METHODOLOGY

Modelling with ANIMO: between Fuzzy Logic and Differential Equations

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

[*]Correspondence: r.langerak@utwente.nl
[1]Formal Methods and Tools, Faculty of EEMCS, University of Twente, P.O. Box 217, 7500AE Enschede, The Netherlands
Full list of author information is available at the end of the article
[#]Equal contributor

Abstract

Background: Experimental biologists need to deal with large amounts of data on a daily basis. This makes computational support essential in order to reason on the dynamics of biological systems. The software tool ANIMO (Analysis of Networks with Interactive MOdelling) provides such computational support and allows insight into the complex networks of signalling events occurring in living cells. Biology experts are able to exploit the power of the underlying formal methods via a user interface specifically tailored for biological applications. Knowledge of the underlying formalisms is not necessary to use ANIMO.

Results: ANIMO models show a degree of precision intermediate between ordinary differential equations (ODEs) and fuzzy logic. We show this by presenting ANIMO models for two case studies: *Drosophila melanogaster* circadian clock, and signal transduction events downstream of TNFα and EGF in HT-29 human colon carcinoma cells. The models were originally developed with ODEs and fuzzy logic, respectively.

Conclusions: ANIMO models replicate with good precision the results of both the ODE and fuzzy logic models. Moreover, ANIMO models require less parameters than ODEs and are more precise than fuzzy logic. For this reason we position the modelling paradigm of ANIMO between ODEs and fuzzy logic.

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

Background

Modelling in cell biology

Executable biology is a young subfield in computational modelling, aimed at constructing models that mimic biological phenomena *in silico*. It provides an interesting paradigm to enhance network diagrams with an underlying formal description of network components and their interactions. Executable models typically contain descriptions of the direct interactions of network components, requiring the modeller to think in terms of "cause and effect". Biologists reason in similar ways about the molecular mechanisms of network interactions, and this makes the construction of an executable model an intuitive process. ANIMO (Analysis of Networks with Interactive MOdelling, [1, 2]) provides an enabling technology to increase the use of computational models by experimental biologists using their domain-specific language. ANIMO enriches the normally static biological network diagrams with dynamic information, which is then used to automatically produce formal models representing the biological network. Such models are indispensable for formally

comparing experimental data with prior knowledge, or for structuring experimental findings into a new theory. When dealing with complex biological networks, executable biology models are particularly useful to understand the non-linear dynamics and the entailed emergent properties of the networks. In those cases, an ANIMO model can be used as a support to obtain insight based on abstract representations of the interactions occurring inside living cells. Other applications of ANIMO models include performing *in silico* experiments and the storage and transfer of knowledge on biological networks.

A brief introduction to ANIMO

The user interface of ANIMO is displayed in Figure 1, where we present an example of a biological network enriched with dynamic information. ANIMO is implemented as a plug-in to Cytoscape [3], a software tool developed to represent biological networks. On top of the static topological information displayed in Cytoscape, ANIMO represents biological interactions starting from the basic concept of *activity*: each biological entity in an ANIMO model is considered to be either active or inactive. Activity is to be interpreted in a very broad sense: for example, an active gene is being transcribed, an active kinase can perform phosphorylations, etc. Each node in an ANIMO network represents both active and inactive entities of the same type, with the relative amount of active entities (the *activity level*) represented by the node colour on a user-configurable scale. Interactions among nodes define how the biological entities in a network influence each other's activity. Only nodes whose activity level is larger than 0 (such nodes are called *active*) can have an influence on their downstream targets, and only if that influence is not counterbalanced by intervening opposite interactions. For example, the interaction $A \rightarrow B$ (read "A activates B") indicates that node A, if active, will increase the activity level of node B. If we add an additional interaction to the example, $C \dashv B$ ("C inhibits B"), with C also active, then the activity level of B will change depending on the activity levels of A and C, and on their quantitative influence. The influence of an ANIMO interaction is quantified by a parameter k, which defines the speed at which that interaction occurs: higher values of k give faster interactions. These k-values are the only parameters needed in an ANIMO model, and can be given in either a quantitative (as real numbers) or a qualitative way, choosing among self-explanatory descriptions such as "very slow", "slow", "medium", "fast", "very fast".

