A Model and Analysis of the AKAP Scaffold

Oana Andrei[1] Muffy Calder[2]

Department of Computing Science
University of Glasgow
Glasgow, United Kingdom

Abstract

We study the biochemical processes involved in scaffold-mediated crosstalk between the cAMP and the Raf-1/MEK/ERK pathways. We model the system by a continuous time Markov chain with levels and analyse properties using Continuous Stochastic Logic and the symbolic probabilistic model checker PRISM. We consider two kinds of properties of the model, causal events and pulsating behaviour, and, in order to formulate these properties, we enrich the model with trend formulas. The system is currently under wet-lab investigation and our approach was developed in collaboration with the experimentalists.

Keywords: scaffold proteins, CTMC with levels, stochastic model checking, trend formulas

1 Introduction

Intracellular signal transduction pathways are the mechanism whereby signals are transmitted from a receptor to the nucleus, where they determine a response such as cell growth or apoptosis. Scaffold proteins can play a major role in these pathways as they anchor particular proteins into specific locations for receiving signals or transmitting them. Under certain circumstances, a scaffold can increase the output of a signalling cascade or decrease the response time for a faster output. While individual pathways have a specific signalling role, they can also interact with each other, called cross-talk. In this paper we consider how to model and reason about the scaffold-mediated crosstalk between the cyclic adenosine monophosphate (cAMP) and the Raf-1/MEK/ERK pathways, an interaction that has an important role in the regulation of cell

[1] Email: oandrei@dcs.gla.ac.uk
[2] Email: muffy@dcs.gla.ac.uk

This paper is electronically published in
Electronic Notes in Theoretical Computer Science
URL: www.elsevier.nl/locate/entcs

proliferation, transformation and survival. The scaffold protein is *A-kinase anchoring protein*, usually abbreviated to AKAP.

The behaviour of AKAP is complex and is the topic of current wet-lab investigation; we have worked with life scientists at the University of Glasgow in the development of this model, which we believe to be the first formal model. As experiments are ongoing, some aspects of the behaviour are still unknown and the subject of conjecture.

Since some information is incomplete, our approach to modelling AKAP is based on a stochastic, computational and concurrent view, using continuous-time Markov chains (CTMCs). In particular, we use an abstraction of CTMCs based on discrete levels of concentrations, namely CTMC with levels [6]. The formal language used for modelling is the state-based language underlying the probabilistic symbolic model checker PRISM [9]. We express temporal properties of the model in Continuous Stochastic Logic (CSL) [1], focusing on causality properties and pulsations. In order to express these properties, we define *trend* formulas for variables in the PRISM model which compute for every state the probability that the value of a specific variable increases or decreases during a transition.

The main contributions of this paper are the following:

- a novel, formal model of AKAP scaffold-mediated crosstalk between the cAMP and the Raf-1/MEK/ERK pathways, developed in collaboration with wet-lab experimentalists,
- validation of pulsation behaviour through model checking with rewards,
- introduction of trend formulas and the use of transient temporal properties in CSL to prove formally causality and pulsation properties.

The paper is organised as follows. In the next section we give an overview of AKAP scaffold behaviour. In Sect. 3 we give an overview of the PRISM model. Section 4 contains a description of some properties of interest and the results of analysis. Conclusions and directions for future work follow.

2 The AKAP Scaffold

In intracellular signal transduction pathways, scaffolds are proteins that play mainly an organisational role rather than a signalling role [8]. Scaffolds have two functions:

- *anchoring function* by placing particular proteins in specific intracellular locations for receiving signals or transmitting them;
- *catalytic function* by increasing the output of a signalling cascade or decreasing the response time for a faster output under certain circumstances.

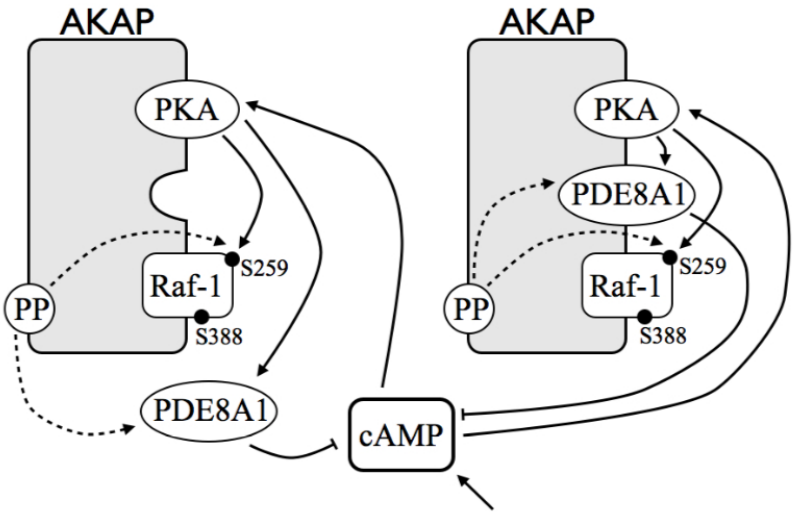

Fig. 1. Interactions between cAMP, unfilled AKAP scaffold, free PDE8A1 and filled scaffold

