Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Modelling and Simulation of the TLR4 Pathway with Coloured Petri Nets

C. Täubner, B. Mathiak, A. Kupfer, N. Fleischer, and S. Eckstein
Technical University of Braunschweig,
Institute of Information Systems

Abstract— **This paper demonstrates the first steps of an automation process to develop models of signaltransduction pathways using discrete modelling languages. The whole approach consists of modelling, validation, animation, linking databases to simulation tools and also the qualitative analysis of the data. In this paper, we detail the modelling and simulation of the TLR4 pathway with a Coloured Petri Net simulation tool and the validation of this model against the semantic and mechanistic map from a biological database. These graphical maps contain all necessary reactions as a figure.**

We start with an UML class diagram to understand the static structure of molecules involved in the TLR4 pathway. Afterwards we model and simulate each "pathway step reaction" - one after another - to get the behaviour of the final system. The result is a model of the pathway which can be used in simulations, derived solely from basic chemical reactions in the database. Also, it is a lesson on critical points where human decision-making is needed, because not all the required information is stored directly in the database.

I. INTRODUCTION

Modelling and simulation of pathways can be found in many areas of bioinformatics. Information about molecules, chemical reactions and associated networks are gathered in different kinds of databases and are mostly freely available. Unfortunately, such information can be long, self-contradictory, incomplete, and mistakable [11]. For this reason databases like TRANSPATH [14] provide maps or pathwaybuilder to visualise the information as directed graphs or figures. Using this information and figures to develop and analyse the behaviour of molecules and reactions can bring new insights into pathways.

For the description of dissolution, activation and translocation in pathways, dynamic modelling languages like Petri Nets are especially suited. To understand the static structure of pathways, we first modelled a UML (Unified Modelling Language) [1] class diagram [3]. It involves classes describing the molecules and associations between the classes describing the activation, inhibition or binding between molecules.

Petri Nets (PNs) are a well established modelling and simulation technique [12] and can be used for qualitative and quantitative analysis [5]. For qualitative analysis, we can learn more about the properties of the system by identifying invariants [6], the presence of bounderies [16] or the liveness in a system of biochemical reactions [9]. In the quantitative analysis the behaviour may be observed by simulation and obtaining the concentration graphs [4], [7] or testing the reachability of a steady state [15]. In all these cases PN simulation tools were not connected to a database, and modelling and simulation were done by hand.

Various extensions of PNs have been used for modelling molecular biological systems and it is possible to identify modelling goals for each [5]. The advantage of using Coloured Petri Nets (CPNs) is the possibility to analyse biological system properties and behaviour with a freely available software [2] (see sections III and IV).

Peleg, Yeh and Altman [11] studied the development of biological processes using PNs models, coming to the following conclusions: The model should represent a

- static-structural view: molecules that participate in the system, their properties, and the relationships among them;
- dynamic view: sequential, parallel, conditional and iterative processes;
- functional view: show the actors, e. g. enzymes that perform each function, the substrates and products of each function. Also the cellular location of the substrates and products should be specified.

We have modelled, simulated and validated the TLR4 pathway using CPNs on the basis of these items in order to improve the automatic generation process of pathway PN models out of a pathway database. The whole analysis process and the difficulties during the first steps of an automation are shown in this paper.

II. BIOLOGICAL BACKGROUND

The example used in this paper is dedicated to the infection process of *Pseudomonas aeruginosa* and lung epithelial cells. *Pseudomonas aeruginosa* is an opportunistic pathogene, meaning that it exploits some break in the host defence to initiate an infection. The infection is a serious problem in patients hospitalized with cancer, cystic fibrosis, and burns. The TLR4 signaltransduction pathway starts after the infection of the lung epithelial cells on the surface of the cells.

This work was funded by the German Federal Ministry of Education and Research (BMBF) for the Bioinformatics Competence Center 'Intergenomics' (Grant No. 031U210C).

