Process algebras for quantitative analysis: Lecture 8 — Systems Biology

Stephen Gilmore

School of Informatics
The University of Edinburgh
Scotland

February 12, 2010

Acknowledgements

M. Calder, A. Duguid, S. Gilmore and J. Hillston,

Stronger computational modelling of signalling pathways using both continuous and discrete-state methods,

Computational Methods in Systems Biology,

Trento, Italy, 2006.

Outline

Outline

Motivation and Background

Over the last few years we have been experimenting with the use of the stochastic process algebra PEPA to model and analysis biochemical signalling pathways.

Motivation and Background

Over the last few years we have been experimenting with the use of the stochastic process algebra PEPA to model and analysis biochemical signalling pathways.

In order to test the expressiveness of PEPA and our *reagent-centric* style of modelling we wanted to undertake a large case study.

Motivation and Background

Over the last few years we have been experimenting with the use of the stochastic process algebra PEPA to model and analysis biochemical signalling pathways.

In order to test the expressiveness of PEPA and our *reagent-centric* style of modelling we wanted to undertake a large case study.

Choosing an example from the literature meant that we could also validate our analysis techniques against published results.

Motivation and Background

Over the last few years we have been experimenting with the use of the stochastic process algebra PEPA to model and analysis biochemical signalling pathways.

In order to test the expressiveness of PEPA and our *reagent-centric* style of modelling we wanted to undertake a large case study.

Choosing an example from the literature meant that we could also validate our analysis techniques against published results.

The aim was not to make grand discoveries about this particular signalling pathway — more to explore the boundaries of modelling biological systems with PEPA.

Outline

Performance Evaluation Process Algebra (PEPA)

Performance Evaluation Process Algebra (PEPA)

PREFIX: $(\alpha, r).S$ designated first action

Performance Evaluation Process Algebra (PEPA)

PREFIX:	$(\alpha, r).S$	designated first action
CHOICE:	$S + S$	competing components (determined by race policy)

Performance Evaluation Process Algebra (PEPA)

PREFIX:	$(\alpha, r).S$	designated first action
CHOICE:	$S + S$	competing components (determined by race policy)
CONSTANT:	$A \stackrel{def}{=} S$	assigning names

Performance Evaluation Process Algebra (PEPA)

PREFIX:	$(\alpha, r).S$	designated first action
CHOICE:	$S + S$	competing components (determined by race policy)
CONSTANT:	$A \stackrel{def}{=} S$	assigning names
COOPERATION:	$P \underset{L}{\bowtie} P$	$\alpha \notin L$ concurrent activity (*individual actions*) $\alpha \in L$ cooperative activity (*shared actions*)

Performance Evaluation Process Algebra (PEPA)

PREFIX:	$(\alpha, r).S$	designated first action
CHOICE:	$S + S$	competing components (determined by race policy)
CONSTANT:	$A \stackrel{def}{=} S$	assigning names
COOPERATION:	$P \underset{L}{\bowtie} P$	$\alpha \notin L$ concurrent activity (*individual actions*) $\alpha \in L$ cooperative activity (*shared actions*)

Performance Evaluation Process Algebra (PEPA)

PREFIX:	$(\alpha, r).S$	designated first action
CHOICE:	$S + S$	competing components (determined by race policy)
CONSTANT:	$A \stackrel{def}{=} S$	assigning names
COOPERATION:	$P \underset{L}{\bowtie} P$	$\alpha \notin L$ concurrent activity (*individual actions*) $\alpha \in L$ cooperative activity (*shared actions*)
HIDING:	P/L	abstraction $\alpha \in L \Rightarrow \alpha \rightarrow \tau$

Mapping biological systems to process algebra

There has been much work on the use of the stochastic π-calculus and related calculi, for modelling biochemical signalling within cells

Mapping biological systems to process algebra

There has been much work on the use of the stochastic π-calculus and related calculi, for modelling biochemical signalling within cells

This work treats a molecule in a pathway as corresponding to the component in the process algebra description.

Mapping biological systems to process algebra

There has been much work on the use of the stochastic π-calculus and related calculi, for modelling biochemical signalling within cells

This work treats a molecule in a pathway as corresponding to the component in the process algebra description.

In the PEPA modelling we have been doing we have experimented with more abstract mappings between process algebra constructs and elements of signalling pathways.

Mapping biological systems to process algebra

There has been much work on the use of the stochastic π-calculus and related calculi, for modelling biochemical signalling within cells

This work treats a molecule in a pathway as corresponding to the component in the process algebra description.

In the PEPA modelling we have been doing we have experimented with more abstract mappings between process algebra constructs and elements of signalling pathways.

For example, we focus on species (c.f. a type rather than an instance, or a class rather than an object) and use local states to capture discretized levels of concentration.

Reagent-centric modelling

Reagent role	*Impact on reagent*	*Impact on reaction rate*
Producer	decreases concentration	has a positive impact, i.e. proportional to current concentration
Product	increases concentration	has no impact on the rate, except at saturation
Enzyme	concentration unchanged	has a positive impact, i.e. proportional to current concentration
Inhibitor	concentration unchanged	has a negative impact, i.e. inversely proportional to current concentration

Deriving quantitative data

PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

Deriving quantitative data

PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

The language may be used to generate a Markov Process (CTMC).

Deriving quantitative data

PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

The language may be used to generate a Markov Process (CTMC).

SPA
MODEL

Deriving quantitative data

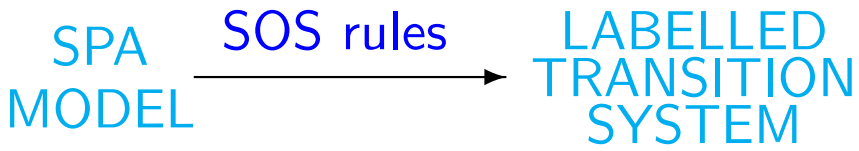
PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

The language may be used to generate a Markov Process (CTMC).

Deriving quantitative data

PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

The language may be used to generate a Markov Process (CTMC).



Deriving quantitative data

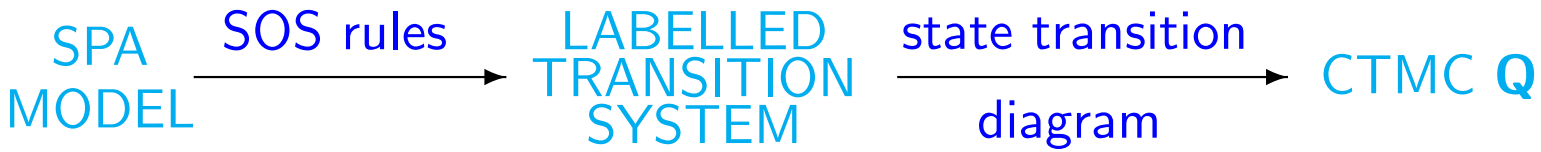
PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

The language may be used to generate a Markov Process (CTMC).

