Additional Materials

Table of Contents

Appendix A: Naming conventions

Table S1 explains the abbreviations used in the paper.

Table S1: Explanation of the abbreviated names referring to molecular species in the main text.

Abbreviation	Full name	UniProt ID
Akt	protein kinase B	P31749
AP-1	activator protein 1	heterodimer of c-Jun and c-Fos
Casp3	caspase 3	P42574
Casp8	caspase 8	Q14790
c-Fos	proto-oncogene-protein c-fos	P01100
c-Jun	Jun activation domain binding protein	P05412
CLK	clock	O61735
CRY	cryptochrome	O77059
CWO	clockwork orange	Q9VGZ5
CYC	cycle	O61734
CYC/CLK	cycle-clock complex	
DBT	double-time kinase	O76324
DISC1	death-inducing signaling complex 1	
DISC2	death-inducing signaling complex 2	
EGF	epidermal growth factor	P01133
EGFR	EGF receptor	P00533
ERK	extracellular regulated kinase	P27361
FKHR	forkhead box protein O1	Q12778
IKK	inhibitor of nuclear factor kappa-B kinase	O14920
IL-1a	interleukin 1 α	P01583
IL-1R	interleukin 1 receptor	P14778
IL-1ra	interleukin 1 receptor antagonist	P18510
IRS1 (S)	insulin receptor substrate 1 (Serine 636)	P35568
IRS1 (Y)	insulin receptor substrate 1 (Tyrosine 896)	P35568
JNK1	c-Jun N-terminal kinase 1	P45983
MEK	MAPK ERK kinase	Q02750
MEKK1	MAPK/ERK kinase kinase 1	Q13233

Continued on next page

Table S1 - continued from previous page

Abbreviation	**Full name**	**UniProt ID**
MK2	mitogen-activated protein kinase-activated protein kinase 2	P49137
MKK3/6	dual specificity mitogen-activated protein kinase kinase 3/6	P46734 / P52564
MKK4/7	dual specificity mitogen-activated protein kinase kinase 4/7	P45985 / O14733
NF-kB	nuclear factor kappa-B	P19838
p38	mitogen-activated protein kinase p38	Q16539
PDP1	par-domain protein 1	Q9TVY0
PER	period	P07663
PER/TIM-p	phosphorylated period-timeless complex	
RAF	Raf	P04049
RAS	Ras GTPase-activating protein	P01112
TGFα	transforming growth factor α	P01135
TNFα	tumor necrosis factor-α	P01375
TNFR	TNF receptor	P19438
TIM	timeless	P49021
VRI	vrille	O18660

Appendix B: Normalizing experimental data for use with ANIMO

Document S1 in the Supplemental Data of [11] contains three tables, named Replicates, Averages and DPLSR dataset. The data we use to compare the results computed with ANIMO are based on the values from the Averages table. In particular, we compute activity data by performing a normalization on a $0 \ldots 100$ scale using this formula

$$v_{\text{norm}} = \frac{v}{v_{\text{max}}} \times 100$$

where v is the datum to be normalized taken from the column v in the Averages table, v_{norm} is the normalized value and v_{max} is the maximum value over the whole column v in the Replicates table. For each series, we compute also the standard deviation using the triplicate measurements present in table Replicates. The standard deviation is also normalized using the formula presented for the average. Each table produced with this process contains a subset of the columns from the Averages table, and refers to one treatment condition only. A column with time references is added to the table in first position. Finally, a column named *Number_of_levels* containing only the value 100 (see the instructions in the ANIMO manual available online at `http://fmt.cs.utwente.nl/tools/animo/content/Manual.pdf`) is added at the rightmost position. The tables are all exported in .csv format to be used with ANIMO, and are included in the *Model_and_data.zip* file in the additional materials of the present work.

Appendix C: ANIMO and Timed Automata

C.1 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.

As we have shown in the Results Section and in [1, 2, 7], the abstraction described here preserves ample expressivity to capture the dynamic behaviour of a biological network.

C.2 ANIMO

The modelling approach described in Section C.1 and in the Methods Section is implemented in the software tool ANIMO (Analysis of Networks with Interactive MOdelling, [1]) as a plug-in to the network visualization tool Cytoscape [3]. 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 [6]. 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 Figure S1. 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 (all the graphs in this paper were made directly in ANIMO). The ANIMO user workflow and the features described above are illustrated in a video on ANIMO's web page `http://fmt.cs.utwente.nl/tools/animo`.

