METHODOLOGY

Useful biological models with no formal training

Stefano Schivo[1#], Jetse Scholma[2#], Brend Wanders[3], Ricardo A. Urquidi Camacho[2], Paul E. van der Vet[3], Marcel Karperien[2], Rom Langerak[1], Jaco van de Pol[1] and Janine N. Post[2*]

[*]Correspondence:
j.n.post@utwente.nl
[2]Developmental BioEngineering, MIRA Institute for Biomedical Technology and Technical Medicine, University of Twente, 7522NH Enschede, The Netherlands
Full list of author information is available at the end of the article
[#]Equal contributor

Abstract

Background: The large amount of data with which an experimental biologist must deal makes it largely impossible to reason on the dynamics of complex systems without relying on computational support. In order to be fully profitable, computational modelling tools must be both accessible and powerful, providing a means to formalise and analyse experimental outcomes without requiring additional training.

Results: We show that the software tool ANIMO (Analysis of Networks with Interactive MOdelling) can be used in different modelling tasks, and illustrate how this can be an asset for the biological researcher. ANIMO models for *Drosophila melanogaster* circadian clock and signal transduction events downstream of TNFα and EGF in HT-29 human colon carcinoma cells are used as case studies to illustrate the possibilities offered by the tool.

Conclusions: The software tool ANIMO allows the biologist to couple the power of a formalism with the usability of a user friendly interface, achieving useful insight without the need to expand their knowledge of mathematical modelling tools.

Keywords: modelling; signalling pathway; timed automata; dynamic behaviour

Background

In living cells, processes are regulated by networks of interacting molecules. Aberrations in these networks underlie a wide range of pathologies. The development of new therapies requires a thorough insight in the functioning of these networks. Obtaining such insight can be a challenging task. Feedback loops and crosstalk between pathways lead to an intricate wiring of the network. Hence, it is necessary to study the ensemble of molecules involved, because the behaviour of individual molecules is not sufficient for a complete understanding. Since the human brain is ill-suited to grasp the non-linear dynamics of these complex networks and the entailed emergent properties, the role of computational support is increasing in molecular biology.

As models are a formalization of knowledge or theories, an underlying formalism is needed to express this knowledge. Different formal methods have been successfully applied to construct representations of biological systems. Among these methods are Boolean logic [1, 2], ordinary differential equations (ODEs, reviewed by [3]), interacting state machines [4, 5], process calculi [6, 7], Timed Automata [8, 9, 10] and Petri nets [11, 12]. Most of these formal methods have been implemented into software tools to aid the process of modelling. Due to the lack of such a supporting tool, Timed Automata have remained a less frequently applied method.

Timed Automata have been developed to model the dynamic behaviour of systems with processes running in parallel [13]. As such, Timed Automata have been

applied in communication protocols and industrial control engineering [14, 15, 16]. The parallels between these application areas and regulatory processes in cells have triggered the step towards their use in biology. [8] use Timed Automata to extend a classical modelling paradigm [17], allowing to add temporal dynamics to gene network models. [9] describe a model of biological oscillators and test synchronization properties in this dynamic system. A discretization of ODEs to Timed Automata is proposed by [10], applying a translation between the two formalisms to an example gene regulatory network. Two different approaches to transforming a Petri net model into Timed Automata are presented by [18], who also address the important issue of state space explosion in their paper. Finally, [19] propose an *ad hoc* Timed Automata model of a radiation treatment system, which is then validated through the analysis tool UPPAAL [20].

Each of these approaches has been successfully validated, demonstrating the potential of Timed Automata in biological applications. However, these approaches were all limited to simple or specific examples and none of these modelling methods has led to a tool implementation of the proposed method to encourage a broader use of Timed Automata in molecular biology. We have recently introduced the software tool ANIMO (Analysis of Networks with Interactive MOdelling, [21]), which aims at making available the power of Timed Automata to the experimental biologists without requiring them to be fluent in any mathematical formalism.

In this paper we present ANIMO 2.0, an updated version of ANIMO which allows for a considerably more efficient analysis of models. ANIMO 2.0 also introduces new features, such as the possibility to interrogate an ANIMO model with model checking queries represented as human language sentences instead of mathematical formulae. The case studies presented in the following sections are centered around possible usage scenarios for ANIMO, and aim at illustrating how a biologist can profit from the tool in its latest iteration. An installation manual for ANIMO 2.0 is available at `http://fmt.cs.utwente.nl/tools/animo/manual.html`.