Deriving quantitative data

PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

The language may be used to generate a Markov Process (CTMC).



Deriving quantitative data

PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

The language may be used to generate a system of ordinary differential equations (ODEs).

Deriving quantitative data

PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

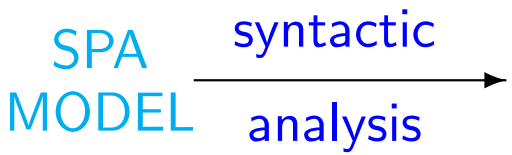
The language may be used to generate a system of ordinary differential equations (ODEs).

SPA
MODEL

Deriving quantitative data

PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

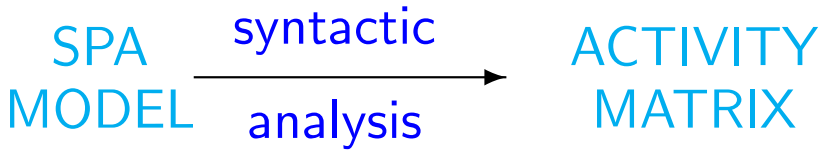
The language may be used to generate a system of ordinary differential equations (ODEs).



Deriving quantitative data

PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

The language may be used to generate a system of ordinary differential equations (ODEs).



Deriving quantitative data

PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

The language may be used to generate a system of ordinary differential equations (ODEs).

Deriving quantitative data

PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

The language may be used to generate a system of ordinary differential equations (ODEs).

Deriving quantitative data

PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

The language also may be used to generate a stochastic simulation.

Deriving quantitative data

PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

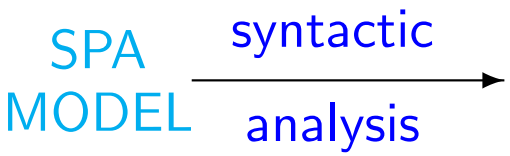
The language also may be used to generate a stochastic simulation.

SPA
MODEL

Deriving quantitative data

PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

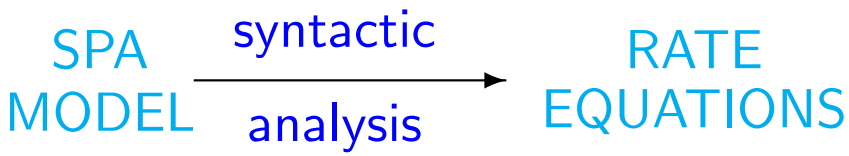
The language also may be used to generate a stochastic simulation.



Deriving quantitative data

PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

The language also may be used to generate a stochastic simulation.



Deriving quantitative data

PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

The language also may be used to generate a stochastic simulation.

Deriving quantitative data

PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

The language also may be used to generate a stochastic simulation.

Deriving quantitative data

PEPA models can be analysed for quantified dynamic behaviour in a number of different ways.

The language also may be used to generate a stochastic simulation.

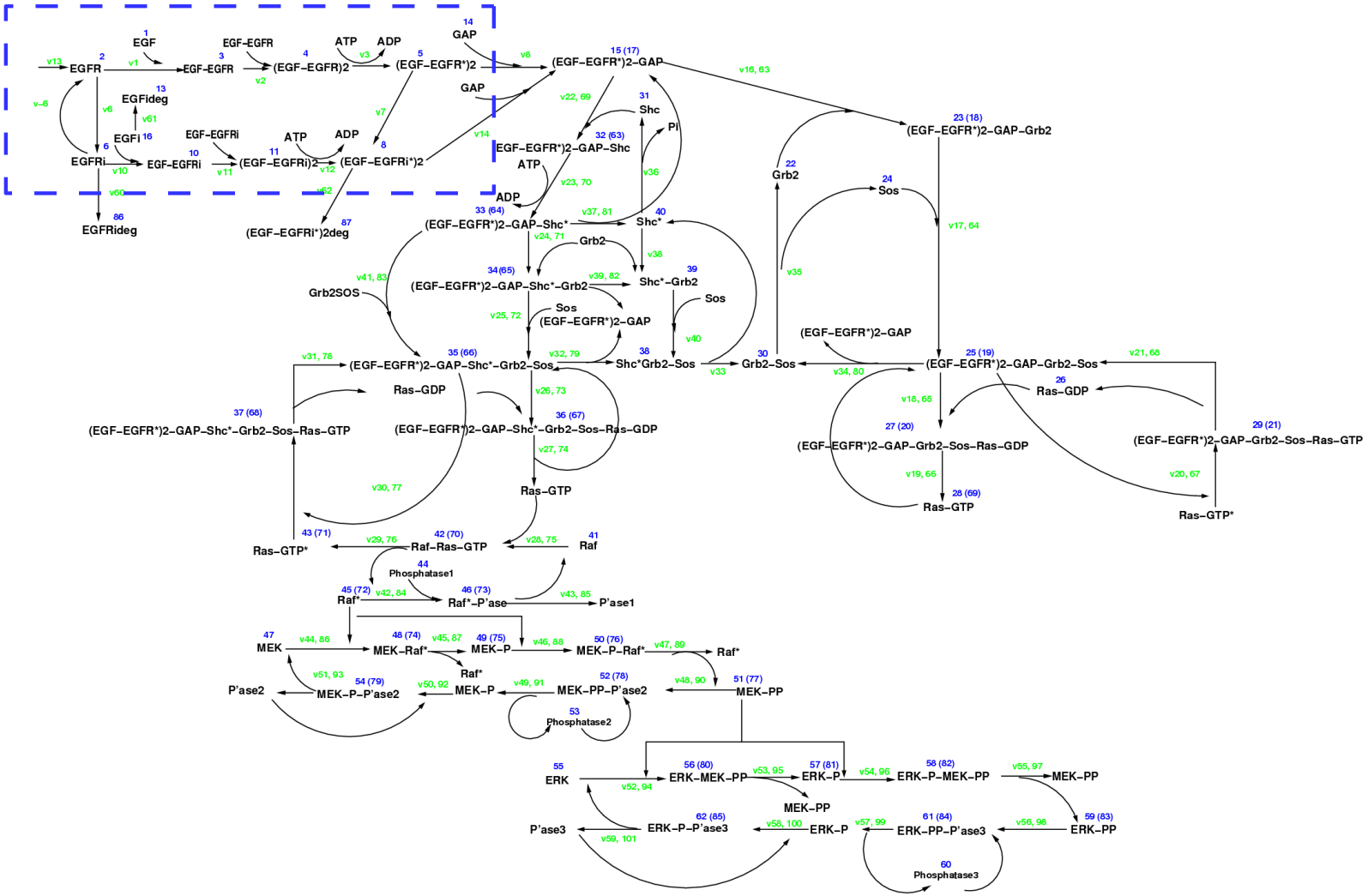
Each of these has tool support so that the underlying model is derived automatically according to the predefined rules.

