Modelling intracellular fate of FGF receptors with BioAmbients

S. van Bakel [2] I. Khan M.G. Vigliotti [1,2]

Department of Computing, Imperial College,
180 Queen's Gate, London SW7 2BZ, UK

J.K. Heath [3]

CRUK Growth Factor Group,School of Bioscience,
University of Birmingham, Egbaston, B15 2TT UK

Abstract

In this paper we consider a model for different sorting of receptors of Fibroblast Growth Factor via the endocytotic pathway. In order to accurately model the relocation in the different compartments of the cell by the ligand-receptor complex, we use the stochastic version of Bioambients. The stochastic simulation is carried out using BAM (BioAmbient Machine), which is a Java implementation of BioAmbients via Gillespie's Algorithm. Our model and the associated results of the simulation shed light on different mechanisms that influence the spatial distribution of the different components in the pathway.

Keywords: Stochastic Process algebra, system biology, FGF pathway, stochastic simulation.

1 Introduction

In recent years it has been recognised that simulation of biochemical reactions using stochastic process algebra is adequate for modelling biological systems [6,21,4,2,19,14]. The process-algebraic approach -different from mathematical techniques based on sets of differential equations- forces a rigourous description of the interaction of biological components. This way of modelling yields further insight into the dynamics of biological phenomena and, at the same time, allows to derive quantities of interest. It has been shown that results derived 'in silico', like experiments using process algebra, are consistent with both results derived from real experiments and with the Ordinary Differential Equations (ODEs) approach [2,20,11,10,13,3]. The latter models well the average behaviour of large quantities of molecules over time, but does not give an insight on the evolution of components to which those quantities refer [25]. Modelling biological complex systems in process algebra

[1] Supported by EPSRC grant SPARTACOS (EP /D047587/1).

[2] Email: svb@doc.ic.ac.uk, mgv98@doc.ic.ac.uk

[3] Email: J.K.Heath@bham.ac.uk

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

presents several advantages: models can be compositionally built, offering the opportunity to compose parts of the model that are developed at different times by different people; models can be easily manipulated by simply changing some components and evaluating the impact of those changes over the behaviour of the whole model. Needless to say, for models built in the process algebraic way, *in silico* experiments can be repeated -i.e. several runs of the same model can be performed- with different parameters allowing a simple and effective sensitivity analysis.

In recent years, many known stochastic process algebrae have been used to model biochemical reactions: PEPA [12], stochastic π-calculus [18], BioAmbients [21], to name a few. Process calculi have been successfully applied to represent biochemical pathways. It is useful to think of pathways as protocols, where the participants are the chemicals and the rules according to which the protocol evolves are prescribed by the chemical reactions present in nature; pathways describe the way in which cells communicate. Similar to protocols in computer science, pathways are very complex and present concurrent behaviour, which makes it virtually impossible to understand the temporal evolution of the whole system by simply analysing each chemical reaction in isolation. Classic process algebra is especially suited to model biochemical pathways because it naturally describes the causal dependencies among events and the concurrent behaviour among different competing components. The ultimate goal of understanding the causal dependencies among chemical reactions in pathways is to help to develop new therapies by targeting specific components.

In recent years, it has become clear that reactions in pathways can vary according to the location of components [16,9,8,23]. A typical example of this kind of pathway are the receptor mediated endocytotic pathways [15]. Among the best known pathways of this kind we mention the Epidermal Growth Factor (EGF) pathway and the Fibroblast Growth Factor (FGF) pathway [15]. To model these pathways there is a need to explicitly represent different compartments of the cell to finely describe the relocation of components. The work we present in this paper describes the endocytotic pathway of the FGF [9,8] via BioAmbients. Endocytosis is a common communication mechanism in eukaryote cells. It is a mechanism by which the cell membrane envaginates to form a membrane limited vesicle. Vesicles relocate in different compartments inside the cell. Eukaryotic cells continually engage endocytosis to supply the cell with nutrients. There are different causes to endocytosis, however, if initiated by external proteins binding to receptors located on the cell, we speak of *receptor meditated endocytosis*. The extra-cellular protein that initiates the endocytosis is called a *ligand*. The route taken by the vesicle in the receptor meditated endocytosis is well documented in the literature [15]. The vesicle containing the complex ligand-receptor moves to the *sorting endosome* and then to the *late endosome*. At this point the fate of receptors varies: either they are degraded into the *lysosome*, or they reach the membrane via the *recycling endosome*. Receptors are inactive unbound, yet the binding with the ligand activates a chemical signal which in turn could be considered the cause of cell's activity such as stimulation to divide, to migrate or to differentiate into a different cell type. Over-stimulation of such signal is deemed to be responsible for several diseases such as cancer. Thus it is important to understand very well the causes of the activation and deactivation of the receptor's signal.

It is believed that signalling of receptors stops in the lysosome as the receptor degrades. However, in recent years new insight has been gained in the FGF endocytotic pathway [9,8,23]. It is important to recall that FGF hormones are a family of twenty proteins that

share similar structures. Similarly, the four different FGF receptors are known to share similar structures. Very little is known about the different roles of the four receptors as they relocate inside the cell. In recent studies [9,8] it was shown that the distribution of the four receptors inside the lysosome and the recycling endosome varies quite dramatically. In this paper we model the different fate of receptors in the endocytotic pathway of the FGF. The key point of the paper is the relocation of the complex (FGF:FGFR) in different parts of the cell. To this aim we use the stochastic version of BioAmbients, because it naturally represents both compartments and their movements; other process caluculi such as π-calculus or PEPA do not have primitives in their language to directly model compartments. We will show that our model in BioAmbient faithfully reproduces the results present in [9,8], and will shed some light on causes for the sorting of the different receptors in the cell. In short the contributions of the paper are:

- Theoretical development in BioAmbients of the FGF endocytotic pathway focuses mostly on relocation and sorting of receptors. We simulate 'in silico' some of the experiments produced in [8]. The results of the simulations are proved consistent with the data from experiments. To the best of our knowledge, this is the first time that such a model has been proposed.
- We tested the model in two ways:

(i) We run an 'in silico experiment' of the global model with the four receptors. This shows that, for a limited period of time, the sorting of the receptors is consistent with the experiments carried out in isolation.