We are interested in the AKAP mediated crosstalk between cyclic AMP and the Raf-1/MEK/ERK pathways. Figure 1 illustrates AKAP and its anchoring role as positions are filled and unfilled. We focus here on the following species:

- cyclic adenosine monophosphate (cAMP);
- protein kinase A (PKA) which is the main cAMP effector;
- Raf-1 with two phosphorylation sites of interest, Serine 338 (S338) and Serine 259 (S259);
- phosphodiesterase 8 (PDE8A1);
- phosphatase PP.

An overview of AKAP's behaviour is as follows:

Activation and inhibition of Raf-1. If the concentration level of cAMP rises above the basal one, cAMP activates PKA by binding to its regulatory subunits. When PKA becomes active, its catalytic subunits catalyse the transfer of ATP terminal phosphates to the phosphorylation site S259 of Raf-1. The site S338 of Raf-1 is inhibited when S259 is phosphorylated. Only when S338 gets phosphorylated, the pathway Raf-1/MEK/ERK is activated and the signalling cascade begins.

Downregulation by PDE8A1. The catalytic function of PKA sometimes couples with the AKAP, by binding PKA together with phosphodiesterase PDE8A1 on the scaffold to form a complex that functions as a signal module. Under these conditions, as the cell is stimulated, cAMP activates PKA, and then PKA is responsible for the activation of PDE8A1 (by phosphorylation). PDE8A1 converts cAMP to AMP by hydrolysis. If phosphorylated, PDE8A1 degrades more cAMP, hence rapidly reducing the amount of cAMP that can

activate PKA; this leads to a feedback mechanism for downregulating PKA.

The inhibition of Raf-1 at S338 is correlated with a high activity of PKA. At the beginning, cAMP synthesis is induced, causing a rise of PKA's activity, which then causes the inhibition of Raf-1.

In Fig. 1 we use three types of arrow to distinguish between different types of interactions:

- A activates or phosphorylates B : $A \longrightarrow B$
- A dephosphorylates B : $A \dashrightarrow B$
- A degrades B : $A \longrightarrow\!\!\!\!\!\!\mid\ \ B$

The arrow with no source and with target cAMP represents a diffusion of cAMP from the environment.

Scaffold notation. The AKAP scaffold has three positions to be filled by PKA, Raf-1 and PDE8A1 respectively. Hereafter we use a binary representation of the states of each position: 1 for activation or phosphorylation and 0 otherwise. The second position concerns the state of the site S259 of Raf-1. If the scaffold position for PDE8A1 is not filled, we only represent the first two positions of the scaffold. For instance $S100$ stands for a filled scaffold with active PKA and unphosphorylated site S259 and PDE8A1, whereas $S01$ for an unfilled scaffold with inactive PKA and phosphorylated S259.

2.1 Biochemical Reactions

In Fig. 2 we describe the biochemical reactions of the model. Each reaction is given in pseudo-chemical notation, with explicit reference to the scaffold positions (the underlying reactions have mass action kinetics). We associate reaction rate constants (from r_1 to r_{26}) with each biochemical reaction.

Currently, we do not have good experimental data concerning rates for the reactions. However, we have some information on the ratio between the rate of PKA phosphorylating Raf-1 at site S259 and PDE8A1 (either on the scaffold or not). On unfilled scaffolds, PKA phosphorylates two or three times less unscaffolded PDE8A1 than Raf-1 at site S259 from the same scaffold. On filled scaffolds, PKA phosphorylates Raf-1 at S259 and PDE8A1 at the same rate. Consequently the relation between constant rates of the reactions involving PKA phosphorylating either PDE8A1 or Raf-1 is: $r_4 = r_5 = r_6 = r_{10} = r_{11} = 3 * r_{12} = 3 * r_{13}$. In addition, phosphorylated PDE8A1 degrades about three times more cAMP than PDE8A1 does, hence we deduce the following ratios between the constants rates of the reactions where PDE8A1 degrades cAMP : $r_{19} = r_{20} = r_{21} = r_{22} = r_{23} = r_{24} = 3 * r_{25} = 9 * r_{26}$.

Finally, when PKA and PDE8A1 form a complex on the scaffold, PKA's activity becomes more efficient.

