The University of Reading

Computer modelling of early Focal Adhesion dynamics

Thesis submitted for the degree of Doctor of Philosophy

(Ph.D.)

By

Jonathan Rudge

School of Biological Sciences

July 2018

Declaration of original authorship

Declaration: I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Signature: ……

Name: Jonathan Rudge

Date: ……

Table of Contents

Table of Tables

Table of Figures

Abstract

Focal adhesions are plasma membrane-associated protein complexes that connect the cellular cytoskeleton to the extracellular matrix via discrete “focal” points on the plasma membrane. As such they are critical to the movement of cells and therefore of great interest in the study of organismal development, immunity and inflammation, and in cancer metastasis. Therefore it is critical to better understand the structure of focal adhesions and the underlying processes behind their assembly and disassembly, as well as how assembly and disassembly are regulated. Modelling these processes mathematically, using computers, can offer insights not available to other approaches. Here differential equations (both ordinary and partial) are used to model various aspects of focal adhesion dynamics to produce compartmental and spatial models, respectively. By these methods new insights have been derived, unveiling the huge complexity of focal adhesion dynamics, but also pointing to the key factors that principally determine overall rates of focal adhesion assembly and disassembly, as well as levels of individual isoforms.

Overall, the key finding of this project is that, for focal adhesion dynamics to occur over a physiologically realistic timescale (typically 5-20 minutes), numbers of focal adhesion-associated proteins must be highly enriched in the immediate vicinity of where such dynamics are taking place. This implies a higher level of organisation in both the plasma membrane and cytosol than is usually assumed in reaction-diffusion models such as this. Only by assuming that concentrations of such proteins are much higher locally than they would be if uniformly distributed throughout these and other relevant cellular compartments, can the required interactions between them occur at a rate consistent with what is observed physiologically. This finding is likely to extend to many other cell-critical behaviours.

Keywords: Focal adhesions, mathematical modelling, computational modelling, differential equations, ODEs, PDEs, FAs, dynamics, modelling, lamellipodia, outside-in signalling, inside-out signalling, compartmental model, spatial model, reaction-diffusion system.

Chapter 1. Introduction and aims of the project

Before describing the aims of the project, it is necessary to explain what focal adhesions (FAs) are, what cellular role they perform and how they perform this role in terms of their underlying structure and dynamics. By "dynamics", in this context, is meant changes in overall FA structure in response to various events (including various forms of external stimulation, and changes within FAs themselves, such as component protein phosphorylations and conformational changes) and how these changes are regulated.

Since this project also involves modelling an important aspect of FA dynamics, it is also first necessary to explain what exactly mathematical modelling is, what forms it can take (and therefore what options are available for modelling FA dynamics) and how it can be used to advantage in the study of cell behaviour, including FA-associated behaviour.

Having outlined the above, this chapter will conclude with a description of the overall aims of the project, together with a brief summary of how it is intended to achieve these aims. Subsequent chapters (Chapters 2, 3 and 4) will describe the three key phases of the model development project. Each chapter will include a description of the context of the problem, model design considerations and subsequent model implementation details. A further description of the results from running the model will be followed by a discussion of these results and any insights arising from them.

The last chapter (Chapter 5) will summarise these results and insights, evaluate the model and any lessons learnt from developing it, and suggest future improvements and extensions of it. This will be followed by a discussion concerning what the writer has learned about the modelling process in general. This will include issues that have arisen that are typical of what modellers encounter generally, why this might be and what needs to change so that the quality of models

can improve. Finally, there will be a brief commentary on the current status of mathematical modelling within cell biology and why it matters.

1.1 FAs

1.1.1 FAs in context

Before describing FAs in detail it is helpful to understand the wider cellular and extracellular context in which they develop and function. For the most part it can be said that the majority of cells within complex organisms (including humans and other animals) do not exist surrounded by empty space. Typically such cells will be organised into various tissues and organs, meaning that most are closely-bound to other cells. Any non-cellular regions in the immediate vicinity of such tissues and organs will often consist of a protein-rich complex, called the extracellular matrix (ECM), laid down by specialist cells called fibroblasts, as well as by cells in the immediately adjacent tissues[1].

Whilst the composition of ECM varies widely, discounting proteoglycans (which, in hydrated form, provide the gel-like matrix characteristic of ECM [2]), it tends to be dominated by just a few fibrous protein types, four of which, elastins , collagens, fibronectin and laminin are most commonly seen [1]. So, for instance, one form of the ECM, called the basement membrane (BM, a thin layer of ECM that separates sheets of epithelial cells from the largely acellular stroma, whilst also providing a form of scaffolding), tends to be rich in laminin and collagen IV [3,4], whereas interstitial matrix or stroma (more generally known as connective tissue, and providing more general support for the organ or tissue [4]) typically contains more of the other collagen isoforms, as well as fibronectin and elastin [4,4].

The widespread presence of ECM provides both a means and a barrier for cell locomotion. Given that locomotion is a primary requirement of such important biological processes as organismal

development, inflammation and immunity, and cancer metastasis (describing the process in which tumour-originating cells spread to other parts of the affected body), the means by which cells engage with the ECM and move through it has been of great interest to cell biologists, as well as specialists investigating these particular aspects of biology.

Although cells(other than flagellated cells, such as spermatozoa) are known to be capable of an amoeboid form of locomotion by blebbing[5], it is thought that they primarily move by engaging collagens, fibronectin and laminin (as well as less common proteins, such as vitronectin[1]) of the ECM. This allows them sufficient traction to pull themselves in a particular direction, commonly determined by a chemical gradient, such as provided by a growth factor or chemokine.

For instance, it is clear that metastasis requires substantial digestion of ECM and BM proteins by proteolytic cleavage, principally by secretion of matrix metalloproteinases (MMPs)[6]. However, it is also clear that many of the ECM and BM proteins must remain in order for metastatic cells to gain traction. Similar considerations must obviously also apply to inflammation, immunity and development, wherever existing ECM lies between a cell and its intended destination. Given the widespread presence of ECM there has been growing interest in FAs, protein complexes whose primary role is to engage most of the major ECM proteins, linking them to the cell's contractile actin- and myosin-based cytoskeleton, thus providing traction. These are described in the next section.

1.1.2 Overall focal adhesion structure

As stated, FAs are plasma membrane-associated protein complexes that principally connect the cellular cytoskeleton to the ECM via discrete "focal" points on the plasma membrane (PM), giving them an important role in cell anchorage, as well as migration. FAs are highly dynamic in structure, ranging from small structures known as focal complexes (composed of integrins and a

few other associated proteins such as paxillin and talin) up to fully mature multi-molecular adhesions, made up of multiple cell types and bound to stress fibres, which are themselves bound within the cellular cytoskeleton [7](Figure 1.1).

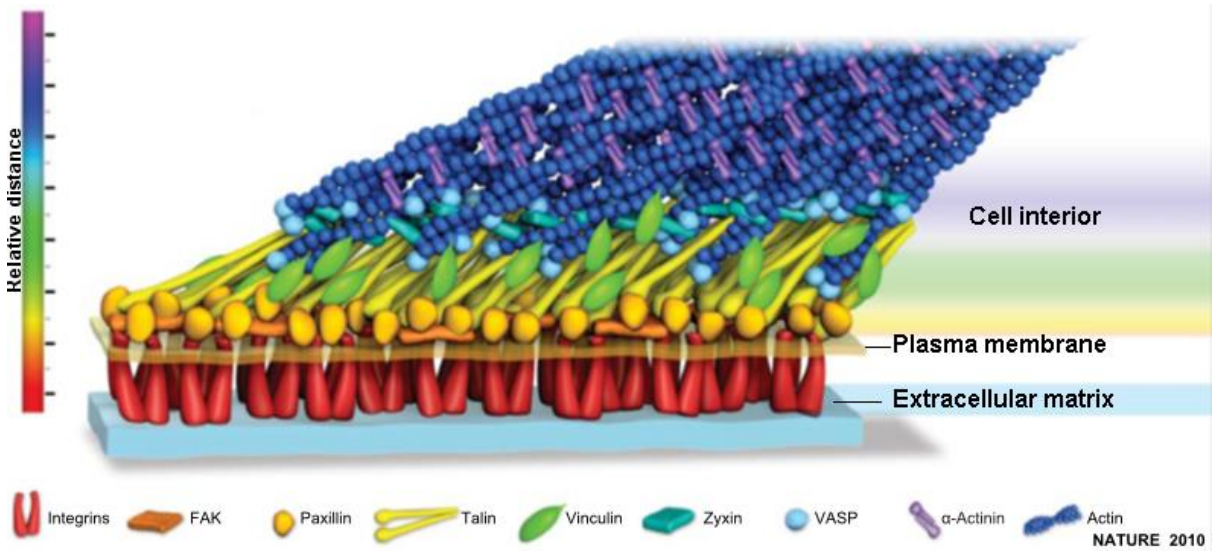

Figure 1.1.A hypothetical representation of a mature FA with principal component proteins. As determined by three-dimensional super-resolution fluorescence microscopy. (No stoichiometry implied). Source: [8].

Such mature FAs are thus found in abundance within quiescent cells that occupy a fixed position, especially those bound together within tissues, such as epithelial cells. By contrast more motile cells, such as mesenchymal-derived fibroblasts and mesenchymal-like metastatic cells, are characterised by the dynamic formation and breakdown of less mature FA structures, most of which fail to progress beyond the focal complex stage, sufficient to bind local elements of the actin cytoskeleton and provide short-term traction, but not positional stability[9,10].

More than 160 proteins have now been reliably associated with mature FAs[11]. This includes transmembrane receptors (principally the integrins and growth factor receptors), adaptor (sometimes referred to as scaffold) proteins such as paxillin and vinculin, and kinases such asp21-activated kinases (PAKs [12]) and Rho kinase (also known as Rho-associated kinase or ROCK1)[13].

Some FA-associated proteins combine several functions. For instance, focal adhesion kinase (FAK) functions both as a kinase and as an adapter protein[14].

Current understanding is that FAs typically begin with binding events at the cell membrane, between integrins and their external ligands, triggering the formation of "nascent adhesions" around the cytoplasmic tail of integrins[15]. These nascent adhesions are believed to be recruited from a few different proteins, many of which are rapidly and extensively phosphorylated after recruitment[16]. Nascent adhesions are known to include paxillin [17]but may include other proteins, such as FAK, Src and talin[8,15], known to be early components, at least in mature adhesions. For this reason it makes sense to describe each of these proteins in turn, before looking at other proteins that are known, or believed, to be associated with more mature forms of FAs. Clearly no model can incorporate all the hundred or so proteins known to be involved in FA formation and it makes sense to start at the beginning when there are thought to be just a handful involved.

1.1.3 Integrins

The integrins are a family of heterodimeric transmembrane glycoproteins, normally found concentrated in FAs, i.e. at discrete sites on the plasma membrane[18]. As heterodimers they are composed of two distinct polypeptides, referred to as the alpha and beta subunits (with 18 alpha and 8 beta subunits distinctly identified in mammals [19]), which differ in their domain composition (Figure 1.2)[19,20]. This difference is most obvious in their N-terminal region. Here, the beta subunit has an (insertion or interaction) domain(also referred to as an A or I/A domain), dominated by a divalent metal ion-coordinating site, which binds highly specific ligands, including ECM proteins, such as fibronectin and collagen, and membrane-bound proteins of the immunoglobulin superfamily, such as ICAM-1[19,20] (Table 1.1). Although some alpha subunits also contain an I-domain(which dominates ligand-binding in any heterodimers in which it appears[19]) most do not,

their N-terminal regions being dominated by a so-called propeller domain, a seven-bladed structure that contributes to ligand-binding[19,21](Figure 1.2).

Ligand	Integrins	Notes
Collagens	α1β1, α2β1, α10β1, α11β1	Principally receptors for collagens, integrins α1β1 and α2β1 may bind other ligands with a lesser affinity.
Laminins	α3β1, α6β1, α6β4, α7β1	Principally receptors for laminins, integrin α3β1 may bind other ligands with a lesser affinity.
'RGD'-containing proteins	α5β1, α8β1, αIIbβ3, αvβ1, αvβ3, αvβ5, αvβ6, αvβ8	Each integrin binds its own subset of the RGD-containing ligands (which include fibronectin, fibrin and vitronectin). Some bind additional ligands in an RGD-independent manner.
Immunoglobulin superfamily proteins	αDβ2, αEβ7, αLβ2, αMβ2, αXβ2, α4β1, α4β7, α9β1	In addition to specific IGSF ligands (VCAM, ICAMs), many of these integrins bind additional ligands. For example, integrin α4β1 binds to fibronectin in an RGD-independent manner, while αMβ2 and αXβ2 bind fibrin.