Results

Modelling oscillation

Results obtained with ANIMO are comparable to results with other modeling approaches. To demonstrate this, Figure 3 represents an ANIMO model of the circadian clock in *Drosophila Melanogaster*, based on the work by [22], where ordinary differential equations (ODEs) were used. The cyclic behavior of the circadian clock is based on the alternating formation and destruction of the CYC/CLK protein complex. Concentration levels of this complex are in turn regulated by a series of proteins which are produced as a consequence of CYC/CLK formation. The CWO protein is central to the functioning of the network, as it degrades the mRNA for most of the involved proteins. As such, CWO act as an inhibitor that counterbalances the effect of CYC/CLK. The positive influence of the light-regulated cryptochrome CRY on the degradation of TIM is a consequence of the passage between day and night, allowing the circadian clock to synchronize to a time zone (see Suppl. Sect. **??**).

The output of the ANIMO model in Figure 3 closely matches the original ODE model. In particular, the oscillations in both models show the same periods and

phases (see Suppl. Fig **??**). Due to the compositional nature of Timed Automata ANIMO allows for intuitive *in silico* knock-out experiments, by right-clicking a node in the model and disabling it. Such experiments have been done before [22] and give similar results in our model.

Exploiting formal methods: model checking

To show why one should use a formal model like Timed Automata, we demonstrate the application one of the biggest advantages of Timed Automata: state space-based analyses such as model checking [23] provided by UPPAAL [20].

The technique of model checking consists in verifying the truth of some properties that describe the behaviour of a model. This is usually done by exploring every possible evolution of the model starting from an initial state. The set of all states that can be reached by a model is called *state space*, and formulae describing the evolution of a system are usually described by (a subset of) computation tree logic (CTL, [24]).

Using ANIMO to generate hypotheses

In order to validate our modeling approach, we constructed a larger model of the signaling network downstream of TNFα and EGF, formalizing the crosstalk that takes place between the pathways at different levels of cellular regulation. We first modeled the two pathways in isolation (Suppl. Figs. **??**, **??**), using information on protein interactions from the KEGG [25] and phosphosite [26] databases. These models were fitted to experimental data from studies by [27] and [28]. We then merged the two pathways into a single model and added autocrine crosstalk between the pathways that has been described by [28]. Briefly, stimulation with TNFα leads to a rapid release of TGFα (TGFa in the model), which activates the EGF receptor (EGFR). This activation causes secretion of IL-1α (IL-1a) at later time points. The effect of IL-1α is down-regulated by the secretion of IL-1 receptor antagonist (IL-1ra) downstream of TNFα. The resulting model (Fig. 4A) was compared to the experimental data for treatments with 100 ng/ml TNF alone and 100 ng/ml EGF alone (data not shown) [27].

At this point, the behavior of the model deviated from the data for some of the nodes. This is an interesting situation, as it requires modifications to the model, that can be interpreted as new hypotheses. Below, we give two examples and show how adaptation of the model can be used to generate novel testable hypotheses.

Experimentally, treatment with TGFα alone does not lead to secretion of IL-1α. Instead, a co-stimulation with TGFα and TNFα is required [28]. However, in the first version of the model, treatment with TGFα was sufficient for IL-1α expression (Fig. 4B). Given the time delay until secretion of IL-1α, it can be expected that *de novo* synthesis of IL-1α is required and that both TNFα and TGFα are needed to activate transcription of the IL-1α gene. JNK1 and ERK signal downstream of TNFα and TGFα, respectively, and are known to affect the activity of multiple transcription factors. We altered the model to make activation of IL-1α expression dependent on both JNK1 activity and ERK activity (Suppl. Fig. **??**, arrows linking JNK1 and ERK to IL-1a gene). After this modification to the model, IL-1α was no longer secreted upon stimulation with TGFα alone, which greatly improved the

fit between the measured IL-1α levels and the model (Fig. 5B). This hypothesis could now be used to design a new experiment to validate IL-1α as a target of combined JNK1 activity and ERK activity in HT-29 cells. For example, kinase inhibitors specific to JNK1 and ERK could be used to confirm that activity of both kinases is required for expression and secretion of IL-1α. Performing the experiment is beyond the scope of this study, but this hypothesis finds support in literature. Transcription factors c-Jun and c-Fos together form a heterodimer known as AP-1 and are activated by JNK1 and ERK, respectively [29, 30]. AP-1 has been reported to bind to the promoter of IL-1α, providing evidence for a role in the regulation of IL-1α expression [31]. Based on these findings in literature we included c-Jun and c-Fos in our model as transcriptional activators of IL-1α (Fig.5A).

