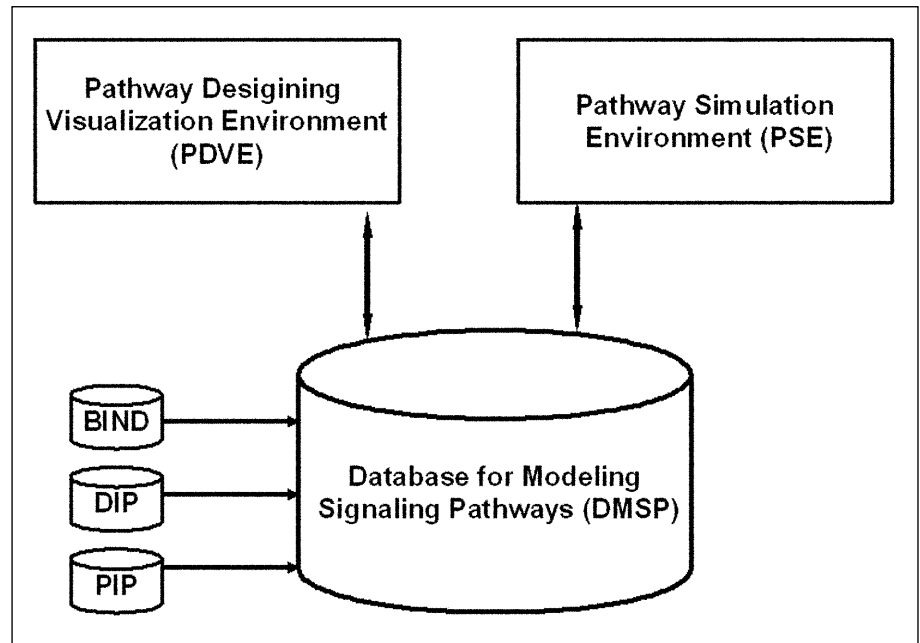


Fig. 1 The architecture of the system was designed as a 3-tier system that includes a pathway designing-visualization environment and simulation environment as the front-end layers coupled with a relational database as the back-end layer.

relying on fundamental assumptions. Kinetics of enzymatic reactions for instance can be modeled based on the Michaelis-Menten model of enzyme kinetics. A signal transduction system usually behaves as a slowly time-varying non-linear system during a reaction period based on biological observations.

Further it is assumed that a cell keeps the concentration of each signaling protein constant before and after each signaling, that is, the concentration of these proteins returns to a steady state after the reaction. Through these methods of enzyme kinetics the ordinary differential equations (ODEs) for the mathematical models are derived and stored for simulation of signaling pathways. Considering the fact that the steady state of enzymes or proteins depends upon the local environment in the cell, there are two possibilities how to consider these dynamics in a mathematical model. The first possibility is to model the rates of generation and degradation of the enzymes or proteins as a function of time. The role of these rates is to maintain the steady-state concentration in the model. Figure 2 gives a graphical illustration of how the kinetic data are considered for simulation based on a pathway

model designed in PDVE. A circle represents a state for the concentration, a bar represents a rate of reaction, and directed arrows connect the circles and the bars. Here the bars represent the rate of formation and breakdown of the complex or product. The relations between the rate of catalysis and the change of concentration for the substrate, the enzyme, the complex, and the product are represented by the set of ordinary differential equations (ODEs). It gives an appropriate function for a dynamic description of the time-dependent change of concentrations of the involved components.

In this example an enzyme (E) combines with a substrate (S) to produce an enzyme-substrate (ES) complex and this complex dissociates to produce a product (P) and an enzyme. This process of association and disassociation is controlled by a set of kinetics that is involved and is described in the form of ODEs. The system of differential equations explains the process of association and disassociation of proteins and protein complexes for this model, simulation and analysis of the mathematical model is performed by applying standard numerical integration methods [11]. The main goal is that the modeling information of the reac-