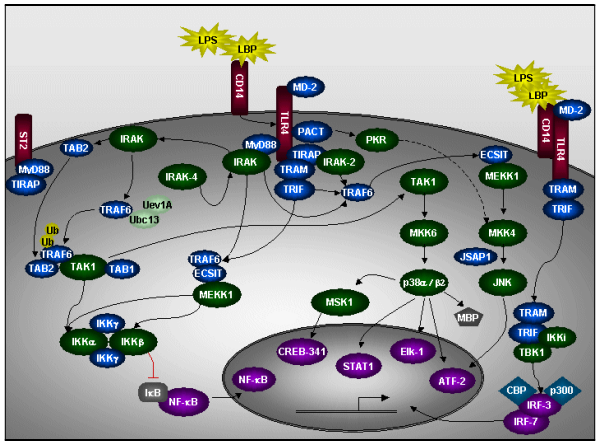

Figure 1: TLR4 pathway as a map

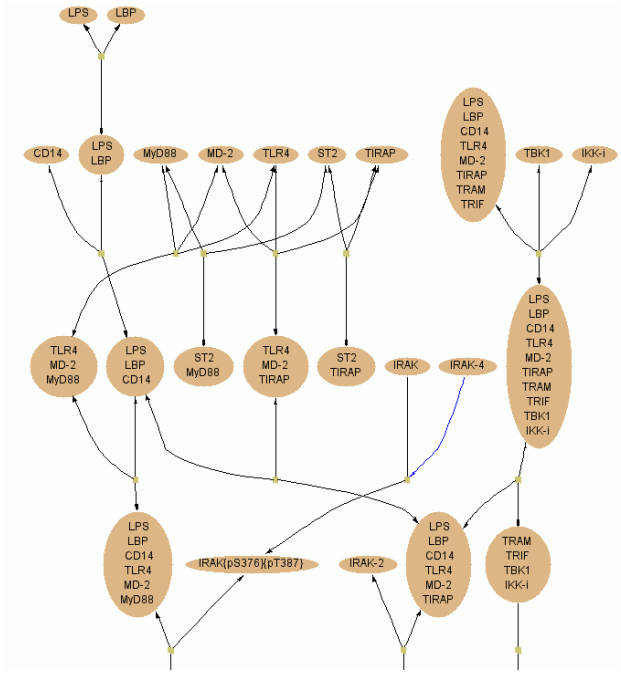

Figure 2: A part of the TLR4 pathway modelled with the pathwaybuilder

Figure 1 and *Figure 2* demonstrate the workflow of this pathway: *Figure 1* presents the semantic view of TLR4, *Figure 2* shows a part of the mechanistic view generated with the pathway builder of TRANSPATH on the basis of the chemical reactions. A complete description of the TLR4 pathway does not exist even though it is the best annotated pathway in TRANSPATH. On the basis of these two figures, the references and the reactions in TRANSPATH we were able to develop the signaltransduction of TLR4 with CPN. The abbreviations of the enzymes and proteins involved in TLR4 can be found in [14].

III. CPN Tools

CPN Tools [2] is a tool package supporting the use of CPN. It consists of four integrated tools for modelling and simulation:

- an editor for construction, modification and syntax check of models,
- a simulator for interactive and automatic simulation,
- a graph tool for construction and analysis of occurence graphs and
- a performance tool for simulation based on performance analysis.

CPN Tools supports CPN models with complex data types (colour sets) and data manipulations (arc expressions and guards) - both specified in the functional programming language Standard ML [13]. The package also supports hierarchical CPNs, so its possible to model a set of separate modules (subnets) with well-defined interfaces.

Our model consists of 3 modules, 17 ressources, 74 complexes, 7 variables, 75 places and, 47 transitions. Compared to a typical industrial model, which consists of 50-200 modules each with 10-50 different places and transitions it is a very small model. Regarding the size of the pathways stored in TRANSPATH there are no problems to handle these with CPN Tools.

IV. Coloured Petri Nets

CPN is a graphical language for design, specification, simulation and verification of systems. It is in particular well-suited for systems that consist of a number of processes which communicate and synchronise. Typical examples of application areas are communication protocols, distributed systems, automated production systems, work flow analysis and VLSI chips.