Table 1.1. Primary ligands of integrins of different alpha-and beta-subunit composition. "RGD" refers to arginine-glycine-aspartic acid amino acid sequences, found in various proteins that form a class of integrin ligands. Table features all 18 alpha- and all 8 beta-integrins identified in mammals. Source: [22].

At the other end of integrins, both alpha and beta subunits have short cytoplasmic "tails". These contain binding sites that are critical to integrin functioning, by binding to kinases and actin-associated scaffold proteins, amongst others [19,21].

Linking these important N-and C-terminal regions are long, mostly linear, "legs", made up of various domains, including, in beta subunits, 4 repeat EGF domains, all rich in cysteine residues[19,20,21], together with a hybrid and PSI (plexin-semaphorin-integrin) domain [24]. Aside from connecting the N- and C-terminal regulatory regions, it is thought that these transmembrane leg regions help to transmit information, in an outside-in and inside-out manner, by adopting a straight or bent conformation [19,20,24,25] (Figure 1.3). Thus, following external ligation, both alpha and beta subunits are believed to adopt an upright conformation, which forces apart the cytoplasmic tail regions of both subunits, making it easier for intracellular binding partners, such as talin and FAK to bind to these regions [19,20].

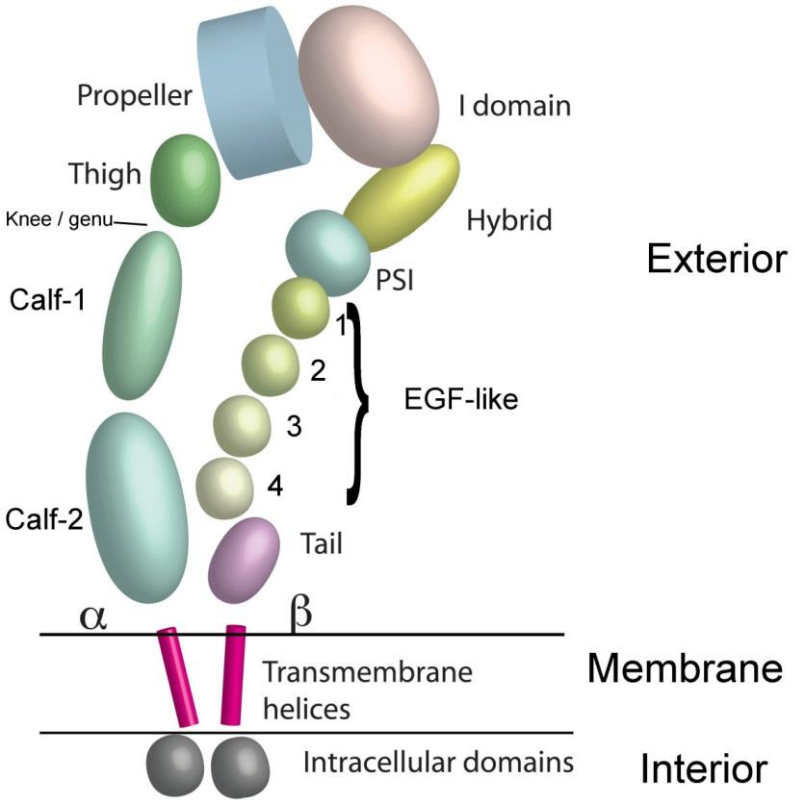

Figure 1.2.Typical domain structure of integrins. Key: I-domain = insertion/interaction domain; PSI = plexin-semaphorin-integrin domain; EGF-like = epidermal growth factor-like domain. Some alpha subunits have an I-domain above the propeller domain, providing the principal ligand-binding site for the integrin dimer. Figure shows integrin dimer in its typical fully-activated, extended form. In inactive form the alpha subunit bends at the Knee (often referred to as the ”genu”) domain, whilst the beta subunit bends at the second and third EGF-like domains. Adapted from [23].

Similarly, binding by such intracellular ligands stabilises the separation of the subunit tail regions, which is thought to also stabilise the external integrin structure in an upright position, in which it has a much higher affinity for its extracellular ligand[19,21]. Thus external ligation promotes internal ligation (“outside-in signaling”) and vice-versa (“inside-out signalling”, Figure1.3)[19,20,25].

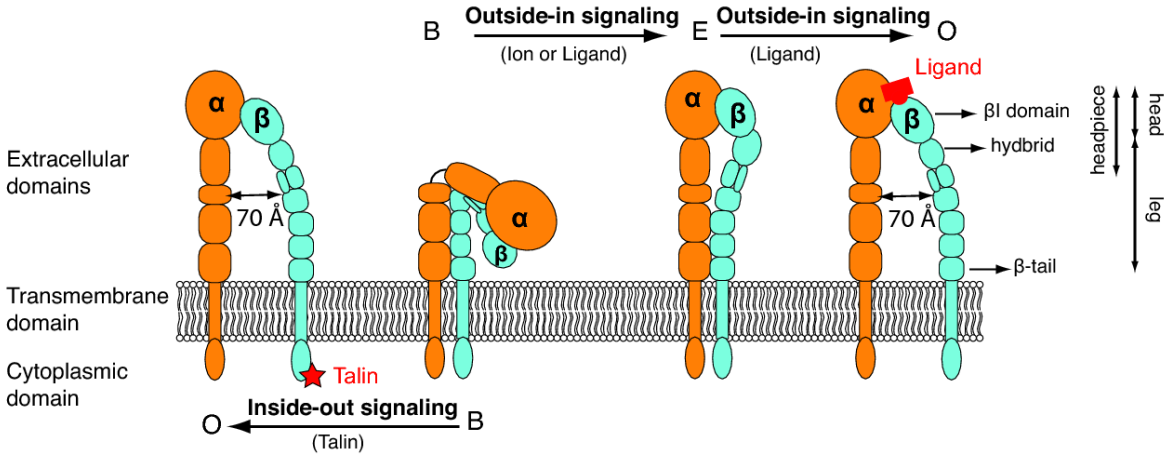

Figure 1.3. Integrin inside-out and outside-in signaling. In inside-out signaling, binding of talin (driven by exposure to negatively charged phospholipids) and other proteins to integrin cytoplasmic tails results in tail separation and adoption of an upright conformation by integrin dimer, with resulting high affinity for external ligands. In outside-in signaling, external ligation leads to the same tail separation and upright conformation. Source:[25]

However, the reality is rather murkier than the above explanation suggests, as this model is still in need of further experimental support and there remain questions as to whether the differences in extracellular ligand affinity of the bent and straight conformations are as great as claimed[19,25]. Furthermore, there are still a number of unanswered questions as to how some of the more important FA proteins, such as talin, FAK, alpha-actinin and paxillin bind to the cytoplasmic tail region, what their binding affinities are, and in what order they tend to bind.

1.1.4 Talin

Talin is a high-molecular-weight (235-kDa) scaffold protein whose principle role appears to be as a link between focal complexes and the actin cytoskeleton, which it is able to achieve through a combination of binding sites (Figure 1.4) [26]. Principal among these is an N-terminal FERM (Band 4.1 homology ezrin radixin moesin) domain[27,28], which binds the cytoplasmic tail of integrin beta subunits, thereby helping to stabilise integrin heterodimers in the high-affinity ligand-binding conformation[27,29] (although the same FERM domain can also bind actin instead). Further down (in

the so-called C-terminal rod domain) talin contains a mixture of F-actin-, integrin- and vinculin-binding sites[26,27,30] (Figure 1.4), vinculin being another scaffold protein, which as well as binding talin, contains a number of actin binding sites[31].

However, both integrin and vinculin/actin binding require talin to adopt favourable conformations. Firstly, integrin-binding requires talin to be in an open, "active", conformation, as opposed to the default closed ("inactive") conformation, in which parts of the C-terminal rod domain are folded back on the N-terminal region, masking the FERM domain[32,33]. This can be relieved by exposure to negatively charged phospholipids, especially phosphatidylinositol 4,5-bisphosphate (PIP2)[33,34]. In this respect it is interesting to note that talin FERM domains also bind PIP5K1C, a kinase (of the P15K family) which catalyses PIP_2 formation, and that high-affinity binding of the talin FERM domain to the integrin beta tail requires that the heterodimer is bound within negatively-charged phospholipids[27,33,35,36].

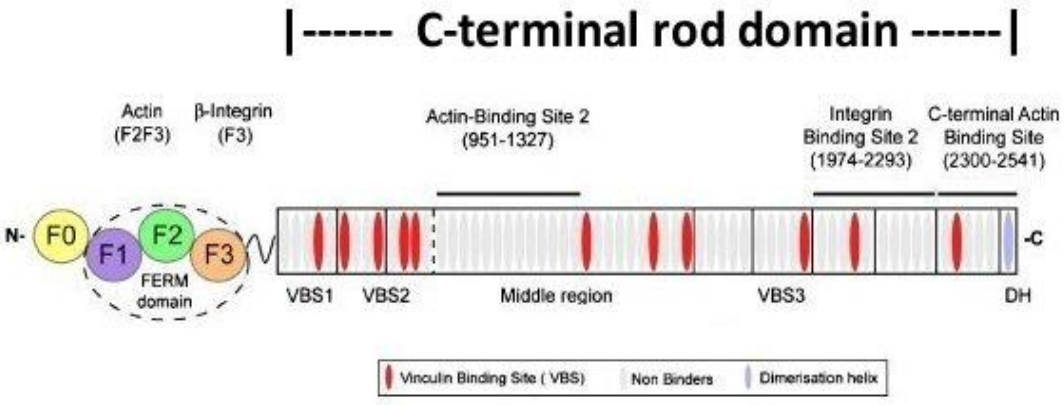

Figure 1.4.Talin domain structure. Key: FERM = Band 4.1 homology ezrin radixin moesin domain; VBS = vinculin-binding site; DH = dimerisation helix. Talin shown in open, active conformation in which the FERM domain and all vinculin and actin binding sites are exposed and thus available for binding. Dimerisation helices of two talin monomers form an anti-parallel dimer. Adapted from[32].

Similarly, the closed talin conformation hides critical vinculin-binding sites, which, it is thought, require mechanical stretching to be exposed[27,30]. Thus talin-mediated FA maturation appears to

require both specific local membrane conditions and the exertion of mechanical force (exerted either by extracellular components or by stress fibres, or both) in order to occur.

1.1.5 FAK

Another key protein in FA formation is FAK, which shares many features with talin. This arises from the fact that it also has an FERM domain (Figure 1.5) that is known to be able to bind to the same cytoplasmic region of the integrin beta tail, as well as to several growth factor tyrosine kinase receptors[37].

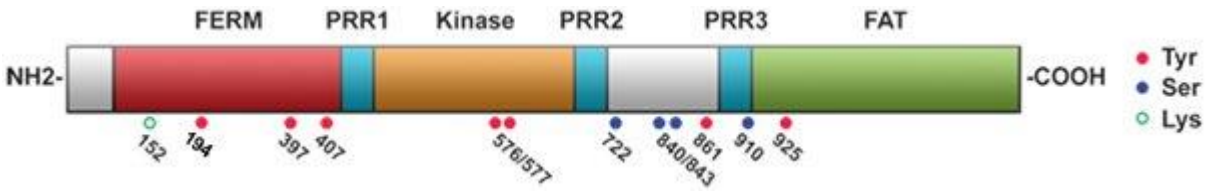

Figure 1.5. Focal adhesion kinase (FAK) domain structure with principal phosphorylation/sumoylation sites. Key: FERM = Band 4.1 homology ezrin radixin moesin domain; PRR1-3 = Proline-Rich Regions; FAT = Focal Adhesion Targeting domain; Tyr = tyrosine residue (a phosphorylation site); Ser = serine residue (a phosphorylation site); Lys = lysine residue (a sumoylation site). Adapted from [38].

However, whereas talin is predominantly a scaffold protein, holding other proteins together in a complex, FAK also has a critical role as a protein tyrosine kinase and can, less commonly, function as a transcription factor[37].