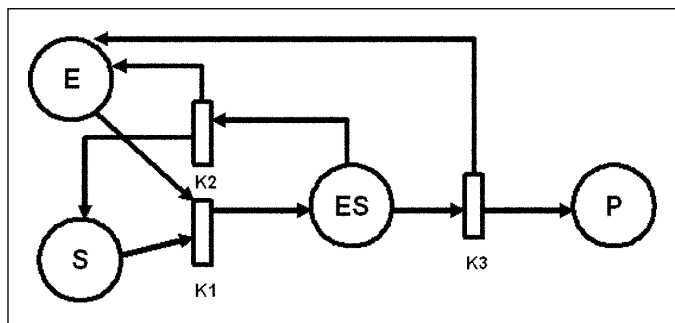


Fig. 2
A graphical illustration of a simple enzyme kinetic reaction model

tion is now combined with the biological information of the proteins and interactions.

The ontology of modeling pathway [12] is considered by the modeling dataset of our proposed database. The curation of the modeling dataset is based on a literature analysis concerning mathematical models of signaling pathways. This literature analysis was conducted by us. In the following, the main modeling elements that are required to model signaling pathways were identified. Based on this analysis we built the modeling dataset of DMSP. In all the modeling dataset includes kinetic constants, kinetic equations and initial concentrations. The dataset is related to the components and to the reactions of different signaling pathways. Note that the modeling knowledge also has to consider the complex building for signaling pathways. So we identified the required modeling elements from literature analysis for different signaling pathways. This process cannot be done automatically. The modeling knowledge needs to be carefully extracted from literature about different signaling pathway models.

The quantitative level of understanding signaling pathways is currently a new dimension in molecular cell biology. One of the prominent challenges is to integrate protein interaction data for various signaling pathways with the mathematical models concerning them. DMSP database curation and development is being done along this line. So we built quantitative models by integrating qualitative data that is obtained from protein interaction databases and primary literature studies concerning various mathematical models for signaling pathways.

The protein interaction data obtained from various databases provides us with the qualitative information related to various proteins with their respective interactions for

signaling pathways. In order to understand the deeper mechanism between protein interaction data for signaling pathways we need more mathematical information like its kinetic rate laws, etc. associated with the protein interaction maps. This will help us to clearly understand the role and significance of protein interactions. Because there is less kinetic information, not all of the kinetic parameters are available to be downloaded directly from the literature studies. For this purpose, literature curation was done to identify parameters for specific pathways and we implemented them in our system and studied their behavior. Hence we were able to get more quantitative information for different signaling pathways. We also have to consider the point that the quantitative information related to protein interactions for signaling pathways is much lesser than that of metabolic pathways. So we have to build up the quantitative knowledge step by step for signaling pathway models from a qualitative level. More and more we are able to describe quantitative protein interactions, aiming to a better understanding of signaling pathways. Therefore this procedure improves the way of studying protein interaction maps and helps us to validate the model by comparing the results with the experimental data for a concerned pathway.

The curation of the biological datasets is based on the data presented in different protein interaction databases. The information associated with each protein and its reaction information was obtained from aforementioned online databases which are available as XML files [13] or in Flat File format. These formats provide an effective way through which specific datasets can be extracted according to the DMSP specification and integrated into the database. The

datasets are downloaded individually from their respective data sources in a specified format. These downloaded files are stored in dummy tables based on their XML schema specification. Later on they are converted by matching the individual attributes with respect to our database schema specification and integrated into our system. The logical implementation and the process involved in integrating the biological knowledge based on external data sources into DMSP were published recently [27].

In order to make sure that the information is not duplicated during data integration specific verifications are done. For example, protein interaction data sets are normally downloaded e.g. based on the component and reaction specification from databases like BIND or DIP. At first they are stored in a temporary storing area (dummy tables) and then they are converted according to the respective specification of concerned tables and implemented into DMSP. Usually when there is a new interaction data set to be implemented, it is checked based on its unique identification. If the interaction is not present then it is implemented instantly. If the interaction has already been incorporated into the database, then it is checked based on the protein's synonym name, protein sequence and its experimental conditions. If the experimental condition matches our need then it is stored, otherwise it is ignored.