Outline

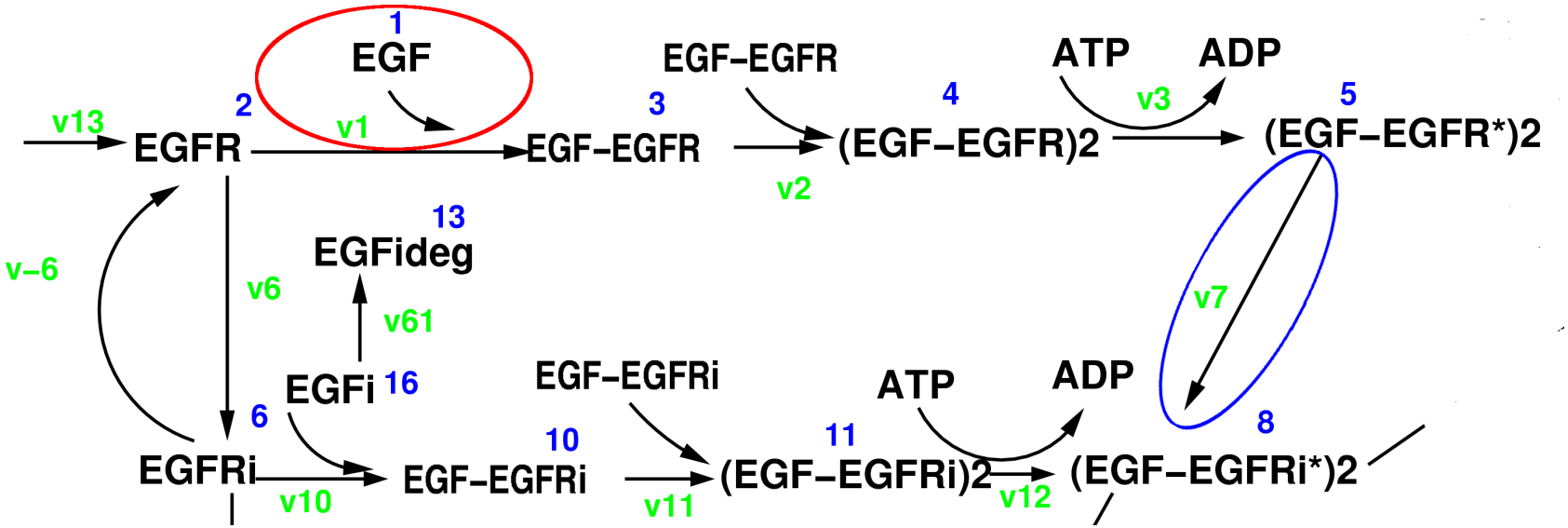
Schoeberl *et al.*'s model of the MAP Kinase Cascade

- Published in *Nature Biotechnology* 20:370-375 in 2002.
- Influential, cited by more than 500 subsequent published papers (as of 2010).
- Consists of 94 reagent species involved in 125 reactions.
- Substantial ODE model consisting of 94 state variables and 95 parameters.
- Original model constructed "by hand", with help of a graphical representation.
- Original analysis based on numerical integrators of the Matlab numerical computing platform.