(cAMP diffusion)
$\_ \rightarrow_{r_1} cAMP$

(PKA activation)
$S000 + cAMP \rightarrow_{r_2} S100$
$S00 + cAMP \rightarrow_{r_3} S10$

(S259 phosphorylation)
$S100 \rightarrow_{r_4} S110$
$S101 \rightarrow_{r_5} S111$
$S10 \rightarrow_{r_6} S11$

(S259 dephosphorylation)
$PP + S010 \rightarrow_{r_7} PP + S000$
$PP + S011 \rightarrow_{r_8} PP + S001$
$PP + S01 \rightarrow_{r_9} PP + S00$

(cAMP release)
$S111 \rightarrow_{r_{17}} S011 + cAMP$
$S11 \rightarrow_{r_{18}} S01 + cAMP$

(PDE8A1 phosphorylation)
$S100 \rightarrow_{r_{10}} S101$
$S110 \rightarrow_{r_{11}} S111$
$S10 + PDE8A1 \rightarrow_{r_{12}} S10 + pPDE8A1$
$S11 + PDE8A1 \rightarrow_{r_{13}} S11 + pPDE8A1$

(PDE8A1 dephosphorylation)
$PP + S001 \rightarrow_{r_{14}} PP + S000$
$PP + S011 \rightarrow_{r_{15}} PP + S010$
$PP + pPDE8A1 \rightarrow_{r_{16}} PP + PDE8A1$

(cAMP degradation)
$S011 + cAMP \rightarrow_{r_{19}} S011$
$S001 + cAMP \rightarrow_{r_{20}} S001$
$S100 + cAMP \rightarrow_{r_{21}} S100$
$S110 + cAMP \rightarrow_{r_{22}} S110$
$S010 + cAMP \rightarrow_{r_{23}} S010$
$S000 + cAMP \rightarrow_{r_{24}} S000$
$pPDE8A1 + cAMP \rightarrow_{r_{25}} pPDE8A1$
$PDE8A1 + cAMP \rightarrow_{r_{26}} PDE8A1$

Fig. 2. Biochemical reactions occurring during scaffold-mediated crosstalk between the cAMP and the Raf-1/MEK/ERK pathway. The notation $Sv_1v_2v_3$ represents a filled scaffold with v_1, v_2, v_3 denoting the activation state of the bound PKA, PDE8A1 and S259 respectively, i.e., 0 for inactive and 1 for active or phosphorylated. Similarly, Su_1u_2 represent a filled scaffold with u_1 and u_2 denoting the activation state of the bound PKA and S259 respectively.

2.2 Properties

Our collaboration with life scientists has revealed the following expectations, or conjectures, about AKAP's behaviour.

Causal relation between concentration fluctuations. From discussions with life scientists, informally we define *causality* to mean: assuming $\uparrow x$ ($\downarrow x$) denotes increasing (resp. decreasing) concentration levels for the species x, if we have $\uparrow x \Rightarrow \downarrow y$ then a decrease in y's level is necessarily preceded by an increase in x's level. It is expected that increasing amounts of phosphorylated PDE8A1 leads to a cascade of changes in the concentration levels of the other reactants: decreasing amounts of cAMP and active PKA, and an increase in the activity of Raf-1 – due to lower levels of phosphorylated Raf-1 at site S259. Informally, we express this causality relation by the following:

$$\uparrow \text{pPDE8A1} \Rightarrow \downarrow \text{cAMP} \Rightarrow \downarrow \text{active PKA} \Rightarrow \uparrow \text{active Raf-1}$$

Pulsating behaviour. Time courses from laboratory experiments suggest the presence of a pulsating behaviour in the system. The pulsations ensure that the state of the Raf-1 pathways alternates between active and inactive

(note: very long periods of activity or inactivity may increase the risk of disease). In the current model we do not consider explicitly interactions between cAMP and Raf-1. However, the system is not closed since we include a kind of exogenous interaction concerning the diffusion of cAMP. We conjecture this makes the system exhibit a pulsating behaviour corresponding to the feedback mechanism for the downregulation of PKA, coupled with the diffusion of cAMP. Note that we call such a behaviour pulsating, not oscillating. This is because oscillation assumes fluctuation around a given value; current data does not provide us with such a value, hence our choice of pulsating rather than oscillating behaviour.

We note that these properties require a semi-quantitative analysis since they involve relative rather than exact values and since we have approximate data on the reaction rates and the concentration amounts of each species.

3 The PRISM Model for the AKAP

Continuous-time Markov chain (CTMC) with levels models were introduced in [3,6] as stochastic, population based models that are more abstract than molecular CTMCs – the models underlying stochastic simulation in which states are characterised by counts of molecular species. In a CTMC with levels, states are characterised by concentration *ranges*, discretised into a number of levels, for each species. One advantage of this approach is that it is *semi-quantitative*, allowing us to deal with incomplete or only relative information about molecular concentrations, often the case in experimental settings. Furthermore, in comparison to the CTMC underlying a stochastic simulation, CTMCs with levels have a reduced state space, leading to models that may be amenable to stochastic model checking.