C.3 Timed Automata model

The Timed Automata (TA) model underlying an ANIMO network is generated whenever an analysis is requested by the user. Starting from the network represented in the Cytoscape-based user interface, ANIMO automatically generates a TA model to be used with UPPAAL. The analysis result is then parsed and properly presented to the user, for example as a graph of reactant activity levels. This workflow is described in Figure S1.

Each TA model generated by ANIMO contains one automaton for each interaction (activation or inhibition) in the network. A TA representing an interaction

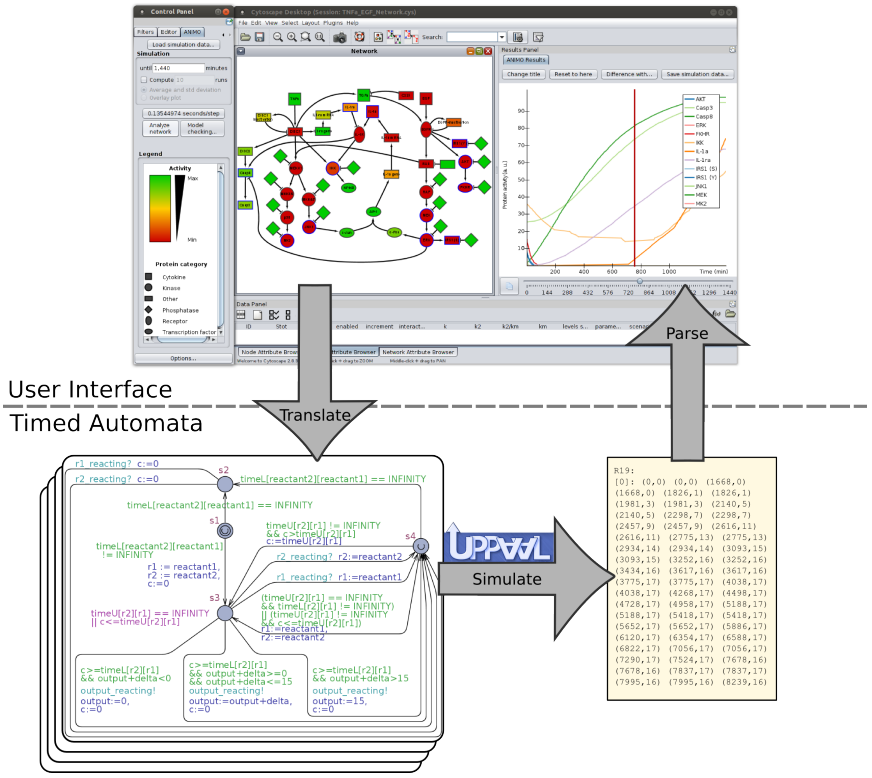

Figure S1: The passages intercurring between the press of the *Analyse network* button and the generation of an activity level graph. The model represented in the Cytoscape visual interface is automatically translated by ANIMO into a network of Timed Automata, which is then simulated using UPPAAL. A simulation produces, for each selected node, a series of pairs (t, a), with t a time instant along the simulation and a an activity level. These data are then parsed and represented as a graph, which is shown in ANIMO's user interface.

performs a cyclic series of steps, continuously updating the target of the interaction it represents, and adapting the timing of the next update according to the user-defined dynamics. Synchronizations between different automata occur when the activity level of a network component (e.g. ERK) changes: this allows the automata depending on that component to update their time settings.

The abstract behaviour of the interaction MEK $\rightarrow$ ERK in the TA model used in ANIMO is described in Figure S2. There, the activity levels of MEK and ERK are represented by variables called, respectively, $\text{MEK}_{\text{activity}}$ and $\text{ERK}_{\text{activity}}$. A more detailed description of the TA model underlying ANIMO was presented in the IEEE Journal of Biomedical and Health Informatics [1].

C.4 Granularity of an ANIMO network node

Figure S3 shows the differences between different choices for the number of levels of a node. This allows to adapt a model to the quality of experimental data.