(ii) We modify the model in such a way that both FGF and FGFRs are randomly created, and each receptor is created at a different rate. The model shows that in the very long run most of FGFR4s end in the cell membrane. Because the rates for the creation of FGFR4 are not known, our model may not be realistic, however it sheds light on the fact that with a specific set of rates it is possible have an over-production of FGFR4s. Further sensitivity analysis is necessary to document different realistic scenarios.

The results of the simulation are obtained by using BAM (BioAmbient Machine), a tool developed at Imperial College London [17]. The tool has been implemented in Java 1.5 and simulates the stochastic behaviour of the BioAmbients via Gillespie's algorithm [7]. Our work sets the basis to understand the complex dynamics of the recycling of FGFRs in the endocytotic pathway. Such understanding could help in developing new therapies for the diseases caused by the over-stimulation of FGF receptors.

The BioAmbient model is clearly not complete, but it describes the main dynamics of the different sorting of the FGFRs. It can be made more accurate in the future by compositionally adding new components. Our work is of course an abstraction with respect to reality in the sense of [22]. In simple words, we do not aim to faithfully represent all the details of the biological system, but focus on the issue of activation/deactivation of signalling in different compartments and on relocation of components. We also do not aim to model accurately a set of chemical reactions involved in the endocytotic pathway. In this respect our work is orthogonal to [10,13,11], which consider in great details the chemical reactions involved in the early stages of binding of the FGF with the FGFR.

The rest of the paper is organised as follows: in Section 2 we review the syntax and the semantics of Bioambients. We define the stochastic semantics along the lines of [12,24]. In Section 3 we review the FGF endocytotic pathway as described in the literature (see

also [8]) and we present a model in Bioambients. In Section 4 we discuss the results obtained by running different simulations and we discuss the predictive capacity of our model. Conclusions follow. We conclude in the paper with an appendix that describes in details what experiments we have simulated [8] and the associated quantitative information.

2 BioAmbients

BioAmbients [21] are a dialect of the Ambient Calculus [5], suitable to model membrane and compartments in biology. In the original paper [21], the calculus was presented with a standard operational semantics expressed in terms of rewriting rules, and the implementation in Prolog using Gillespie's algorithm.

In this section, we introduce both syntax and the stochastic operational semantics of BioAmbients, leaving out the formal description of derivation of the underpinning Continuous Time Markov Chain, which is standard [12,1,18,24]. We slightly modify the syntax of the calculus by using explicit recursion as opposed to replication and by introducing the delay operator τ_δ.

We shall assume the existence of a set of names or channels $\mathcal{N}$, and let the meta-variables $n, m, z, s, \ldots$ range over this set.

Definition 2.1 The set of processes of BioAmbients is given by the following syntax:

$$\begin{array}{rcl} P, Q & ::= & \mathbf{0} \mid P|Q \mid (\mathsf{new}\, n)\, P \mid \mathbf{[}P\mathbf{]} \mid \mathsf{A}\langle \widetilde{x} \rangle \mid \sum_{i\in I} M_i.P_i \\ M, N & ::= & \mathbf{enter}\, n \mid \mathbf{exit}\, n \mid \mathbf{accept}\, n \mid \mathbf{expel}\, n \mid \tau_r \mid \\ & & \mathbf{merge}^+\, n \mid \mathbf{merge}^-\, n \mid \$n(x) \mid \$\overline{n}\langle m \rangle \\ \$ & ::= & \mathbf{s2s} \mid \mathbf{local} \mid \mathbf{p2c} \mid \mathbf{c2p} \end{array}$$

We assume that each name has a unique rate associated to it, and that there is an environment $\rho : \mathcal{N} \to \mathbb{R}$ that formally keeps track of the rate associated to names.

We will now informally explain the syntax of the calculus. In BioAmbients there are different primitives for communication: first of all, communication happens on a channel n by sending on a name -or channel- m; $\$n(x)$ stands for the input, and $\$\overline{n}\langle m \rangle$ stands for output. There are three ways of communicating: channels in the same ambient perform *local communication*, $\mathbf{local}\, n(y)$ for the input on channel n and $\mathbf{local}\, \overline{n}\langle m \rangle$ for the output of m on channel n. Inputs and outputs located in sibling ambients respectively perform *sibling communication*; $\mathbf{s2s}\, n(y)$ stands for such input and $\mathbf{s2s}\, \overline{n}\langle m \rangle$ stands for output. Finally, parent to child communication happens when inputs $\mathbf{p2c}\, n(y)$ and outputs $\mathbf{p2c}\, \overline{n}\langle m \rangle$ are located in parent-child ambients respectively (or vice-versa for $\mathbf{c2p}\, n(y)$ and $\mathbf{c2p}\, \overline{n}\langle m \rangle$). The capabilities such as $\mathbf{exit}\, n$ or $\mathbf{enter}\, n$ give the ambient the power to become active; $\mathbf{enter}\, n / \mathbf{accept}\, n$ allow an ambient to move into a sibling, $\mathbf{exit}\, n / \mathbf{expel}\, n$ allow a child ambient to leave the parent, while $\mathbf{merge}^+\, n$ and $\mathbf{merge}^-\, n$ together fuse two sibling ambients into a single ambient.

As far as processes are concerned, *Nil* represents the inactive process; *Local sum* $\sum_{i\in I} M_i.P_i$ represents the standard choice. Given a set of indexes I and a permutation p on it, we write $\sum_{p(i)\in I} M_{p(i)}.P_{p(i)}$ to represent a reordering of the terms of the summation. We reserve the letters G, C to represent summation as in $\sum_{i\in I} M_i.P_{p(i)} = M_j.P_j + G$ where $G = \sum_{i\in I\ i\neq j} M_i.P_i$. In general, inputs are binding operators on the arguments. This means that in the process $\mathbf{local}\, n(y).P$ the name y is bound in P, and not accessible from

$$
\begin{array}{ll}
P \mid \mathbf{0} & \equiv P \\
P \mid Q & \equiv Q \mid P \\
(P \mid Q) \mid R & \equiv P \mid (Q \mid R) \\
(\mathsf{new}\, n)\, \mathbf{0} & \equiv \mathbf{0} \\
(\mathsf{new}\, m)\, (\mathsf{new}\, n)\, P & \equiv (\mathsf{new}\, n)\, (\mathsf{new}\, m)\, P \\
(\mathsf{new}\, n)\, (P \mid Q) & \equiv P \mid (\mathsf{new}\, n)\, Q \quad \text{if } n \notin \mathsf{fn}(P) \\
(\mathsf{new}\, m)\, \mathbf{[}P\mathbf{]} & \equiv \mathbf{[}(\mathsf{new}\, m)\, P\mathbf{]} \\
\sum_{i \in I} M_i.P_i & \equiv \sum_{p(i) \in I} M_{p(i)}.P_{p(i)} \\
\mathsf{A}\langle \widetilde{m} \rangle & \equiv P\{\widetilde{m}/\widetilde{x}\} \qquad \text{if } A(\widetilde{x}) = P
\end{array}
$$

