Replicating the MAP Kinase Cascade in PEPA

Muffy Calder[1] Adam Duguid[2] Stephen Gilmore[2] Jane Hillston[2]

[1]Department of Computer Science, The University of Glasgow

[2]School of Informatics, The University of Edinburgh

30 June 2006

Why?

Q Why do we want to replicate the MAP Kinase cascade?

Why?

Q Why do we want to replicate the MAP Kinase cascade?

A The MAP Kinase cascade model as defined by Schoeberl *et al.* is a large and complex model. By showing that PEPA can cope with models of this size we reinforce the idea that PEPA is suitable for modelling biological systems.

The aim is not to make grand discoveries about this particular signalling pathway - merely to explore the boundaries of modelling biological systems with PEPA.

The PEPA Language

Prefix	$(\alpha, r).P$ - activity α happens at rate r before transitioning to component of type P.
Choice	$P + Q$
Cooperation	$P \underset{L}{\bowtie} Q$
Hide	P/L
Constant	$A \stackrel{\mathit{def}}{=} P$

The PEPA Language

Prefix	$(\alpha, r).P$
Choice	$P + Q$ - allows either P or Q to occur. Which happens is based on race conditions (see Prefix).
Cooperation	$P \underset{L}{\bowtie} Q$
Hide	P/L
Constant	$A \stackrel{def}{=} P$

The PEPA Language

Prefix $(\alpha, r).P$

Choice $P + Q$

Cooperation $P \underset{L}{\bowtie} Q$ - If both P and Q can perform an activity α where $\alpha \in L$ then they must both make the transition at the same time.

Hide P/L

Constant $A \stackrel{def}{=} P$

The PEPA Language

Prefix	$(\alpha, r).P$
Choice	$P + Q$
Cooperation	$P \underset{L}{\bowtie} Q$
Hide	P/L - The activities in set L are 'hidden' denying co-operation with these activities. Used to ensure correctness.
Constant	$A \stackrel{def}{=} P$

The PEPA Language

Prefix $(\alpha, r).P$

Choice $P + Q$

Cooperation $P \underset{L}{\bowtie} Q$

Hide P/L

Constant $A \stackrel{def}{=} P$ - Assigns the label A to component P.

The PEPA Language

(Reagent-Centric Approach)

Prefix $(\alpha, r).P$

Choice $P + Q$

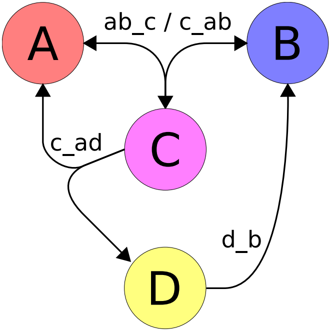
Cooperation $P \underset{L}{\bowtie} Q$

Hide P/L

Constant $A \stackrel{def}{=} P$ - A is now tagged with a H showing a high concentration and L for a low concentration i.e. $\mathrm{A_H}$.

Current implementation has synchronization over common activity names so Cooperation is implicit in naming.
The combinator Hide also has no equivalent behaviour.

Reagent-Centric Approach Example

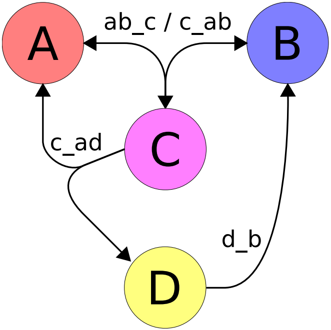


$$
\begin{aligned}
A_H &\stackrel{def}{=} (ab\_c, \alpha).A_L \\
A_L &\stackrel{def}{=} (c\_ab, \beta).A_H + (c\_ad, \gamma).A_H \\
B_H &\stackrel{def}{=} (ab\_c, \alpha).B_L \\
B_L &\stackrel{def}{=} (c\_ab, \beta).B_H + (d\_b, \delta).B_H \\
C_H &\stackrel{def}{=} (c\_ab, \beta).C_L + (c\_ad, \gamma).C_L \\
C_L &\stackrel{def}{=} (ab\_c, \alpha).C_H \\
D_H &\stackrel{def}{=} (d\_b, \delta).D_L \\
D_L &\stackrel{def}{=} (c\_ad, \gamma).D_H
\end{aligned}
$$