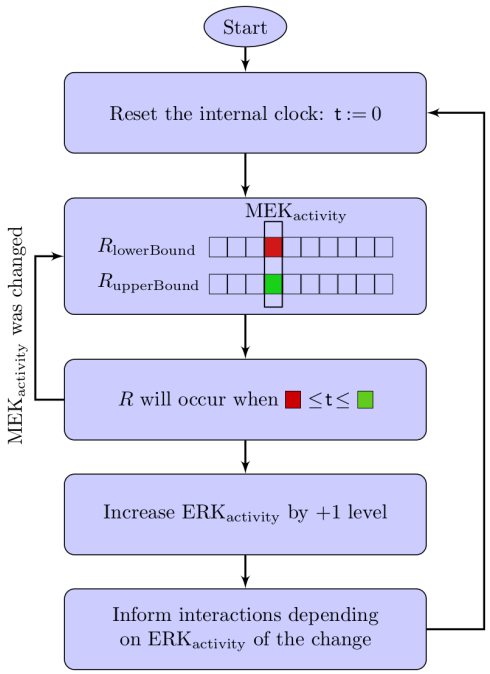

Figure S2: Schematic overview of the steps taken during a simulation run by a Timed Automaton modelling an interaction R that increases $\mathbf{ERK_{activity}}$ and depends only on $\mathbf{MEK_{activity}}$. In this example, MEK is represented with a granularity of 10 activity levels. After resetting the internal clock t, the automaton sets the time constraints for the interaction. $\mathrm{MEK_{activity}}$ is used as the index inside the time tables $R_{\mathrm{lowerBound}}$ and $R_{\mathrm{upperBound}}$, which contain pre-computed lower- and upper-bounds for the interaction timing. Once the bounds have been identified, R can occur when t reaches a value inside the continuous time interval [■, ■]. When it occurs, R increases the value of $\mathrm{ERK_{activity}}$ by 1. All interactions that depend on $\mathrm{ERK_{activity}}$ are notified of the change (via a synchronization on a specific channel), so that the associated time bounds are updated accordingly. After resetting the clock t, the process can restart. If $\mathrm{MEK_{activity}}$ was changed by another automaton before the occurrence of R, the time bounds are updated according to the new activity level of MEK.

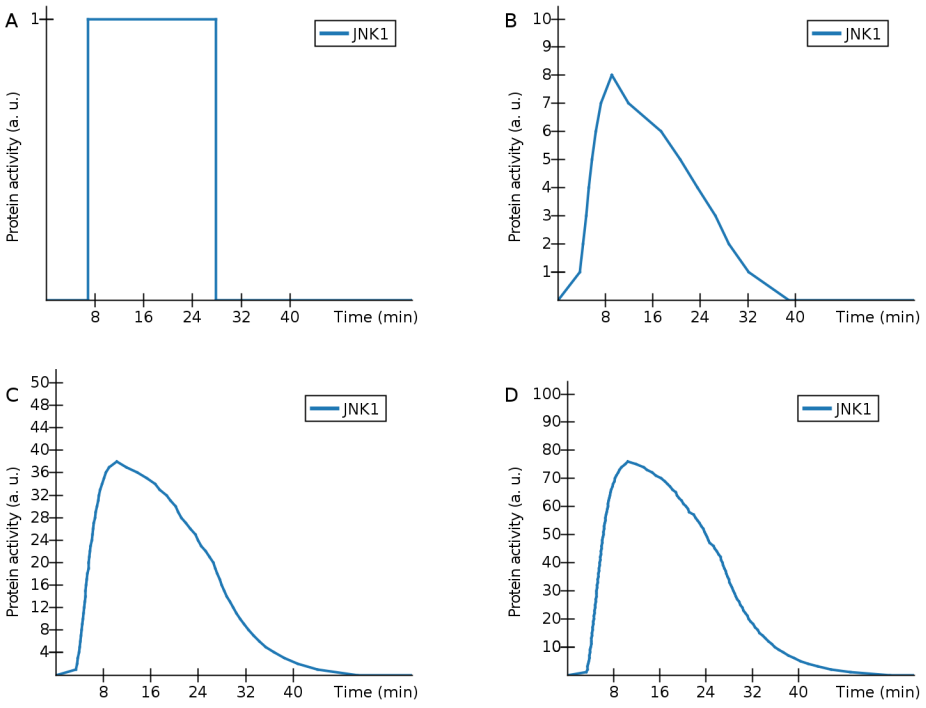

Figure S3: Comparing different reactant granularity settings. (A) 2 levels, **(B)** 10 levels, **(C)** 50 levels, **(D)** 100 levels. The JNK1 series is computed from the model presented in Figure 5, considering 100 ng/ml TNFα as treatment condition over a period of 60 minutes.

Appendix D: Additional notes

D.1 ANIMO model of the Drosophila circadian clock

The model built in ANIMO to represent the circadian clock of *Drosophila Melanogaster* can be seen in Figure S4. Its topology is based on the one presented in [8], while the parameters were set by hand in order to obtain an oscillatory behavior.