Informally, in a CTMC with levels each species is characterised by a number of levels, equidistant from each other, with step size h. We assume that all the species have the same step size. We assign to each species different concentration levels, from 0 (corresponding to null concentration) to a maximum number N. Here, we assume all reactions have mass action kinetics and stoichiometry equal to one.

Definition 3.1 [CTMC with levels] Given a finite set of atomic propositions AP, a *continuous-time Markov chain* (CTMC) is a triple $\mathcal{C} = (S, R, L)$ where S is a finite set of states, $R : S \times S \to \mathbb{R}_{\geq 0}$ a rate matrix, and $L : S \to 2^{AP}$ a labelling of states. For a given state s, there is a race between outgoing transitions from s if there are more than one states s' such that $R(s, s') > 0$. The probability that a transition from s to s' completes within t time units when $R(s, s') > 0$ is determined according to the memoryless negative distribution and is equal to $1 - e^{-R(s,s')\cdot t}$.

A *CTMC with levels* for a biochemical system is a CTMC where the states represent levels of concentrations of the species and the reactions between species define transitions and transition rates between states defined as follows. For n different species $(A_i)_{i=1..n}$, a state is a tuple $s = (l_1, l_2, \ldots, l_n)$ with l_i the discrete concentration level for the species A_i, for all $i = 1..n$. Let us consider a reaction in general form $A_{i_1} + \ldots + A_{i_k} \xrightarrow{r} A_{j_1} + \ldots + A_{j_l}$ with $i_1, \ldots, i_k$ and $j_1, \ldots, j_l$ ranging from 1 to n pairwise distinct, and r the constant reaction rate. We associate to this reaction the characteristic species vectors *pre* and *post* of size n for the set of reactants and the set of products respectively. Then the reaction can be fired from a state s if $s - pre \geq \mathbf{0}$ and $s - pre + post \leq (N, \ldots, N)$. If a transition from s is taken according to this reaction, then we move to the state $s' = s - pre + post$. The rate of the transition $R(s, s')$ is equal to the product of the concentration levels of the reacting species $\frac{r}{h} \cdot l_{i_1} \cdot \ldots \cdot l_{i_k}$.

Following the style adopted in [4], we define a PRISM model[3] of AKAP using CTMCs with levels as follows. A molecular species is represented as a PRISM process (module) and its behaviour by labelled transitions. The reactions are represented by multi-way synchronisations between transitions based on common labels. There is one module for each species, i.e. for cAMP, scaffold, unscaffolded PDE8A1 and PP, each with corresponding variables representing levels of concentrations. In particular, the module for the scaffold has a variable for each possible combination of scaffold positions (S000, S100, S101, S110, S011, S010, S001, S111, S00, S10, S01, S11). Commands in the modules correspond to reactions, which are synchronised on each participating module (i.e. consumers and producers in the chemical reaction) based on common labels. Additionally, we define diffusion of cAMP from time to time.

As an example, consider the reaction r_2 from Fig. 2 indicating that cAMP activates PKA when the level of cAMP is above the basal level. Then in the module describing cAMP we add the command:

```
[activate_PKA] (cAMP > basal_camp) -> (cAMP) : (cAMP' = cAMP-1);
```

in the module describing the scaffold we have the coupling command:

```
[activate_PKA] (S000 > 0) & (S100 < scaffold_max) ->
               (S000) : (S100' = S100+1) & (S000' = S000-1);
```

and in the module with all constant rates we add the command with the same label as for the other two commands:

```
[activate_PKA] true -> (r2/h) : true;
```

where h is the step size obtained by dividing the maximal molar concentration of the species by the number of levels chosen for the discretisation. Through

[3] The full PRISM model of AKAP is available at http://www.dcs.gla.ac.uk/~muffy/akap.

synchronisation on the common label, the transition rate will then be the product of $r2/h$ and the concentration levels of cAMP and PKA.

Unless otherwise stated, we assume the number of levels $N = 3$. We consider maximum N levels of filled/unfilled scaffolds, $2 * N$ levels of phosphatase PP, and around $N/2$ levels of unscaffolded PDE8A1. Whereas for cAMP, since it is diffused in the system, we allow a greater concentration of cAMP, maximum $10 * N$. The state space of the model with $N = 3$ has the size of order $1.6 * 10^6$ with $1.4 * 10^7$ transitions. The time for model construction was 2.355 seconds on a 2 GHz Intel Core 2 Duo processor with 4GB memory.

4 Model-Checking the AKAP Model

In this section we formalise the properties described in Sect. 2.2 and we use PRISM [9] to verify their satisfaction (or not). First we use rewards to compute the expected level of concentration at a particular time. However this analysis is not sufficient to confirm the causality relation between events and the pulsating behaviour. So second, we consider transient properties. We extend the set of state formulas of the CTMC model to include trend formulas for some variables of interest. The aim is to reason in terms of ascending/descending (positive/negative) trends for particular concentration levels of populations. Prior to this, we review very briefly the syntax and semantics of CSL and the PRISM model checker.