$$((A_H \underset{\{ab\_c, c\_ab\}}{\bowtie} B_H) \underset{\{ab\_c, c\_ab, c\_ad\}}{\bowtie} C_L) \underset{\{c\_ad, d\_b\}}{\bowtie} D_L$$

ab_c,	$A + B \to C$	$, \alpha$	c_ab,	$C \to A + B$	$, \beta$
c_ad,	$C \to A + D$	$, \gamma$	d_b,	$D \to B$	$, \delta$

Reagent-Centric Approach Example

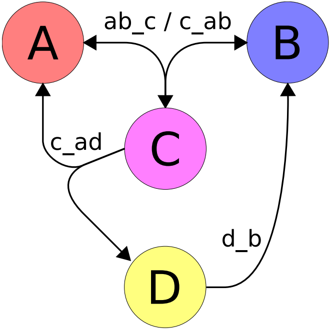


$$
\begin{aligned}
A_H &\stackrel{def}{=} (ab\_c, \alpha).A_L \\
A_L &\stackrel{def}{=} (c\_ab, \beta).A_H + (c\_ad, \gamma).A_H \\
B_H &\stackrel{def}{=} (ab\_c, \alpha).B_L \\
B_L &\stackrel{def}{=} (c\_ab, \beta).B_H + (d\_b, \delta).B_H \\
C_H &\stackrel{def}{=} (c\_ab, \beta).C_L + (c\_ad, \gamma).C_L \\
C_L &\stackrel{def}{=} (ab\_c, \alpha).C_H \\
D_H &\stackrel{def}{=} (d\_b, \delta).D_L \\
D_L &\stackrel{def}{=} (c\_ad, \gamma).D_H
\end{aligned}
$$

$$((A_H \underset{\{ab\_c, c\_ab\}}{\bowtie} B_H) \underset{\{ab\_c, c\_ab, c\_ad\}}{\bowtie} C_L) \underset{\{c\_ad, d\_b\}}{\bowtie} D_L$$

$ab\_c,$	$A + B \rightarrow C$	$, \alpha$	$c\_ab,$	$C \rightarrow A + B$	$, \beta$
$c\_ad,$	$C \rightarrow A + D$	$, \gamma$	$d\_b,$	$D \rightarrow B$	$, \delta$

Reagent-Centric Approach Example

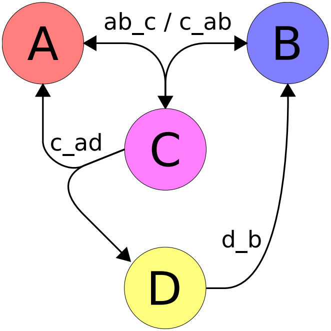


$$
\begin{aligned}
A_H &\stackrel{def}{=} (ab\_c, \alpha).A_L \\
A_L &\stackrel{def}{=} (c\_ab, \beta).A_H + (c\_ad, \gamma).A_H \\
B_H &\stackrel{def}{=} (ab\_c, \alpha).B_L \\
B_L &\stackrel{def}{=} (c\_ab, \beta).B_H + (d\_b, \delta).B_H \\
C_H &\stackrel{def}{=} (c\_ab, \beta).C_L + (c\_ad, \gamma).C_L \\
C_L &\stackrel{def}{=} (ab\_c, \alpha).C_H \\
D_H &\stackrel{def}{=} (d\_b, \delta).D_L \\
D_L &\stackrel{def}{=} (c\_ad, \gamma).D_H
\end{aligned}
$$

$$((A_H \underset{\{ab\_c, c\_ab\}}{\bowtie} B_H) \underset{\{ab\_c, c\_ab, c\_ad\}}{\bowtie} C_L) \underset{\{c\_ad, d\_b\}}{\bowtie} D_L$$

ab_c,	$A + B \rightarrow C$	$, \alpha$	c_ab,	$C \rightarrow A + B$	$, \beta$
c_ad,	$C \rightarrow A + D$	$, \gamma$	d_b,	$D \rightarrow B$	$, \delta$