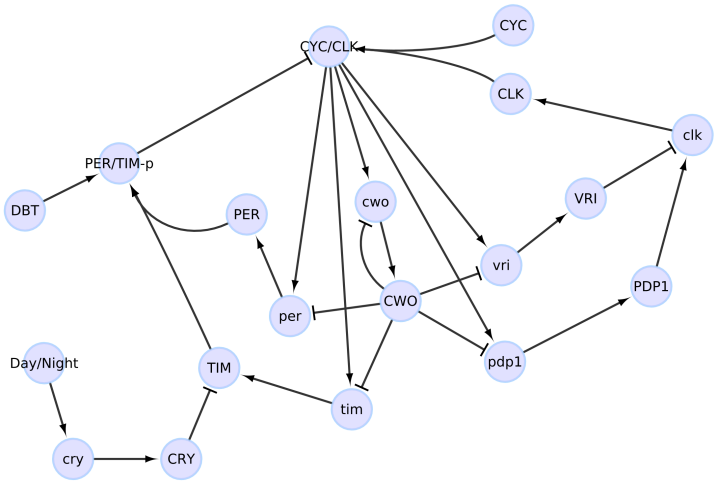

Figure S4: ANIMO model of the circadian clock in *Drosophila Melanogaster*. Auto-regulatory negative feedback loops are present on each of the nodes of the network, following the original model in [8]. These feedback loops ensure that protein levels decrease over time when activating inputs are absent. The feedback loops are not represented here for cosmetic reasons only. Naming conventions follow the same rules as in the original model, with lower-case names representing mRNA, and upper-case names indicating proteins.

D.2 Simulating the day-night cycle

The model presented in Figure S4 contains a node labelled Day/Night. That node abstracts our representation of the cyclic alternation of day and night, which causes the variations in cryptochrome (cry): these oscillations allow the network to synchronize to a time zone. Note that the network oscillates also when the node cry is not included in the model.

The alternation between day and night is represented in our model with a repressilator-like [28] subnetwork, as can be seen in Figure S5. In the model in [8] a specific function was introduced in the equations to approximate the experimental data from [29].

D.3 Comparing the models of the Drosophila circadian clock

We compared the simulation results from the ANIMO model presented in Figure S4 with the ODE model described in [8]. The raw data coming from the two models were aligned to have a roughly close initial point, and all amplitudes were normalized following the procedure described in Suppl. Sect. B. The results of this comparison can be seen in Figure S6.

Most of the molecules represented in the two models evolve with the same period and phase. CLK and clk in the ANIMO model have a small oscillation range (their values change by around 10% during a simulation), so their behaviour match the continuous model less precisely.

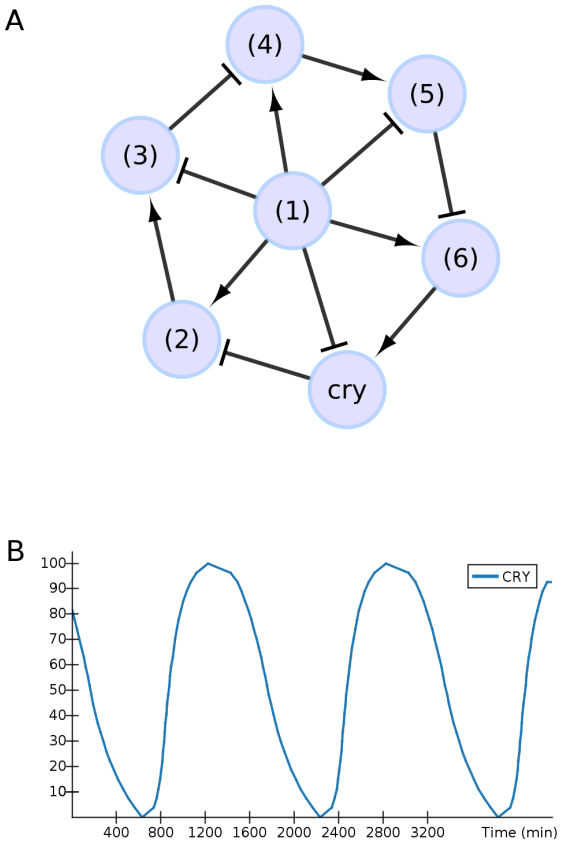

Figure S5: Modelling the dany/night cycle in ANIMO.(A) The repressilator-like subnetwork used to represent the alternation between day and night that cause the oscillations in CRY concentrations in the network modelled in the Results Section. **(B)** A graph plotting the oscillations in CRY along a period of three days.