Continuous Stochastic Logic (CSL) [1] is a stochastic extension of the Computational Tree Logic (CTL) allowing one to express a probability measure of the satisfaction of a temporal property in either transient or in steady-state behaviours. The formulas of CSL are state formulas and their syntax is the following:

$$\begin{array}{lrcl} \textit{State formula} & \Phi & ::= & \textit{true} \mid a \mid \neg\Phi \mid \Phi \wedge \Phi \mid \mathcal{P}_{\bowtie p}[\phi] \mid \mathcal{S}_{\bowtie p}[\phi] \\ \textit{Path formula} & \phi & ::= & \mathbf{X}\,\Phi \mid \Phi\,\mathbf{U}^{\mathrm{I}}\,\Phi \end{array}$$

where a ranges over a set of atomic propositions, $\bowtie \in \{\leq, >, \geq, >\}$, $p \in [0, 1]$, and I is an interval of $\mathbb{R}_{\geq 0}$. There are two types of CSL properties: transient (of the form $\mathcal{P}_{\bowtie p}[\phi]$) and steady-state (of the form $\mathcal{S}_{\bowtie p}[\phi]$). For this current work we are only interested in transient or time-dependent properties. A formula $\mathcal{P}_{\bowtie p}[\phi]$ is true in state s if the probability that ϕ is satisfied by the paths starting from state s meets the bound $\bowtie p$. The path formulas are constructed using the $\mathbf{X}$ (next) operator and the $\mathbf{U}^I$ (time-bounded until) operator. Intuitively the path formula $\mathbf{X}\,\Phi$ is true if Φ is satisfied in the next state, whereas $\Phi_1\,\mathbf{U}^{\mathrm{I}}\,\Phi_2$ is true if Φ_2 is holds at some time instant in the interval I and at all preceding time instants Φ_1 holds. This is a minimal set of operators for CSL. The operators *false*, disjunction and implication

can be derived using basic logical equivalences. Two more path operators are available as syntactic sugar:

- the *eventually* operator $\mathbf{F}$ (future) where $\mathbf{F}^{\mathrm{I}}\,\Phi \equiv \mathit{true}\,\mathbf{U}^{\mathrm{I}}\,\Phi$, and
- the *always* operator $\mathbf{G}$ (globally) where $\mathbf{G}^{\mathrm{I}}\,\Phi \equiv \neg(\mathbf{F}^{\mathrm{I}}\,\neg\Phi)$.

The PRISM language supports *rewards* (or cost) structures for extending the set of state formulas in CSL [9]. In this paper we are only interested in instantaneous rewards which have the form $\mathcal{R}_{\bowtie r}[I^{=t}]$ with $\bowtie \in \{\leq, <, >, \geq\}$, $r, t \in \mathbb{R}_{\geq 0}$. Such a formula, from a state s, is true if the expected state reward at time instant t meets the bound $\bowtie r$.

The PRISM probabilistic model checker has a property specification language based on the temporal logics PCTL, CSL, LTL and PCTL*, including extensions for quantitative specifications and rewards. PRISM allows one to express a probability measure that a temporal formula is satisfied. The bound $\bowtie p$ may not be specified, in which case a probability is calculated in PRISM. We can check the satisfaction of properties of a signalling pathway like stability of a protein by steady state analysis (the concentration of a protein becomes stable at a particular level and stays there for some reaction rates), or transient behaviour of proteins (concentration peak within a given time interval or conditioned by a specific state of another protein, monotonic increase of protein concentration), and many more. A model can be extended with information about rewards or costs such that one can analyse expected values of the rewards cumulated up to a specific state or a time instant, at a specific time instant, or in a steady-state.

4.1 Reward-based Analysis of the AKAP *Model*

For each species of interest: phosphorylated PDE8A1, free cAMP (not bound to some PKA), active PKA and phosphorylated Raf-1 at site S259, we associate a reward structure which evaluates to the expected level of concentration at a particular time. For example, the following instantaneous reward associated with pPDE8A1, named `phosphopde8`, computes the sum of the expected levels of unscaffolded and phosphorylated pPDE8A1 and the expected levels of scaffolded and phosphorylated PDE8A1, at each time instant:

```
rewards "phosphopde8"
    true : pPDE8A1 + S101 + S111 + S001 + S011;
endrewards
```

We use the temporal query $\mathcal{R}_{\text{"}phosphopde8\text{"}=?}[I^{=T}]$ for the reward `phosphopde8` with T an integer variable ranging from 0 to 30 time-units. In Fig. 3 we plot the expected levels of concentration for each species. We observe delays in peak successions for all variable values as expected. However, this does not prove that the properties expressed in Sect. 2 are satisfied, it does not prove