Currently we try to combine the most prominent protein interaction databases like BIND, DIP, and PIP. But it is also possible to combine various other existing databases like KEGG, Rctome and MIPS. If these databases were incorporated they would be converted according to DMSP specification and all the interactions would be checked based upon the experimental conditions.

Currently the database considers mathematical models concerning the EGF, Wnt, NF κ B and TNF α pathways. These mathematical models have been combined with their respective biological knowledge integrated from the online protein interaction databases. Before we developed the database schema we had done a detailed study about TNF α and its related protein interaction maps. Then we looked at mathematical models concerning the TNF α pathway,

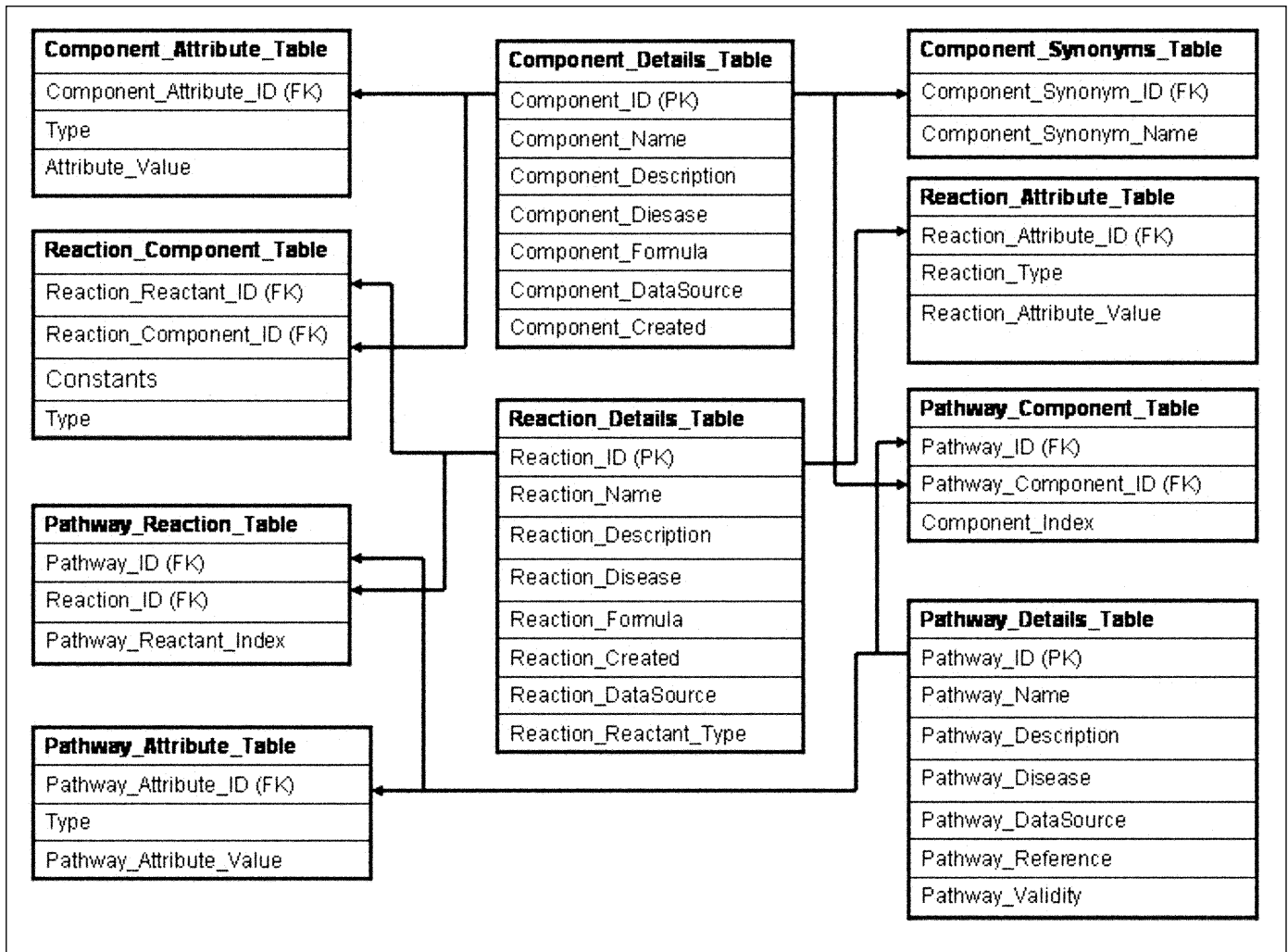


Fig. 3 An instance of the database schema representation. The database consists of tables that store the component, reaction, and pathway data that includes biological and modeling information.