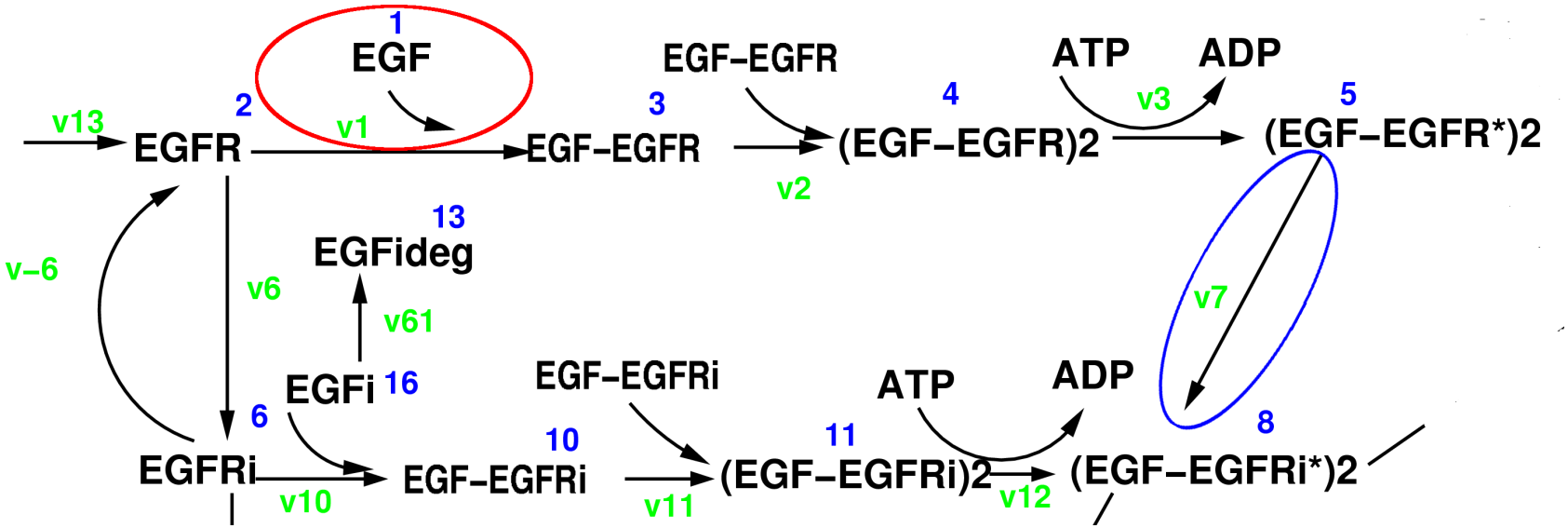
The MAP Kinase Cascade



The MAP Kinase Cascade

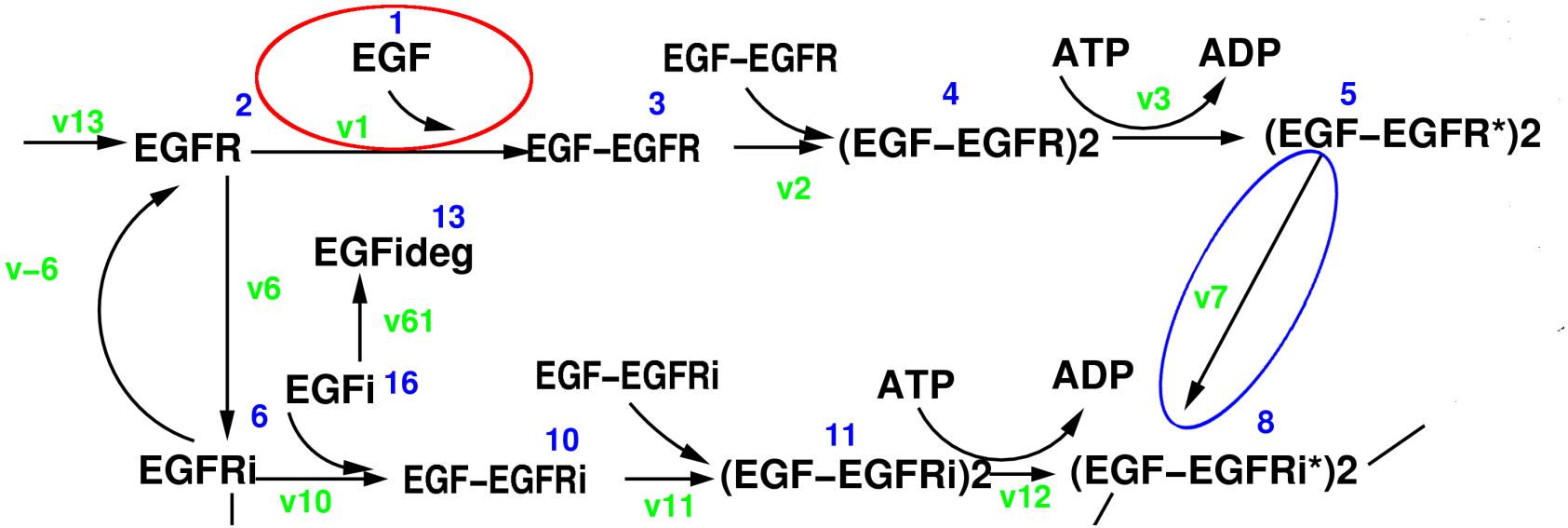


The MAP Kinase Cascade



There are many ambiguities in the graphical representation, e.g.

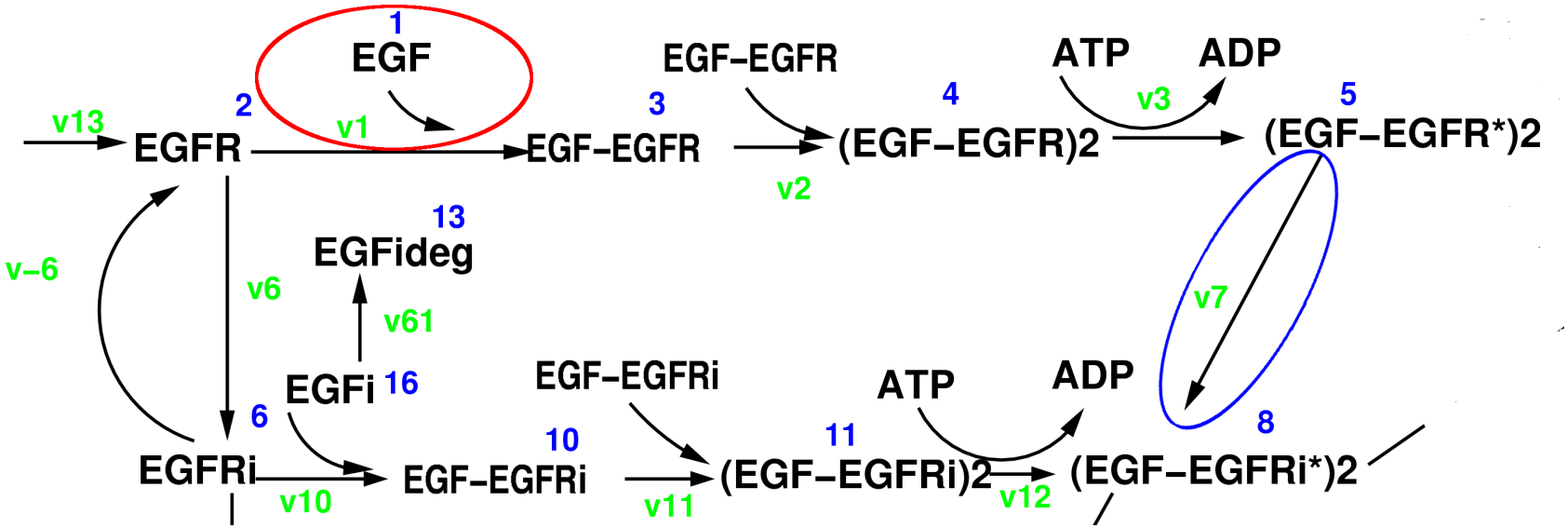
The MAP Kinase Cascade



There are many ambiguities in the graphical representation, e.g.

- An infinite supply of EGF is assumed;

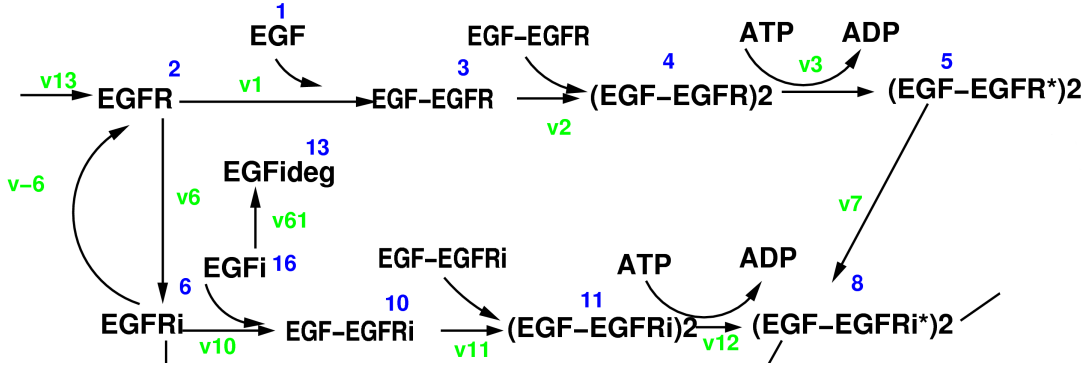
The MAP Kinase Cascade



There are many ambiguities in the graphical representation, e.g.