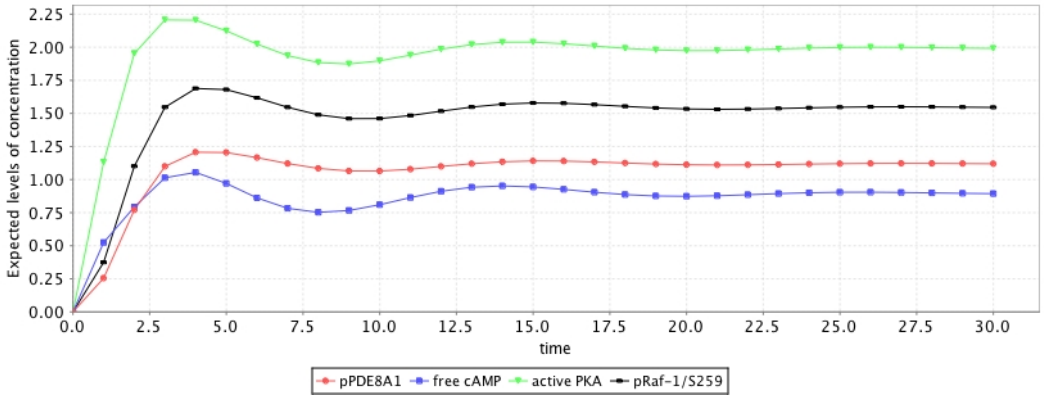

Fig. 3. Expected levels of concentrations of phosphorylated PDE8A1, free cAMP, active PKA, and phosphorylated Raf-1 at site S259 after 30 time-units

there is a causality relation between all pulsations. In order to prove causality we consider transient properties based on trend formulas.

4.2 Trend-based Analysis of the AKAP *Model*

We compute the *positive* or *ascending* (*negative* or *descending*) *trend* of a variable in a CTMC model based on the probability of any state where the value of the variable increases (decreases) being the next state to which a transition is made from the current state. For R a set of biochemical reactions and A a species involved in some reactions in R, the probabilities for positive and negative trends in a state s are computed by the following formulas:

$$P_{\uparrow A}(s) = \begin{cases} \sum\{\rho(r,s) \mid r \in R, pre_A(r) < post_A(r)\}/E(s), & \text{if } E(s) \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

$$P_{\downarrow A}(s) = \begin{cases} \sum\{\rho(r,s) \mid r \in R, pre_A(r) > post_A(r)\}/E(s), & \text{if } E(s) \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

where $\rho(r,s)$ returns the transition rate of the reaction r if possible in state s, or 0 otherwise, and $E(s)$ is the exit rate of state s, i.e., $E(s) = \sum_r \rho(r,s)$. We choose a confidence threshold of ξ for the probabilities to indicate a positive or negative trend which we choose to take the value 0.6 for our model. Then the formula $\uparrow A$ is a state formula for the CMTC with levels model and it says that the species A follows an ascending trend in state s if $P_{\uparrow A}(s) \geq \xi$. Similarly, the formula $\downarrow A$ is true if $P_{\downarrow A}(s) \leq \xi$. The trend formula for a variable in the CTMC model is the stochastic counterpart of the sign of the first-order derivative for the same variable in the associated ODE model.

We add trend formulas to the PRISM model for cAMP, active PKA (PKA*) and scaffolded phosphorylated PDE8A1 (pPDE8A1) using PRISM formulas and labels. We illustrate in the following the procedure for computing the ascend-

ing trend formula for active PKA. The concentration level of active PKA is increased only by reactions r_2 and r_3. The rates of the transitions corresponding to these reactions are computed by formulas `rate_r2` and `rate_r3` respectively with respect to the current state. In case a reaction is not possible in the current state, the transition rate takes the value 0. The rates for all transitions induced by the reactions depicted in Fig. 2 are computed in a similar manner and the formula `exit_rate` represents their summation. The probability $P_{\uparrow \mathsf{PKA\_A}}(s)$ in the current state s is given by formula `PKA_A_up`, while the state label `trend_PKA_A_up` will be used in the CSL formula for reasoning over the ascending trend of active PKA, denoted by $\uparrow \mathsf{PKA}^*$.

```
formula cond_r2 = (cAMP>basal_camp)&(S000>0)&(S100<scfld_max);
formula rate_r2 = (cond_r2 ? (r2/H)*(cAMP*H)*(S000*H) : 0);
formula cond_r3 = (cAMP>basal_camp)&(S00>0)&(S10<u_scfld_max);
formula rate_r3 = (cond_r3 ? (r3/H)*(cAMP*H)*(S00*H) : 0);
formula rate_PKA_A_up = rate_r2 + rate_r3;

formula PKA_A_up = (exit_rate=0 ? 0 : rate_PKA_A_up/exit_rate);
label "trend_PKA_A_up" = (PKA_A_up>=threshold);
```

In the following we formalise the temporal properties of Sect. 2.2.

Necessarily Preceded.