As with talin, FAK has a folded default conformation, which prevents its FERM domain from binding to integrin beta-subunits tails[39], and this can also be relieved by exposure to the phospholipid PIP_2[40], as described earlier, principally found in the inner leaflet of the plasma membrane. However, this autoinhibitory conformation can also be relieved by phosphorylation of the FAK Tyr-194 residue by growth factor receptor tyrosine kinases such as Met[37,41], or by sumoylation of the Lys-152 residue[42], or it may require both together. Relief of the autoinhibitory

conformation, as well as exposing the FERM domain, also exposes FAK's kinase domain, allowing it to phosphorylate other FA-related proteins nearby, such as paxillin and p130^{Cas}[43,44]. This autoinhibitory relief also frees the kinase domain to autophosphorylate FAK itself, via its Tyr-397 residue, thus activating FAK[39].

FAK autophosphorylation leads to the recruitment of Src and other Src-family kinases, which bind to the phosphorylated Tyr-397, via their SH2 domains[45], as does PI3 kinase [46], with Src binding leading to further phosphorylation of FAK[45]. Amongst the FAK tyrosine residues phosphorylated by Src are Tyr-576 and-577, both necessary for full FAK activation [45], andTyr-861, required for association of FAK with p130^{CAS}[47]. Also phosphorylated by Src is Tyr-925, this phosphorylation being necessary for FAK association with Grb2[48,49] and dynamin[48,50], as well as being associated with increased levels of phosphorylated (as opposed to unphosphorylated) paxillin, enhanced nascent adhesion formation, greater FA turnover and higher rates of cell migration[51].

Additional FAK scaffolding functionality, provided in the form of three proline-rich regions (PRR1-3, Figure 1.5) allows the recruitment of other proteins, principally via SH3 domains[52]. One example of this is endophilin A2[37].Finally, the FAT domain of FAK (Figure 1.5) is known to provide additional binding sites for proteins, of which the best-known are paxillin [53](whose significance in FA formation is outlined later), talin [54] (which possibly requires initial transport by FAK to the vicinity of nascent adhesions before binding directly to integrin tails) and VEGFR3 [55], a growth factor receptor known to be involved in cell survival, lymphatic angiogenesis and metastasis. In common with the previously mentioned phosphotyrosine sites, the principal regulatory influence of such binding sites results from how they encourage phosphorylation by recruiting target proteins to the growing kinase nexus located at maturing FAs[56].

Whereas tyrosine phosphorylation promotes both the kinase and scaffold roles of FAK, thus directly affecting FA dynamics, other forms of post-translational modification depend upon trans-

locating FAK for their regulatory effect. Thus, serine phosphorylation at Ser-732 by Cdk5 translocates FAK to the centrosome, where it appears to regulate microtubule organisation[57,58], whilst sumoylation (whether at Lys-152, or elsewhere)[42] translocates FAK to the nucleus, where it enhances cell proliferation and reduces inflammation by promoting the polyubiquitination and subsequent degradation of (respectively) p53 and GATA4, as well as regulating general gene transcriptional activity[59,60].

1.1.6 Src

Src is a non-receptor tyrosine kinase (as well as the name for the family of at least 11 Src-related kinases[61]), in which overactivity has been linked to various forms of cancer-associated behaviour, including proliferation, survival, angiogenesis and invasiveness [61,62,63]. In particular, it is known to have an early and important role in FA formation, amongst other things determining adhesion strength [63].

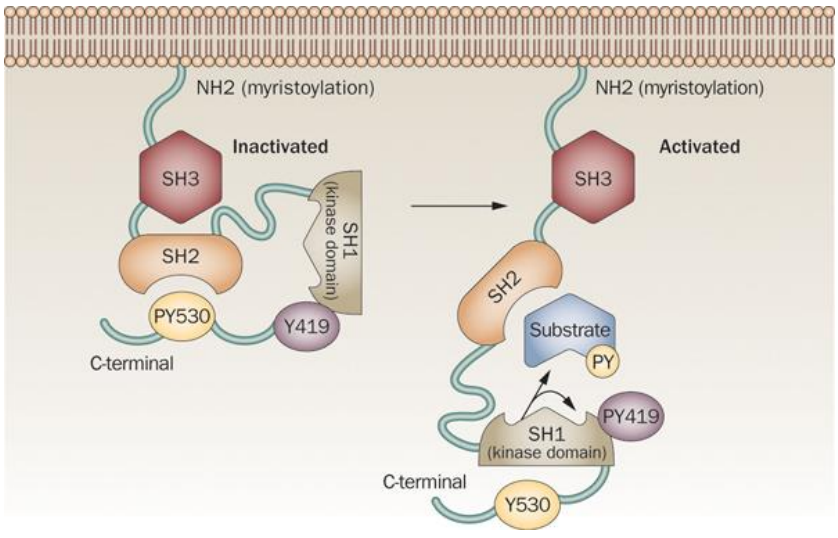

Figure 1.6. Src domain structure, with principal tyrosine phosphorylation sites (Y419 and Y530) and activation process. Src is made up of 3 (in some accounts, 4) so-called Src Homology (SH) domains. SH1 is a protein tyrosine kinase domain; SH2 binds to phosphotyrosine sites within proteins; SH3 binds to proline-rich protein regions. The amino terminal end (sometimes referred to as SH4) can be myristoylated, allowing insertion into membranes. See main text for details of Src activation.Source:[64].

Src has four principal domains, known as Src homology (commonly abbreviated to SH) domains[64], (Figure 1.6). SH1 is the protein tyrosine kinase domain. On the opposite (N-terminal) end, the SH4 domain (often omitted from Src domain classifications) contains the membrane-targeting region, including a Gly-2 residue which must first be myristoylated before Src can attach itself to the cell membrane[64,63](Figure 1.6).In between, the SH2 domain enables Src attachment to phosphorylated tyrosine residues in other proteins, whereas the SH3 domain typically binds to proline-rich regions in respective binding partners[64,63]. In addition, there are two important Src tyrosine residues, Tyr-530 and Tyr-419, whose phosphorylation state is critical to Src activity. Phosphorylation of Tyr-530 (e.g. by CSK) results in the Src SH2 domain binding to it, thereby permitting the SH3 to bind to a proline-rich region in between the SH1 and SH2 domains and inactivating Src (Figure 1.6). In this stable, folded-up conformation the SH1 kinase domain is unable to bind and phosphorylate other proteins, and so Src is effectively inactivated[64,63].

Dephosphorylation of Tyr-530 (by protein tyrosine phosphatases) results in the Src SH2 domain detaching, with subsequent opening up and activation of Src, since the SH1 kinase domain is now free to bind and phosphorylate other proteins[64,63]. Amongst these target proteins is Src itself, with autophosphorylation of Tyr-419 (or phosphorylation by other tyrosine kinases) moving this residue away from the active site, thereby improving access for other proteins[6463]. Another protein interaction that helps activate Src, occurs when its SH2 domain binds to phosphotyrosine residues in other proteins (such as platelet-derived growth factor receptor [65] and FAK [45]), holding Src in its active conformation and thus preventing the SH2 domain from binding to newly-phosphorylated Tyr-530[63].

With respect to FAs, it has already been described above how a number of FAK tyrosine residues are phosphorylated by Src (or by Src-family kinases or SFKs), once Src has bound to the FAK phosphotyrosine at position 397, using its SH2 domain[45]. Such phosphorylation leads, directly or indirectly, to full FAK activation, association of FAK with proteins such as p130^{CAS}, dynamin and

Grb2, and, in the case of Tyr-925, to increased nascent adhesion formation and mature FA turnover, leading to greater cell mobility [50,51]. This increase in FA turnover appears to be partly the result of Tyr-925 being contained within the FAT domain of FAK, which includes unphosphorylated paxillin as a binding partner.

Thus "crosstalk" between Src, FAK and paxillin, and other proteins known to be involved in initial FA formation, is seen to be critical to FA-related processes with obvious relevance to cell migration. In fact, the influence of Src phosphorylation extends much beyond such proteins, but a more detailed account of Src phosphorylation (with a description of how phosphorylation alters each phosphorylated protein's subsequent behaviour) is beyond the scope of this report.

1.1.7 Paxillin

As indicated previously, paxillin is an early component of nascent adhesions [17]and this scaffold protein is thought to play a critical role in adhesion turnover, together with FAK and Src. Paxillin is able to perform this role because it has a number of binding domains to which a variety of FA-related proteins can bind [17]. Four LIM domains (Figure 1.7), each composed of a double zinc-finger motif, are located in the C-terminal half of paxillin [17,66]. Of these, LIM2 and LIM3 are known to be essential for targeting of paxillin to FAs, this targeting requiring that these domains be phosphorylated [17,66,67].

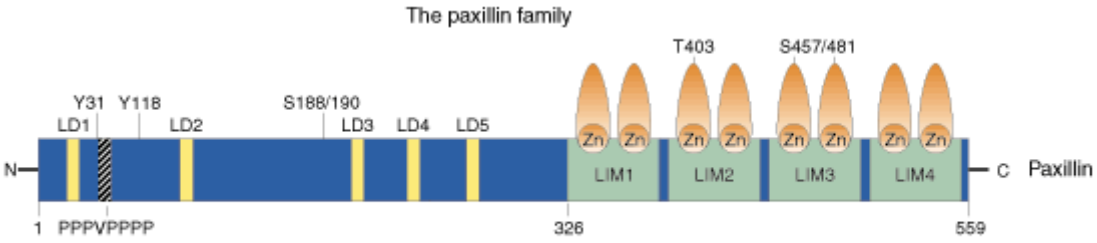

Figure 1.7. Paxillin domain structure, with principal tyrosine (Y) and serine/threonine(S) phosphorylation ites and protein-binding domains. Key: LD1-5 =Leucine-Aspartate Repeat domains; LIM1-4 = double zinc finger LIM domains; PPPVPPPP = proline-rich SH3 domain-binding region. Source:[68].

Despite much work there is still some uncertainty about which proteins act as the main docking partners for paxillin, when locating to FAs. Although one study in 1995 reported paxillin binding to peptides mimicking beta integrin cytoplasmic domains, and another in 2000 [68]showed paxillin co-immunoprecipitating with integrins [69], there has been little subsequent evidence to suggest a direct connection between paxillin and integrins. However, there is stronger experimental evidence that paxillin LIM domains provide binding sites for several structural and regulatory FA-related proteins, such as tubulin[70,71] (a component of microtubules) and the protein-tyrosine phosphatase PTP-PEST (also known as PTPN12)[17,72], thus playing an important role in regulating FA dynamics.

Furthermore, the N-terminal domain is composed of 5 leucine- and aspartate-rich LD domains (Figure 1.7), which interact with a number of very important FA-related proteins, including FAK[73,74,75], GIT1/2[75,76] and vinculin[17,75,77]. In addition LD domains may contribute to paxillin binding to phosphorylated cytoplasmic domains of α4 integrins[78,79].

Allied to its ability to bind many different kinds of proteins, paxillin has many potential phosphorylation sites, and this capacity for variable phosphorylation, adds additional binding sites such as phosphotyrosines, which can be bound by SH2 domains in other proteins), accounting for its importance in FA regulation, particularly in early FA formation[17]. For instance, upon integrin ligation of collagen or fibronectin, paxillin becomes tyrosine phosphorylated, principally on the Tyr-31 and Tyr-118 residues, by FAK and Src[17,80,81]. Such phosphorylation is known to promote Rac1 and Cdc42 activity and inhibit that of RhoA, via the association of these phosphorylated residues with the CrkII-DOCK180-ELMO complex and with p120RasGAP, respectively[17,82,83,84].Rac1 and Cdc42 are activated as a result of DOCK180-ELMO cooperating to act as protein-specific guanine nucleotide exchange factors, whereas the removal p120RasGAP from the GTPase-activating protein p190RhoGAP frees the latter to inhibit RhoA. As explained later, these

differential effects on these members of the Rho family of small GTPases are likely to promote lamellipodia formation, FA turnover and cell migration.