- An infinite supply of EGF is assumed;
- Reaction $v7$ is uni-directional whereas all others are reversible.

Extracts from the MAP Kinase PEPA model



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

$$\mathrm{EGFR_H} \stackrel{def}{=} (v_1, k_1).\mathrm{EGFR_L} + (v_6, k_6).\mathrm{EGFR_L}$$

$$\mathrm{EGFR_L} \stackrel{def}{=} (v_{\text{-}1},\ k_{\text{-}1}).\mathrm{EGFR_H} + (v_{\text{-}6},\ k_{\text{-}6}).\mathrm{EGFR_H} + (v_{13}, k_{13}).\mathrm{EGFR_H}$$

$$\text{EGF-EGFR}_\mathrm{H} \stackrel{def}{=} (v_2, k_2).\text{EGF-EGFR}_\mathrm{L} + (v_{\text{-}1}, k_{\text{-}1}).\text{EGF-EGFR}_\mathrm{L}$$

$$\text{EGF-EGFR}_\mathrm{L} \stackrel{def}{=} (v_1, k_1).\text{EGF-EGFR}_\mathrm{H} + (v_{\text{-}2}, k_{\text{-}2}).\text{EGF-EGFR}_\mathrm{H}$$

The PEPA model

Similar PEPA definitions were constructed for each of the 94 species in the pathway.

The PEPA model

Similar PEPA definitions were constructed for each of the 94 species in the pathway.

This was tedious, but not difficult, although care was needed to handle the points of ambiguity in the graphical representation.

The PEPA model

Similar PEPA definitions were constructed for each of the 94 species in the pathway.

This was tedious, but not difficult, although care was needed to handle the points of ambiguity in the graphical representation.

In order to complete the model we also needed to capture the interactions (i.e. cooperations) between the reagents. In this case we assumed that whenever reagents participated in reactions with the same name they did so in cooperation. The system equation was then automatically generated.

The PEPA model

Similar PEPA definitions were constructed for each of the 94 species in the pathway.

This was tedious, but not difficult, although care was needed to handle the points of ambiguity in the graphical representation.

In order to complete the model we also needed to capture the interactions (i.e. cooperations) between the reagents. In this case we assumed that whenever reagents participated in reactions with the same name they did so in cooperation. The system equation was then automatically generated.

The size of the model precluded numerical analysis of an underlying Markov chain for even the coarsest granularity of discretization of the concentrations of the reagents.

Validation of the PEPA model

- Once the PEPA model was constructed, we wanted to ensure that it was generating the same mathematical representation of the system.

Validation of the PEPA model

- Once the PEPA model was constructed, we wanted to ensure that it was generating the same mathematical representation of the system.
- In the first instance we derived a set of ODEs in a format suitable for Matlab.

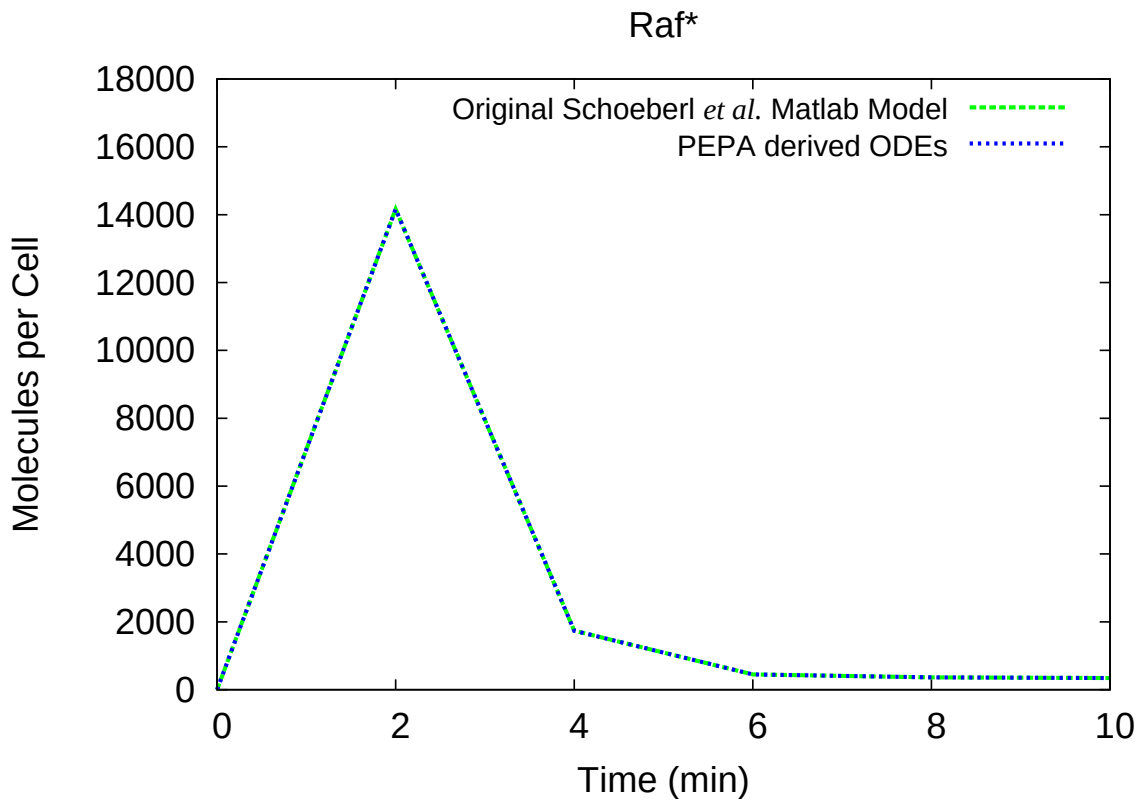
Validation of the PEPA model

- Once the PEPA model was constructed, we wanted to ensure that it was generating the same mathematical representation of the system.
- In the first instance we derived a set of ODEs in a format suitable for Matlab.
- These could not be compared directly with Schoeberl *et al*'s ODEs due to different representations being used, but we compared them empirically in terms of the results.