As a second example, we considered the behaviour of JNK1 and MK2. In the model, both proteins were located downstream of TNFα but not TGFα or EGF. Hence, the model did not show an effect of C225, a pharmacological inhibitor of ligand-EGFR binding, on activation of JNK1 or MK2 after stimulation with TNFα. However, experimental data show that C225 strongly reduces activation of JNK1 and MK2 upon stimulation with TNFα [28]. This fact is indicative of a role for EGFR in activation of JNK1 and MK2. Since both JNK1 and MK2 are located downstream of MEKK1, we hypothesized that activation of MEKK1 is dependent on both TNFα-signalling and TGFα-signalling. In the model we added a new hypothetical node Hyp 2 (hypothesis 2) to link EGFR to MEKK1 (Suppl. Fig. **??**). This addition led to an improved fit of the model to the data upon treatment with TNFα + C225: activation of both MK2 and JNK1 was strongly suppressed by C225 (Fig. 5C). Stimulation with EGF alone did not lead to activation of JNK1 and MK2. These data support the validity of the modification to the model. Further support for a link between EGFR and MEKK1 was found in literature. Specifically, Ras has been reported as a direct activator of MEKK1 [32]. EGFR is a well-known and potent activator of Ras, which is why it was already in our network [25]. Other studies also report activation of JNK1 and phosphorylation of c-Jun downstream of Ras, which is consistent with an interaction between Ras and MEKK1 [30, 33]. Based on these findings, we adapted our model by removing the Hyp 2 node and creating a direct interaction between Ras and MEKK1 (Fig. 5A). Experimentally, the role of Ras could be confirmed by using a pharmacological inhibitor of Ras activity, and measuring the effect of this inhibitor on the activation of JNK1 and MK2. Together, our model suggests that EGFR activity is required but not sufficient for activation of JNK1 and MK2 in HT-29 cells.

There are other nodes for which the experimental data deviates from the model in one or more of the experimental conditions. A comparison between model and experimental data can be found in Figures **??**, **??** and **??**. A complete deciphering of the signalling events in this biological system is outside the scope of this paper. Instead, we illustrated how interactive modelling of the dynamic behaviour of a signal transduction network can be used to extend previous pathway topologies and can lead to the generation of novel hypotheses.

Discussion

The ultimate aim of research projects is to solve a problem or get the answer to a scientific question. In biology, an in-depth understanding of the relevant biological

system is an important step towards this goal. Successive repetitions of the empirical cycle result in a stepwise increase in understanding, until the goal is reached. For complex biological systems, computational modelling is indispensable in this process. The mere act of creating a computational model based on prior knowledge, experimental data and hypotheses assists in gaining more insight in the system.

We developed a new modelling approach by proposing a series of abstractions from the detailed molecular mechanisms of biological systems. These abstractions reduce the need for kinetic parameters, while preserving enough expressivity for a useful description of the dynamic behaviour of biological networks. A novel modelling tool, ANIMO, allows effective use of this approach and enables an intuitive construction of formal models.

ANIMO is not the first modelling tool to provide an interface to a modelling formalism. Such interfaces exist in many other tools (see Suppl. Tab. **??**). With its focus on user-friendliness and intuitive modelling, ANIMO's main contribution lies in making computational modelling more accessible to experts in biology. Making use of the visual interface provided by Cytoscape, network representations subscribe to biological conventions. Model parameters are kept to a minimum and can be directly accessed by mouse-clicking on nodes and edges. Because of the automatic translation of the network topology and user-defined parameters into an underlying formal model, training in the use of formal methods is not needed. In Supplementary Section **??**, a more in-depth comparison between ANIMO and other modeling tools is given. For this comparison we selected a tool for each of the most commonly used formalisms, and used criteria with a strong focus on user-friendliness.

In Section , we described the construction of an ANIMO model of the ciracadian clock in *Drosophila Melanogaster*. This model captured the dynamics of the regualtory network and led to similar conclusions as an ODE model that had been published previously [22]. This finding supports the use of the series of modelling abstractions that we proposed. The biggest difference between the construction of these models is that the model by [22] is constructed by writing a system of mathematical equations, together with an algorithm for simulation. In ANIMO, instead, a number of network nodes is drawn for the molecules involved. These nodes are then linked by directed interactions that represent cause-and-effect relationships, with a single parameter that defines the strength of each interaction. This is a more intuitive approach to construct a model. Further contributing to an interactive modelling process is the compositionality of the model. Each node in the network can be disabled at any time by the user, or extra nodes can be added, without having to change any of the existing interactions.

In Section , we showed the construction of an executable model of signalling events downstream of TNFα and EGF in human colon carcinoma cells. This data set has been used for previous modelling studies, based on partial least-squares regression and fuzzy logic [34, 35]. The partial least-squares regression model describes an abstract data-driven model that uses statistical correlations to relate signal transduction events to various cellular decisions. This type of modelling is very useful in uncovering new and unexpected relations. It is also successful in making predictions, but gives little direct in the dynamic behaviour of the network. Fuzzy logic analysis led to a model that gave a better fit to the dynamic network behaviour

than discrete logic (Boolean) models. Inspection of the inputs to the logical gates that were used to model protein behaviour led to the prediction of novel interactions between proteins, showing the usefulness of this approach. For most of the proteins, such as JNK1, time was used as an input parameter. For example, the logical gates "if TNFα is high *AND* time is low, then JNK1 is high" and "if TNFα is high *AND* time is high, then JNK1 is low" were used to describe the dynamic behaviour of JNK1. Although this leads to a representative description of the dynamic behaviour of JNK1, peaks in protein activity at early time points were not reproduced by the fuzzy logic model. Moreover, it gives no insight in the molecular interactions that are involved in activation or inhibition.