Reagent-Centric Approach Example

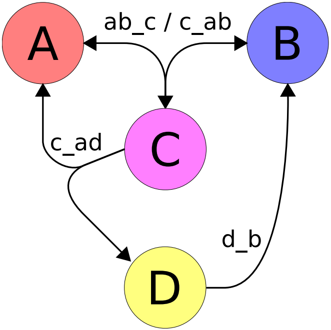


$$\begin{aligned}
A_H &\stackrel{def}{=} (ab_c, \alpha).A_L \\
A_L &\stackrel{def}{=} (c_ab, \beta).A_H + (c_ad, \gamma).A_H \\
B_H &\stackrel{def}{=} (ab_c, \alpha).B_L \\
B_L &\stackrel{def}{=} (c_ab, \beta).B_H + (d_b, \delta).B_H \\
C_H &\stackrel{def}{=} (c_ab, \beta).C_L + (c_ad, \gamma).C_L \\
C_L &\stackrel{def}{=} (ab_c, \alpha).C_H \\
D_H &\stackrel{def}{=} (d_b, \delta).D_L \\
D_L &\stackrel{def}{=} (c_ad, \gamma).D_H
\end{aligned}$$

$$((A_H \underset{\{ab\_c, c\_ab\}}{\bowtie} B_H) \underset{\{ab\_c, c\_ab, c\_ad\}}{\bowtie} C_L) \underset{\{c\_ad, d\_b\}}{\bowtie} D_L$$

$ab\_c,$	$A + B \to C$	$, \alpha$	$c\_ab,$	$C \to A + B$	$, \beta$
$c\_ad,$	$C \to A + D$	$, \gamma$	$d\_b,$	$D \to B$	$, \delta$

Reagent-Centric Approach Example

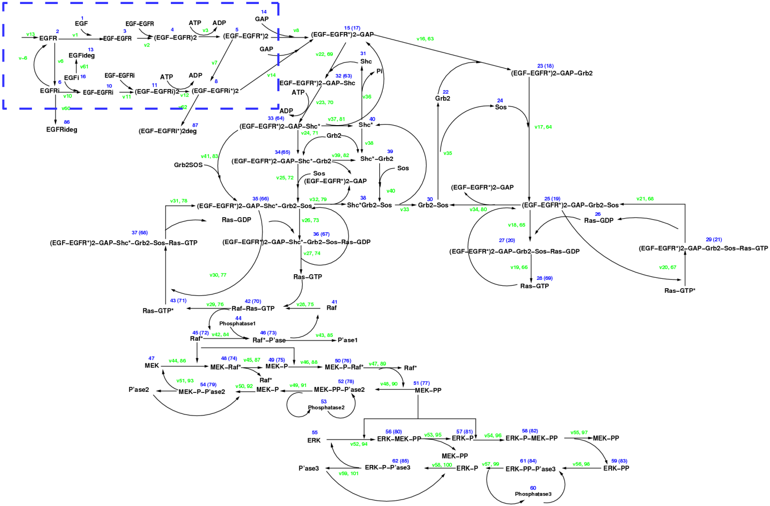


$$\begin{aligned}
A_H &\stackrel{def}{=} (ab_c, \alpha).A_L \\
A_L &\stackrel{def}{=} (c_ab, \beta).A_H + (c_ad, \gamma).A_H \\
B_H &\stackrel{def}{=} (ab_c, \alpha).B_L \\
B_L &\stackrel{def}{=} (c_ab, \beta).B_H + (d_b, \delta).B_H \\
C_H &\stackrel{def}{=} (c_ab, \beta).C_L + (c_ad, \gamma).C_L \\
C_L &\stackrel{def}{=} (ab_c, \alpha).C_H \\
D_H &\stackrel{def}{=} (d_b, \delta).D_L \\
D_L &\stackrel{def}{=} (c_ad, \gamma).D_H
\end{aligned}$$