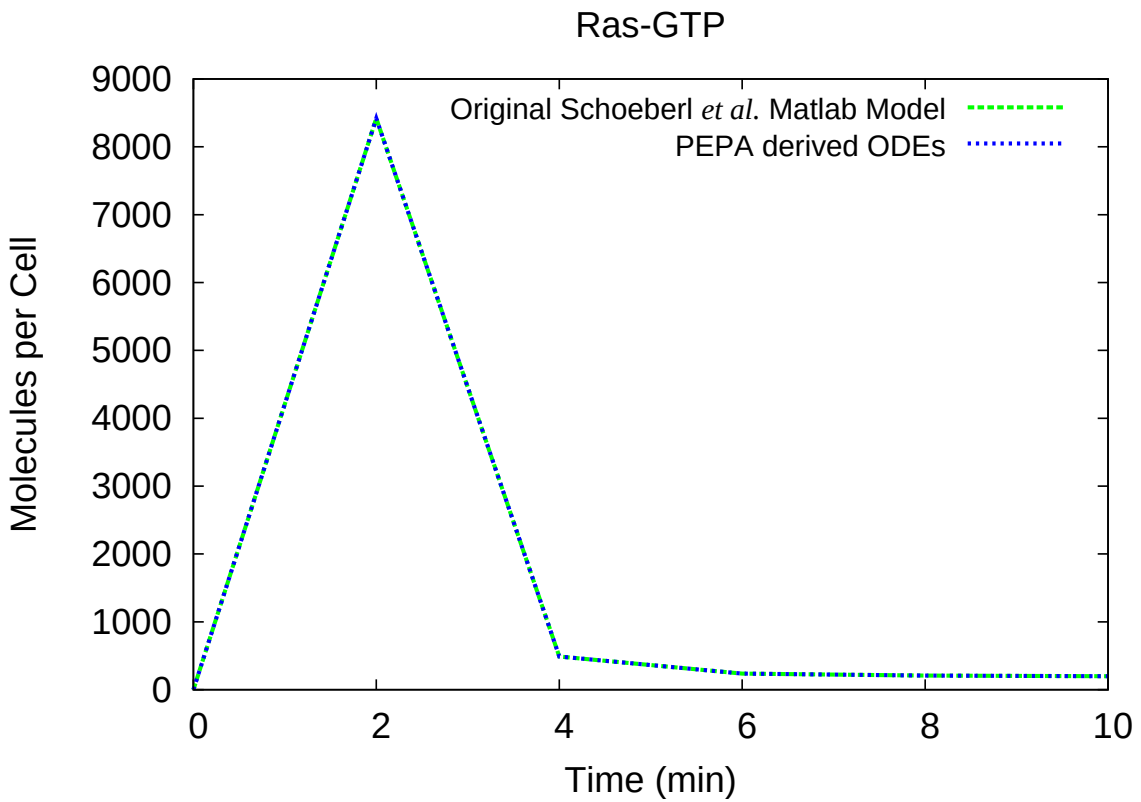
Validation of the PEPA model

- Once the PEPA model was constructed, we wanted to ensure that it was generating the same mathematical representation of the system.
- In the first instance we derived a set of ODEs in a format suitable for Matlab.
- These could not be compared directly with Schoeberl *et al*'s ODEs due to different representations being used, but we compared them empirically in terms of the results.
- Then we used an alternative mapping from the PEPA to generate a stochastic simulation of the system, and compared the results again.