In this study, we used the same data set and performed a single round of the empirical cycle. This cycle starts off with the experiments carried out by [27]. We used the resulting experimental data, together with knowledge from curated databases [25, 26] to construct an executable model of the biological system. In contrast with the two approaches described above, ANIMO is aimed at the construction of more mechanistic models, mimicking biochemical interactions *in silico*. This way of modelling gives a different type of insight. In the process of model construction, we extended a prior-knowledge network with time-dependent extracellular crosstalk that has been reported previously [28]. To come up with possible explanations for a disagreement between the model and the experimental data, two additional layers of crosstalk were introduced, at the signal transduction and transcriptional level. These modifications improved the fit of the model to the data and can be interpreted as novel testable hypotheses. Finally, we proposed new experiments that could be carried out to test these hypotheses, closing the empirical cycle. Together, our model sheds more light on the intricate entanglement between the TNFα and EGF pathways at multiple cellular levels. But above all, the model provides an excellent starting point for further investigation. Every new round in the empirical cycle will lift the understanding of the system to a higher level, leading to an incremental build-up of knowledge and an upward empirical spiral. Being intuitively accessible, ANIMO models facilitate sharing knowledge within and between groups and encourage collaborations.

Conclusions

Methods

Modelling abstractions

In living cells, cascades of chemical and physical interactions enable propagation of signals through molecular networks. In this process, the activity of upstream molecules induces a change in the concentration or activity of downstream molecules. For many reactions, the values of the kinetic parameters are unknown or difficult to collect. This lack of knowledge hampers the feasibility of computational models that describe molecular networks in fine mechanistic detail, especially for larger networks. As a solution to this problem, we propose the construction of models at a higher level of abstraction, thereby reducing the number of parameters involved. In choosing a suitable abstraction level, it is important to retain enough descriptive power to give a meaningful formal description of the topology and the associated dynamic behaviour of biological networks.

As a first abstraction in ANIMO models, the active and inactive forms of each network component are represented together by a single node in the network. Each of these nodes is characterized by its *activity level*, which represents the fraction of active molecules of that molecular species. When a molecule is known to be constitutively active, changes in concentrations of that molecule are treated as changes in its activity level. Activity levels are discretized into integer variables with a user-defined granularity, ranging from Boolean (2 levels) to near-continuous (100 levels).

Detailed biochemical reaction mechanisms are abstracted to *interactions*, which can represent either activations ($\rightarrow$) or inhibitions ($\dashv$). This aggregation of elementary reactions into single interaction steps reduces the number of kinetic parameters involved, while preserving cause-and-effect relationships. For example, consider a reaction in which enzyme E phosphorylates and activates substrate S, transferring a phosphate group from a molecule of ATP to a molecule of S. Biochemically, this reaction can be represented as

$$E + S + ATP \rightleftarrows ES + ATP \rightarrow ES^P + ADP \rightleftarrows E + S^P + ADP,$$

with conservation condition $S + S^P = \text{constant}$ and $ATP + ADP = \text{constant}$.
Under the assumption of ATP constantly being replenished by the cell, this reaction is abstracted in ANIMO to the corresponding interaction

$$E \rightarrow S.$$

Each occurrence of the interaction $E \rightarrow S$ will increase the activity level of S by one discrete step. Since the activity level is defined as the active fraction of a molecular species, an increase in the active fraction implies a decrease in the inactive fraction. Hence, the original conservation condition is automatically satisfied. The interaction rate, R, depends on the activity levels of the reactants involved and on a single kinetic parameter k that is set by the user. The three available interaction scenarios can be interpreted as abstracted kinetic rate laws:

1. $R = k \times [E]$: the interaction rate depends only on the activity level of the upstream node.
2. $R = k \times [E] \times [1 - S]$ (activations) or $R = k \times [E] \times [S]$ (inhibitions): the rate depends on the activity levels of both the upstream and downstream participants. Activations depend on the presence of inactive substrate, *[1 - S]*, whereas inhibitions depend on the level of active substrate, *[S]*.
3. $R = k \times [E_1] \times [E_2]$: this scenario can be used when the activation or inhibition of a downstream node depends on the simultaneous activity of two upstream nodes. This scenario is comparable to an *AND-gate* in Boolean logic.

We will show in Section that the abstraction proposed here preserves ample expressivity to capture the dynamic behaviour of a biological network.