Fig. 1. Structural congruence

outside P. A similar argument applies to the other inputs in the communication primitives. The process $\tau_r.P$ represents the delay for an amount of time that is exponentially distributed with rate r. *Ambient* $\mathbf{[}P\mathbf{]}$ represents a compartment with an active process P. *Parallel composition* $P|Q$ means that P and Q are running in parallel. *Restriction* $(\mathsf{new}\, a)\, P$ of the name a makes that name private and unique to P: the name a becomes bound in P. *Recursion* $\mathsf{A}\langle \widetilde{x} \rangle$ models infinite behaviour by assuming the existence of a set of equations of the form $A(\widetilde{x}) \stackrel{\text{df}}{=} P$ such that $\{\widetilde{x}\} \subset \mathsf{fn}(P)$, where $\mathsf{fn}(P)$ stands for the usual free names of P. The definition of $\mathsf{fn}(P)$ is standard, taking into account that the only binding operators are inputs and restriction. We write $P\{y/m\}$ to mean the substitution of every occurrence of the name y by m in P. Similarly we write $P\{A/Q\}$ to mean the substitution of every occurrence of the process A by Q in P.

Formally, steps of computation are represented by a *reduction relation* which is defined in Figure 2. The reduction relation specifies how terms evolve syntactically, and the rate yielded by the environment ρ is sufficient information to specify how the process evolves over time as well. We assume that each transition involving the name is exponentially distributed with parameter given by $\rho(n)$ (for the name n involved in the transition). That means that we can associate to each term a random variable over any interval Δx for $x \in \mathbb{R}$ and regard the evolution of terms as a stochastic process.

The underpinning model turns out to be a Continuous Time Markov Chain. We omit in this paper the formal description on derivation of Continuous Time Markov Chains given a stochastic process algebra since it is standard [12,1,18,24]. The definition of reduction relation involves *structural congruence*, $\equiv$. This is the smallest congruence relation as defined in Figure 1.

3 Modelling the intracellular trafficking of FGFRs BioAmbients

Cells (in complex organisms) communicate with neighbouring cells and their environment via receptors situated in the membrane. Cell receptors are classified into families, based upon similarity in structure, ligand binding and the biological response they induce [9]. *Receptor Tyrosine Kinases* (RTK) are receptors located on the surface of the cell. Their activity is induced by the corresponding signalling RTK protein generally located outside the cell.

$$
\begin{aligned}
G + \tau_r.P &\xrightarrow{r} P \\
[(G + \mathbf{enter}\, n.P) \mid Q] \mid [(G' + \mathbf{accept}\, n.R) \mid S] &\xrightarrow{\rho(n)} [[P \mid Q] \mid R \mid S] \\
[[(G + \mathbf{exit}\, n.P) \mid Q] \mid (G' + \mathbf{expel}\, n.R) \mid S] &\xrightarrow{\rho(n)} [P \mid Q] \mid [R \mid S] \\
[(G + \mathbf{merge}^{+}\, n.P) \mid Q] \mid [(G' + \mathbf{merge}^{-}\, n.R) \mid S] &\xrightarrow{\rho(n)} [P \mid Q \mid R \mid S] \\
[(C + \mathbf{local}\, n(y).P) \mid (C' + \mathbf{local}\, \overline{n}\langle m\rangle.Q)] &\xrightarrow{\rho(n)} [P\{y/m\} \mid Q] \\
[(C + \mathbf{s2s}\, n(y).P)] \mid [(C' + \mathbf{s2s}\, \overline{n}\langle m\rangle.Q)] &\xrightarrow{\rho(n)} [P\{y/m\}] \mid [Q] \\
[[(C + \mathbf{c2p}\, n(y).P)] \mid (C' + \mathbf{p2c}\, \overline{n}\langle m\rangle.Q)] &\xrightarrow{\rho(n)} [[P\{y/m\}] \mid Q] \\
[[(C + \mathbf{c2p}\, \overline{n}\langle m\rangle.Q)] \mid (C' + \mathbf{p2c}\, n(y).P)] &\xrightarrow{\rho(n)} [[Q] \mid P\{y/m\}]
\end{aligned}
$$

$$
\frac{P \xrightarrow{r} P'}{P \mid R \xrightarrow{r} P' \mid R} \qquad \frac{P \xrightarrow{r} P'}{(\mathsf{new}\, n)\ P \xrightarrow{r} (\mathsf{new}\, n)\ P'}
$$

$$
\frac{P \xrightarrow{r} P'}{[P] \xrightarrow{r} [P']} \qquad \frac{P \equiv P' \xrightarrow{r} Q' \equiv Q}{P \xrightarrow{r} Q}
$$

Fig. 2. Reduction Relation

FGFs and the four related *Tyrosine Kinase Fibroblast* growth factor receptors (FGFRs) play a significant role in the regulation of many key cellular responses in wound healing and embryonic development. We repeat below the findings of experiments reported on in [9].

After binding, the FGF rapidly internalises into the cell as a complex i.e. the endocytosis has begun. The receptors are deemed to be active, i.e. produce signalling while bound. A small number of receptors are internalised into the cell without having a ligand attached to them; these are deemed inactive. Upon internalisation, the receptors appear in early/sorting endosomes. This is the first main branch point in the receptor-mediated endocytotic pathway. Molecules in the sorting endosomes can be sorted to late endosomes, or those which are not retained in the sorting endosomes recycle either directly or via the endocytotic recycling compartment back to the cell surface. After extensively studying the EGF receptors, it was found that the ligand-free receptors are recycled to the cell surface whereas the ligand-occupied receptors are routed to the lysosomes to be degraded. However, a small fraction of the ligand-occupied receptors recycle to the surface via the recycling endosome. The routing of the ligand-receptor complex to the late endosomes from the early/sorting endosomes aims to terminate the signalling. From here the complex may recycle to the cell surface through via recycling endosome or degrade by entering the lysosomes [9,8].