We express the causality property stated in Section 2.2 as a temporal query using the *necessarily preceded* or *requirement* pattern [10]. This pattern represents an ordering relation between two events, the occurrence of the later being conditioned by the occurrence of the former: *a state ϕ is reachable and is necessary preceded all the time by a state ψ*. This pattern is expressed as the CTL formula $\mathbf{EF}\phi \wedge (\mathbf{AG}((\neg\psi) \Rightarrow \mathbf{AG}(\neg\phi)))$.

Assume the following two state formulas $\phi_1 = (\downarrow \mathsf{cAMP} \wedge \downarrow \mathsf{PKA}^*)$ and $\psi_1 = \uparrow \mathsf{pPDE8A1}$ with ϕ corresponding to a state where the levels of cAMP and active PKA are decreasing, and ψ to a state where the level of phosphorylated PDE8A1 is increasing. Employing basic propositions equivalences, we translate the requirement pattern into CSL to obtain the following formula which was checked as **true** for our PRISM model:

$$\mathcal{P}_{>0}[\mathbf{F}\phi_1] \wedge \mathcal{P}_{\geq 1}[\mathbf{G}((\neg\psi_1) \Rightarrow \mathcal{P}_{\leq 0}[\mathbf{G}\phi_1])] \longrightarrow \mathbf{true}$$

We express a tighter causality relation between increasing concentration levels of pPDE8A1 ($\psi_2 = \uparrow \mathsf{pPDE8A1}$) and decreasing levels of cAMP ($\phi_2 = \downarrow \mathsf{cAMP}$) using the following formula checked as **true** for our PRISM model:

$$\mathcal{P}_{\geq 1}[\mathbf{F}((\neg\psi_2 \wedge \neg\phi_2)\mathbf{U}(\mathcal{P}_{\geq 1}[(\psi_2 \wedge \neg\phi_2)\mathbf{U}\phi_2]))] \longrightarrow \mathbf{true}$$

This formula stands for $\uparrow$ pPDE8A1 $\Rightarrow \downarrow$ cAMP in the notation from Sect. 2.2 and it states that there is a continuous time period where the trend of pPDE8A1 is not descending and the trend of cAMP is not descending until the concentration level of pPDE8A1 starts increasing and soon after the level of cAMP starts decreasing. This CSL pattern can also be employed in order to show that $\downarrow$pPDE8A1 $\Rightarrow \uparrow$cAMP and $\downarrow$ cAMP $\Rightarrow \downarrow$ PKA*.

Pulsating behaviour.

An oscillating behaviour concerns fluctuation around a given value k. Oscillation and its expression as temporal formulas in CTL and PCTL has been studied in [2] and informally described as *always in the future, the variable x departs from and reaches the values k infinitely often.* The corresponding CTL formula is $\mathbf{AG}(((x = k) \Rightarrow \mathbf{EF}(x \neq k)) \wedge ((x \neq k) \Rightarrow \mathbf{EF}(x = k)))$. In the context of BIOCHAM [5], a weaker form of oscillation properties expressed in CTL is used with the symbolic model checker NuSMV; the oscillating behaviour is approximated by the necessary but not sufficient formula $\mathbf{EG}((\mathbf{EF}\neg\varphi) \wedge (\mathbf{EF}\varphi))$.

We are interested in pulsating behaviour, i.e. no fixed k. We therefore consider oscillations (around 0) of the values of some variables. We refer to this approximate oscillating behaviour as pulsation. Note that we can observe a pattern corresponding to a pulsation in Fig. 3: we repeatedly have the situation where the level of phosphorylated PDE8A1 increased whereas the levels of cAMP and active PKA decreased, and then the level of phosphorylated PDE8A1 decreased whereas the levels of cAMP and active PKA increased. Assume the two state formulas $\phi = (\uparrow$pPDE8A1 $\wedge \downarrow$cAMP$\wedge \downarrow$PKA$^*)$ and $\psi = (\downarrow$ pPDE8A1 $\wedge \uparrow$ cAMP$\wedge \uparrow$ PKA$^*)$. The CSL formula describing a pulsation involving the two state formulas ϕ and ψ is the following and it was checked as **true** for our model using PRISM:

$$\mathcal{P}_{\geq 1}[\mathbf{G}((\phi \Rightarrow \mathcal{P}_{>0}[\mathbf{F}\psi]) \wedge (\psi \Rightarrow \mathcal{P}_{>0}[\mathbf{F}\phi]))] \longrightarrow \mathbf{true}$$

Similarly we can show individual pulsations for pPDE8A1, cAMP and PKA*.

5 Conclusion and Future Work

We have developed a formal model of the behaviour of the AKAP scaffold and scaffold-mediated crosstalk between the cAMP and the Raf-1/MEK/ERK pathways. The model is a CTMC with levels and is implemented in the PRISM language.