D.4 Note on the parameters in the TNFα-EGF model

The parameters in the models in Figures 4 and 5 have been set by fitting the model to the experimental data for the treatment conditions with 100 ng/ml TNFα, 100 ng/ml EGF and 100 ng/ml TNFα + 100 ng/ml EGF. In the model we have set the activity of the corresponding TNFa and EGF nodes either at their maximum value or at 0, depending on the modelled treatment. This level is a dimensionless quantity that indicates either the maximum or minimum activity level in the data set. However, we found that in order to model the condition with 5 ng/ml TNFα we needed to set the activity of the TNFa node to a value higher than 5 out of 100. We believe that this has to do with the fact that 100 ng/ml is a highly supra-physiological concentration of TNFα, that will rapidly cause activation of all receptors present. Fitting the model to this experimental condition may have resulted in slight deviations in the parameter values. Nevertheless, the modelling results illustrate that building a model with basic kinetic rate laws can give useful predictions over a range of concentrations. Figure 4C was obtained using the model in Figure 4A and setting the activity of TNFa to 12 out of 100 (lower values gave 0 activity for both JNK1 and MK2), while Figure 5C was obtained from the model in Figure 5A with TNFa set at 8 out of 100.

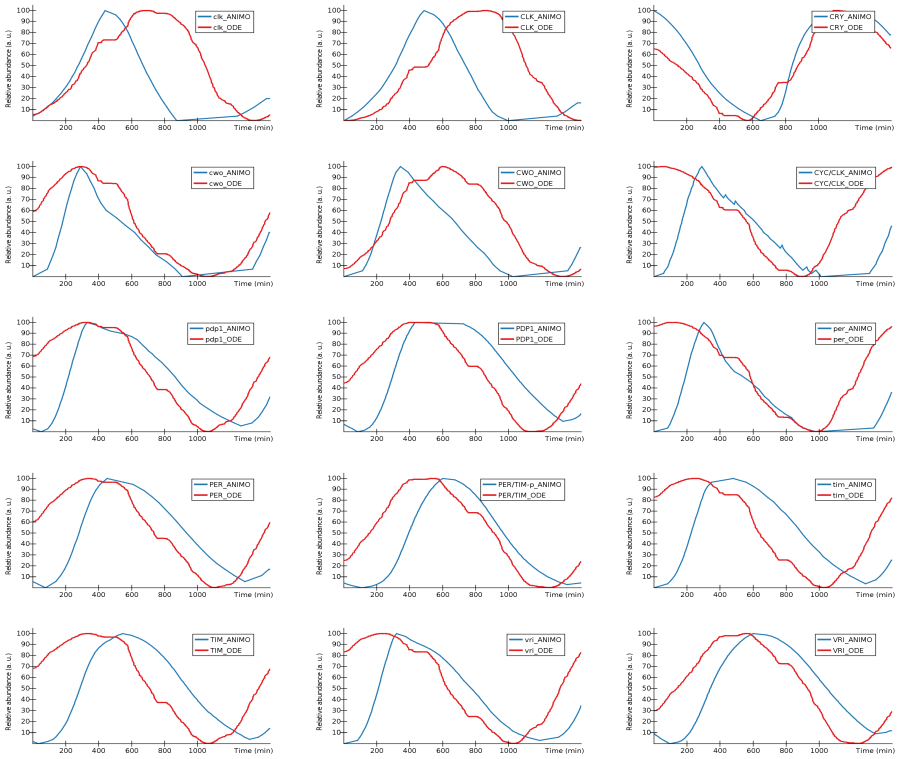

Figure S6: Comparing the simulation results from the ANIMO and ODE models of *Drosophila Melanogaster* circadian clock. Blue lines (_ANIMO series) represent data from the ANIMO model in Figure S4, while red lines (_ODE series) represent data from the ODE model in [8].