later we started to develop the DMSP schema combining biological knowledge related to protein interactions as well as storing the basic modeling knowledge required to do simulations. Since the TNF α pathway contains various complex components like NF κ B which play an important role in various other pathways like EGF, Wnt, NF κ B, we also included these specific pathways.

3. Results

The DMSP database schema was designed to represent different levels of data description. The core of the database schema is based on the workspace of the currently

used designing-visualization environment. They are mainly classified into three distinct levels of entities: components, reactions and pathways (Fig. 3). Hence, the tables in the database are separated into component table entities, reaction table entities and pathway table entities. These database table entities inherit both the biological and modeling information from a biological and modeling point of view for signaling pathways. Each entry in these tables includes an identifier for identification of components and its involvement in a reaction and their occurrence in a pathway.

In the database the component table entities (*Component_Details_Table*, *Component_Attributes_Table*, *Component_Synonyms_Table*) store a wide range of in-

formation that includes biological descriptions about the components. A component can be a protein, a gene, a DNA, a sub-molecule, etc. A component can take part in a disease that is stored in *Component_Details_Table* as a *Component_Disease*-column entity. The *Component_Formula* column entity of the *Component_Details_Table* stores the chemical formula or protein complex information. A component can also have one or more synonyms and this information is stored in the *Component_Synonym_Table*. Component table entities also store information about the components involvement in different reactions for different pathways. Each entry in these tables includes a unique identifier for the component.