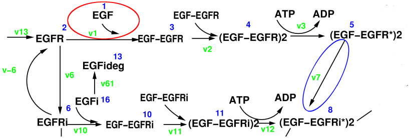
$$((A_H \underset{\{ab\_c, c\_ab\}}{\bowtie} B_H) \underset{\{ab\_c, c\_ab, c\_ad\}}{\bowtie} C_L) \underset{\{c\_ad, d\_b\}}{\bowtie} D_L$$

$ab\_c,$	$A + B \longrightarrow C$	$, \alpha$	$c\_ab,$	$C \longrightarrow A + B$	$, \beta$
$c\_ad,$	$C \longrightarrow A + D$	$, \gamma$	$d\_b,$	$D \longrightarrow B$	$, \delta$

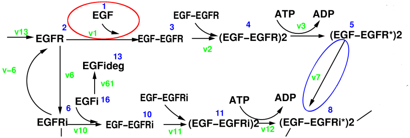
The MAP Kinase Cascade



Complexities with Graphical Representations



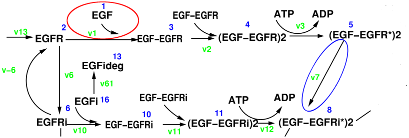
Complexities with Graphical Representations



Reaction v_7 is uni-directional. Almost all other arrows define bi-directional reactions.

Complexities with Graphical Representations

Reaction v_1 looks like every other reaction. In the original model EGF is not 'consumed' in this particular binding.



Reaction v_7 is uni-directional. Almost all other arrows define bi-directional reactions.

Binding of EGF to EGFR

Diagram The diagram seems to suggest the binding has EGF as a typical reactant.

Binding of EGF to EGFR

Diagram The diagram seems to suggest the binding has EGF as a typical reactant.

Supp. Tab. 1 [EGFR] + [EGF] $\rightarrow$ [EGF-EGFR]
EGF still would appear to act as a reactant.

Binding of EGF to EGFR

Diagram: The diagram seems to suggest the binding has EGF as a typical reactant.

Supp. Tab. 1: [EGFR] + [EGF] $\rightarrow$ [EGF-EGFR]
EGF still would appear to act as a reactant.

rates.dat: $v_1 = k_1 \times \text{EGF} \times \text{EGFR}$ - $k_{\text{-}1} \times \text{EGF-EGFR}$ (from Matlab definition)
v_1 definition is no different to any other defined reaction.

Binding of EGF to EGFR

Diagram: The diagram seems to suggest the binding has EGF as a typical reactant.

Supp. Tab. 1: [EGFR] + [EGF] → [EGF-EGFR]
EGF still would appear to act as a reactant.

rates.dat: $v_1 = k_1 \times$EGF$\times$EGFR - $k_{\text{-}1} \times$EGF-EGFR (from Matlab definition)
v_1 definition is no different to any other defined reaction.

goEFG.m: SM(1,:)=0;

Line 84 quietly deletes the equation for EGF

Binding of EGF to EGFR with PEPA

The claim: Building biological systems in PEPA leads to a clearer model. So how does PEPA handle EGF?

Binding of EGF to EGFR with PEPA

The claim: Building biological systems in PEPA leads to a clearer model. So how does PEPA handle EGF?

- EGF can be redefined as a constant, used to alter the rate of the $v1$ reaction.