Growth factors bind to more than one receptor. In many cases it is unclear what the different roles for the separate receptors in signal transduction are. The intracellular trafficking of ligand bound receptors for FGFs was studied in [9,8] to determine whether intracellular sorting of ligand-receptor complexes may modulate the signalling.

There are four tyrosine kinase FGFR (FGFR1-FGFR4) and about twenty FGFs. It was found that FGF1 binds equally well to any receptor, so only FGF1 was used in the experiments in [9,8]. It is known that the endocytosis of the FGFRs utilises different mechanisms

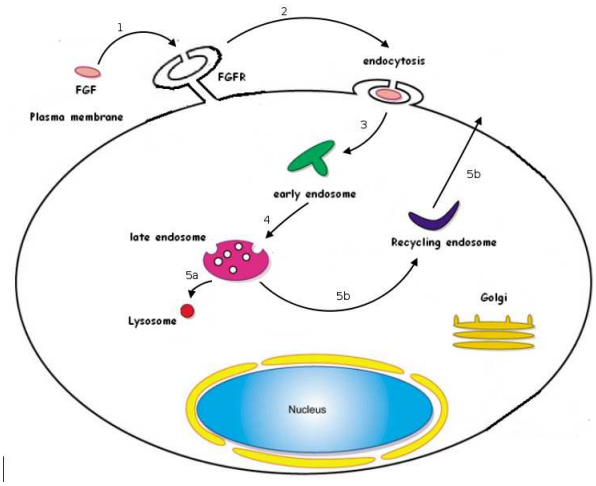

Fig. 3. Endocytosis route

for internalisation and this also varies between cell types. In the experiment reported in [9,8] HeLa cells (cells derived from cervical cancer cells) were used and transfected with any one of the receptors to test whether the signalling was modulated depending upon the ligand-receptor complexes formed. We present the experiments in Appendix A by stating their purpose, in a brief overview of the results of the experiments [9,8]. The experiments overall aimed to understand where the FGF was located once it had been internalised into the cell. It was found that depending on which receptor was used to internalise the FGF, the probability of it locating to different compartments within the cell varied.

In this paper we have taken the approach of simulating the different experiments as described in Appendix A. We have then compared the results, and run a simulation of the full model by considering all the four receptors at once. We will use the pathway defined above to model the intracellular sorting. Figure 3 shows the cell with the movements of the FGF:FGFR complex once it has entered the cell.

The experiments were carried out to determined the fate of FGF1 after binding with one of its receptors FGFR1-FGFR4. The receptors are located on the cell surface. FGF1 binds with the FGFR to form a complex which is then internalised into the cell. Experiment 2 shows that once the complex is internalised, the receptor and ligand both go to the same compartments giving rise to the idea that they are both still bound and do not separate once they have entered the cell. Once the ligand-receptor complex has entered the cell, the complex goes to the endosomal compartment. Experiment 3 was carried out to prove this happens. Experiment 4 was carried out to find where the complex goes after two hours and to determine whether the routing differs depending on which receptor the FGF1 binds to. In two hours it was found that a greater percentage of FGF1:FGFR1 had routed to the late endosomal compartment than the FGF1:FGFR4 complex, suggesting the rates differ depending on the receptor the ligand binds to. After the complex has entered the late endosomes, Experiment 5 was carried out to determine what happens next. It was found that the FGF1:FGFR1-FGFR3 route to the lysosomes where they are digested and degraded. The FGF1:FGFR4 complex was found to route to the recycling compartment.

The final few experiments were carried out to calculate the rates at which the ligand and receptors degrade. By analysing the results from the experiment we can clearly see that two different routes of movement within the cell exist, depending on which FGFR the FGF1 had

bound with. From the experiments it has been found that FGF1 internalised by FGFR1-FGFR3 is generally routed to the lysosomes for degradation and that FGF1 internalised by FGFR4 is mainly routed to the recycling compartment.

From the experiments, we have obtained two extremely important pieces of information:

(i) The movement of the FGF:FGFR complex inside the different compartments in the cell depends on which FGFR receptor it has bound with.

(ii) Rate information detailing the different rates of degradation and recycling depends on which receptor the FGF bound with.

In our model we aim to obtain a high-level view of the movements of the complex, as opposed to the finer details of the chemical reactions. That is, we abstract away from representation of chemical reactions and we just concentrate of localisation of the complex.

The model we have implemented in BioAmbient is the following:

(i) FGF is found outside cells and FGFRs are located on the cell's surface. The FGF binds with the FGFR and forms a complex;

(ii) The FGF:FGFR complex enters the cell through endocytosis;

(iii) The FGF:FGFR complex moves to the early/sorting endosomal compartment;

(iv) The FGF:FGFR complex is routed to the late endosomes;

(v) From this point there is a choice as to what may happen next depending on which FGFR was used to internalise the FGF1:
   (a) FGFR1-FGFR3: The complex is routed to the lysosomes where it degrades;
   (b) FGFR4: The complex moves to the endosomal recycling compartment to be recycled back to the cell surface.

To formally model the intracellular sorting pathway we need to make several assumptions.

(i) For the binding of each FGF protein, exactly one FGFR receptor is required.

(ii) When the FGF:FGFR complex enters the sorting endosome, we assume they remain bound. It is known from the literature that the complex decomposes in the early endosome, and follows the route of recycling. We do not model the unbinding directly, but model the FGFR moving in other compartments of the cell.

(iii) When the FGFR complex enters the lysosomes, we assume they are completely degraded and digested and the ligand and receptor are both destroyed. We do not model directly the degradation of the receptor.

(iv) When the FGFR enters the recycling compartment it returns to the cell surface, ready to be reused.

(v) We do not model the endocytosis of inactive receptors.

A recent paper [23] has shown that if FGF:FGFR does not decompose in the sorting endosome then the complex recycles back to the cell surface directly. We leave for future work to compositionally increment our model to take into account this new findings, and to explicitly model the biological switch that causes the unbinding of the FGF:FGFR. The model presented in this paper is the starting point to model the behaviour of FGFRs, whose

over-production is deemed to play a key in role in the development of cancer.