Modelling interactions with Timed Automata

Timed Automata have been shown to be a powerful formalism to model biological processes [8, 9, 10]. A timed automaton consists of locations and transitions between these locations (see Fig. 1), and a system of timed automata can be used to model

a system of interacting molecules. At any time, each automaton is in a specific location, and together these locations represent the current state of the biological system. Each timed automaton can have one or more local clocks associated to it, allowing temporal control of transitions between locations. These transitions are used to represent interactions between molecules. Fast interactions take less time than slow interactions to perform an activation or inhibition step. We have previously described in detail how the scenarios presented in Section can be used to calculate the timing of molecular interactions to give a description of network dynamics [21]. Figure 1 presents a small example that illustrates the basic properties of TA. This model describes the activation of ERK by MEK[1].

ANIMO

The modelling approach described in Section and Section is implemented in the software tool ANIMO (Analysis of Networks with Interactive MOdelling, [21]) as a plug-in to the network visualization tool Cytoscape [36]. The visual interface of Cytoscape makes the construction, expansion and rewiring of a network topology a fast and user-friendly process.

When a new node is added to the network, it has to be initialized with the number of activity levels and its initial activity. For each interaction, a scenario needs to be selected, together with the corresponding kinetic parameter and the interaction type: activation or inhibition. All settings can be readily adapted by double clicking, or via a table of nodes or interactions.

ANIMO automatically translates the user input to a TA model, which is then simulated with the model checking tool UPPAAL [20]. The results are subsequently parsed and translated to a graph that shows the dynamic behaviour of nodes in the network. A schematic overview of this process is given in Supplementary Section **??**. No training or prior knowledge on the use of TA or UPPAAL is needed in order to benefit from ANIMO. Nevertheless, the TA model and the model checking process in UPPAAL can be accessed when desired by the user.

The dynamic behaviour of a model can be interactively explored by moving a time slider underneath the graph to highlight time points in a simulation. In the network view, each node will be coloured according to its activity level at the selected time point. Experimental data can be compared to the model by importing and superposing these data upon an output graph from the model (Figure 2 B,D,F). The ANIMO user workflow and the features described above are illustrated in Suppl. Video 1.

Competing interests
The authors declare that they have no competing interests.

Author's contributions
S.S. designed and performed the experiments, developed the Cytoscape integration, wrote the manuscript; J.S. conceived, designed and performed the experiments, wrote the manuscript; B.W. developed the Cytoscape integration, and designed and developed the time slider UI; P.E.vdV. initiated the study, conceived the Cytoscape implementation, supervised the project; R.A.U.C. performed experiments; M.K. designed experiments, analyzed data and wrote the manuscript; R.L. Contributed to the methodology and supervised the project; J.vdP. contributed to the strategy and methodology in the manuscript, in particular the connection with formal methods; J.N.P. designed experiments, analyzed data sets, contributed in particular to the application of ANIMO for large biological data, wrote the manuscript, and supervised the project.

[1] All acronyms used in this paper and their corresponding UniProt IDs are listed in Suppl. Sect. **??**.

Acknowledgements

We would like to thank Christof Francke for valuable discussions.

Author details

[1]Formal Methods and Tools, Faculty of EEMCS, University of Twente, 7522NH Enschede, The Netherlands. [2]Developmental BioEngineering, MIRA Institute for Biomedical Technology and Technical Medicine, University of Twente, 7522NH Enschede, The Netherlands. [3]Human Media Interaction, Faculty of EEMCS, University of Twente, 7522NH Enschede, The Netherlands.