However, just as important is the stable recruitment of the GIT1/2-PIX-PAK-NCK complex to FAs, via the paxillin LD4 motif, with this complex promoting both Rac1 and Cdc42 activity[17]. If LD4 is not binding this complex, it and LD2 can both interact with FAK, via the latter's C-terminal FAT domain[17,85]. Moreover, phosphorylation of paxillin at Ser-273 causes the LD4 domain to bind preferentially with GIT1, whereas unphosphorylated Ser-273 results in the LD4 domain adopting a conformation preferential to FAK binding[17,87]. Since FAK must bind to both LD4 and LD2 for full activation[88] and downstream signalling, and since Tyr-31 and Tyr-118 phosphorylation is also known to enhance FAK binding to adjacent LD motifs[73], this suggests that differential phosphorylation/dephosphorylation may provide a useful switch for fine tuning local RAC1 and FAK activity, thus modifying FA turnover, actin dynamics and related functions. In support of this a large number of kinases (including FAK, Src, ROCK1, JNK, CDK5, PAK and ERK) and phosphatases (such as PTP-PEST, PP2A and SHP-2/PTPN11) are known to affect cell adhesion, migration and protrusion in a paxillin-mediated way, with such effects seen to be accompanied by modifications in RAC1 and RhoA behaviour[17].

Certainly the regulation of Rho-family GTPase signaling and cell migration appears to be very well conserved evolutionarily, with paxillin-like proteins seen to modulate Rho-family behaviour and affect cell adhesion and migration in a wide range of organisms, including yeast, slime moulds, drosophila, amphibians and zebrafish[17]. However, another critical role for paxillin in FA regulation is suggested by its affinity for tubulin, and its association with various parts of the microtubule cytoskeleton in adherent cells [70], especially given that microtubule localisation and stabilisation at the leading edge (almost certainly close to FAs) are known to be early events in cell migration[89].

1.1.8 Vinculin

Structurally, vinculin can be divided into three distinct regions, an acidic N-terminal region, globular in shape, a basic C-terminal region and, in the middle, a proline-rich hinge region (Figure 1.8). The opposite charges of the two terminal regions means that vinculin exists in an equilibrium between two conformations (Figure 1.8), an inactive state in which the head and tail domains are bound together and an active state in which one or both of the terminal domains is bound to other partners. Given that these alternative binding partners are either FA component proteins or other membrane-associated macromolecules, active vinculin also tends to be membrane-associated, whereas, in its inactive state it is primarily to be found in the cytosol [92].

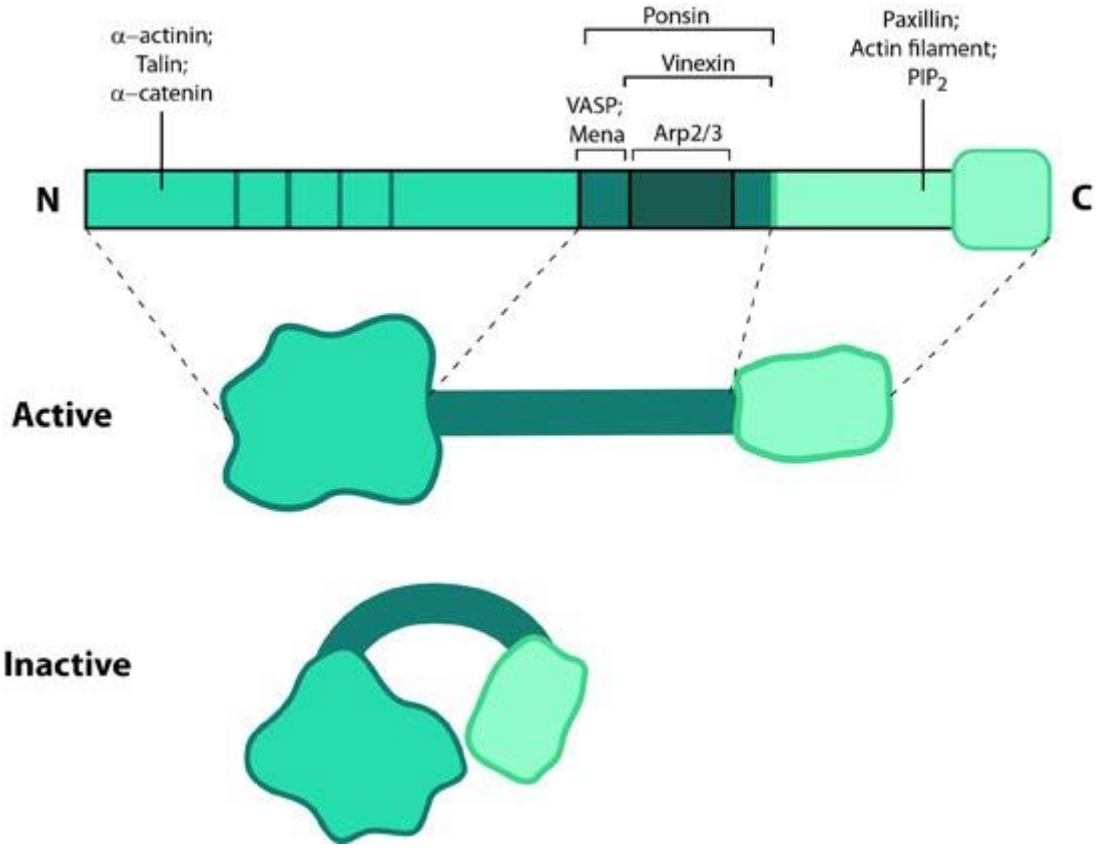

Figure 1.8. Structure and conformations of vinculin [Source:[90]].

In terms of specific binding partners, the N-terminal region contains binding sites for talin and α-actinin, while the C-terminal region contains binding sites for F-actin (actin in filamentous form), paxillin and acidic phospholipids[91,92](Figure 1.8). Talin is known to be a primary activator of

vinculin, often in conjunction with phospholipids (particularly PIP2, which, as described above, is a primary activator of talin itself[27]). In fact, talin contains 11 putative vinculin binding sites (VBSs), but, as explained above, these are believed to be only fully exposed as a result of mechanical stretching, the inference being that talin acts as a form of force-induced switch [30].

Vinculin, in its turn, is believed to act like a "molecular clutch". The idea behind this is that filopodia (finger like projections, projecting ahead of the cell front) and lamellipodia (wider cell membrane protrusions, usually associated with filopodia) depend upon the cell membrane being pushed forward by elongating actin filaments[87]. However, these actin fibres are not bound to anything at their opposite ends. This means that, without further intervention, their elongation will serve only to extend their length rearwards, without generating any forward force on the cell membrane. By binding to such elongating actin fibres, it is thought that FA-associated vinculin generates the necessary traction for the actin fibres to push against, thereby driving the cell membrane forward[93].

This raises the question of how force-activation of talin fits in with vinculin's putative molecular clutch role. Since talin has its integrin-binding site in its N-terminal FERM domain, whereas it has numerous actin and VBSs in its C-terminal rod domain (Figure 1.4), it requires to be bound to an ECM-ligated integrin at one end, and rearward-moving vinculin or actin fibres at the other end to be stretched sufficiently to reveal all of its VBSs. Therefore, as more elongating actin fibres are recruited to FAs (either directly to actin-binding sites on talin, or by vinculin acting as a bridge between talin and actin), more force is exerted on talin, stretching it to reveal more VBSs. In this way, force-mediated talin stretching can be seen to provide a positive feedback mechanism for recruiting actin to FAs.

Activation of vinculin, as well as making the above-mentioned binding sites available for binding to their respective partners, also reveals binding domains in the proline-rich hinge region joining these two head and tail regions. These bind components of actin-associated proteins of

theArp2/3 complex, as well as Ena/VASP proteins (Figure 8, [92]). These are known to be involved in promoting and regulating actin polymerisation [94], adding a further positive feedback mechanism for actin elongation and recruitment to FAs [95].

1.1.9Rho-family GTPases

As indicated above the Rho-family of GTPases play an important role in the dynamics of the actin cytoskeleton adhesions and of actin-associated dynamics[96,97,98], with RhoA, RAC1 and Cdc42 being particularly important in this respect. Like all GTPases, those of the Rho-family depend on being in a GTP-bound conformation in order to interact with and activate downstream target proteins, but this requires prior stimulation by proteins known as guanine nucleotide exchange factors (GEFs) to first remove GDP, otherwise they will remain in an inactive GDP-bound conformation. Similarly, active, GTP-bound GTPases can be inactivated by proteins (known as GTPase activating proteins, or GAPs), which promote GTP hydrolysis, leaving the GTPase bound to GDP and therefore in its inactive conformation [97]. Thus, in common with most GTPases, those of the Rho-family are activated (or "switched on") by GEFs and inactivated ("switched off") by GAPs, with examples including p50RhoGAP and p190RhoGAP[99], which inactivate RhoA.

Regarding the principal Rho-family GTPases,RAC1 promotes lamellipodia and other forms of broad sheet-like projection of the cell leading edge, accompanied by membrane ruffling[100,98]. Like RAC1, Cdc42 stimulates the formation of nascent adhesion complexes and drives cell polarity[87]. However, unlike RAC1 , Cdc42 promotes the formation of filopodia, needle-like leading edge projections, that often accompany lamellipodia[100,98].

In contrast to both the above, RhoA stimulates stress-fibre (contractile actin-myosin filament) formation[98] and promotes both the maturation of leading edge FAs and the disassembly of trailing edge adhesions. Thus RhoA has an important role in coordinating directed cell migration,

something most apparent from its inhibition, or that of its most immediate downstream effectors. Such inhibition of RhoA leads to increased RAC1 activity (since RhoA normally inhibits RAC1, and vice-versa[101]),resulting in increased lamellipodia formation[102,103]. However, these lamellipodia all tend to pull in competing directions[103]. This, together with the absence of RhoA-mediated trailing-edge release[104105], means that inhibition of RhoA (or of ROCK1, a serine/threonine kinase, which acts as a major effector protein of RhoA[13]) prevents directed cell migration[103].

In support of this, studies have shown that extensive RhoA or ROCK1 inhibition leads to the formation of many lamellipodia around the periphery of the host cell, accompanied by a lack of directional migration[102,103]. These inappropriate lamellipodial protrusions strongly correlate with increased integrin adhesion, together with raised levels of integrin-dependent tyrosine phosphorylation[103]. RhoA/ROCK1 inhibition is known to be associated with disruption of the cortical actin cytoskeleton (the peripheral part of the cytoskeleton immediately beneath the cell membrane), leading to increased clustering (and therefore total avidity) of integrin molecules[103,106,107]. Integrin adhesion events are known to be associated with increased tyrosine phosphorylation of paxillin and Pyk-2 (a protein tyrosine kinase of the same superfamily as FAK)[84,103,108]with this tyrosine phosphorylated paxillin and Pyk-2 both being known to promote lamellipodia formation[84,103,108]. Thus one interpretation of the above is that RhoA/ROCK1 are required to prevent excessive integrin clustering events, which would otherwise lead to excessive tyrosine phosphorylated paxillin and Pyk-2, leading in turn to over-extensive, competing lamellipodial protrusions[103,107].

ROCK1 is known to promote actin fibre formation by activating LIMK1, a protein serine/threonine kinase, which inactivates (via phosphorylation of Ser-3) the F-actin depolymerizing factor cofilin[13,110] (Figure 1.9). It also promotes actomyosin contractility by phosphorylating myosin II regulatory light chain (MLC) and myosin light chain phosphatase [13]. Thus it can prevent excessive integrin adhesion events either by stabilising the cortical actin cytoskeleton (and thus limiting

lateral integrin movement) or by promoting the maturation of stress fibres and their associated FAs, and so preventing integrin release and recycling[13,103].

As ROCK1 ‘s upstream activator, RhoA is known to have similar effects, enhanced by its direct activating influence on mDIA (a major downstream effector of RhoA, also known as mammalian diaphanous and by other names), which facilitates profilin-mediated nucleation and polymerisation of actin[13,111], (Figure 1.9). Thus both RhoA andROCK1 can be seen as dampeners, preventing an excessive, RAC1-driven, lamellipodial response to the presence of ECM and other external ligands, but, in the right quantities, permitting purposeful, directed migration[103].