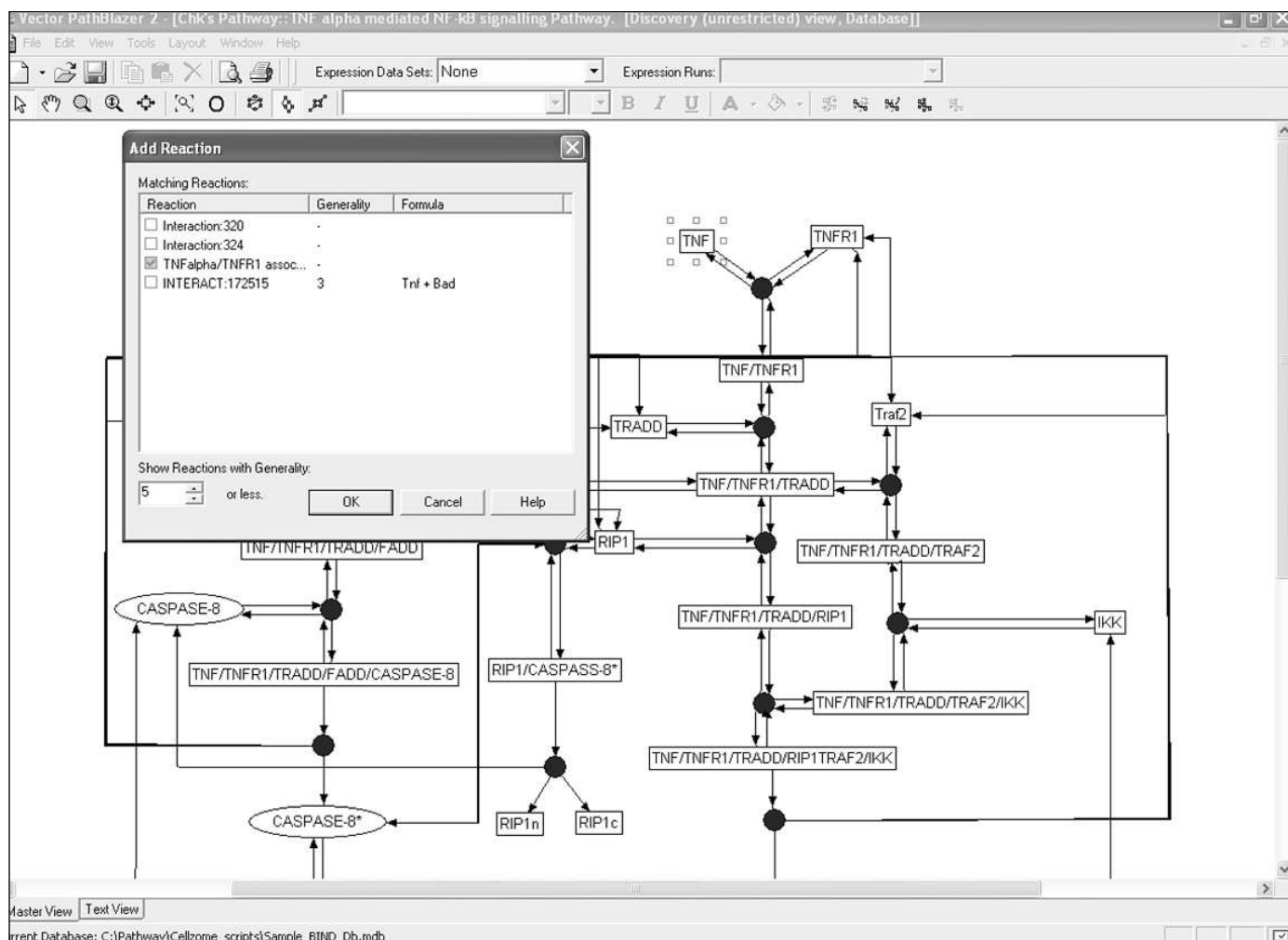


Fig. 4 A TNF α pathway model that was implemented and designed using the designing-visualization environment based on the information about the pathway stored in the database

The reaction table entities contain further entry specification details relating to reactions. The reaction tables include information about the components that are involved in a particular reaction for a pathway. The *Reaction_Details_Table* table can store the name of the reaction and the reaction formula. In particular, the *Reaction_Attribute_Table* considers different types of attributes relating to a particular reaction. For instance, one or more components can take part in a reaction with various kinetic constants. Every reaction is identified by a unique reaction identifier.

The tables for storing pathways are the next level of organization. These tables contain information about different signaling pathway models. They are identified by

using a specific pathway identifier that differentiates between different models. These tables include the structure of the pathway that contains different reactions and the set of components that are involved in a reaction of a pathway.

As an example we present the TNF α pathway that was used to test the requirements of our proposed integrative DMSP database. The biological data were extracted from the online databases DIP and BIND and integrated in our database with respect to the TNF α pathway. The modeling data that were integrated in the database was extracted from literature studies of [14, 15].

The TNF α pathway model implemented in the database contains 31 components that include protein and protein complexes.

These components are involved in 19 different reactions of the pathway. We stored the biological description about the protein, protein complexes and genes that take part in the TNF α signaling pathway. This information is stored in the components table entities. For example cIAP is a protein that takes part in TNF α pathway. Therefore, we implemented the protein information of the cIAP protein into our database by gathering specific information from the online databases.

The reactions that occur in the pathway model are represented in the reaction tables together with the components that participate in the reaction. The kinetic equations as well as kinetic constants that were implemented for a particular reaction of the TNF α

4. Discussion and Future Plans

There are many database projects with wide diversity and objectives that deal with signaling pathways. The database projects like DOQCS [9] and TRANSPATH [16] have extensive information about chemical interactions in signaling pathways. The missing objective was to facilitate modeling data and simulation results with biological data. More and more modeling data with biological data are essential for a better understanding of signaling pathways. Protein interaction databases DIP and BIND for instance provide protein interaction data about different signaling pathways.