ANIMO produces graphs showing how the activity levels of selected nodes change over time, allowing the user to obtain a view on the dynamic behaviour of their network. In order to obtain these results, a model defined in ANIMO is automatically translated into its corresponding representation as a network of Timed Automata [4] and then analysed *behind the scenes* with the software tool UPPAAL [5]. The formal language of Timed Automata allows to represent and analyse complex behaviours precisely and efficiently, but the user does not need to directly manipulate Timed Automata or UPPAAL, as the analysis process is made transparent. A fastidious user can still access the underlying models and perform other analyses in UPPAAL, but that is not required in order to fully profit from ANIMO.

A detailed description on how the Timed Automata models defined by ANIMO work, and how the results are obtained, can be found in [1]. The choice of parameters

for ANIMO models is described in [6]. Additional guidance on the design of ANIMO models and how to best profit from biological experimental data can be found in [2].

Figure 2 shows the position of ANIMO in a spectrum of modelling methods: justification for this placement is given in the Results section.

Results

Modelling oscillation in Drosophila Melanogaster circadian clock

To demonstrate that results obtained with ANIMO are comparable to results from widely used modelling approaches, we present an ANIMO model of the circadian clock in *Drosophila Melanogaster* (Suppl. Fig. S4). This ANIMO model is based on [7], where ordinary differential equations (ODEs) were used. The cyclic behaviour 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 acts 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. D.2).

The output of the ANIMO model (see Suppl. Fig.S4) matches the original ODE model. In particular, both models achieve an oscillatory behaviour with similar periods and phases (see Fig. 3 for an example, and Suppl. Fig S6 for the complete comparison). Details of the comparison between the ANIMO model and the original ODE model are given in Supplementary Section D.3.

Using ANIMO to generate hypotheses in human colon carcinoma cells

We now present a comparison with an existing fuzzy logic model, which we use also as an example of how ANIMO can be used to create reference models and help to obtain insight into complex biological events.

We constructed a model of the signalling network downstream of TNFα and EGF in HT-29 human colon carcinoma cells, formalizing the crosstalk that takes place between the pathways at different levels of cellular regulation. We first modelled the two pathways in isolation (Suppl. Figs. S7, S8), using information on protein interactions from the KEGG [8] and phosphosite [9] databases. These models were fitted to experimental data from previous studies [10, 11]. We then merged the two pathways into a single model and added autocrine crosstalk between the pathways that has been described in [11]. Briefly, stimulation with TNFα (TNFa in the model) leads to a rapid release of TGFα (TGFa), 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) [10].

At this point, the behaviour of the model deviated from the data for some of the nodes. Changing the parameters of the model was not enough to reproduce

the behaviour shown by experimental data. This is an interesting situation, as it requires changes in the topology of the model [6], reflecting the formulation of a new hypothesis on the structure of the model. 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 [11]. 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. S9, 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 [12, 13]. 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 [14]. 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 (Fig. 4C). However, experimental data show that C225 strongly reduces activation of JNK1 and MK2 upon stimulation with TNFα [11]. 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. S9). 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 [15]. EGFR is a well-known and potent activator of Ras, which is why it was already in our network [8]. 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 [13, 16]. 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 Supplementary Figures S10, S11 and S12. Comparing these results with the ones from [17] shows a better fit of the ANIMO model, which is also intrinsically more precise, being more mechanistic in nature (see Fig. 2). 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

Final remarks on the models

We first described the construction of an ANIMO model of the circadian clock in *Drosophila Melanogaster*. This model captured the dynamics of the regulatory network and led to similar conclusions as an ODE model that had been published previously [7]. This finding supports the use of the series of modelling abstractions on which ANIMO is based. The biggest difference between the construction of these models is that the model in [7] is constructed by writing a system of mathematical equations, together with an algorithm for simulation. In ANIMO 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 quantitative effect of each interaction. This approach to construct a model is closer to biological practice. 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. All this comes at the price of lower model precision: the curves representing oscillation of protein activities in the ANIMO model are not as precise as those obtained from the original ODE model (see Suppl. Fig.S6).

We also 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 [18, 17]. 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 insight in the dynamic behaviour of the network (see Fig. 2). 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, as measured in wet-lab experiments, were not reproduced by the fuzzy logic model. Moreover, the fuzzy logic model gave no insight in the molecular interactions that are involved in activation or inhibition.