References

1. Mendoza, L., Thieffry, D., Alvarez-Buylla, E.R.: Genetic control of flower morphogenesis in arabidopsis thaliana: a logical analysis. Bioinformatics **15**(7), 593–606 (1999)
2. Shmulevich, I., Zhang, W.: Binary analysis and optimization-based normalization of gene expression data. Bioinformatics **18**(4), 555–565 (2002)
3. de Jong, H.: Modeling and simulation of genetic regulatory systems: A literature review. Journal of Computational Biology **9**, 67–103 (2002)
4. Fisher, J., Piterman, N., Hubbard, E.J.A., Stern, M.J., Harel, D.: Computational insights into caenorhabditis elegans vulval development. Proceedings of the National Academy of Sciences of the United States of America **102**(6), 1951–1956 (2005)
5. Efroni, S., Harel, D., Cohen, I.R.: Toward rigorous comprehension of biological complexity: Modeling, execution, and visualization of thymic t-cell maturation. Genome Research **13**(11), 2485–2497 (2003)
6. Dematté, L., Priami, C., Romanel, A.: Modelling and simulation of biological processes in BlenX. SIGMETRICS Perform. Eval. Rev. **35**, 32–39 (2008)
7. Ciocchetta, F., Hillston, J.: Bio-PEPA: A framework for the modelling and analysis of biological systems. Theor. Comput. Sci. **410**, 3065–3084 (2009)
8. Siebert, H., Bockmayr, A.: Temporal constraints in the logical analysis of regulatory networks. Theor. Comput. Sci. **391**(3), 258–275 (2008)
9. Bartocci, E., Corradini, F., Merelli, E., Tesei, L.: Model checking biological oscillators. Electronic Notes in Theoretical Computer Science **229**(1), 41–58 (2009). Proceedings of the Second Workshop From Biology to Concurrency and Back (FBTC 2008)
10. Batt, G., Salah, R.B., Maler, O.: On timed models of gene networks. In: Proceedings of the 5th International Conference on Formal Modeling and Analysis of Timed Systems. FORMATS'07, pp. 38–52. Springer, Berlin, Heidelberg (2007)
11. Reisig, W.: Petri nets. In: Koch, I., Reisig, W., Schreiber, F., Dress, A., Vingron, M., Myers, G., Giegerich, R., Fitch, W., Pevzner, P.A. (eds.) Modeling in Systems Biology. Computational Biology, vol. 16, pp. 37–56. Springer, London (2011)
12. Bonzanni, N., Krepska, E., Feenstra, K.A., Fokkink, W., Kielmann, T., Bal, H.E., Heringa, J.: Executing multicellular differentiation: quantitative predictive modelling of c.elegans vulval development. Bioinformatics/computer Applications in The Biosciences **25**, 2049–2056 (2009)
13. Alur, R., Dill, D.L.: A theory of timed automata. Theor. Comput. Sci. **126**(2), 183–235 (1994)
14. Bengtsson, J., Griffioen, W.O.D., Kristoffersen, K.J., Larsen, K.G., Larsson, F., Pettersson, P., Yi, W.: Automated verification of an audio-control protocol using UPPAAL. The Journal of Logic and Algebraic Programming **52-53**(0), 163–181 (2002)
15. Hessel, A., Pettersson, P.: Model-based testing of a wap gateway: an industrial case-study. In: Proceedings of the 11th International Workshop, FMICS 2006 and 5th International Workshop, PDMC Conference on Formal Methods: Applications and Technology. FMICS'06/PDMC'06, pp. 116–131. Springer, Berlin, Heidelberg (2007)
16. Marques Jr., A.P., Ravn, A.P., Srba, J., Vighio, S.: Model-checking web services business activity protocols. International Journal on Software Tools for Technology Transfer **15**(2), 125–147 (2013)
17. Thomas, R.: Boolean formalization of genetic control circuits. J. Theor. Biol. **42**(3), 563–585 (1973)
18. Nakano, S., Yamaguchi, S.: Two modeling methods for signaling pathways with multiple signals using UPPAAL. In: Proceedings of the 2nd International Workshop on Biological Processes & Petri Nets (BioPPN2011), pp. 87–101. CEUR-WS.org, Aachen (2011)
19. Man, K.L., Krilavicius, T., Wan, K., Hughes, D., Lee, K.: Modeling and analysis of radiation therapy system with respiratory compensation using UPPAAL. In: Proceedings of the 9th IEEE International Symposium on Parallel and Distributed Processing with Application - ISPA 2011, pp. 50–54 (2011)
20. Larsen, K.G., Pettersson, P., Yi, W.: UPPAAL in a nutshell. International Journal on Software Tools for Technology Transfer (STTT) **1**, 134–152 (1997)
21. Schivo, S., Scholma, J., Wanders, B., Camacho, R.A.U., van der Vet, P.E., Karperien, M., Langerak, R., van de Pol, J., Post, J.N.: Modelling biological pathway dynamics with Timed Automata. IEEE Journal of Biomedical and Health Informatics **XX**(XX), (2013)
22. Fathallah-Shaykh, H.M., Bona, J.L., Kadener, S.: Mathematical model of the drosophila circadian clock: Loop regulation and transcriptional integration. Biophysical Journal **97**(9), 2399–2408 (2009)
23. Clarke, E.: Model checking. In: Ramesh, S., Sivakumar, G. (eds.) Foundations of Software Technology and Theoretical Computer Science. Lecture Notes in Computer Science, vol. 1346, pp. 54–56. Springer, Berlin / Heidelberg (1997)
24. Emerson, E.A., Clarke, E.M.: Using branching time temporal logic to synthesize synchronization skeletons. Science of Computer Programming **2**(3), 241–266 (1982)
25. Kanehisa, M., Goto, S.: KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Research **28**(1), 27–30 (2000)
26. Hornbeck, P.V., Chabra, I., Kornhauser, J.M., Skrzypek, E., Zhang, B.: PhosphoSite: A bioinformatics resource dedicated to physiological protein phosphorylation. Proteomics **4**(6), 1551–1561 (2004)