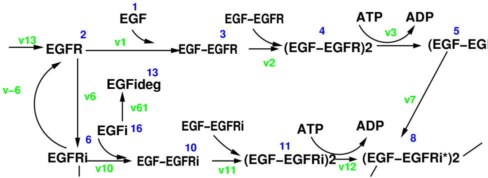
Binding of EGF to EGFR with PEPA

The claim: Building biological systems in PEPA leads to a clearer model. So how does PEPA handle EGF?

- EGF can be redefined as a constant, used to alter the rate of the $v1$ reaction.
- EGF can be defined as a catalyst.

$$\mathrm{EGF}_{\mathrm{H}} \stackrel{def}{=} (v_1, k_1).\mathrm{EGF}_{\mathrm{H}}$$

Extracts from the PEPA model of the MAP Kinase Cascade



$$\mathrm{EGF_H} \stackrel{def}{=} (v_1, k_1).\mathrm{EGF_H}$$

$$\mathrm{EGF\text{-}EGFR_H} \stackrel{def}{=} (v_2, k_2).\mathrm{EGF\text{-}EGFR_L} + (v_{\text{-}1}, k_{\text{-}1}).\mathrm{EGF\text{-}EGFR_L}$$

$$\mathrm{EGF\text{-}EGFR_L} \stackrel{def}{=} (v_1, k_1).\mathrm{EGF\text{-}EGFR_H} + (v_{\text{-}2}, k_{\text{-}2}).\mathrm{EGF\text{-}EGFR_H}$$

$$\mathrm{(EGF\text{-}EGFR)2_H} \stackrel{def}{=} (v_{\text{-}2}, k_{\text{-}2}).\mathrm{(EGF\text{-}EGFR)2_L} + (v_3, k_3).\mathrm{(EGF\text{-}EGFR)2_L}$$

$$\mathrm{(EGF\text{-}EGFR)2_L} \stackrel{def}{=} (v_2, k_2).\mathrm{(EGF\text{-}EGFR)2_H} + (v_{\text{-}3}, k_{\text{-}3}).\mathrm{(EGF\text{-}EGFR)2_H}$$

Stoichiometric Information

We just saw how reaction $v2$ was defined for components EGF-EGFR and (EGF-EGFR)2.

But the PEPA doesn't encode the fact that $v2$ is the dimerization of EGF-EGFR, and requires two EGF-EGFR to form one (EGF-EGR)2.

This is where stoichiometry is required. We augment the PEPA with knowledge of how many of each component are required for each reaction.

Automatically Generating ODEs from PEPA

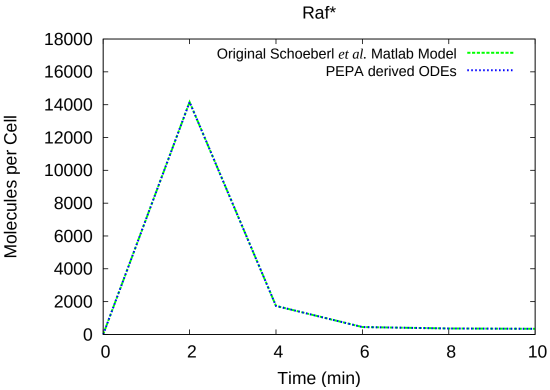
$$
\begin{aligned}
\text{EGF-EGFR}_{\mathrm{H}} &\stackrel{def}{=} (v_2, k_2).\text{EGF-EGFR}_{\mathrm{L}} + (v_{\text{-}1}, k_{\text{-}1}).\text{EGF-EGFR}_{\mathrm{L}} \\
\text{EGF-EGFR}_{\mathrm{L}} &\stackrel{def}{=} (v_1, k_1).\text{EGF-EGFR}_{\mathrm{H}} + (v_{\text{-}2}, k_{\text{-}2}).\text{EGF-EGFR}_{\mathrm{H}}
\end{aligned}
$$

$$
\begin{aligned}
(\text{EGF-EGFR})2_{\mathrm{H}} &\stackrel{def}{=} (v_{\text{-}2}, k_{\text{-}2}).(\text{EGF-EGFR})2_{\mathrm{L}} + (v_3, k_3).(\text{EGF-EGFR})2_{\mathrm{L}} \\
(\text{EGF-EGFR})2_{\mathrm{L}} &\stackrel{def}{=} (v_2, k_2).(\text{EGF-EGFR})2_{\mathrm{H}} + (v_{\text{-}3}, k_{\text{-}3}).(\text{EGF-EGFR})2_{\mathrm{H}}
\end{aligned}
$$