Unfortunately, these databases do not have modeling information that relates to the signaling pathways. In contrast, GeneNet [17] or AFCS (Alliance for Cellular Signaling) [18] have stated to include quantitative pathway models. They are excellent at providing enzyme level information. In this context, the protein interaction databases try to provide qualitative datasets, but they are not efficient at describing mathematical modeling data that are required for simulating signaling pathways.

Recently, there have been efforts in this direction from DOQCS [9] and SigPath [19] to try to maintain mathematical models of cellular signaling pathways with minimum or no biological annotations. The quantitative modeling of signaling pathways is currently a new dimension in molecular cell biology. The challenge is to integrate biological information into a quantitative model based on current sources of information, such as primary literature, not searchable with a computer in a way that would allow the setup of a precise model by combining biological knowledge based on protein interaction databases for signaling pathways. The main challenge is that both qualitative and quantitative information can be stored. This will make it possible to add kinetic rate constants to reactions concerning pathways.

DMSP contains the basic qualitative information gathered from protein interaction databases and the more specific quantitative mathematical knowledge gathered from the scientific literature. The

protein interaction data, gathered from databases like BIND and DIP, provided us with a detailed explanation about specific components and their interaction characteristics. The mathematical models provide us with the more detailed look towards the components and reaction functionality within a pathway based upon certain experimental conditions. Most of these protein interaction databases do not provide us with such information, whereas DMSP tries to answer this task. This was successfully done for the TNF α pathway model that was implemented and stored in our approach. However, there are similar projects like SigPath [19] and Biomodels.net [20] wherein they try to build pathway models and do annotations to the models. SigPath e.g. stores biochemical information with the details for quantitative modeling in specific aspects of the cellular machinery. SigPath [19] contains data based on background information imported from primary specific databases (SwissProt plus mammalian and Trembl for proteins).

Biomodels.net [20] provides access to published, peer-reviewed, quantitative models of biochemical and cellular systems. Human curators annotate and cross-link components of the models to other relevant data resources. This allows users to precisely identify the components of models, and helps them to retrieve appropriate models, which they can then visualize. This means that all the models that are published in the public domain are freely available for everyone.

DMSP has been developed as a pathway database like Reactome [23]. The aim of DMSP is very specific. It looks at more quantitative mathematical information that is required to model and to validate pathways. Reactome covers a more global view of providing more qualitative and quantitative information in terms of more biological information. In DMSP we have currently incorporated the basic qualitative information relating to the protein-protein interaction and modeling data concerning signaling pathways. This information is stored at different levels in the pathway data model, from components level, to reaction level, to pathway level.

The TNF α pathway was implemented to test the specification of our integrative environment. It helped us store and modify pathway models based on the obtained simulation results. When simulation and analysis was performed some of its parameters had a high influence on the course of concentration of its components. During simulation and analysis for the TNF α pathway model eight kinetic parameters were identified, which showed high sensitivity regarding the behavior of the general system. A comparison of the eight identified parameters with the mathematical model of the TNF α pathway shows that these parameters are mostly situated at very important locations of the pathway. Most of these highly sensitive kinetic parameters are involved in interactions, which directly connect the three different modeled modules of the mathematical model. It seems to be reasonable that a change in one of these parameters has a large influence on the behavior of the system. This process of analysis has also helped us integrate other pathway models and do systematic analysis using DMSP in an integrative environment.

The focus of DMSP is not only towards doing biological and modeling annotation towards published pathway models like Biomodels. The emphasis of DMSP is towards supporting the research community to implement and build new models that are not published using our database and perform biological and mathematical annotation to their models before publishing locally. Our focus is to concentrate mainly on providing the biological knowledge that has been integrated from different online databases with a possible way to store mathematical modeling knowledge using our proposed database schema. This will help in doing simulations and visualizations of pathway models. This is achieved by providing the whole dataset online as well as creating a plug-in to open source tools for simulation and visualizations like BIOUML [21]. BIOUML is also an SBML-based tool for visualization and simulation for pathways. Through this way the end user can download the whole datasets that include the published TNF α pathway model. After implementing the database schema the users can implement their own pathway models by combining model-

ing and biological knowledge respectively and visualize using BIOUML. This process will help us to store their models using DMSP, as well as share in SBML [24]. Hence we use BIOUML for simulation and visualization of pathway models, which is also based on SBML like CellDesigner [25].