In BioAmbients,

- we model the compartment such as early endosome, late endosome, lysosome, recycling endosome;
- we do not explicitly model the cell membrane or the perimeter of the cell;
- we model the FGFR as a compartment because it is natural to think of it as moving inside and outside of different compartments;
- we model the binding with the receptor as communication;
- we do not model the degradation of either the FGF or the FGFR directly, we simply model the different routes among the different complexes.

The core implementation of the endocytosis of FGF in BioAmbients is shown in Figure 4. The ligand FGF sends a message to ambient FGFR, which, after having received it, can move to the early endosome EN, then to the late endosome LE and then either the FGFR ends in the ambient of the lysosome LYSO, where nothing further happens or it enters the recycling compartment RECYCLE where it is routed out again ready to be used. It must be noted that in this implementation we assume that each compartment allows a finite number of vesicles to enter or exit. The core model assumes that there are about one-thousand and three hundred receptors and one thousand ligands. We run four different experiments by changing the rate of the channel $lyso$ and $lendo$. This is meant to simulate the different experiments in [8]. Lastly, we run the full model with the four different receptors at the same time. The core model presented in Figure 4 is easy to manipulate and we have run a few 'experiments' which results are reported in Section 4.

4 Results

In this section we report on the findings of running the 'in silico' experiments.

First of all, we have build a basic model, where we consider essentially one kind of receptor only. The behaviour of this model simulates the FGFR1 as can be seen in Figure 5 where 80% of the receptors end in the lysosome (lysosome in Figure 5), while 20% of the receptors end in the recycling compartment (recycling). In this model we tag also the binding to the receptor (p2c fgfbind! $\{$bind$\}$), the movement in the early endosome (in! endo) and in the late endosome (in! lendo). We assume that the cell had initially no receptors in any in of the compartments. In this sense, it seems we mimic the experiments made in [9,8].

Furthermore, we have examined each receptor in isolation, as shown in the graphs in Figure 6, 7, and 8; our results are consistent with the findings in the paper [9,8]. Finally, we have run a model with the four receptors and the findings are reported in Figure 9. It must be noted that the graph shown in Figure 9 can be interpreted as our prediction of what a real experiment of this kind would look like in real life. We can see in that experiment that the measures produced by the each single previous experiment is preserved.

We have also interrogated our model in different ways. We have tried to understand where there are other factors that induce a change of distribution over the lysosome and recycling. As can be expected, if there is a variation in concentration or size of in either the lysosome or the recycle, then the proportion of receptors that gets recycled or degraded in the lysosome varies. This is shown in Figure 10, which reports the result of an experiment

$$
\begin{aligned}
\textsc{fgf} &= \mathbf{p2c}\,\overline{\text{fgfbind}}\langle bind\rangle \\
\mathsf{L} &= \mathbf{accept}\,lyso.\mathsf{L} \\
\textsc{lyso} &= [\underbrace{\mathsf{L} \mid \ldots \mid \mathsf{L}}_{25}] \\
\textsc{en} &= \mathbf{accept}\,endo1.\textsc{en} + \mathbf{expel}\,endo2.\textsc{en} \\
\textsc{endo} &= [\underbrace{\textsc{en} \mid \ldots \mid \textsc{en}}_{45}] \\
\textsc{le} &= \mathbf{accept}\,lendo1.\textsc{le} + \mathbf{expel}\,lendo2.\textsc{le} \\
\textsc{latendo} &= [\underbrace{\textsc{le} \mid \ldots \mid \textsc{le}}_{40}] \\
\mathsf{R} &= \mathbf{accept}\,recycle1.\mathsf{R} + \mathbf{expel}\,recycle2.\mathsf{R} \\
\textsc{recycle} &= [\underbrace{\mathsf{R} \mid \ldots \mid \mathsf{R}}_{35}] \\
\mathsf{C} &= \mathbf{p2c}\,\text{fgfbind}(x).\mathbf{enter}\,endo1.\mathbf{exit}\,endo2.\mathbf{enter}\,lendo1.\mathsf{C}1 \\
\mathsf{C}1 &= \mathbf{exit}\,endo2.(\mathbf{enter}\,lyso + \mathbf{enter}\,recycle1.\mathbf{exit}\,recycle2.\mathsf{C}) \\
\textsc{fgfr} &= [\mathsf{C}] \\
\textsc{cell} &= \textsc{endo} \mid \textsc{latendo} \mid \textsc{lyso} \mid \textsc{recycle} \mid \underbrace{\textsc{fgfr} \mid \ldots \mid \textsc{fgfr}}_{1300} \\
\textsc{System} &= \underbrace{\textsc{fgf} \mid \ldots \mid \textsc{fgf}}_{1000} \mid \textsc{cell}
\end{aligned}
$$

Fig. 4. Implementation of endocytosis of FGFin BioAmbients

where we have tripled the volume of lysosome. Clearly, the fate of the FGFR4 changes quite dramatically.

It should be appreciated that our experiments involve only a finite number of biological components and run for a limited period of time, as the experiment reported in [9,8]. We wanted to see what happens in case both the FGF and FGFR get reproduced by the cell. We have implemented reproduction in such a way that a new FGF is immediately created after the binding, while each receptor is created after a small delay. Each receptor is created at a different rate. We ran the model with four hundred FGF and one thousand and seven hundred receptors. The result of this is shown in Figure 11.

The model shows that in the very long run most of FGFR4 end in the cell membrane. Because the rates for the creation of FGFR4 are not known, our model may not be realistic, however it sheds light on the fact that with a specific set of rates it is possible have an over-production of FGFR4. Further sensitivity analysis is necessary to document different realistic scenarios

5 Conclusions

In this paper we have carefully studied the FGF endocytotic pathway both from the literature [15] and from the research paper [8]. We have simulated the results of some of the experiments in [8]. We have chosen to simulate the experiments that highlight the different