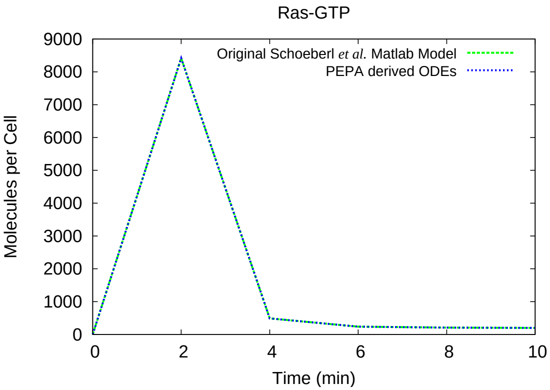
These PEPA definitions can be used to form ODEs such as the ones shown here [EGF-EGFR is $y(3)$ and (EGF-EGFR)2 is $y(4)$].

$$
\begin{aligned}
\frac{\mathrm{d}y(3)}{\mathrm{d}t} &= k_1 y(1) y(2) - k_{\text{-}1} y(3) - 2k_2 y(3)^2 + 2k_{\text{-}2} y(4) \\
\frac{\mathrm{d}y(4)}{\mathrm{d}t} &= k_2 y(3)^2 - k_3 y(4) - k_{\text{-}2} y(4) + k_{\text{-}3} y(5)
\end{aligned}
$$

Comparing Original Results and PEPA Derived ODEs

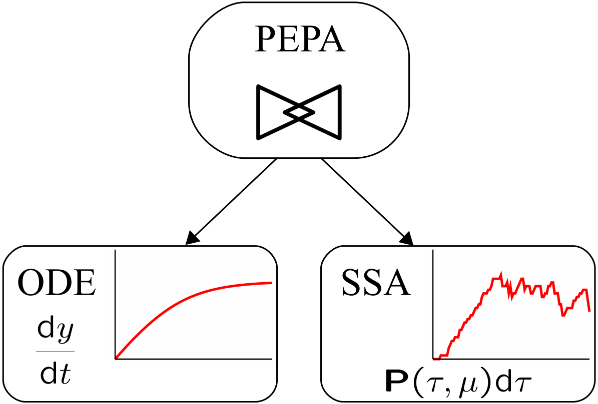


Comparing Original Results and PEPA Derived ODEs

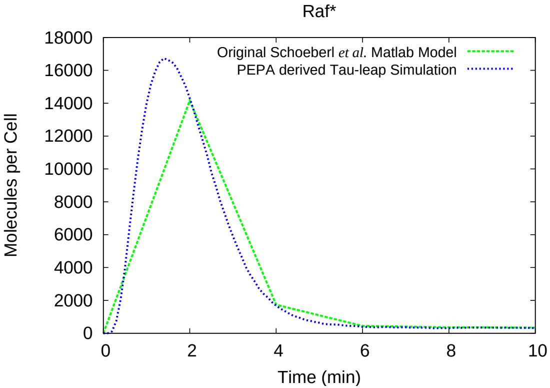


The PEPA derived ODEs return the same results as the Schoeberl *et al.* Matlab model.