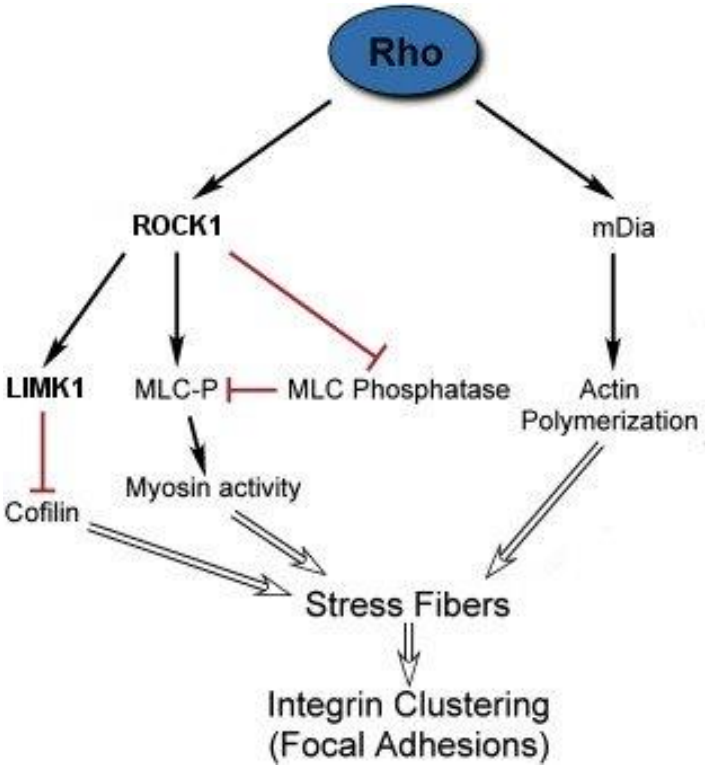

Figure 1.9. RhoA-mediated control of stress fibre formation and stabilisation. Key: Rho = RhoA; ROCK1 = Rho kinase; mDia = mammalian diaphanous; LIMK1 = LIM kinase 1; MLC-P = phosphorylated myosin II regulatory light chain; MLC phosphatase = myosin light chain phosphatase. Adapted from [112].

As explained earlier, tyrosine phosphorylation of paxillin (at the Tyr-31 and Tyr-118 residues, typically following integrin ligation) leads to enhanced RAC1 and Cdc42 activity[17,82,83,84]. These two Rho-family GTPases then promote (respectively) lamellipodia and filopodia formation by acting on

the ARP2/3 complex which plays a major role in regulating the actin cytoskeleton[97]. However, whereas RAC1 acts upon the ARP2/3 complex through the WAVE protein complex[97], Cdc42 acts through WASP (and its neuronal counterpart N-WASP)[97], which may account for the obvious differences between lamellipodia and filopodia (Figure 1.10). However, RAC1 and Cdc42 also act on other proteins. For instance RAC1 also acts onPOR1[112,113], which is known to be involved in membrane ruffling[114], and both RAC1 and Cdc42 act on IQGAP1, which promotes cell polarisation in migrating cells by recruiting microtubules to their leading edge, linking it to the actin cytoskeleton[115]. These may better account for the differences between RAC1 and Cdc42.

Furthermore, both RAC1 and Cdc42 activate Diaphanous-related formins (other than mDia), as well as PAK proteins (PAKs)[97,116,117,118,119].This is of interest since both formins and PAKs promote F-actin polymerisation, the former directly as actin nucleators[120], whereas PAKs

(certainly PAK 1-3) act via LIMKs (LIMK1 and LIMK2), in a similar manner to ROCK1[97] (Figure 1.10).

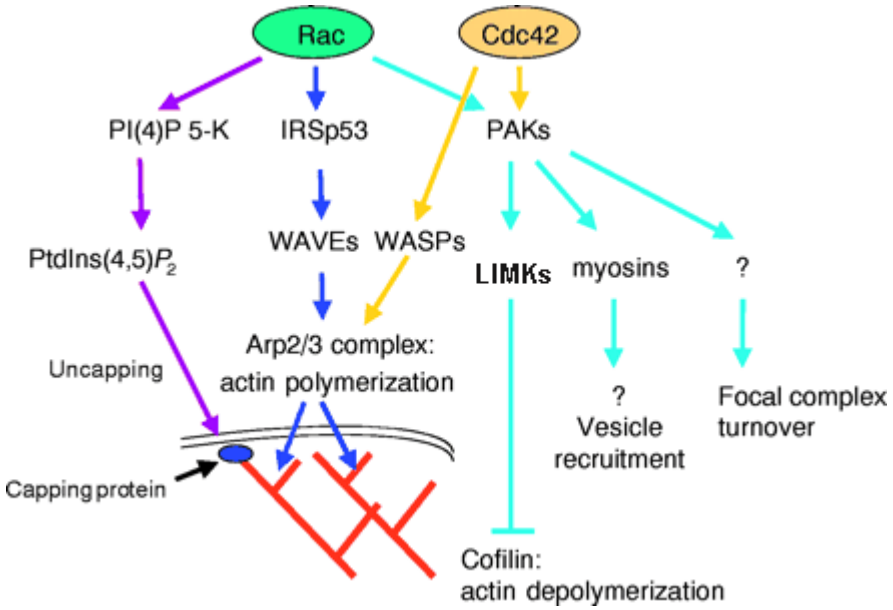

Figure 1.10. Major downstream pathways of RAC1 and Cdc42.Adapted from[97].

In summary, it is clear that these three Rho GTPases have an important role in directly regulating FA maturation and disassembly, as well as indirectly by promoting lamellipodia and filopodia

(both of which are known to be associated with the formation of nascent adhesions). An obvious example of direct regulation is RhoA-mediated promotion of stress-fibre formation and general maturation of leading edge FAs, together with disassembly of trailing edge adhesions. However, there are clearly many more proteins involved in mature FA assembly, including, most obviously, those of which mature FAs and stress fibres are composed.

1.1.10 Other principal proteins in mature FAs

As shown in Figure 1.1, the principal component proteins in mature FAs, other than integrins, FAK, paxillin, talin and vinculin are zyxin, VASP (vasodilator-stimulated phosphoprotein), α-actinin and actin. (Stress fibres are mainly composed of actin and myosin, but are attached at one end to mature FAs, binding them to the ECM.)

The role of zyxin has yet to be fully elucidated but the evidence points to it being associated with areas of stress fibre damage, being necessary both for repair (helping to recruit α-actinin and VASP to compromised areas), as well as for the generation of traction force [121]. As such it is often considered a marker of mature FAs.

As suggested by the above, the evidence suggests that whatever function VASP serves in FA assembly involves actin, both in promoting actin nucleation, filamentous actin formation and in binding to actin filaments[122].However, there is also some evidence that it is associated with nascent adhesions, in an unknown capacity [123].

As indicated by its name and by its role in repairing stress fibres, α-actinin is primarily associated with actin fibres, cross-binding individual fibres[124] (as shown in Figure 1.1). Therefore it would appear to be mainly associated with mature FAs, although there is some evidence that it may also play a role in FA maturation, by linking integrins to actin fibres [125]. Finally, the importance of the protein actin, both in the formation of actin filaments and stress fibres, and thus in mature FAs, is

obvious. As most of the details of actin filament formation are unrelated to FA formation, a long, detailed, account of the former would be inappropriate here, and it is summarised in Figure 1.11 below.

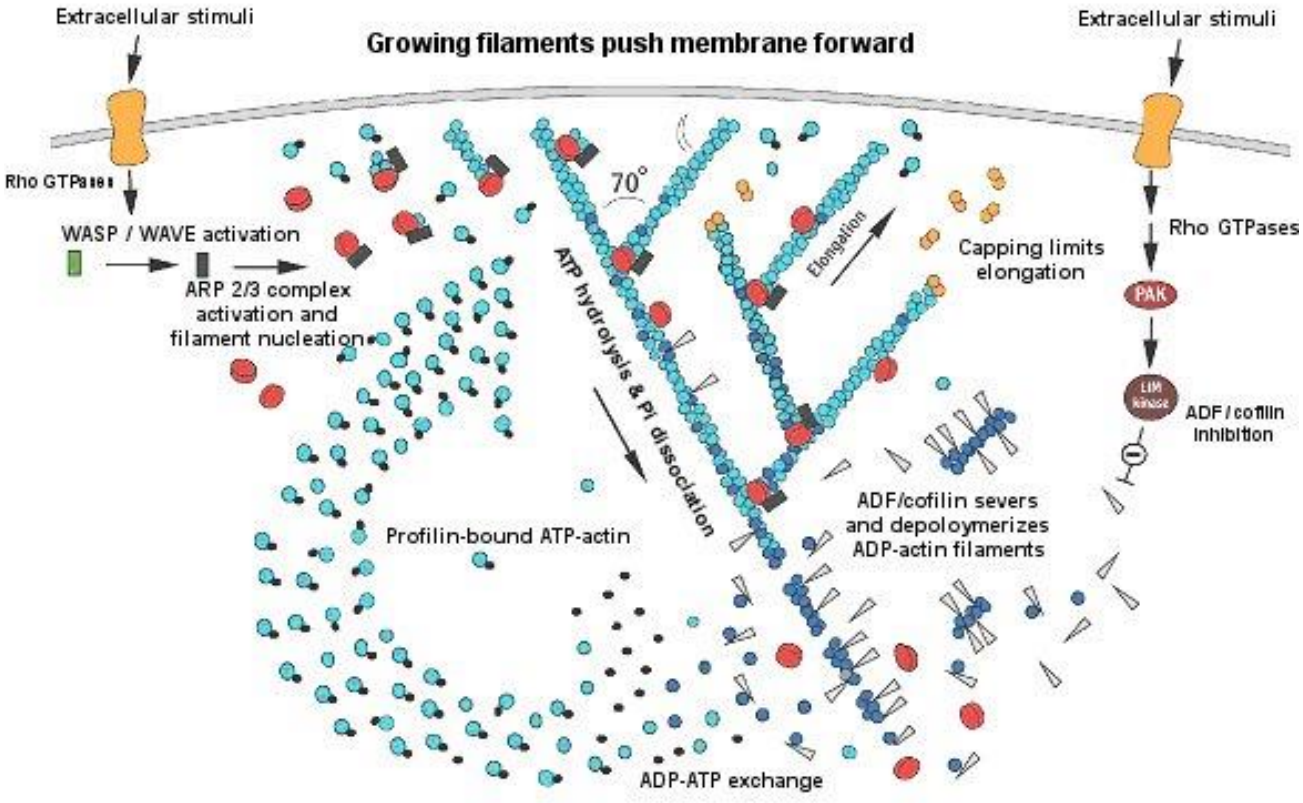

Figure 1.11. Actin filament dynamics, showing actin-associated proteins and Rho GTPases. Cofilin (also known as ADF) severs ADP-bound actin monomers from actin filaments, promoting F-actin depolymerisation, whereas profilin helps catalyse the exchange of actin-bound ADP for ATP and recruits these actin-ATP monomers to the growing (barbed) filament ends. Arp2/3 protein complexes (activated by WASP, WAVE and similar proteins) nucleate new branched filaments (at about 70° angles) from existing ones, whereas capping proteins prevent elongation of existing filaments. Role of Rho GTPases explained earlier. Key: cyan circles = actin-ATP monomers; dark blue circles = actin-ADP monomers; red circles = Arp2/3 complex; white triangles = ADF/cofilin; black circles = profilin; orange circles = capping proteins; green/black rectangles = inactive/active WASP/WAVE proteins. Adapted from[126]. Rho family GTPases shown here being externally stimulated, but as explained above, may also be activated by FA proteins (especially Paxillin in outside-in signalling).

1.2 System modelling

1.2.1 Reasons why system modelling may be needed

The above account reflects how much has been learnt over the last few decades about the detailed protein composition of FAs and about the various processes by which they are regulated. However, there is still a great deal to be determined and discovered. Advances in microscopy and other forms of imaging, and new experimental techniques, will no doubt further our understanding of FA structure and behaviour. Nevertheless, it is likely that large areas of uncertainty will remain. This is because imaging is always likely to be hindered by difficulties in seeing events in real-time beyond, but in the immediate vicinity of, the cell membrane, whilst simultaneously correctly identifying all the significant actors, in ways that do not seriously affect their normal physiological behaviour.

Similarly, new laboratory experimental approaches are unlikely to overcome traditional concerns, not least doubts about the binding-specificities of inhibitors, antibodies or other markers, and perennial questions, such as whether the observed effects of inhibiting or silencing the expression of some protein (or of knocking in or out some gene of interest)can be explained solely in terms of the specific behaviours ascribed to it, rather than being simply experimental artefacts, or having other more general explanations.