CPN are used for three different - but closely related - purposes. First of all, a CPN model is a description of the modelled system, and thus can be used as a specification or while presenting and explaining the model to other people. Also the dynamic behaviour of a system can be investigated in detail. Furthermore, the behaviour of a CPN model can be analysed, either by means of simulation, like debugging or program execution or by program verification. Finally, the process of creating the description and performing the analysis usually gives the modeller an understanding of the modelled system - and it is often the case that this is more valuable than the description and the analysis results themselves. An overview about Petri Nets properties, the analysis possibilities and applications can be found in [10].

A small example of a CPN modelled with CPN Tools is shown in Fig. 3. It describes the activation of the TRAF6:TAB2-complex and the building to the TRAF6:TAB2:TAK1:TAB1-complex (TRIKA2-complex) in the TLR4 pathway. The ellipses are called places and describe the states of the system. The rectangles are called transitions and describe the actions of the system. The arrows between the places and transions are called arcs. The arc expressions describe how the state of the CPN changes when the transitions occur. Each place contains a set of markers called tokens which carry a data value from a given type. In our model the places symbolise several molecules or molecule complexes, the actions are reactions

in the cell and the tokens are the types of molecules, e.g. enzyme. As an example the place in the upper left corner of Fig. 3 has 5 tokens in the initial state. All the token values belong to the type Enzyme. All the 1´s in front of the small backslashes indicate that there is exactly one for all 5 types of tokens.

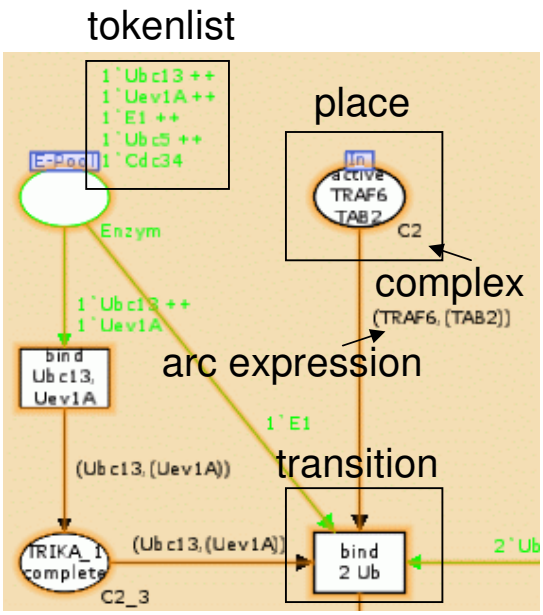

Figure 3 : TRIKA

In order for a transition to take place (the transition fires), there have to be sufficient tokens on its input places, and these tokens must have token values that match the arc expressions. When the transition fires, it removes tokens from its input places and adds some at all of its output places. The transition “bind 2 Ub” for example can only fire if TRAF6 and Tab2 are active, Ubc13 and Uev1A connected to “TRIKA_1 complete” and, only one E1 token and 2 UB tokens are available (see all the input arcs of the “bind 2 ub” transition). The brackets demonstrate the order of binding molecules.

During the construction of the TLR4 CPN, simulations were used to validate the CPN model, for example to check that the model has the expected behaviour. In the first modelling-steps the TLR4 model was simple, covering only selected parts of the system and ignoring some aspects of the final system. Later, the scope of the model was extended and more details were added step by step. So we found critical points that we want to discuss in the next section.

V. Modelling and Simulation with CPN

Modelling and simulation of signaltransduction pathways (STP) with CPN is a process with four steps: modelling a pathway-step-reaction with a simulation tool, simulate the model, compare this model with experimental data, and in case of discrepancies adjust the model otherwise extend the model with the next pathway-step-reaction. The resulting model of this “Cycle of Development” (CoD) is an exact mapping of the information in the database.

This section focuses on critical points in the CoD where human decision-making was needed to complete the whole pathway. As an example we detail two pathway-step-reactions of the TLR4 pathway modelled and simulated with CPN Tools [2].