Here we used a data set based on the wet-lab experiments described in [10]. We used the resulting experimental data, together with knowledge from curated databases [8, 9] to construct an executable model of the biological system. In contrast to the two approaches described above (partial least-squares regression and fuzzy logic), 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 [11]. 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.

User experience: ANIMO and other modelling tools

Different formalisms are in use in the field of computational modelling of biological systems, each with their specific characteristics. Many of these formalisms have been implemented into software tools to support modelling efforts. To compare ANIMO with existing tools, we have selected a number of mathematical formalisms, each connected to a supporting tool. With an emphasis on the modelling process rather than the final model, and in an attempt to evaluate the degree of "interactivity" of these tools, we compared them on the basis of the following parameters:

1. **Domain-specific interface:** the underlying formalism is manipulated through an interface targeted towards the biological domain
2. **Visual modelling:** the tool allows the user to model using a visual interface, and is not exclusively founded on formula-, text- or table-based input forms
3. **Qualitative parameters:** parameters for reactions can be input as approximated estimations, and not exclusively as numbers
4. **Tight coupling with topology:** models are tightly and clearly coupled to the networks they represent, showing the visual representation of the model in a shape similar or comparable to the representation currently used by biologists for signalling pathways
5. **User-chosen granularity:** if discretization is applied during the modelling process, the user can change the granularity with which such discretization is made, possibly for each component of the model separately

Table 1 shows the comparison between ANIMO and the selected tools.

As related work, we also signal the powerful tool CellNOpt [19]. CellNOpt is a software which can work with logic descriptions (Boolean, fuzzy) and differential equations, and automatically suggests the best network topologies to match a given data set. Thanks to the CytoCopteR plug-in for Cytoscape [20], which provides an accessible user interface, CellNOpt can be used in tandem with ANIMO: after computing the most likely network topologies with CytoCopteR, the biologist can carry on the analysis process with ANIMO, working on new hypotheses to explain the experimental data.

Going beyond the user interface, there are a number of "pros and cons" for using ANIMO and Timed Automata in the biological context. First and foremost, as Timed Automata is an executable formal language, a state space can be derived from a Timed Automata model. This means that state space-related analyses such as model checking can be performed on Timed Automata: this can be done directly in ANIMO, as ANIMO acts as an intermediary towards the powerful model checking tool UPPAAL.

The smaller amount of parameters needed to build an ANIMO model, if compared e.g. to a differential equation-based approach, makes the modelling process faster, allowing the biologist to obtain interesting results in less time. A less parameter-intensive modelling approach may be more appropriate to some applications: in our case we abstract sequences of biochemical reactions into activation or inactivation interactions, simplifying any relation between two entities in a network as either activating or inhibiting. This allows for some simplifications in the networks and the underlying formal models, as it does not represent in full detail more specific concepts such as reactant concentration or reaction affinity.

For each reactant modelled in ANIMO we assume that the total amount of active and inactive molecules remains constant. While this assumption is not always applicable, it encourages abstract thought: many biological processes can be represented as networks driven by *activity-based* interactions (see Methods Section). Even if with a limited scope, ANIMO can be applied also in the analysis of metabolic processes, considering the production of protein A as the activation of node A, with degradation being represented as inhibition. As stated before, it cannot be expected from such models to be a completely realistic representation of their target biological processes, but they can still be a useful tool. This can be seen for example in the circadian clock model in the Results section, where mRNA and protein concentrations were abstracted to activity-driven processes in the ANIMO model.

Conclusions

In this paper we discussed the placement of ANIMO among other modelling paradigms and tools, highlighting ANIMO's strong points.

From the point of view of model precision, we position ANIMO between fuzzy logic and ODEs. Being less parameter-intensive than ODEs and more precise than logic-based models, ANIMO models are flexible enough to be useful in many occasions.

ANIMO adds a dynamic component to the static networks already familiar to biologists, allowing the domain experts to build formal executable models of complex biological networks. ANIMO is not the first tool to provide an interface to

a modelling formalism: as shown in Table 1, such interfaces exist in many other tools. Focusing on user-friendliness and interactive modelling, ANIMO makes computational modelling more accessible to experts in biology. Thanks to 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.