In any case, such is the level of complexity already uncovered that other approaches are necessary if we are to arrive at a coherent explanation of how FAs are composed, of how they assemble, mature and disassemble, and of how such assembly, maturation and disassembly are regulated.

One such alternative approach is modelling, the use of mathematical, logical and computational descriptions in order to explain complex systems, to study the interactions of their components and their effects, and to enable predictions about system behaviour. Without such approaches there is a danger that the study of FAs does not advance beyond a set of competing and contradictory narratives, limited in scope and lacking internal logical consistency. By contrast, modelling provides a mechanism by which various explanations and hypotheses can be tested and compared, allowing refinement of the most successful explanations, so that they increasingly match what is observed under the microscope and within experiments.

1.2.2 System models and what they can do

A system model can be viewed as "a representation of the essential aspects of an existing system … which presents knowledge of that system in usable form" [127]. Or, according to a more detailed definition, it can be viewed as "a simplifying abstraction [that] generates predictions of system behavior under different conditions (as reflected by observations) and illuminates the roles of various system components in these behaviors"[128].From such definitions it is clear that a worthwhile model is founded upon an adequate understanding of the system being studied, including detailed knowledge of the major system components and how they relate to each other. Without this foundation it is impossible to know what system details can be safely excluded from the model without changing its behaviour so much that it no longer sufficiently reflects the behaviour of the original system to be useful.

That achieved, a good model offers a number of benefits, most obviously that it allows the modeller to test current system explanations and to uncover gaps in understanding. Where problems are uncovered, it may be possible, through alterations to the model (such as changes in parameter values or alterations in model component behaviour), to arrive at suitable alternative solutions which can then be incorporated into a new understanding of the system.

Through repeated running with altered settings, a good model can highlight poorly-appreciated relationships between system components, suggest components, lead to new predictions about the overall system. None of these benefits might arise from using another approach. In a similar way, repeated runs of such a model allows detailed analysis of system behaviour, including exploration of equilibrium states, as well as tests of system stability and of its sensitivity to changes in the values of variables and parameters.

Arguably, though, the chief benefit of a good system model is that it encourages the replacement of essentially qualitative system descriptions with more rigorous quantitative alternatives, based on the quantitative properties of the model itself. These are usually more straightforward to test experimentally, to prove or disprove, as the case may be. However, of course, this requires a good model in the first place and to achieve this involves a number of considerations on the part of the modeller.

1.2.3 Modelling considerations and approaches

One of the first decisions any system modeller must make concerns what aspect of the system to model. In reaching this decision the modeller will be guided by how much is known about that system aspect, and how complex its behaviour is known to be. Clearly it is much easier to model a relatively simple system, about which much is already known, than to model a complex system about which little is known. However, the benefits of modelling the former system are likely to be small, given that much is already known about it, whereas the great challenges of adequately modelling a complex system may be justified if there is the possibility of great advances in knowledge about the system. This can be especially beneficial if there is no easier way of obtaining the same information, such as through laboratory experiments.

Before deciding what biological system aspect to model, it is useful to first conduct a detailed cost-benefit analysis. This must be guided by considerations as to what it is one wishes to learn about the particular system aspect, as well as what form that knowledge must take. For instance, having decided that useful knowledge about a complex biological system can only be obtained by modelling its behaviour over a long, continuous timescale, it may, at a late stage, be decided to abandon the attempt because it would require excessive computational resources. Similar considerations are required in determining the best modelling approach to take.

For any system there are a number of different modelling approaches that can be adopted. This choice will be determined mainly by the properties of the system itself. That is to say, its scale and complexity, whether it receives inputs from outside the system or is entirely self-contained, how much is known about the system and what it is that one still seeks to learn about it. For instance, a deterministic model may be most appropriate where a great deal of information is known about the system and chance plays little or no role, whereas a stochastic, probabilistic, model may be a better approach for modelling systems with a lot of unknown and uncertain variables. Similarly, different modelling approaches may be appropriate if one merely seeks to determine how two variables co-relate to each other, as opposed to determining any deeper causal relationships between them.

Other factors to consider, prior to model implementation, include the length of the time-course the model must cover, whether this time period can be modelled in discrete steps or continuously, and also what variables, independent and dependent, need to be incorporated and measured within the model. Related to this latter point, one must also consider if one needs to include any additional parameters (values that may vary between different runs of the model but which remain constant during any particular run) that may affect the overall results.

The answers to all these questions will help determine how best to implement the proposed model. Essentially all system models are mathematical in nature, at least in the widest sense of that term. What mathematical form they take will principally depend on what results need to be produced, what data is known about the individual components of the model, and what computational resources are available.

1.2.4 Types of system modelling approaches

1.2.4.1 Differential equation modelling

In their most complex form system models may be heavily mathematical, taking the form of a set of differential or algebraic equations (e.g. ordinary or partial differential equations, commonly referred to as ODEs and PDEs, respectively), although they may also be constructed primarily as vectors and matrices, drawing heavily on aspects of linear algebra. Usually such models are solidly deterministic in nature, although stochastic elements may be added, and, by their nature, they are particularly suitable for yielding results over a continuous timescale.

The choice between using ODEs or PDEs will typically depend on whether one wants a compartmental or spatial model. By this is meant whether one wants the model species (for example, individual proteins and other macromolecules in a model of a particular cellular behaviour) to be characterised only by their concentration (or density) within a particular volume (or area or length, depending on the dimensionality of the model), or also by their spatial position within that compartment.

For instance, if one is only interested in changes in cytosolic concentration of certain proteins over time, then a compartmental model will normally suffice, requiring just one compartment (representing the cytosol). Assuming that this is a three-dimensional representation, all that is then required is that a volume is specified for this cytosolic compartment, typically one that is

consistent with cytosolic volumes observed in real cells. Since one is only interested in one model variable for each species, that is to say changes in concentration over time, ODEs will suffice for such compartmental models. The implicit assumption here is that, within each compartment, model species are uniformly distributed and it is only the overall compartmental concentration that one is interested in.

If, however, one is also interested in how species concentrations vary spatially within their compartment, then PDEs will be necessary as there is now more than one variable of interest (changes in concentration over time, plus changes in concentration over each available spatial dimension, depending on how many such dimensions the model has). Obviously, such models will require that each compartment is defined spatially (i.e. allocated some sort of shape and size), not just specified as a simple concentration. Collectively this set of PDEs and spatial definitions is referred to as a spatial model.

Numerous examples of such models are to be found in systems biology and in biology generally, including in the study of enzyme kinetics, metabolic control analysis, molecular modelling, and in population dynamics[129]. For example, generalised mass action (GMA) models, commonly used in studies of biochemical pathways, are typically composed of coupled sets of ODEs, incorporating rate constants, kinetic orders and similar properties of individual reactions, which can then be run simultaneously to reveal overall pathway dynamics [130]. Other commonly used examples are the SIR compartmental model, used to model infectious disease (with "S" representing susceptible, "I" representing infected and "R" representing removed, i.e. dead or immune, individuals within a given community), and Lotka-Volterra population models, primarily used to model predator-prey or parasite-host dynamics. Again these are normally composed of coupled ODEs [131,132].

GMA and Lotka-Volterra models are just two examples of canonical models, the result of an essentially bottom-up approach that arrives at a set of equations, each such equation carefully mirroring component behaviours of the system. However, this is not always necessary and the modeller may simply choose a mathematical form that mirrors overall system dynamics, without any obvious reason why this should be the case. As an example of such a so-called ad hoc model, one might choose to use a version of the Michaelis-Menten function to model some biological phenomenon completely unrelated to enzyme kinetics, or the logistic function, simply because, in this instance, it produces the right results.

The advantages of deterministic mathematical models of this kind include the ability to achieve a high level of granularity, which makes them highly suitable for studies over continuous timescales. Thanks to their popularity there is a wide range of software tools to aid the user, including general environments, such as Matlab[133] and Mathematica[134], as well as more specific tools such as COPASI[135] and Virtual Cell [136] (both used in GMA and other biological process modelling) and numerous examples of molecular modelling applications, e.g. DOCK[137]and CHARMM [138].

However, there are also disadvantages to this modelling approach, most obviously that their fine granularity often results in such models being much more demanding of computational resources than other modelling approaches. Also (and partly for the same reasons), this form of modelling generally requires much higher levels of information about the underlying system. In the absence of all such information there is a high risk of “overfitting”, in which application of the model ends up describing random error or noise, rather than genuine behaviour of the underlying system.

1.2.4.2 Stochastic modelling approaches

For these reasons it is often more appropriate to use other modelling approaches, including the use of stochastic models. These, by their nature, tend to handle noise in the data more robustly than more strictly deterministic mathematical approaches. Alternatively, one might choose to employ less mathematically sophisticated techniques that are less computationally demanding or because they operate over discrete, as opposed to continuous, timescales.

Stochastic modelling approaches may have other advantages, of which the most useful, perhaps, is the possibility of inferring previously unknown information about the system, such as hidden states. Indeed, this inferential capacity is often the primary purpose for using them, since, in practice, it is usually possible to modify deterministic models, such as coupled ODEs or Boolean networks, to incorporate stochastic elements, rather than requiring the use of strictly stochastic modelling techniques.

Typically, stochastic modelling approaches involve the use of Markov chains, Hidden Markov models (HMMs) or Bayesian Networks, which, in turn, are normally implemented as graphs, vectors and matrices, containing probability values relating to various state transitions. Examples include the TMHMM bioinformatics tool[139], which uses a HMM approach to predict transmembrane regions in proteins, and the widespread use of probabilistic graphical models, such as Markov Networks, Bayesian Networks and Chain Graphs, to infer signaling pathways and other cellular molecular networks, from high-throughput protein expression data or other datasets [128140].

For example, one particularly successful use of this approach involved the Bayesian network analysis of flow cytometry data to infer part of the intracellular signaling network in human primary naïve CD4þ T cells (Figure 1.12) , following system perturbations with various protein-specific inhibitors and activators [140]. This analysis was able to uncover 17 connections between signalling proteins, including two that had not been widely reported at the time. A further

connection was indicated, which subsequently turned out to be in the wrong direction. However three known connections were missed, including two involved in feedback signaling. This indicates one of the weaknesses of such Bayesian and other stochastic network inference, that it does not work well in networks containing numerous cyclic references (where individual nodes have connections, or "edges", to themselves) or where there are numerous connections in the reverse direction.

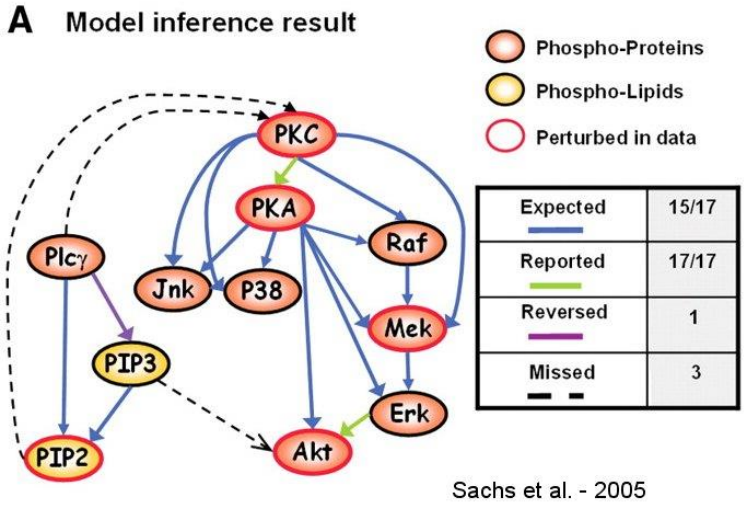

Figure 1.12. Signaling pathways in human primary naïve CD4þ T cells, inferred by Bayesian network analysis. Source: [140].

Another weakness of stochastic models generally is that they tend to cope poorly with dynamically changing networks, altered probabilities and similar changes. This is most obviously the case in Markov models, which, by definition, do not allow model behaviour in the future to be altered by past or present behaviour, this "memoryless" property being a requirement for inference to be possible. A further problem is that they can only realistically incorporate discrete timescales.