Through this way the end users can download the whole datasets with its schema and test their own models by relating to specific signaling pathways. This will help them perform biological and mathematical annotations for their pathways before publishing their results. On the other hand, this process will help us to store their models using DMSP and perform simulation and analysis before publishing a model.

Now when we compare SBML [24] and BioPAX [26] they are defined with an aim of being standards that can be used for exchanging pathway data. There also exist various tools that support these standards. The representation of interactions in all formats has at least one entity for representing subjects. Representing reactions in these formats has been handled in different ways. SBML has several subtypes for representing reactions. The BioPAX and KGML also have only one way of representing reactions between interactors. The principal structure of all formats is similar; they reflect the structure of a pathway graph. The information is structured in representation of the interacting subjects. So SBML represents the mathematical description of pathways and BioPAX focuses on molecule interactions in metabolic pathways.

After looking at all these kinds of formats we developed our database schema to be as simple as possible according to our requirements and the specification. To represent such a huge pool of heterogeneous information concerning pathways we need to have a simple method. Hence in our database format we incorporated both the biological knowledge from these biological databases, which are based on community standards like PSI MI and modeling knowledge required for doing simulations and analysis.

The focus of our on-going research is to expand the database schema by incorporating the simulation results we obtained for different pathways including TNF α path-

way. This process will help us in identifying specific and important components and reactions of a signaling pathway. We also plan to expand the DMSP schema by interpreting information which helps to design protein complex building in signaling pathways. We further plan to include experimental data of specific experiments that are conducted for different signaling pathways.

Most of the biological knowledge available online is wrapped using XML [13]. XML makes it easy to exchange data between different data models. Presently the DMSP database schema and its data are available to other research groups in XML or in MySQL file format for testing if requested. For a more global visualization we are developing plugins to open source tools like BIOUML [21] or Cytoscape [22] through which the dataset of the DMSP can be visualized with regard to signaling pathways.

5. Conclusions

DMSP is an initial step towards the aim of combining modeling and biological knowledge concerning signaling pathways. Pathway models can be designed, visualized and simulated based on the knowledge stored in the DMSP. The user can download the whole dataset and build pathway models using the knowledge stored in our database. This will help them to perform biological and mathematical annotations in relation to their pathway models before publishing their models. The pathway models can be validated by incorporating more experimental data into DMSP. To summarize, this database is mainly developed to gain a better understanding of signaling pathways that will yield in a quantitative description of protein interaction networks. Simulation results enable the interpretation of biological systems from a quantitative and system-theoretic point of view.