Appendix E: Supplementary figures

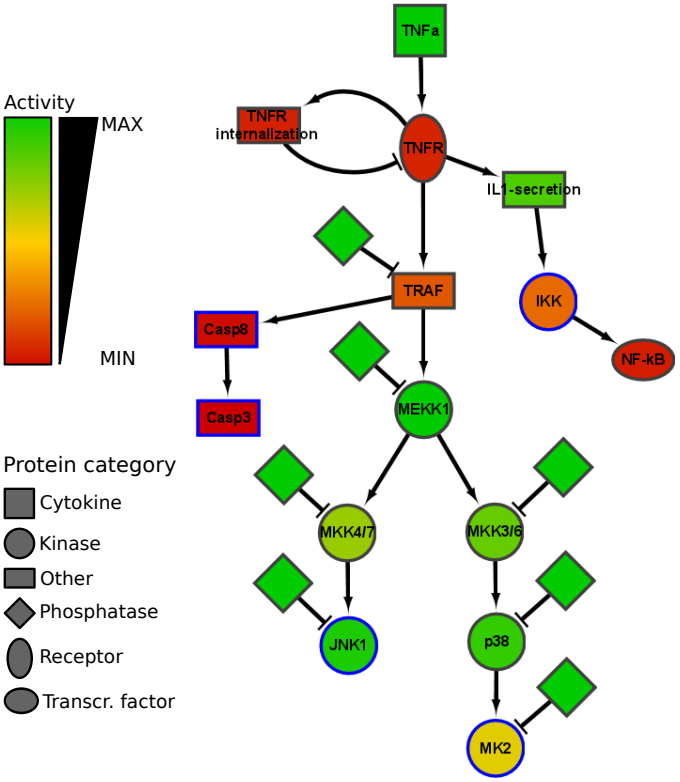

Figure S7: The model for the TNFα pathway in isolation. Node colours represent the activity level of the corresponding modelled reactants at time $t = 10$ minutes after a stimulation of 100 ng/ml TNFα.

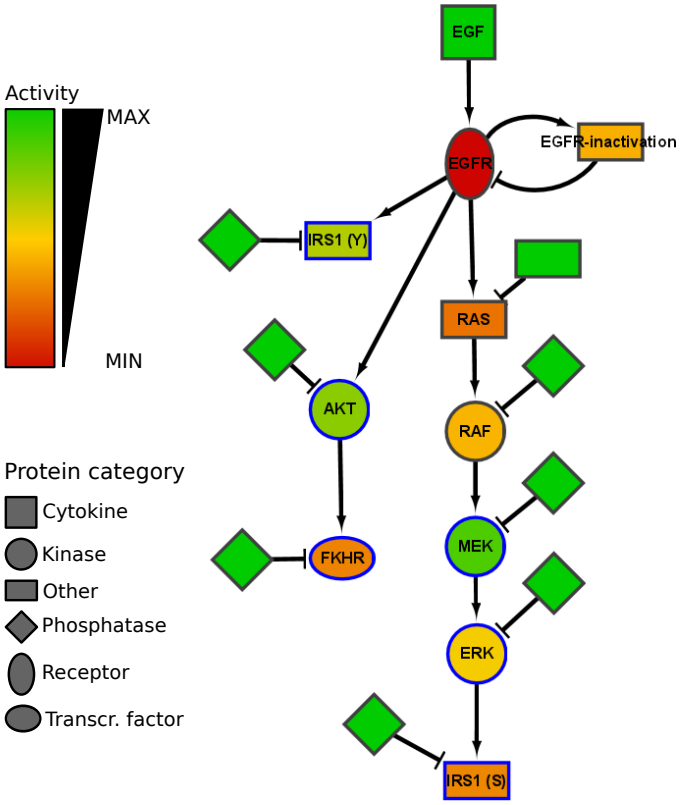

Figure S8: The model for the EGF pathway in isolation. Node colours represent the activity level of the corresponding modelled reactants at time $t = 5$ minutes after a stimulation of 100 ng/ml EGF.