The TLR4 pathway starts with the binding of the lipopolysaccharide-binding protein (LBP), the "adapter-molecule" for LPS-binding to the cell surface receptor CD14 (see *Figure 2*). The ligand-receptor pair LPS/LBP-CD14 seems to be one important element in the TLR4 pathway and is responsible for the signal transmission. Following the pathway it splits into two possibilities of reactions. One transmits the signal to the MyD88-dependent pathway, another to the TIRAP-dependent pathway. The way of signaltransmission is dependent on this binding. Both start with the binding between MD-2 and TLR4.

The first pathway-step-reaction we would like to detail is the activation of PKR (see *Figure 4*, starting with the “build TLR4-Complex”-transition). Following the TIRAP-dependent pathway TIRAP binds to TRAM and TRIF or IRAK-2 and in both cases PACT activated PKR. The way and the time of activation is unknown and is neither found in the chemical reactions nor in the papers attatched to the pathways information (see [14]).

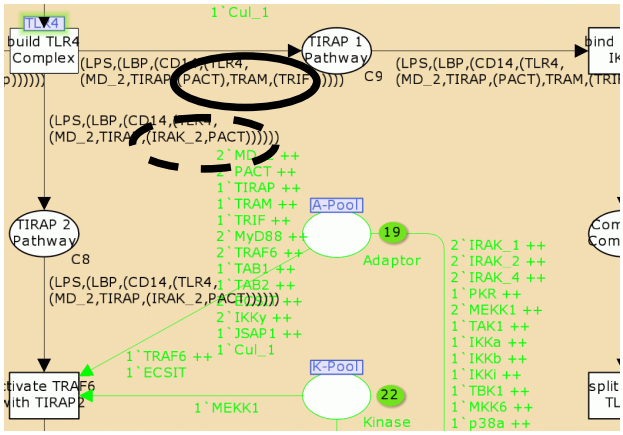

Figure 4: TIRAP-dependent pathway

As we modelled this pathway-step-reaction, we took the “PACT-activates-PKR” reaction for granted, omitted PKR and modelled only PACT as an arc expression (see the handmade ellipses in *Figure 4* above). The reason for doing this was that we did not found information either in the mechanistic view nor in the experimental references for TLR4, but in other pathways in TRANSPATH itself [8]. Ito et al. found an ubiquitously expressed PKR-associated protein named RAX. RAX has a high sequence homology to human PACT, which activates PKR in vitro. This experiment has indicated a general association between PACT and PKR. But the appropriate information between TLR4, PACT and PKR are missing in TRANSPATH. In addition PACT is not connect to TLR4 but to TIRAP (see *Figure 1*).

Another possibility of signaltransmission is the binding of MKK4, jnk and JSAP1 to the MEKK1 Complex (see *Figure 5*). "jnk" is compatible to JNK1, JNK2, JNK3. All three molecules are rendered interchangeable and modelled as an "if-then-else-construct". Compared to *Figure 1* JSAP1 and MEKK1:jnk have not a covalent bond. After the binding of MEKK1:MKK4:jnk:JSAP1 the whole complex splits into several molecules. The next steps are unknown and not found in other pathways information. For this reason all molecules are modelled with "Dead Ends".

There is no obvious attribute in the database to differentiate between an open end in a pathway where information is missing and a "reaction end" where nothing follows. For this reason we did not distinguished between these two possibilities and modelled both with Dead Ends.

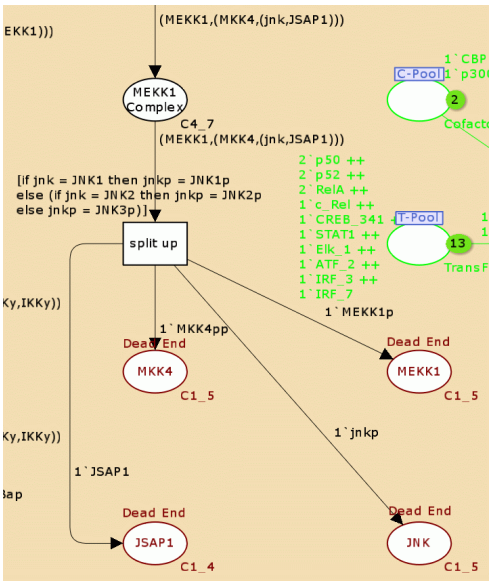

Figure 5: Dead Ends

The whole modelling and simulation process can be found in [2].