Comparing Original Results and PEPA Derived ODEs

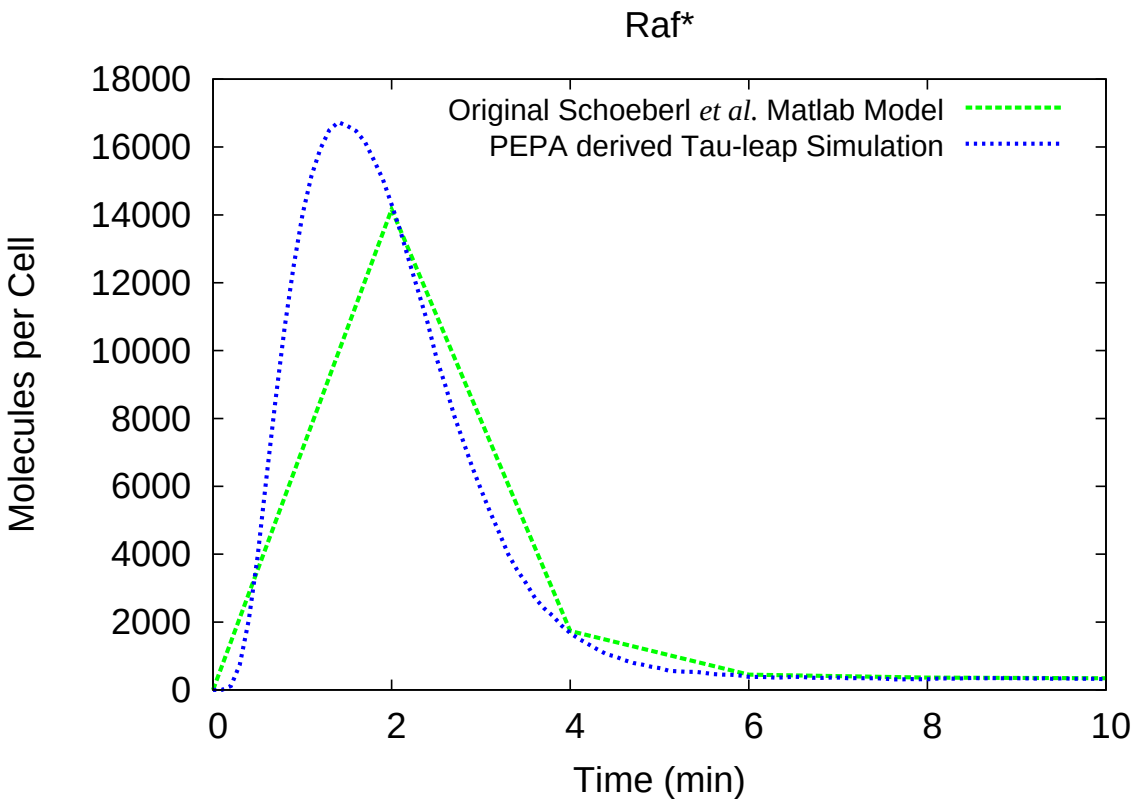


Comparing Original Results and PEPA Derived ODEs

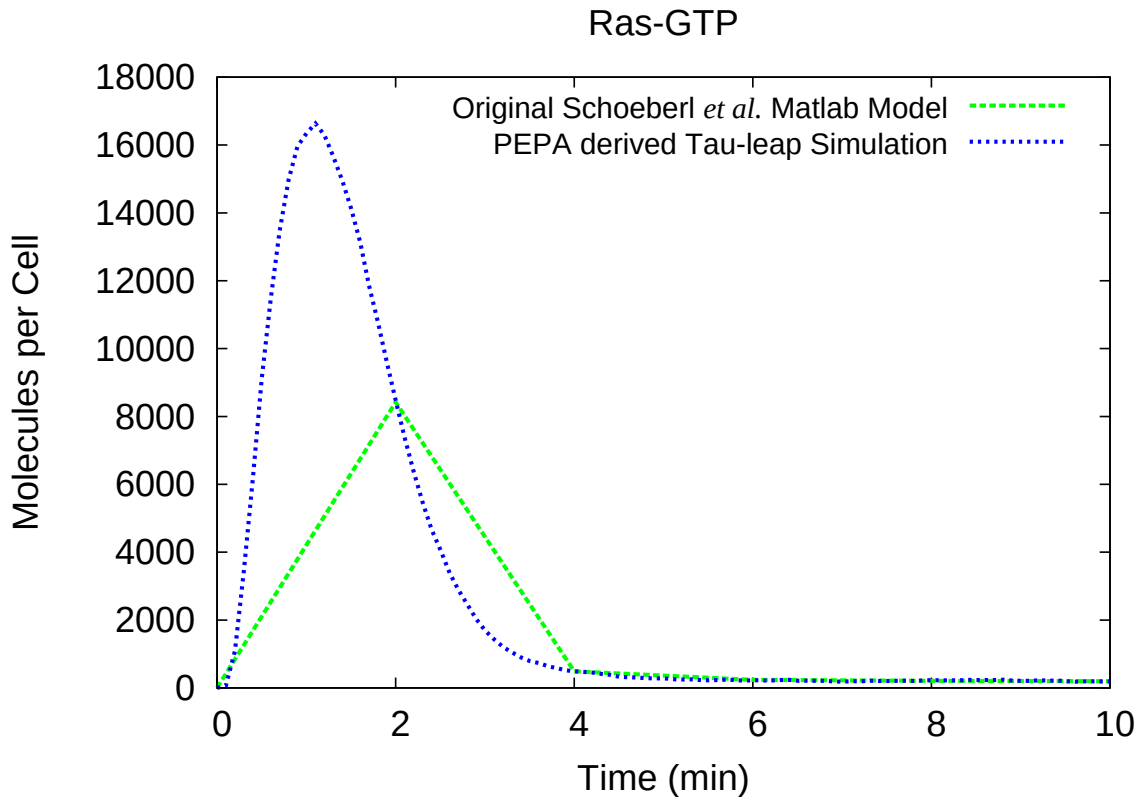


The PEPA derived ODEs return the same results as the Schoeberl *et al.* Matlab model.