Methods

Modelling biological interactions with Timed Automata

Timed Automata have been shown to be a powerful formalism to model biological processes [21, 22, 23]. A timed automaton consists of locations and transitions between these locations (see Fig. 6), 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 clocks associated to it, allowing temporal control of transitions between locations. The 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 approximated reaction kinetics [1] can be used to calculate the timing of molecular interactions (see also Suppl. Sect. C.1). Figure 6 presents a small example that illustrates the basic properties of TA. This model describes the activation of ERK by MEK.

Example: building a model based on data

To illustrate the use of ANIMO in a practical environment, we will demonstrate the generation of a basic version of the model described in the Results Section. The model is based on a literature compendium of signal transduction events in HT-29 human colon carcinoma cells [10]. This data set comprises triplicate measurements of 11 different protein activities or post-translational modification states at 13 time points after treatment with different combinations of tumour necrosis factor-α (TNFα), epidermal growth factor (EGF) and insulin. The data set contains relative protein levels and activities, and no absolute quantities, which is the typical situation in biochemistry. To start, we normalized measurements for each protein to the maximum value in the complete experiment, resulting in a nondimensionalized data set that is suitable for use with ANIMO (see Suppl. Sect. B).

In Figure 7 we show the stepwise construction of a model of a small part of the network that is able to account for measured variations in activity of inhibitor of nuclear factor kappa-B kinase (IKK), c-Jun N-terminal kinase-1 (JNK1), mitogen-activated protein kinase-activated protein kinase 2 (MK2), Caspase 8 (casp-8) and Caspase 3 (casp-3) upon stimulation with 100 ng/ml TNFα. In this example we aimed for inclusion of a minimum number of nodes in the network, while preserving biological relationships. Multi-step cascades were aggregated into a single step when possible. Parameters for all reactions were set manually, resulting in a close match between the model and the patterns observed in the dataset.

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; P.E.vdV. initiated the study, conceived the Cytoscape implementation, supervised the project; M.K. designed experiments, analysed data and wrote the manuscript; J.N.P. designed experiments, analysed data sets, contributed in particular to the application of ANIMO for large biological data, wrote the manuscript, and supervised the project; J.vdP. contributed to the strategy and methodology in the manuscript, in particular the connection with formal methods; R.L. Contributed to the methodology and supervised the project.

Acknowledgements
We would like to thank Christof Francke for valuable discussions.

Author details
[1]Formal Methods and Tools, Faculty of EEMCS, University of Twente, P.O. Box 217, 7500AE Enschede, The Netherlands. [2]Developmental BioEngineering, MIRA Institute for Biomedical Technology and Technical Medicine, University of Twente, P.O. Box 217, 7500AE Enschede, The Netherlands. [3]Human Media Interaction, Faculty of EEMCS, University of Twente, P.O. Box 217, 7500AE Enschede, The Netherlands.

References

1. Schivo, S., Scholma, J., Wanders, B., Urquidi Camacho, R.A., 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 **18**(3), 832–839 (2014)
2. Scholma, J., Schivo, S., Urquidi Camacho, R.A., van de Pol, J., Karperien, M., Post, J.N.: Biological networks 101: Computational modeling for molecular biologists. Gene **533**(1), 379–384 (2014)
3. 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)
4. Alur, R., Dill, D.L.: A theory of timed automata. Theor. Comput. Sci. **126**(2), 183–235 (1994)
5. Larsen, K.G., Pettersson, P., Yi, W.: UPPAAL in a nutshell. International Journal on Software Tools for Technology Transfer (STTT) **1**, 134–152 (1997)
6. Schivo, S., Scholma, J., Karperien, H.B.J., Post, J.N., van de Pol, J.C., Langerak, R.: Setting parameters for biological models with ANIMO. In: André, E., Frehse, G. (eds.) Proceedings 1st International Workshop on Synthesis of Continuous Parameters, Grenoble, France. Electronic Proceedings in Theoretical Computer Science, vol. 145, pp. 35–47 (2014)
7. 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)
8. Kanehisa, M., Goto, S.: KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Research **28**(1), 27–30 (2000)
9. 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)
10. 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)
11. 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)
12. Davis, R.J.: Signal transduction by the JNK group of MAP kinases. Cell **103**(2), 239–252 (2000)
13. 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)
14. 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)
15. 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)
16. 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)
17. 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)
18. 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)
19. Terfve, C., Cokelaer, T., Henriques, D., MacNamara, A., Goncalves, E., Morris, M., Iersel, M.v., Lauffenburger, D., Saez-Rodriguez, J.: Cellnoptr: a flexible toolkit to train protein signaling networks to data using multiple logic formalisms. BMC Systems Biology **6**(1), 133 (2012)
20. Terfve, C., Cokelaer, T., Henriques, D., MacNamara, A., Goncalves, E., Morris, M., Iersel, M.v., Lauffenburger, D., Saez-Rodriguez, J.: Cellnoptr: a flexible toolkit to train protein signaling networks to data using multiple logic formalisms. BMC Systems Biology **6**(1), 133 (2012)
21. Siebert, H., Bockmayr, A.: Temporal constraints in the logical analysis of regulatory networks. Theor. Comput. Sci. **391**(3), 258–275 (2008)
22. 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)