VI. Conclusion

The advantage of using CPN instead of another extension of Petri Nets is the possibility to distinguish between different molecules via their identifiers (colours). Alternative metabolic paths (e. g. the two possibilities of MyD88- and TIRAP-dependent pathway) have to be distinguished because they often result in different reactions. A further advantage is the existence of a freely available simulation tool (CPN Tools) because it supports the whole modelling, simulation and analysing process.

We described the first steps of a process to model and simulate the TLR4 pathway with CPN in order to improve the automatic generation process we implement at the moment. The basis of the CoD were the semantic and mechanistic information stored in the TRANSPATH database. The systematic approach to model the TLR4 CPN demonstrated, that a simple automatic transformation of chemical reactions from the database to places and transitions in the CPN model is not in all cases sufficient to describe the dynamic behaviour of pathways in an adequate way. E.g. the distributed information about pathways like the TLR4 complicate the automation of modelling and simulation because at times the information needed was not connected to TLR4 itself. So, human knowledge about pathways remains an important part during the CoD. These insights have to be taken into account during the steps of automation, especially during the importation of pathway knowledge from TRANSPATH into different simulation tools.

References

[1] G. Booch, J. Rumbaugh and, I. Jacobson. The Unified Modeling Language User Guide. Addison Wesley, 2005.

[2] CPN Tools url: http://wiki.daimi.au.dk/cpntools-help/_home.wiki

[3] N. Fleischer. Modelling and Simulation of the P. aeruginosa Infection Process with Petri Net. German diploma thesis, TU Braunschweig, 2005.

[4] P. J. E. Goss and J. Peccoud. Quantitative modeling of stochastic systems in molecular biology by using stochastic Petri nets. In *Proceeding of the National Academy of Sciences of the USA*. Vol. 95, 6750-6755, June 1998.

[5] S. Hardy and P. N. Robillard. Modeling and Simulation of Molecular Biology Systems using Petri Nets: Modeling Goals of Various Approach. In *Journal of Bioinformatics and Computational Biology*. Vol. 2, No. 4, 619-637, 2004.

[6] M. Heiner and I. Koch. Petri Net Based Model Validation in Systems Biology. In *Proceeding of the International Information Visualization 2002*. IVS, IEEE, 549-554. July 2002.

[7] R. Hofstädt and S. Thelen. Quantitative Modeling of Biochemical Networks. *In Silico Biology*. Vol. 1, 39-53, 1998.

[8] T. Ito, M. Yang and, WS May. RAX, a cellular activator for double-stranded RNA-dependent protein kinase during stress signaling. In *The Journal of Biological Chemistry*, Vol. 274, No. 22, 15427-15432, 1999.

[9] I. Koch, B. H. Junker and Monika Heiner. Application of Petri net theory for modelling and validation of the sucrose breakdown pathway in the potato tuber. In *Bioinformatics*. Vol. 21, No. 7, 1219-1226, 2005.

[10] T. Murata. Petri Nets: Properties, Analysis and Applications. *In Proceedings of the IEEE*, Vol. 77, No.4, April 1989.

[11] M. Peleg, I. Zeh and, R. B. Altmann. Modeling biological processes using Workflow and Petri Net models: In *Bioinformatics*. Vol. 18, No. 6, 825-837, 2002.

[12] C. A. Petri. Communication with Automata. New York: Griffiss Air Force Base. Tech. Rep. RADCTR-65-377, vol.1, Suppl. 1, 1966.

[13] Standard ML url http://www.cs.cmu.edu/~rwh/introsml/

[14] TRANSPATH url: http://www.biobase.de

[15] K. Voss, M. Heiner and I. Koch. Steady-state analysis of metabolic pathways using Petri net. *In Silico Biology 3*, http://www.bioinfo.de/isb/2003/03/0029/, 2003.

[16] I. Zevedei-Oancea and S. Schuster. Topological analysis of metabolic networks based on Petri net theory. *In Silico Biology 3*. , June 2003.