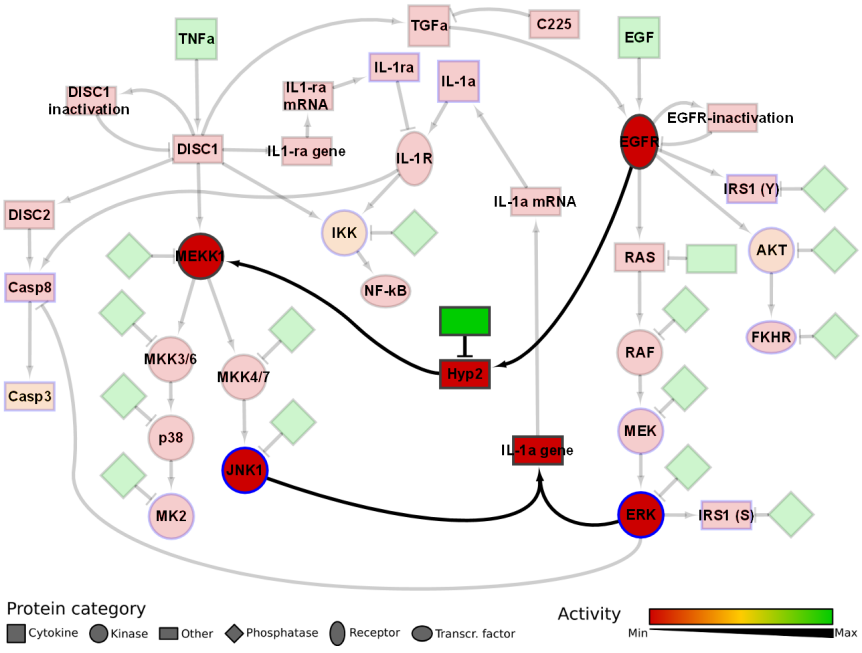

Figure S9: The merged model for the TNFα-EGF pathway with the two hypotheses are highlighted. The first hypothesis is the dependence IL-1α expression on the combined activity of ERK and JNK1. The second hypothesis assumes an as yet unidentified protein (Hyp2) to link EGFR to MEKK1. Node colours represent initial activity levels.

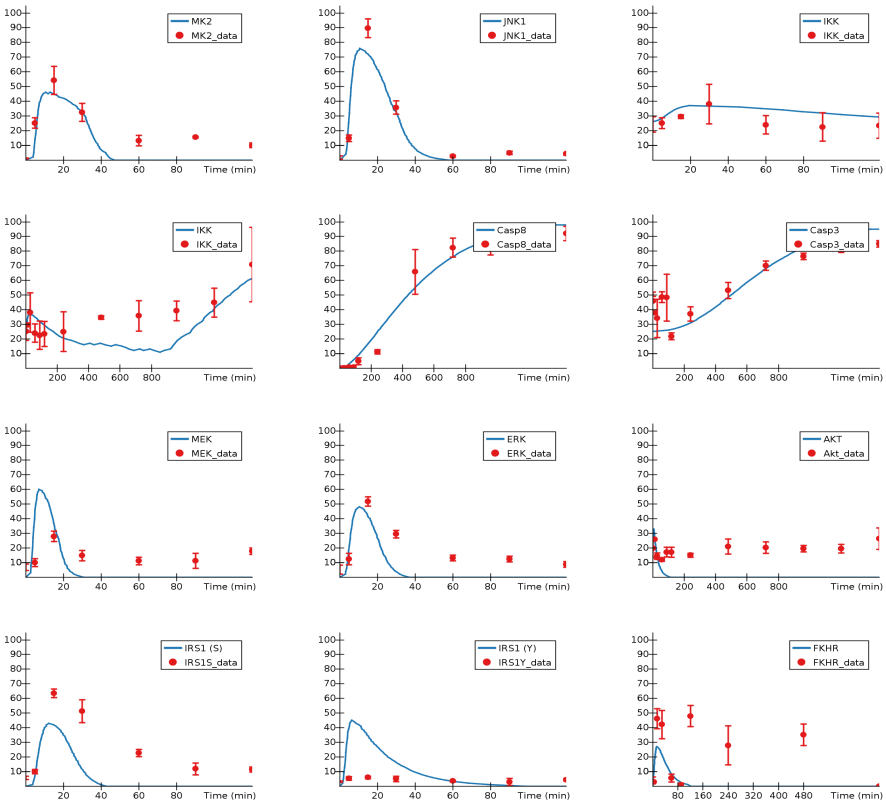

Figure S10: Comparison between the ANIMO model in Figure 5 and experimental data. Treatment condition: 100 ng/ml TNFα. In order to ease the comparison for earlier responses, the time span for those cases is less than 24 hours.