23. 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)
24. Schivo, S., Scholma, J., Wanders, B., Urquidi Camacho, R.A., van der Vet, P.E., Karperien, M., Langerak, R., van de Pol, J., Post, J.N.: Modelling biological pathway dynamics with Timed Automata. In: 2012 IEEE 12th International Conference on Bioinformatics & Bioengineering (BIBE), pp. 447–453. IEEE Computer Society, Los Alamitos, CA, USA (2012)
25. Ciocchetta, F., Duguid, A., Gilmore, S., Guerriero, M.L., Hillston, J.: The Bio-PEPA Tool Suite. International Conference on Quantitative Evaluation of Systems, 309–310 (2009)
26. Nagasaki, M., Saito, A., Jeong, E., Li, C., Kojima, K., Ikeda, E., Miyano, S.: Cell illustrator 4.0: a computational platform for systems biology. Stud Health Technol Inform **162**, 160–81 (2011)
27. Mendes, P., Hoops, S., Sahle, S., Gauges, R., Dada, J., Kummer, U.: Computational modeling of biochemical networks using COPASI systems biology. Methods in Molecular Biology, vol. 500, pp. 17–59. Humana Press, Totowa, NJ (2009). Chap. 2
28. Chaouiya, C., Remy, E., Mossé, B., Thieffry, D.: Qualitative analysis of regulatory graphs: A computational tool based on a discrete formal framework. In: Benvenuti, L., De Santis, A., Farina, L. (eds.) Positive Systems. Lecture Notes in Control and Information Sciences, vol. 294, pp. 830–832. Springer, Berlin / Heidelberg (2003)
29. de Jong, H., Geiselmann, J., Hernandez, C., Page, M.: Genetic Network Analyzer: qualitative simulation of genetic regulatory networks. Bioinformatics **19**(3), 336–344 (2003)

Figures

Figure 1: The Cytoscape user interface running the ANIMO plug-in. The *Network* panel in the centre contains the nodes-edges model of the example TNFα pathway (see Methods Section), with colours indicating node activity levels and shapes representing different protein categories (see the *Legend* on the left). The *Results Panel* on the right contains a graph plotting activity levels of selected nodes during the first 24 hours of simulation of the model. The slider under the graph allows the user to select the time instant (marked as a vertical red line in the graph) on which the colours nodes in the *Network* are based. The series with the _data suffix is experimental data from [10], considering a treatment with 100 ng/ml TNFα.
All acronyms used in this paper and their corresponding UniProt IDs are listed in Suppl. Sect. A.

Figure 2: The spectrum of modelling methods, with the addition of ANIMO. The precision of ANIMO models is halfway between fuzzy logic and ODE. Compared to other modelling tools, ANIMO allows for an easier modelling experience thanks to a user-friendly interface based on the widely used network modelling software Cytoscape.

Figure 3: Comparing the result of the ANIMO model of *Drosophila Melanogaster* circadian clock with the model of [7]. 24 hours simulations of the two models were compared against each other, synchronizing their start point as much as possible. The blue line is the ANIMO model (_ANIMO series), while the red line represents the data computed from the original ODE model (_ODE series). The vaules of all series were rescaled on a [0, 100] interval to facilitate comparison.