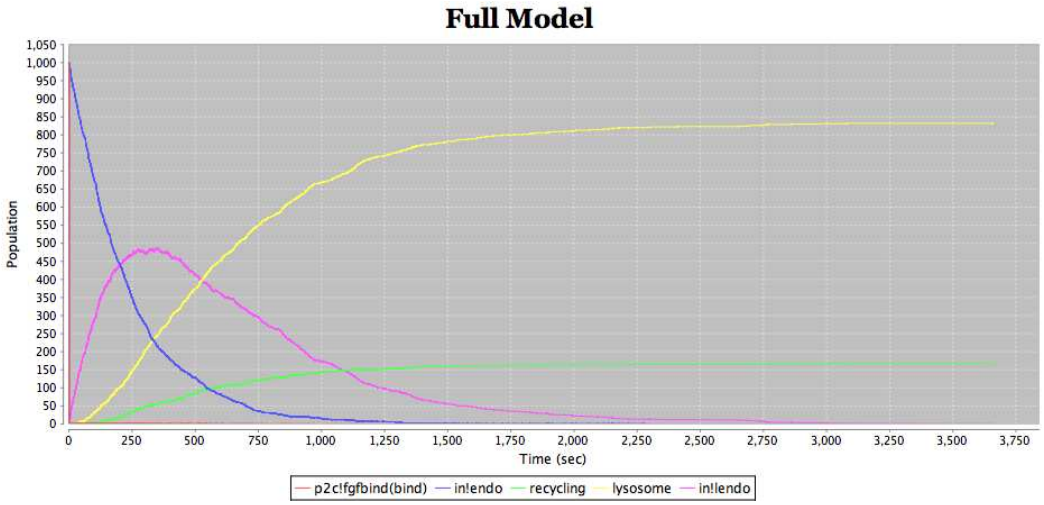

Fig. 5. Full model

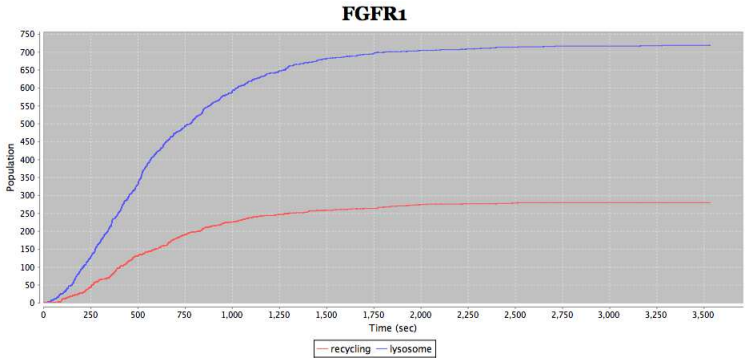

Fig. 6. Description of the fate of FGFR1

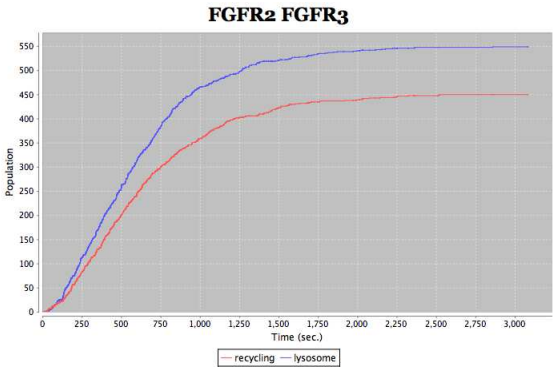

Fig. 7. Description of the fate of FGFR2-FGFR3

fate among the FGFRs; in this paper we have simulated the trafficking on each receptor individually. Our results are consistent with those in the literature. We have run an 'in silico ' experiment with all four receptors at once. We found that sorting of the different receptors were preserved in the full model. We have further interrogated the model to see what are the factors that could determine a different distribution in the sorting. We found that, if we assume a substantial higher concentration -or size- of the lysosome, this could change quite dramatically the result of the sorting. We also found that changing the size of the early or