The behaviour of this scaffold is complex, with feedback, and the model was developed in collaboration with wetlab experimentalists. We have considered the questions and conjectures concerning system behaviour posed by

experimentalists; these include sequentially dependent events and pulsating behaviour. In the context of imprecise and incomplete data, pulsation seems more appropriate than oscillation. We have used rewards and CSL to express the properties, and checked them with the PRISM model checker. In order to express pulsation, we have defined trends. Discussions with the experimentalists confirm their interest and validation of the model and analysis. A stochastic approach seems particularly suited to this problem, given they do not know absolute reaction rates, but some ratios. In this case, they are interested in causal behaviour and trends, rather than precise quantities.

Future work includes reasoning about the amplitude of the pulsation and we will investigate the relation between trends in a CTMC model and the sign of first-order derivatives in the corresponding ODE model. Derivatives have been considered previously in the context of model checking biochemical systems. For example in BIOCHAM [7,11], oscillatory properties are analysed using queries expressed as formulas in LTL with constraints over real numbers. Such formulas are interpreted over traces of states and a state include not only the concentration value of each molecule but the value of its first order derivative as well.

We will also refine the model with data on the rates, as more data become available, and add more detail about relationships within the model (e.g., four molecules of cAMPs are required to activate one PKA).

We note that we have investigated further hypotheses, such as whether the way PDE8A1 decreases the PKA phosphorylation of Raf-1 at S259 can be counterbalanced by the addition of a PDE8A1 inhibitor such as the a drug Dipyridamole. This is the topic of a further paper.

Acknowledgements

We would like to thank Walter Kolch, George Baillie and Kim Brown from the Faculty of Biomedical & Life Science, University of Glasgow, for discussions, guidance and insight into the AKAP scaffold. We also wish to thank Gethin Norman and the anonymous reviewers of this paper for their helpful comments on the work.

This research is supported by the SIGNAL project, funded by the Engineering and Science Research Council (EPSRC) under grant number EP/E031439/1.

References

[1] Baier, C., B. R. Haverkort, H. Hermanns and J.-P. Katoen, *Model-Checking Algorithms for Continuous-Time Markov Chains*, IEEE Trans. Software Eng. **29** (2003), pp. 524–541.

[2] Ballarini, P., R. Mardare and I. Mura, *Analysing Biochemical Oscillation through Probabilistic Model Checking*, Electr. Notes Theor. Comput. Sci. **229** (2009), pp. 3–19.

[3] Calder, M., S. Gilmore and J. Hillston, *Modelling the Influence of RKIP on the ERK Signalling Pathway Using the Stochastic Process Algebra PEPA*, in: C. Priami, A. Ingólfsdóttir, B. Mishra and H. R. Nielson, editors, *Transactions on Computational Systems Biology VII*, Lecture Notes in Computer Science **4230** (2006), pp. 1–23.

[4] Calder, M., V. Vyshemirsky, D. Gilbert and R. J. Orton, *Analysis of Signalling Pathways Using Continuous Time Markov Chains*, in: C. Priami and G. D. Plotkin, editors, *T. Comp. Sys. Biology*, Lecture Notes in Computer Science **4220** (2006), pp. 44–67.

[5] Chabrier-Rivier, N., M. Chiaverini, V. D. F. Fages and V. Schächter, *Modeling and querying biomolecular interaction networks.*, Theoretical Computer Science **325** (2004), pp. 25–44.

[6] Ciocchetta, F., A. Degasperi, J. Hillston and M. Calder, *Some Investigations Concerning the CTMC and the ODE Model Derived From Bio-PEPA*, Electr. Notes Theor. Comput. Sci. **229** (2009), pp. 145–163.

[7] Fages, F., *Temporal Logic Constraints in the Biochemical Abstract Machine BIOCHAM*, in: P. M. Hill, editor, *LOPSTR*, Lecture Notes in Computer Science **3901** (2005), pp. 1–5.

[8] James E. Ferrell, J., *What Do Scaffold Proteins Really Do?*, Sci. STKE **2000** (2000), pp. 1–3.

[9] Kwiatkowska, M. Z., G. Norman and D. Parker, *Stochastic Model Checking*, in: M. Bernardo and J. Hillston, editors, *SFM*, Lecture Notes in Computer Science **4486** (2007), pp. 220–270.

[10] Monteiro, P. T., D. Ropers, R. Mateescu, A. T. Freitas and H. de Jong, *Temporal logic patterns for querying dynamic models of cellular interaction networks*, Bioinformatics **24** (2008), pp. 227–233.

[11] Rizk, A., G. Batt, F. Fages and S. Soliman, *On a Continuous Degree of Satisfaction of Temporal Logic Formulae with Applications to Systems Biology*, in: M. Heiner and A. M. Uhrmacher, editors, *CMSB*, Lecture Notes in Computer Science **5307** (2008), pp. 251–268.