Original Results and PEPA Derived Tau-leap Simulation

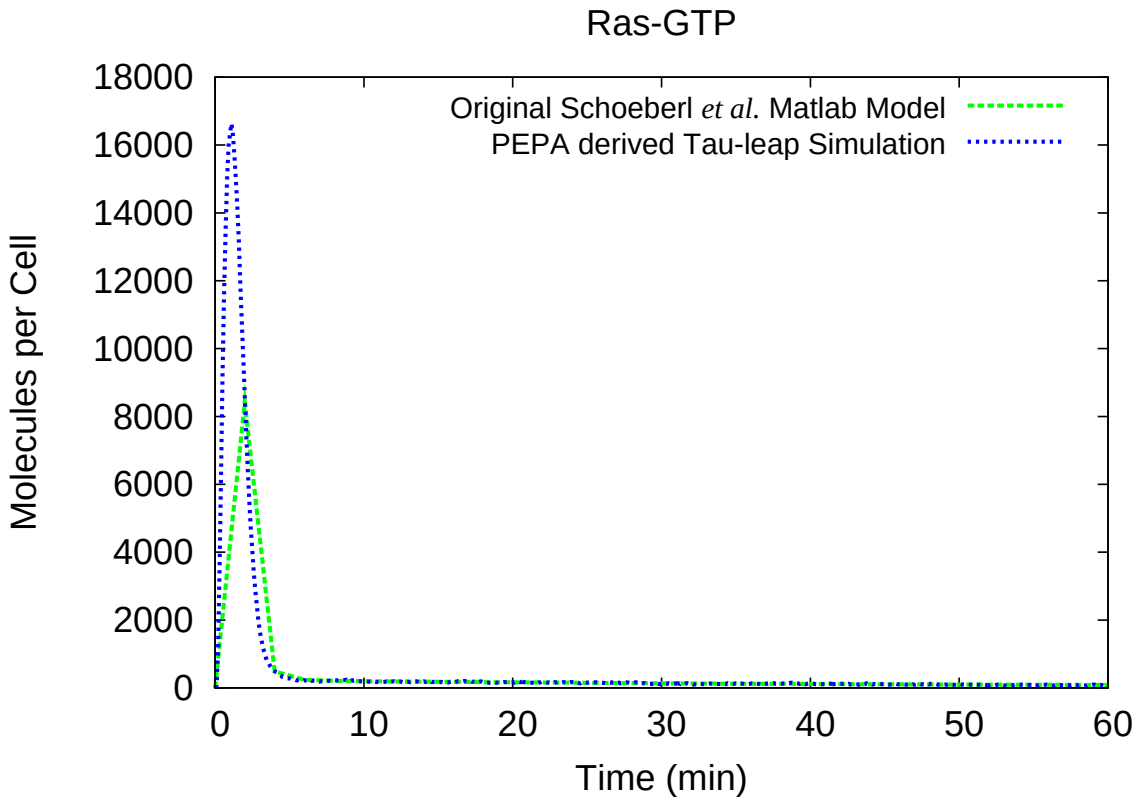


Original Results and PEPA Derived Tau-leap Simulation



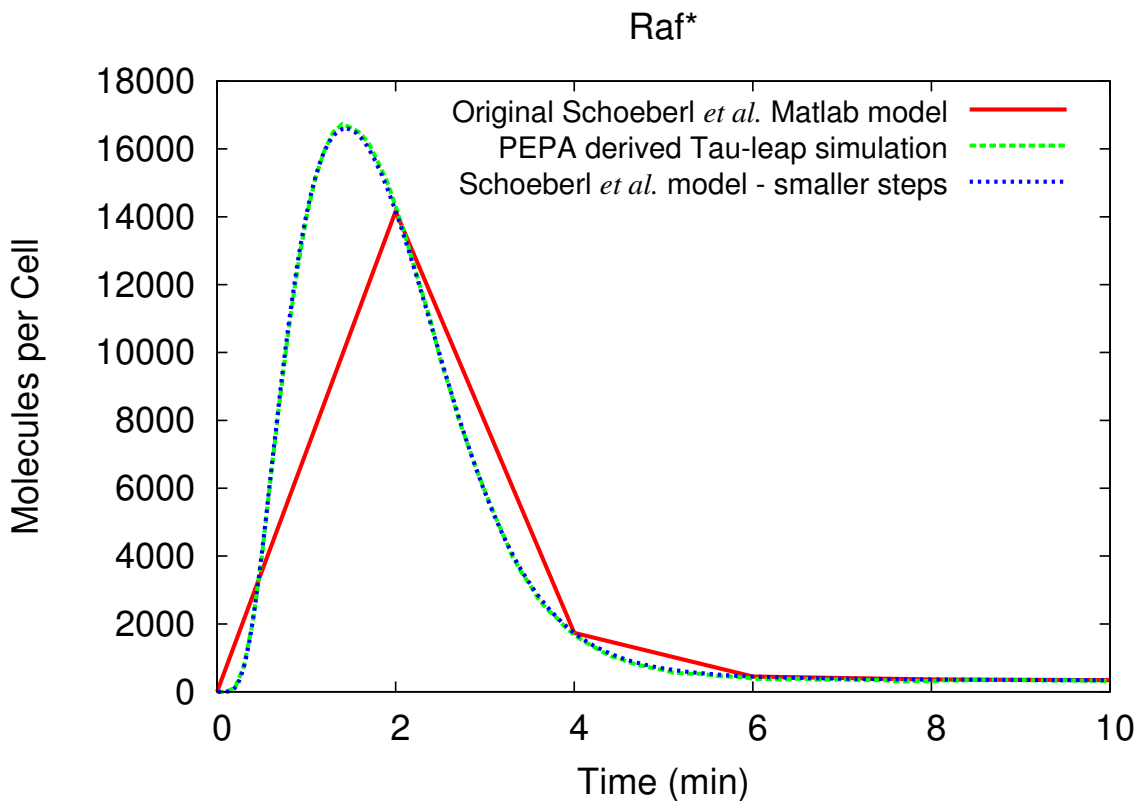
So why the difference between τ leap and the ODEs?

Original Results and PEPA Derived Tau-leap Simulation

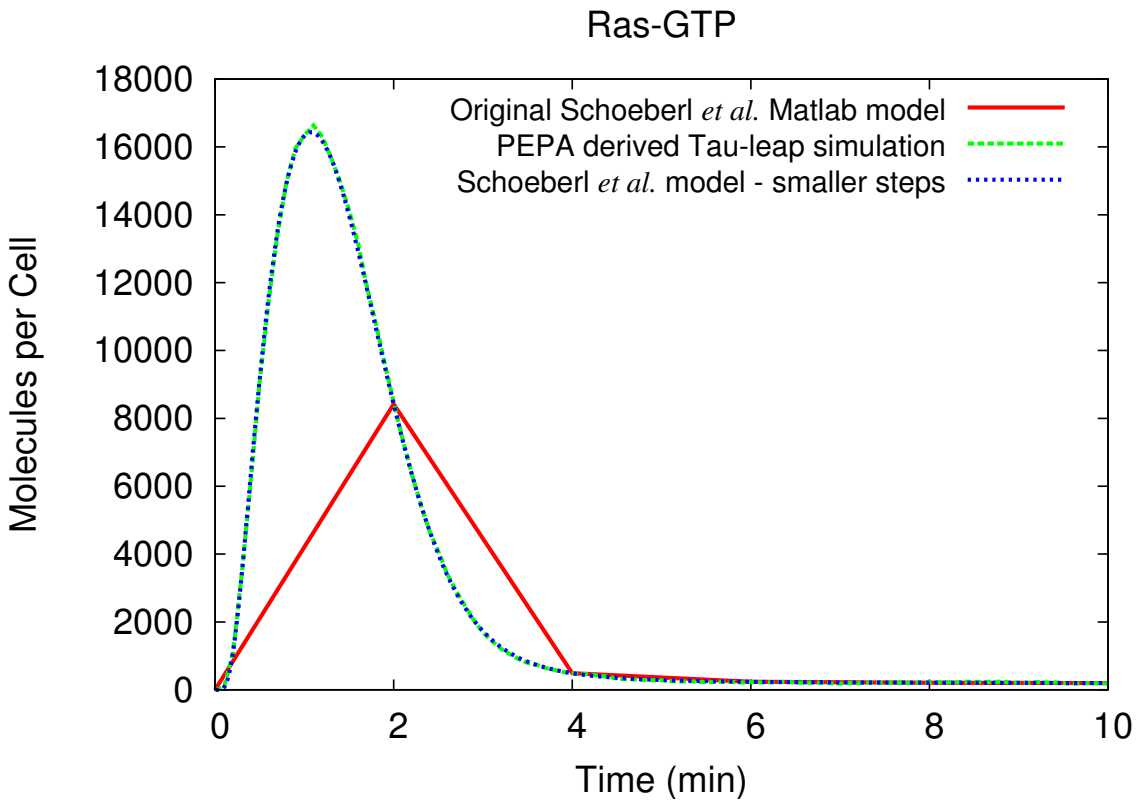


So why the difference between τ leap and the ODEs?

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 in recent literature

Conclusion

- The MAP Kinase cascade is one of the larger biological models in recent literature
- It has been shown that PEPA can cope with models of this size.

Conclusion

- The MAP Kinase cascade is one of the larger biological models in recent literature
- It has been shown that PEPA can cope with models of this size.
- PEPA offers a cleaner, more precise view of the system.

Conclusion

- The MAP Kinase cascade is one of the larger biological models in recent literature
- It has been shown that PEPA can cope with models of this size.
- PEPA offers a cleaner, more precise view of the system.
- Moreover, PEPA allows multiple forms of analysis.

Conclusion

- The MAP Kinase cascade is one of the larger biological models in recent literature
- It has been shown that PEPA can cope with models of this size.
- PEPA offers a cleaner, more precise view of the system.
- Moreover, PEPA allows multiple forms of analysis.
- This ability led to the discovery that the true peaks of Raf* and Ras-GTP concentrations were incorrectly calculated in the original analysis.