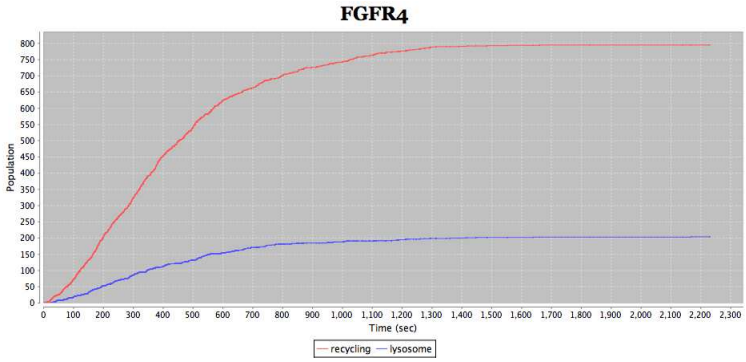

Fig. 8. Description of the fate of FGFR4

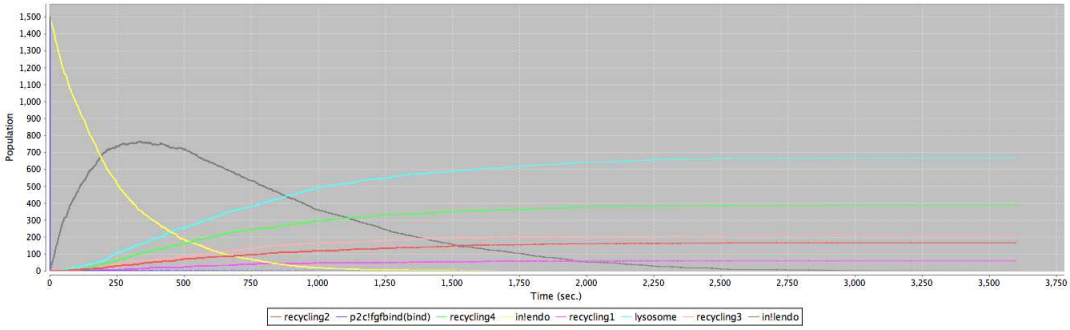

Fig. 9. Full model with four receptors

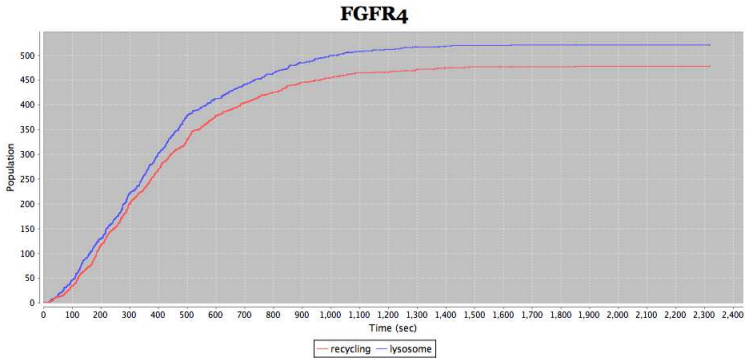

Fig. 10. Fate of FGFR4 with very large lysosome

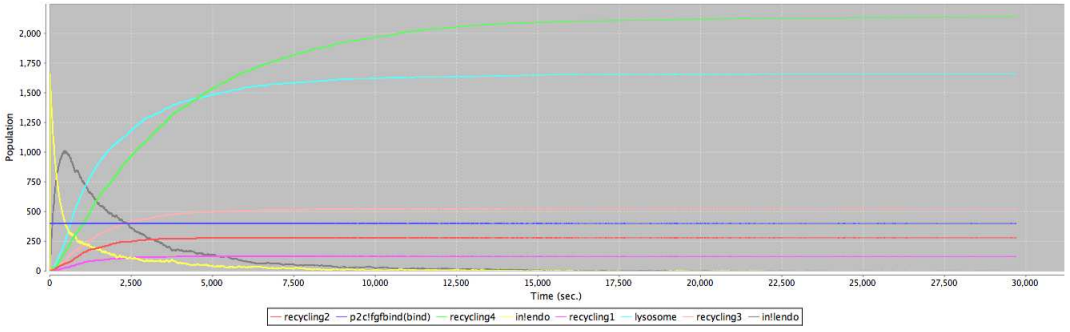

Fig. 11. Full model with increasing population of FGFR and FGF

late endosome does not affect the sorting of the receptors. This result is to some extend to be expected. By looking at the Markov Chain generated by our specifications, it can be seen that changing the size of the lysosome has an impact on the probability of being routed in this compartment. By simply looking at the results in the experiments as described in [8] this could have been both difficult to guess or expensive to verify with an experiment. In the literature there has already been an very successful attempt to model the FGF signalling pathways i.e. the early stages of FGF signal propagation and internalisation using probabilistic model checking PRISM [10]. That work concentrated mostly on modelling chemical reactions and in this respect their purpose is orthogonal to ours. It remains to be seen whether our model could be refined in the long term, in order to model both compartments and chemical reactions, in which case the work carried out in [11] using π-calculus could be directly integrated with ours.

Acknowledgements

We gratefully acknowledge M. Kwiatkowska, G. Norman, and D. Parker for very valuable contributions to the initial discussions on the topic of the present article. We gratefully acknowledge Vinod Muganthan and Andrew Phillips for sharing the enterprise of developing BAM [17]. We finally thank Luca Cardelli for very useful discussions on the topics of this article and on system biology in general.

References

[1] C. Baier, B. Haverkort, H. Hermanns, and J.-P. Katoen. Model-checking algorithms for continuous-time Markov chains. *IEEE Transactions on Software Engineering*, 29(6):524–541, 2003.

[2] Muffy Calder, Stephen Gilmore, and Jane Hillston. Modelling the influence of RKIP on the ERK signalling pathway using the stochastic process algebra PEPA. In Anna Ingolfsdottir and Hanne Riis Nielson, editors, *Proceedings of the BioConcur Workshop on Concurrent Models in Molecular Biology*, London, England, August 2004.

[3] Muffy Calder, Stephen Gilmore, and Jane Hillston. Automatically deriving ODEs from process algebra models of signalling pathways. In Gordon Plotkin, editor, *Proceedings of Computational Methods in Systems Biology (CMSB 2005)*, pages 204–215, Edinburgh, Scotland, 2005.

[4] L. Cardelli. Abstract machine for system biology. *Transactions on Computational Systems Biology.*, 3737:145–168, 2005.

[5] L. Cardelli and A.D. Gordon. Mobile ambients. *Theoretical Computer Science*, 240(1):177–213, 2000.

[6] Vincent Danos and Cosimo Laneve. Formal molecular biology. *Theor. Comput. Sci.*, 325(1):69–110, 2004.

[7] Daniel T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81(25):2340–2361, 1977.

[8] E. M. Haugsten, V. Sørensen, A. Brech, S. Olsnes, and J. Wesche. Different intracellular trafficking of fgf1 endocytosed by the four homologous fgf receptors. *ournal of Cell Science*, 118:3869–3881, 2005.

[9] Ellen Margrethe Haugsten. Intracellular trafficking of fgf1 endocytosed by its four tyrosine kinase receptors. Master's thesis, Department of Molecular Biosciences, University of Oslo, 2004.

[10] J. Heath, M. Kwiatkowska, G. Norman, D. Parker, and O. Tymchyshyn. Probabilistic model checking of complex biological pathways. In C. Priami, editor, *Proc. Computational Methods in Systems Biology (CMSB'06)*, volume 4210 of *Lecture Notes in Bioinformatics*, pages 32–47. Springer Verlag, 2006.

[11] J. Heath, M. Kwiatkowska, G. Norman, D. Parker, and O. Tymchyshyn. Probabilistic model checking of complex biological pathways. *Theoretical Computer Science*,319(3):239-257, 2008.

[12] J. Hillston. *A Compositional Approach to Perfomance Modelling*. PhD thesis, Department of Computer Science, Edinburgh, 1994.

[13] Marta Kwiatkowska, Gethin Norman, David Parker, Oksana Tymchyshyn, John Heath, and Eamonn Gaffney. Simulation and verification for computational modelling of signalling pathways. In *WSC '06: Proceedings of the 38th conference on Winter simulation*, pages 1666–1674. Winter Simulation Conference, 2006.

[14] Paola Lecca and Corrado Priami. Cell cycle control in eukaryotes: A biospi model. *Electron. Notes Theor. Comput. Sci.*, 180(3):51–63, 2007.

[15] H. Lodish, A. Berk, P. Matsudaira, C. A. Kaiser, M. Krieger, M.P. Scott, S. L. Zipursky, and J. Darnell. *Molecular Cell Biology*. W.H. Freeman and Company, 2003.

[16] M. Miaczynska, L. Pelkmans, and M. Zerial. Not just a sink: endosomes in control of signal transduction. *Curr. Opin. Cell.*, 16:400–4006, 2004.

[17] Vinod A. Muganthan, A. Phillips, and M.G. Vigliotti. Bioambient machine. In *Proceedings of 8th International Conference on Application of Concurrency to System Design 2008*, 2008. To appearI.

[18] C. Priami. Stochastic π-Calculus. *The Computer Journal*, 38(7):579–589, 1995.