References

1. Arkin A, Ross J. Computational functions in biochemical reaction networks. *J Biophys* 1994; 67: 560-578.
2. Asthagiri AR. Bioengineering models of cell signaling. *Annual Revision of Biomed Eng* 2000; 2: 31-53.
3. Bhalla US, Iyengar R. Emergent properties of networks of biological signaling pathways. *Sciences* 1999; 283: 381-387.
4. Deville Y, Gilbert D, Helden J, Wodak S. An Overview of Data Models for the Analysis of Biochemical Pathways. *Brief Bioinformatics* 2003; 4 (3): 246-259.
5. Xenarios I, Salwinski L, Duan J, Higney P, Kim S, Eisenberg D. DIP: The Database of interacting proteins. A search tool for studying cellular networks of protein interactions. *Nucleic Acid Res* 2002; 30: 303-305.
6. Bader GD, Hogue CW. BIND – a data specification for storing and describing biomolecular interactions, molecular complexes and pathways. *Bioinformatics* 2000; 16: 465-477.
7. Dongsoo H, Kim HS, Seo J, Jang W. A Domain Combination Based Probabilistic Framework for Protein-Protein Interaction Prediction. *Genome Informatics* 2003; 14: 250-259.
8. Hoebeke M, Chiappello H, Noirot P, Bessières P. SPiD: a subtilis protein interaction database. *Bioinformatics* 2001; 17: 1209-1212.
9. Sivakumaran S, Hariharaputran S, Mishra J, Bhalla US. DOQCS – The database of quantitative cellular signaling: management and analysis of chemical kinetic models of signalling networks. *Bioinformatics* 2002; 19: 408-415.
10. Reshetnikov V, Karpov A, Pot D, Tereshchenko F. Vector PathBlazer: A New Pathway Analysis and Visualization Tool (cited 2007, April 12). Available from: http://www.iscb.org/ismb2003/posters/feodorATinformaxinc.com_468.html
11. Bhalla US. A network within signaling pathways. In: Bower JM, Beeman D (eds). *The Book of Genesis. Exploring realistic neural models within the general neural simulation system*, 1998.
12. Kuroda S, Schweighofer N, Kawato M. Exploration of signal transduction in cerebellar long-term depression by kinetic simulation. *J Neuroscience* 2001; 21: 5693-5702.
13. Achard F, Vaysseix G, Barillot E. XML-bioinformatics and data integration. *Bioinformatics* 2001; 17: 115-125
14. Cho KH, Shin SY, Kolch W, Wolkenhauer O. Experimental Design in Systems Biology Based on Parameter Analysis with Monte Carlo Simulation: A Case Study for TNF α Mediated NF- κ B Signal Transduction Pathway. *Simulation* 2003; 79: 726-739.
15. Cho KH, Wolkenhauer O. Analysis and modeling of signal transduction pathways in systems biology. *J Biochemical Society Transactions* 2003; 31: 1503-1509.
16. Schacherer F, Choi C, Götze U, Krull M, Pistor S, Wingender E. The TRANSPATH signal transduction database: a knowledge base on signal transduction networks. *Bioinformatics* 2001; 17: 1053-1057.
17. Kolpakov FA, Ananko EA, Kolesov GB, Kolchanov NA. GeneNet: a gene network database and its automated visualization. *Bioinformatics* 1998; 14: 529-537.

18. Gilman A. A letter to the signaling community. *AfCS* (cited 2007, April 12). Available from: <http://www.afcs.org/>
19. SigPath (Database) (cited 2007, April 12). Available from: <http://www.sigpath.org/>
20. Le Novère N, Benjamin B, Alexander B, Mélanie C, Marco D, Harish D, Lu L, Herbert S, Maria S, Bruce S, Jacky L, Michael H. BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Res* 2006; 34: D689-D691.
21. Kolpakov FA. BIOUML – Framework for visual modeling and simulation of biological systems. *Proc Int Conf Bioinformatics of Genome Regulations and Structures*; 2002.
22. Shannon P, Markie 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.
23. Joshi G, Gillespie M, Vastrik I, D'Eustachio P, Schmidt E, de Bono B, Jassal B, Gopinath G, Wu R, Matthews L, Lewis S, Birney, E, Stein L. Reactome: a knowledgebase of biological pathways. *Nucleic Acids Res* 2005; 33: D428-432.
24. Finney A. Developing SBML Beyond Level 2: Proposal for Development. *Lec Notes in Computer Science* 2005 (Jan); 3082.
25. Funahashi A, Tanimura N, Morohashi M, Kitano H. CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. *BioSilico* 2003; 5: 159-162.
26. BioPAX (cited 2007, April 12). Available from: <http://www.biopax.org/>
27. Visvanathan M, Breit M, Pfeifer B, Modresprian R, Baumgartner C, Tilg B. Systematic analysis of signaling pathways using an integrative environment. *Methods Inf Med* 2007; 46: 386-391.
28. Cannataro M, Cuda G, Veltri P. Modeling and Designing a Proteomics Application on PROTEUS. *Methods Inf Med* 2005; 44: 221-226.

Correspondence to:

Maresh Visvanathan
Institute for Biomedical Engineering
University for Health Sciences, Medical Informatics
and Technology
Eduard Wallnoefer Centre 1
6060 Hall in Tyrol
Austria
E-mail: mahesh.visvanathan@umit.at