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: Signalling network downstream of TNFα and EGF in human colon carcinoma cells: improved model. (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). The IL-1a (TGFa) series is always 0. **(C)** After the addition of the second hypothesis (activation of MEKK1 downstream of EGFR): activation of JNK1 and MK2 after stimulation with 5 ng/ml TNFα + 10 μg/ml C225 (2 hours). The JNK1 series is always 0. Suppl. Sect. D.4 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.

Figure 6: Abstraction of a biochemical reaction to a TA model. (A) Classical depiction of a well-studied intracellular signal transduction reaction: protein MAPK-ERK kinase (MEK) activates downstream protein extracellular-regulated kinase (ERK). **(B)** A TA model of ERK, consisting of two locations (circles), inactive_ERK and active_ERK, and one transition (arrow) between the locations. This transition will take place when it is possible to synchronize with the corresponding action activate_ERK! in the MEK automaton. **(C)** A TA model of active MEK, consisting of one location and one transition. $t < 20$ is called an invariant on the location, allowing residence in this location as long as clock time t is smaller than 20 units. $t > 18$ is called a guard on the transition, allowing the transition to take place when clock t is greater than 18 units. Together, the invariant and guard in this example ensure that the transition must take place in the (continuous) time interval $18 < t < 20$. When the transition takes place, the action activate_Erk! is performed (thus allowing the ERK automaton to reach the active_ERK location) and the local clock coupled to this automaton is reset, $t := 0$.

Figure 7: Construction of an ANIMO model of signal transduction events in human colon carcinoma cells upon stimulation with 100 ng/ml TNFα. Graphs below show the dynamic behaviour of the corresponding models above, comparing it to the measured activity values from [10] (error bars represent the standard deviation). On the vertical axis, "100" represents the maximum protein activity in the complete experiment. A red vertical line in each graph highlights an arbitrary time point in the time course: nodes in the corresponding model are coloured according to their activity at that time point. **(A, D)** Basic model showing direct activation of JNK1 and MK2 by TNFα. No peak dynamics are observed because no inactivating processes are present. **(B, E)** 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. **(C, F)** The model after addition of IKK, IL1-secretion (abstracting the autocrine IL-1 signalling described in [11]), Casp8 and Casp3, showing the late response to TNFα signalling. 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 $100\% - [\text{pro-Casp3}]$.

Tables

Tool	Formalism	Domain-specific interface	Visual modeling	Qualitative parameters	Tight coupling with topology	User-chosen granularity
ANIMO [24]	Timed Automata	Yes	Yes	Yes	Yes	Yes
Bio-PEPA Workbench [25]	Bio-PEPA	No	No	No	No	Yes
Cell Illustrator [26]	Petri Nets	Yes	Yes	No	Yes	No
COPASI [27]	ODE, stochastic models	No	No	No	No	No
COSBI LAB [1]	BlenX	Yes	Yes	No	Yes	No
GINsim [28]	Boolean Networks	No	Yes	Yes	Yes	Yes [2]
GNA [29]	PLDE	No	Yes	Yes	Yes	Yes [2]
Rhapsody [3]	Statecharts	No	Yes	Yes	No [4]	No

Table 1: Comparison between ANIMO and some existing approaches to modelling biological systems. A "Yes" under a column indicates that the modelling tool (mostly) fulfils the parameter, "No" indicates very limited or no fulfilment.
[1] COSBILab web page `http://www.cosbi.eu/index.php/research/cosbi-lab`
[2] The user can choose the number of levels for each reactant, allowing to define multi-level models based on Boolean reaction dynamics.
[3] IBM Rational Rhapsody web page `http://www-01.ibm.com/software/rational/products/rhapsody/designer`
[4] Statecharts represent more closely the so-called *transition system* of the model as opposed to the components and interactions occurring among them.

Additional Files

Additional file 1 — ANIMO models and data for the case studies

The provided .zip file contains the Cytoscape .cys network files which can be analysed with ANIMO. In addition, the reference values in .csv format are provided, to compare ANIMO's results with the previously existing models. In order to grant a fair comparison, note that ANIMO's results need to be rescaled the same way as the ODE results (i.e. normalize each series w.r.t. to its minimum and maximum) before the comparison.