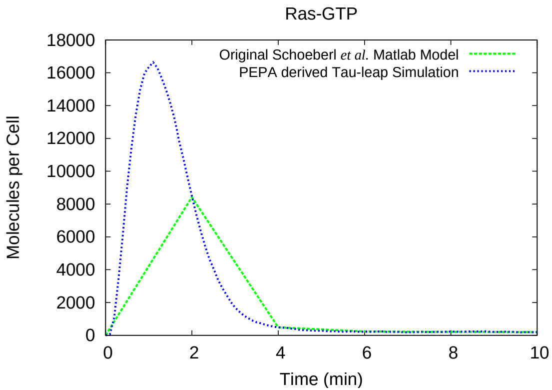
Comparing Original Results and PEPA Derived Tau-leap Simulation



Comparing Original Results and PEPA Derived Tau-leap Simulation

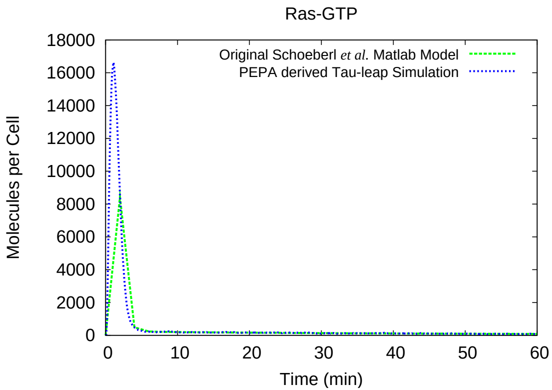


Comparing Original Results and PEPA Derived Tau-leap Simulation



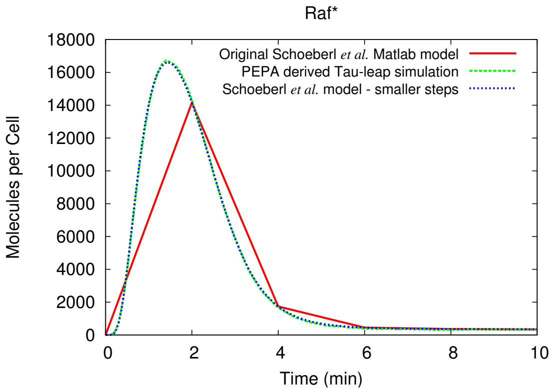
So why the difference between τ leap and the ODEs?

Comparing Original Results and PEPA Derived Tau-leap Simulation

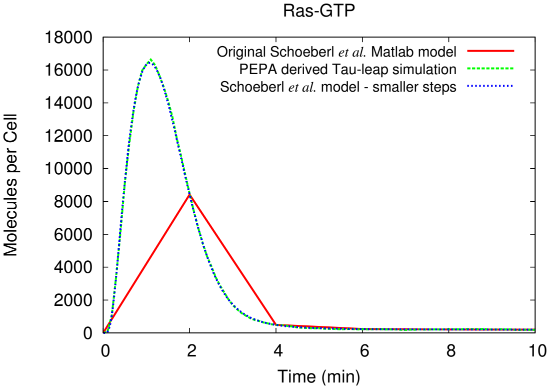


So why the difference between τ leap and the ODEs?

Corrected Time Step in Matlab Model



Corrected Time Step in Matlab Model



The original parameters for the Matlab model stepped over the true peak. The Tau-leap simulation was in fact returning the correct results.

Conclusion

- The MAP Kinase cascade is one of the larger biological models currently in circulation.
- It has been shown that PEPA can cope with models of this size.
- PEPA offers a cleaner, more precise view of the system at hand.
- PEPA allows multiple forms of analysis.
- This ability led to the discovery that the true peaks of Raf* and Ras-GTP were being stepped over by the ODE solver due to a large time step being specified.

Conclusion

- The MAP Kinase cascade is one of the larger biological models currently in circulation.
- It has been shown that PEPA can cope with models of this size.
- PEPA offers a cleaner, more precise view of the system at hand.
- PEPA allows multiple forms of analysis.
- This ability led to the discovery that the true peaks of Raf* and Ras-GTP were being stepped over by the ODE solver due to a large time step being specified.

Thank you