27. Gaudet, S., Janes, K.A., Albeck, J.G., Pace, E.A., Lauffenburger, D.A., Sorger, P.K.: A compendium of signals and responses triggered by prodeath and prosurvival cytokines. Mol Cell Proteomics **4**(10), 1569–1590 (2005)
28. Janes, K.A., Gaudet, S., Albeck, J.G., Nielsen, U.B., Lauffenburger, D.A., Sorger, P.K.: The response of human epithelial cells to TNF involves an inducible autocrine cascade. Cell **124**(6), 1225–1239 (2006)
29. Davis, R.J.: Signal transduction by the JNK group of MAP kinases. Cell **103**(2), 239–252 (2000)
30. Bannister, A.J., Brown, H.J., Sutherland, J.A., Kouzarides, T.: Phosphorylation of the c-Fos and c-Jun HOB1 motif stimulates its activation capacity. Nucleic Acids Res **22**(24), 5173–6 (1994)
31. Bailly, S., Fay, M., Israël, N., Gougerot-Pocidalo, M.A.: The transcription factor AP-1 binds to the human interleukin 1 alpha promoter. European Cytokine Network **7**(2), 125–128 (1996)
32. Russell, M., Lange-Carter, C.A., Johnson, G.L.: Direct interaction between Ras and the kinase domain of mitogen-activated protein kinase kinase kinase (MEKK1). Journal of Biological Chemistry **270**(20), 11757–11760 (1995)
33. Dérijard, B., Hibi, M., Wu, I.-H., Barrett, T., Su, B., Deng, T., Karin, M., Davis, R.J.: JNK1: A protein kinase stimulated by uv light and ha-ras that binds and phosphorylates the c-Jun activation domain. Cell **76**(6), 1025–1037 (1994)
34. Janes, K.A., Albeck, J.G., Gaudet, S., Sorger, P.K., Lauffenburger, D.A., Yaffe, M.B.: A systems model of signaling identifies a molecular basis set for cytokine-induced apoptosis. Science **310**(5754), 1646–1653 (2005)
35. Aldridge, B.B., Saez-Rodriguez, J., Muhlich, J.L., Sorger, P.K., Lauffenburger, D.A.: Fuzzy logic analysis of kinase pathway crosstalk in TNF/EGF/Insulin-induced signaling. PLoS Comput Biol **5**(4), 1000340 (2009)
36. Killcoyne, S., Carter, G.W., Smith, J., Boyle, J.: Cytoscape: a community-based framework for network modeling. Methods in molecular biology (Clifton, N.J.) **563**, 219–239 (2009)