1.2.4.3 Boolean modelling

Another form of modelling approach with some similarities to stochastic network modelling is Boolean modelling, which can involve the use of simple truth tables or more sophisticated implementations, such as Boolean networks [141]. These are commonly used to model signaling pathways [142], gene regulatory networks [143] and similar problems. Typically Boolean networks are implemented as a set of Boolean variables (i.e. variables that can adopt only two values, True or False, or, equivalently, 1 or 0) and a set of functions (variously known as transition or regulation functions) that describe allowed state transitions of these variables (Figure 1.13).

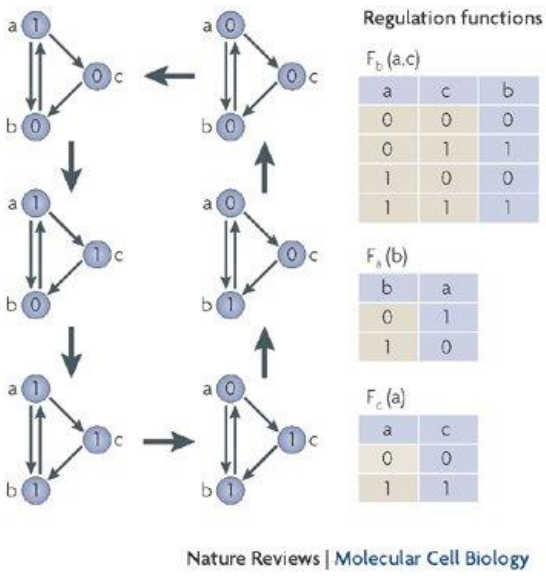

Figure 1.13. Example of a Boolean network. Each of the nodes a, b and c in the network can be in state 0 or 1. State transitions obey the regulation functions shown on the right, which describe the rules of the model. For example, if node a is in state 1 and node c is in state 0, at the next time step the state of node b will be 0. Thin arrows indicate the regulation of each node (i.e. which other current node values affect that node's state after the next time step), while thick arrows represent the time steps. The global state of the model is the combination of the three node states. Source: [143].

Whilst simple in concept, such models are readily scalable, allowing highly complex networks to be modelled, such as the mammalian cell cycle regulatory network [144] and various signaling pathways, such as the one seen in HGF-induced migrating keratinocytes [142](Figure 1.14).

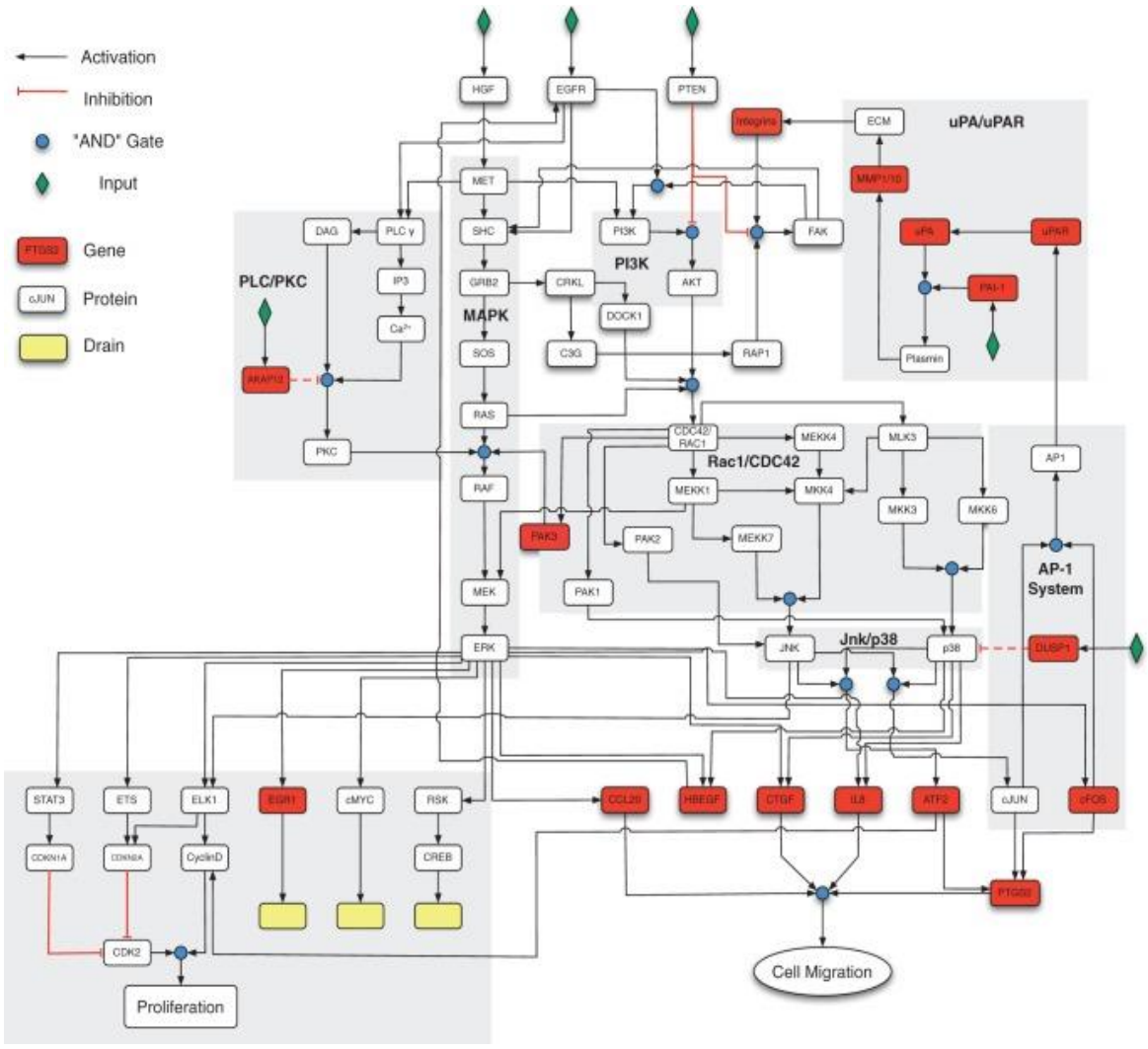

Figure 1.14. Boolean network model of the HGF-induced keratinocyte migration. Source: [142].

Along with being readily scalable, Boolean networks are comparatively easy to set up, certainly compared with deterministic mathematical models and probabilistic networks. The simplicity means that they are also relatively easy to analyse. However, like Markov models, they don't allow dynamic changes in behaviour and they are particularly weak in terms of incorporating time. Continuous timescales are impossible and the synchronisation of model component behaviour can be challenging [144].

1.2.4.4 Process diagrams

Simpler even than Boolean models are state-transition or process diagrams in which the system under study is described as a finite set of possible states, with a defined set of allowed transitions between states. Widely used in the computer industry, in system analysis and design, process diagrams are beginning to be used more widely in cell biology, for instance in modelling signalling pathways (Figure 1.15):

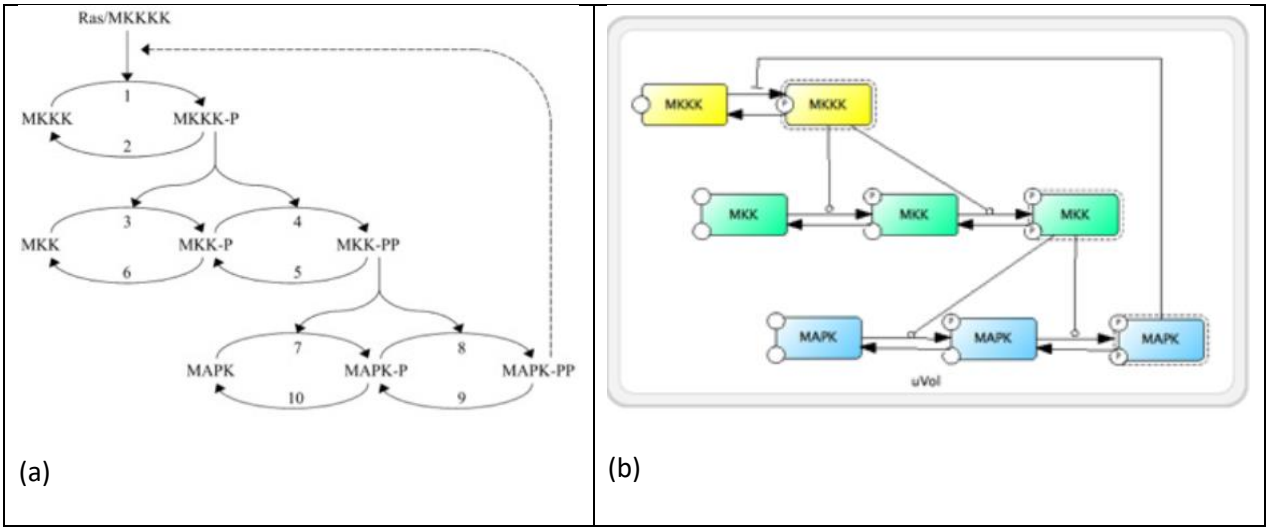

Figure 1.15. Example incorporation of a signalling pathway into a process diagram, using CellDesigner. Source: [145]. (a) Conventional kinetic representation of MAPK signalling cascade. (b) CellDesigner process diagram of same cascade.

Despite their simplicity, process diagrams have the advantage over conventional biological diagrams of being much more rigorous, thereby forcing the modeller to think deeply about the system behaviour being modelled. On the other hand, because of their simplicity, they tend to be computationally very undemanding, compared to other modelling approaches, as well as being relatively easier to implement and scale up.

Whilst the modeller may not feel that they are "doing proper modelling" whilst specifying process diagrams, in practice the comparatively high level of rigour required to implement them can lead to insights comparable to those that might be revealed by Boolean models. Certainly, they are

often very beneficial when used as a first step, prior to employing more complex modelling approaches. By pointing out logical errors and posing previously unconsidered questions, they can lead to better system specifications in subsequent models and prevent much time wasted in avoidable errors. If one considers that the primary point of modelling is to point out problems in our current understanding of systems and pose new questions, it is hard to escape the conclusion that process diagrams are, indeed, "proper modelling".

However, there are a number of disadvantages to process diagrams, including a relative lack of software tools, and, perhaps more seriously, the fact that they do not allow any worthwhile dynamic or quantitative analysis. Essentially they are qualitative, not quantitative, in nature, and essentially static in the sense that they cannot be "run" in the way that other models can. Therefore they are limited in what they can offer, arguably little more than rigorous forms of narrative graphical description.

1.2.4.5 Agent-based modelling, and rule-based modelling in general

A more quantitative approach is provided by agent-based modelling (ABM), in which the system is modelled as a set of rule-based objects with properties. Typically such models are run within a simple logical lattice (Figure 1.16), with only one object allowed to occupy any given lattice cell at any one time, and with object behaviours such as movement between cells, self-copying, self-destruction, etc, being determined by the rules and current properties associated with that object and by the number and types of objects in adjacent cells. Although ABM is better known in other fields, such as economics, it is becoming more popular in biology, with applications in population dynamics[146], and in the study of inflammation [147,148,149]and tumour growth[150,151,152].

Agent-based modelling is a particular form of rule-based modelling, in which the various permitted behaviours of system actors are implemented as a set of rules. Typically such rules result in changes in the properties of one or more system actors, and only involve actors having

certain properties, as specified in the rule. So a rule might say that all white cars must be painted red on Wednesdays, this rule only applying on Wednesdays and only to white cars, whatever other differences there might be between these cars.

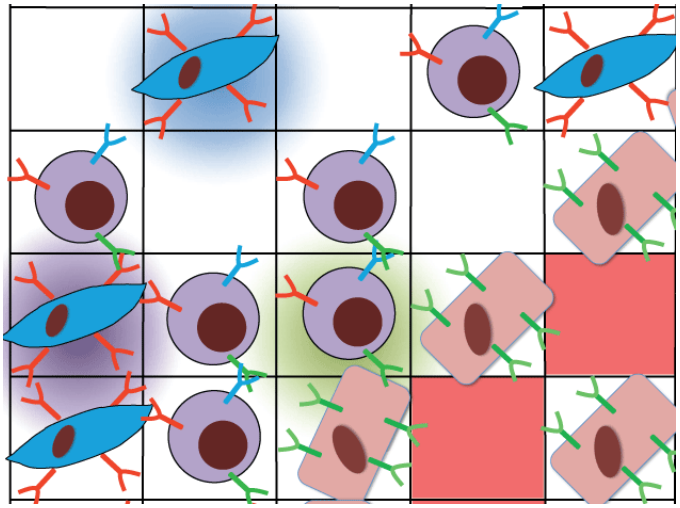

Figure 1.16. A typical agent-based model lattice, here modelling signalling between various cell types within breast tumours. Here, oval-shaped objects represent carcinoma-associated fibroblasts, round objects represent T-cells and rounded rectangles represent carcinoma-associated epithelial cells. White lattice cells represent empty spaces outside the tumour, whilst red lattice cell denote empty spaces within the tumour. Each lattice cell can only be occupied by one object at any time. Signaling can only occur between objects that contain (Y-shaped) receptors of the same colour.