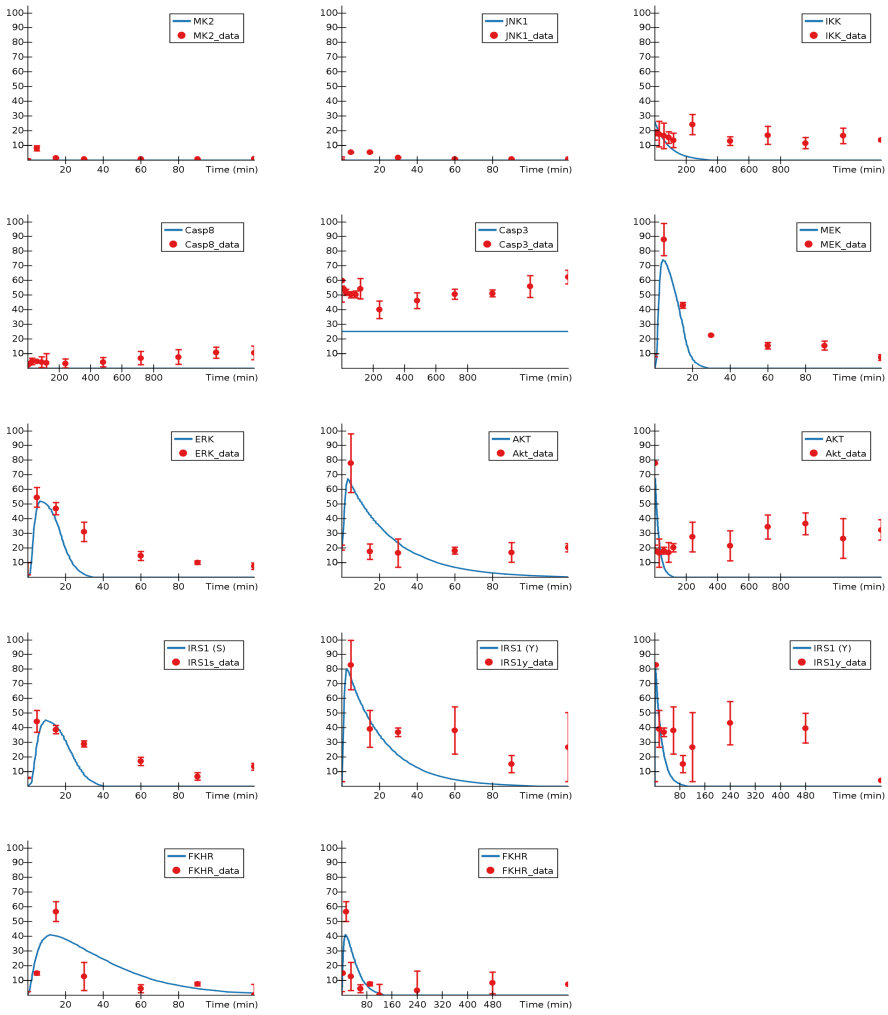

Figure S11: Comparison between the ANIMO model in Figure 5 and experimental data. Treatment condition: 100 ng/ml EGF. In order to ease the comparison for earlier responses, the time span for those cases is less than 24 hours.

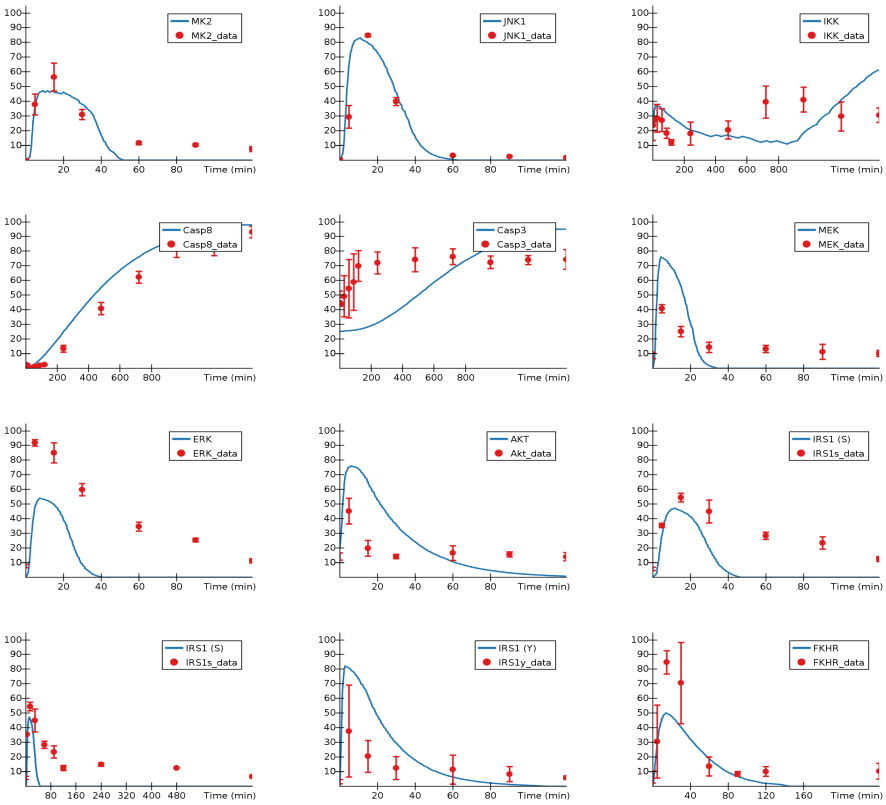

Figure S12: Comparison between the ANIMO model in Figure 5 and experimental data. Treatment condition: 100 ng/ml TNFα + 100 ng/ml EGF. In order to ease the comparison for earlier responses, the time span for those cases is less than 24 hours.