Figures

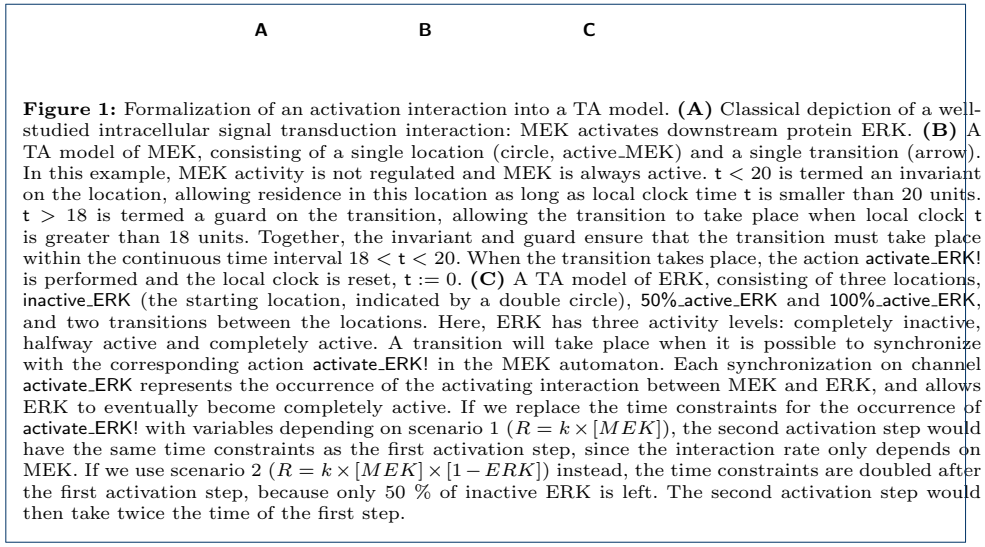

Figure 1: Formalization of an activation interaction into a TA model. **(A)** Classical depiction of a well-studied intracellular signal transduction interaction: MEK activates downstream protein ERK. **(B)** A TA model of MEK, consisting of a single location (circle, active_MEK) and a single transition (arrow). In this example, MEK activity is not regulated and MEK is always active. $t < 20$ is termed an invariant on the location, allowing residence in this location as long as local clock time t is smaller than 20 units. $t > 18$ is termed a guard on the transition, allowing the transition to take place when local clock t is greater than 18 units. Together, the invariant and guard ensure that the transition must take place within the continuous time interval $18 < t < 20$. When the transition takes place, the action activate_ERK! is performed and the local clock is reset, $t := 0$. **(C)** A TA model of ERK, consisting of three locations, inactive_ERK (the starting location, indicated by a double circle), 50%_active_ERK and 100%_active_ERK, and two transitions between the locations. Here, ERK has three activity levels: completely inactive, halfway active and completely active. A transition will take place when it is possible to synchronize with the corresponding action activate_ERK! in the MEK automaton. Each synchronization on channel activate_ERK represents the occurrence of the activating interaction between MEK and ERK, and allows ERK to eventually become completely active. If we replace the time constraints for the occurrence of activate_ERK! with variables depending on scenario 1 ($R = k \times [MEK]$), the second activation step would have the same time constraints as the first activation step, since the interaction rate only depends on MEK. If we use scenario 2 ($R = k \times [MEK] \times [1 - ERK]$) instead, the time constraints are doubled after the first activation step, because only 50 % of inactive ERK is left. The second activation step would then take twice the time of the first step.

Tables

Table 1: Sample table title. This is where the description of the table should go.

	B1	B2	B3
A1	0.1	0.2	0.3
A2	...	..	.
A3	..	.	.

Additional Files

Additional file 1 — Sample additional file title
Additional file descriptions text (including details of how to view the file, if it is in a non-standard format or the file extension). This might refer to a multi-page table or a figure.

Additional file 2 — Sample additional file title
Additional file descriptions text.

Figure 2: Incremental construction of an ANIMO model of signal transduction events in human colon carcinoma cells upon stimulation with 100 ng/ml TNFα. Each construction step (top to bottom) is simulated in ANIMO, giving intermediate feedback useful for the piecewise refinement of the model. The graphs on the right show the dynamic behaviour of the corresponding models on the left, comparing it to the measured activity values by [27] (error bars represent the standard deviation). On the vertical axis, "100" represents the maximum protein activity in the complete experiment. The red vertical line in each graph indicates a selected time point in the time course. Nodes in the corresponding network representation are coloured according to their activity at that time point. All images in this figure are taken from the ANIMO user interface. **(A, B)** Basic model showing direct activation of JNK1 and MK2 by TNFα. No peak dynamics are observed because no inactivating processes are present. **(C, D)** The model after addition of inactivating phosphatases and a negative feedback loop that down-regulates TNFR. Note that adding TNFR internalization or phosphatases alone would not be enough to reproduce activity peaks. **(E, F)** The model after addition of IKK, IL1-secretion (abstracting the autocrine IL-1 signalling described by [28]), Casp8 and Casp3, showing the late response to TNFα signalling.
The _data suffix identifies experimental data; all other series are computed by ANIMO.
As the data set did not contain values for cleaved caspase-3, but only for its non-cleaved precursor pro-caspase-3, we computed the Casp3_data series as 100% − [pro-Casp3].

Figure 3: ANIMO model of the circadian clock in *Drosophila Melanogaster* Autoregulatory negative feedback loops are present on each of the nodes of the network, following the original model by [22]. These feedback loops ensure that protein levels decrease over time when activating inputs are absent. The feedback loops are not represented here for cosmetic reasons and clarity. Naming conventions follow the same rules as in the original model, with lower-case names representing mRNA, and upper-case names indicating proteins.

Figure 4: Signalling network downstream of TNFα and EGF in human colon carcinoma cells. **(A)** The model for the merged TNFα and EGF pathways. Node colours represent the activity level of the corresponding modelled reactants at time $t = 15$ minutes after a stimulation of 100 ng/ml TNFα + 100 ng/ml EGF. **(B)** Modelled production of IL-1α after stimulation with 100 ng/ml TGFα (24 hours). **(C)** Modelled activation of JNK1 and MK2 after stimulation with 5 ng/ml TNFα + 10 μg/ml C225 (2 hours).
The _data suffix identifies experimental data; all other series are computed by ANIMO.

Figure 5: (A) The model for the merged TNFα and EGF pathways after addition of the two hypotheses (highlighted). Hypothesis 1 assumes IL-1α expression to depend on AP-1 activity, which in turn requires both c-Jun en c-Fos to be activated by JNK1 and ERK, respectively. Hypothesis 2 assumes RAS as an activator of MEKK1. Node colours represent the activity levels 15 minutes after stimulation of 100 ng/ml TNFα + 100 ng/ml EGF. **(B)** After the addition of the first hypothesis (activation of IL-1α production depending both on JNK1 and ERK): production of IL-1α after stimulation with 100 ng/ml TNFα (series IL-1a (TNFa)) compared with stimulation with 100 ng/ml TGFα (series IL-1a (TGFa)) (24 hours).
(C) After the addition of the second hypothesis (activation of MEKK1 downstream of EGFR): stimulation with 5 ng/ml TNFα + 10 μg/ml C225 (2 hours). Suppl. Sect. **??** explains how the dosage of 5 ng/ml TNFα was represented in the model.
The _data suffix identifies experimental data; all other series are computed by ANIMO.