[19] Corrado Priami, Aviv Regev, Ehud Shapiro, and William Silverman. Application of a stochastic name-passing calculus to representation and simulation of molecular processes. *Inf. Process. Lett.*, 80(1):25–31, 2001.

[20] Andrew Phillips Ralf Blossey, Luca Cardelli. A compositional approach to the stochastic dynamics of gene networks. *Transactions on Computational Systems Biology IV*, 3939:99–122, 2006.

[21] Aviv Regev, Ekaterina M. Panina, William Silverman, Luca Cardelli, and Ehud Y. Shapiro. Bioambients: an abstraction for biological compartments. *Theoretical Computer Science*, 325(1):141–167, 2004.

[22] Aviv Regev and Ehud Y. Shapiro. Cellular abstractions: Cells as computation. *Nature*, 419(343), 2002.

[23] E. Sandilands, S. Akbarzadeh, A. Vecchione, D.G. McEwan, M.C. Frame, and J.K.Heath. Src kinase modulates the activation, transport and signalling dynamincs of fibroblast growth factor receptors. *Embo Reports*, 2007.

[24] M. Vigliotti and P. Harrison. Stochastic mobile ambients. In A. Di Pierro and H. Wiklicky, editors, *Proc. 4th Int. Workshop Quantitative Aspects of Programming Languages (QAPL 2006)*, volume 164 (issue 3) of *Electronic Notes in Theoretical Computer Science*, pages 169–186. Elsevier, 2006.

[25] Darren J. Wilkinson. *Stochastic Modelling for System Biology*. Chapman & Hall, 2006.

Appendix A

Description of experiments carried out in [8]

5.1 Experimental Data

Experiment 2

Purpose: Determine where FGF locates to once it has bound with FGFR.

Result: The double staining in the experiment showed that both ligand and receptor ended up in the same compartment showing that they remain bound as a complex once they have become internalised into the cell.

Experiment 3

Purpose: This experiment was carried out to follow the endocytotic pathway and identify which compartment FGF1 locates to after internalisation.

Result: This experiment shows that when FGF1 bound with any one of the four FGFR1-FGFR4 receptors they all went to the early/sorting endosomal compartment after internalisation.

Experiment 4

Purpose: This experiment was carried out to find out where the FGF:FGFR complex moves to after entering the early/sorting endosomes.

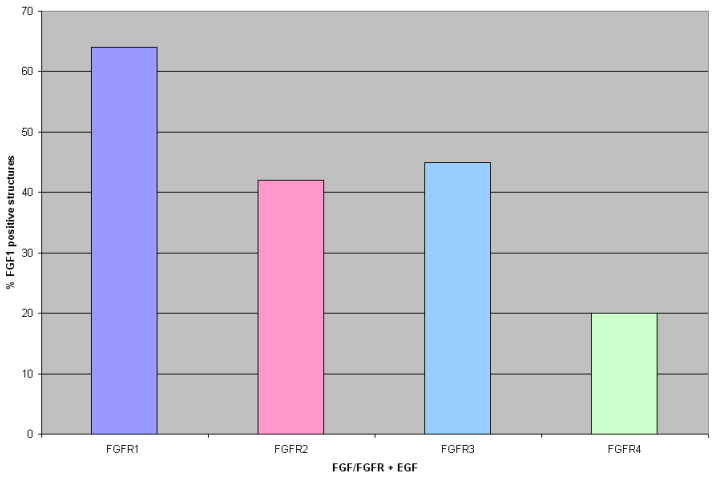

Fig. 12. Percentage of FGF:FGFR routed to lysosomes

Result: After two hours in the presence it was shown that the major part of FGF1 internalised with FGFR1-FGFR3 went to late endosomes. Figure 13 shows the different percentages which were routed to late endosomes for the different receptors after 2 hours. It shows that 90% FGF1-FGFR1 were LAMP-1 positive whereas only 45% were LAMP-1 positive for FGFR4. In the case of FGFR2 and FGFR3 about 70% were LAMP-1 positive.

Experiment 5

Purpose: This experiment was carried out to to determine where the FGF1:FGFR4 complex localised to after late endosomes.

Method: FGF1 was labelled with the fluorescent dye, and EGF (epidermal growth factor) and transferrin were labelled with different colour dyes. From previous experiments it is known that EGF moves to the lysosomes and transferrin to the endosomal recycling compartment (ERC) from the late endosomes. By seeing the overlaps in the dyes of the FGF1 and the transferrin or EGF it could be calculated which compartment the different receptors move to. The experiments were left to run for two hours.

Result: After two hours it was found the majority of FGF1 -FGFR4 complex moved to the endosomal recycling compartment. Figure 12 shows the percentage of each of the receptors which were routed to lysosomes for degradation. We can see that the majority the majority of FGF1- FGFR1-3 complex moves to the lysosomes and only a small amount of FGFR4 does. Figure 13 shows the percentage of each of the receptors which were routed to the recycling compartment to be returned to the surface. We can see that the majority the majority of FGF1-FGFR4 complex moves to the (ERC) and only a small amount of FGFR1-FGF3 does.

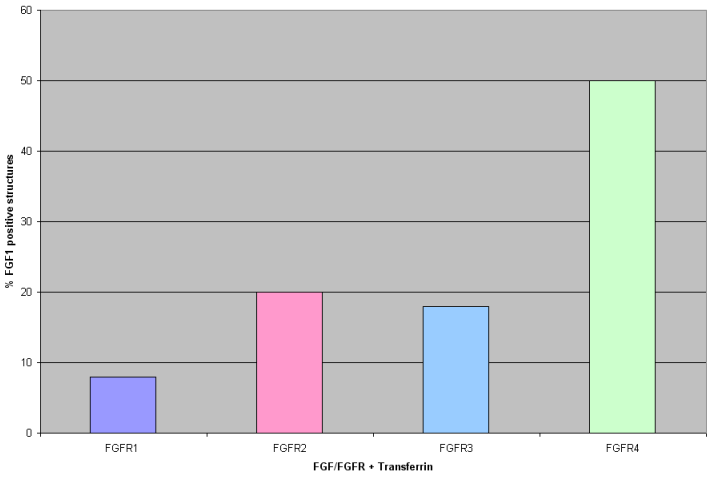

Fig. 13. Percentage of FGF:FGFR routed to the endosomal recycling compartment