The advantage of such rule-based approaches for modelling reaction-diffusion systems is that, unlike conventional differential equation approaches, they avoid the need to specify separate reactions for every model species, regardless of how similar some species may be to each other.

To give an example, if we have two species, A and B, each with two minor variants, it follows that, assuming we are interested in analysing reactions between A and B, these must be modelled as four separate species (call them A1, A2, B1 and B2). Assuming that reactions between either variant, on the one hand, and either B variant, on the other, always result in the same product, then, using conventional differential equation modelling, reactions between them will require four different reaction equations:

A1 + B1 ->Product; A2 + B1 ->Product; A1 + B2 ->Product; A2 + B2 ->Product.

However, it would be much easier to specify the above as a single reaction of the form:

A* + B* -> Product,

where A* and B* represent either variant of A or B. This form of the reaction can be implemented using rule-based modelling (including ABMs) but not using conventional differential equation models.

So, for larger models (say, any model with more than 20 species), one has to be very careful to ensure that, as one adds new species (for example, to denote a new structural variant of an existing species), one remembers to add additional reactions, so that this new species also can react with other model species, in the same way as all the other variants of that species. More generally, as the model develops, it is common for the number of species to increase. As many of these species may require several new reactions to be specified (especially where they are variants of existing species) this can result in an exponential "combinatorial explosion".

For the reasons explained above, this is far more likely to be a problem with conventional differential equation models and far less likely in rule-based equivalents. Given the complexity of FAs and other reaction-diffusion systems, this is clearly something that has to be considered.

Returning specifically to ABMs, these have the additional advantage of generally being relatively easy to implement and scale up. Such models can also lend themselves particularly well to parallel processing. Perhaps their most useful feature is that they are good at uncovering emergent properties. Against that, the difficulty of dynamically regulating ABMs means that they can quickly overwhelm computer resources, and their dynamic properties can be very difficult to analyse mathematically. In particular, they suffer the limitation of only being able to operate over discrete timescales, which are often difficult to relate, in any meaningful way, to real time. Given that each agent behaves autonomously, then, even assuming that each component behaviour is

defined as a time-based rate, it can be computationally unfeasible to coordinate these into a reliable overall system time. Without such a timescale, time-based analysis of system behaviours is difficult if not impossible.

Another disadvantage of ABMs is that, being essentially particle-based, they are not a natural fit with large-scale reaction-diffusion systems, which would seem the most appropriate category for describing FA dynamics. This is because such systems are largely defined in terms of their underlying reactions, which are, in turn, always specified in terms of rates of change in concentrations of reactants and products (i.e. species populations).Reliably translating such concentration rates into appropriate rules for individual particles (modelled as autonomous agents) is hugely challenging, not helped by the time-coordination problem described above.

Overall, given the complexity of FAs, it seems fair to say that ABMs have only limited applicability to their dynamics, for example in modelling smaller-scale phenomena such as integrin clustering. However, there are hybrid approaches that may be more appropriate. For instance, Virtual Cell allows rule-based differential equation modelling, in which rules determine how species can react with each other, but the rate of such reactions are determined at the population/concentration level, using conventional kinetic rate laws, such as mass action and Michaelis-Menten kinetics. Being concentration-based, these are not appropriate for detailed studies at the molecular level (e.g. modelling integrin clustering), but may offer some of the benefits of rule-based modelling to differential equation modelling, whilst avoiding the excessive computational overload, and other disadvantages of agent-based modelling, just mentioned.

1.2.5 Modelling approach adopted for this study of FA dynamics

However, for the purposes of this investigation into FA dynamics, it has been decided to use a straightforward differential equation approach. A primary reason for so doing is the severe limitations of the other approaches for investigating FA dynamics, as described earlier. Another important reason is the availability of software tools (as mentioned above) that are well proven, and that include a number of analytical methods (described later) to help the modeller demonstrate model outputs. In particular these help determine the underlying reasons for particular behaviours and outputs.

It might have been advantageous to adopt a hybrid rule-based differential equation approach, as just described, given the numerous conformational variants, post-translational modifications and other forms of permutation. However, in the writer's experience, this approach can lead to additional challenges in analysing the results (as described also for ABMs earlier), as well as difficulties in conveying model outputs clearly to the reader. For this reason, it was felt that, on balance, the advantages outweighed the disadvantages. However, other modellers might reach a different conclusion in a similar situation.

Having decided this, it remains to describe differential equation modelling in some more detail, so that the reader is clear how such equations are applied to reaction-diffusion systems.

1.2.6 Differential equation modelling

By "differential equations" in this context is meant equations that relate changes in the value of some model species (typically called "the product") to changes in the value of one or more other species (typically called "the reactants"). In reaction-diffusion systems these equations model reactions between species as chemical reactions, with the form of the equation determined by

the kinetic rate laws deemed most appropriate for that reaction. In the next section ODE approaches for modelling reactions are described and illustrated, followed by a discussion of PDE/spatial methods in the following section.

1.2.6.1 Ordinary differential equations: their use in modelling reaction diffusion systems

The most common kinetic rate laws encountered in cell biology are mass action and Michaelis-Menten, with the former commonly used to model various binding reactions between proteins and other macromolecules. Michaelis-Menten kinetics, by contrast, are most commonly used to model enzymatic reactions, in which one or more molecules ("the substrate(s)") are transformed by an enzyme into one or more products that are different to the substrate, while the enzyme itself remains unchanged.

Differential equations for irreversible mass action reactions take the generalised form:

$$d\frac{[P]}{dt} = k.\prod_{I=1}^{n}[Ri]$$

where [P] designates the concentration of the product, [RI] the concentration(s) of the reactant(s), and k is some characteristic constant of proportionality called the "rate constant", which, together with the reactant concentration(s) determines the rate at which the reaction proceeds. (That is to say, together they determine the rate at which the reactants are transformed into product.)

In such cases, where only the concentration of one reactant is rate-determining, the reaction is described as a "first order reaction". (Most commonly first-order reactions have one reactant, but

there are rare exceptions to this rule.) By extension, for second order reactions, the reaction rate is determined by the concentrations of two reactants, for the reversible form, represented as:

$$A \Leftrightarrow B + C$$

which breaks down into 3 component reactions, one for each species:

$$d\frac{[A]}{dt} = k_f.[B].[\mathrm{C}] - k_r.[\mathrm{A}]$$

$$d\frac{[B]}{dt} = k_r.[\mathrm{A}] - k_f.[B].[\mathrm{C}]$$

$$d\frac{[C]}{dt} = k_r.[\mathrm{A}] - k_f.[B].[\mathrm{C}]$$

where k_f represents the rate constant for the (second-order) forward reaction (i.e. for the association of A from B and C) whereas k_r represents the rate constant for the (first-order) reverse reaction (i.e. for the dissociation of A into B and C). (In many modelling tools, k_f has the alternative notation k_1, with k_r represented as either k_2 or k_{-1}.)

By contrast, Michaelis-Menten reactions are almost always irreversible and take the form:

$$d\frac{[P]}{dt} = kcat.[E].\frac{[S]}{K_M + [S]}$$

where [S] denotes the concentration of the substrate, and kcat denotes the turnover number (similar to the rate constant k in mass action reaction, but denoting the number of substrate molecules each enzyme site converts to product per unit time). K_M, also known as the Michaelis constant, shows the concentration of substrate at which the enzymatic reaction proceeds at half its maximal rate. The lower the K_M value, the faster the enzymatic reaction will tend to proceed. However, obviously, the kcat value is important in this respect also, with the ratio kcat/K_M commonly used as a measure of overall enzymatic efficiency.

This mathematical representation is in contrast to the form of the Michaelis Menten reaction normally seen:

$$d\frac{[P]}{dt} = Vmax.\frac{[S]}{K_M + [S]}$$

where Vmax represents the maximum rate of the enzymatic reaction. However, contrary to common belief, Vmax (at least in this context) is not an absolute value, being dependent entirely on the concentration of enzyme used in the experiment in which it was measured. In other words, it is neither a constant (at least in the sense that it is thought to be), nor can it be said to be uniquely characteristic to that enzyme. (This confusion may arise through a mistaken belief that the Michaelis-Menten equation describes the activity of a single enzyme molecule or active site, rather than the aggregate activity of all available molecules - or active sites - of that particular enzyme. In other words, in common with the other reaction rate laws, it is a concentration-, rather than molecule-based, rate.)

This can be a trap for the unwary, as many modelling tools will ask the user to provide a Vmax value and these are readily available in the literature for many enzymes. But supplying such a value will lead to anomalous results, since, as a moment's glance at this second variant of the Michaelis-Menten equation will confirm, it contains no explicit term for the enzyme concentration. In other words, supplying a numeric Vmax value in this situation will mean that overall enzyme activity (i.e. the aggregate rate of product formation) will not vary with the quantity of enzyme available.

Instead, the modeller must substitute an expression of the form "kcat * [E]" (putting the appropriate kcat value and model species name for the enzyme concentration) in place of the Vmax term, just as if supplying values for the first version of the equation. In this way enzyme activity will vary with enzyme concentration in ones model, in much the same way that it varies with substrate concentration.

Specifying all the reactions between model species as a set of differential equations is clearly a critical activity in modelling any reaction-diffusion system but it is not sufficient alone. Obviously one must also specify what the model species (typically the proteins and other molecules involved in all the reactions) are, as well as their initial quantities. Another important consideration is what physical and logical compartments the model must contain to adequately reflect the underlying physiology. Sometimes a single compartment will suffice but, for instance, where one is trying to model a signal transduction process, whereby an extracellular stimulus invokes some intracellular activity after interacting with a receptor protein in the cell membrane, three compartments (extracellular, membrane and cytosolic) may be judged to be necessary. In almost all cases this will require also specifying a volume (or in some instances an area or other dimension) for each compartment and, in some cases, one will also need to specify diffusion coefficients determining the diffusion rate of each species within its compartment.

In practice, the process of specifying model reactions, model species, those species' quantities (and, if necessary, diffusion coefficients) and the model compartments and sizes, tends to be iterative in nature. Reasons for this include the fact that identical molecules appearing in different compartments have to be modelled as separate species, and the consequent necessity of modelling any process involving these separate species as distinct reactions.

Another reason why model specification may be iterative is because it is only after determining what the size of a particular compartment is that the initial quantity of certain species can be determined. Finally, whereas diffusion of a species within a compartment can often be handled by the modelling software (provided a diffusion coefficient or similar is provided), diffusion between compartments typically requires a separate transport reaction to be specified.

As each new parameter value is specified, the modelling software will add this new information into a complete mathematical description of the system. This will typically include a complete equation for each model species. Given that many species may be involved in a number of

different reactions (often as a reactant in one and a product in another) such combined equations can be quite long for some species. Thus, for the species A, involved in the following two irreversible mass action reactions:

B + C-> A

A + D -> E

with forward rate constants k1 for the first reaction and k2 for the second, the complete equation for A will be:

$$d\frac{[A]}{dt} = k_1.[B].[\mathrm{C}] - k_2.[\mathrm{A}].[D],$$

Here the equation is specifying that the rate of creation is determined by the concentrations of species B and C times the k1 rate constant, with the rate of destruction of A determined by the concentration of itself and D, times the k2 rate constant. In most instances a complete set of these equations (one for each of the model species involved in reactions) will form the largest element of the resulting mathematical model.

Having fully specified the model, it is then a matter of running it in order to study its behaviour. This essentially involves simultaneously calculating all the changes occurring in the modelled system (as determined by all the reaction equations) as time progresses, a process known as integration. Whilst single differential equations can be integrated analytically (using methods such as u-substitution and similar) to provide solutions applicable over continuous timescales, this is not practicable for sets of differential equations, as is the case here. Instead, approximations, applicable over discrete timescales, have been devised that are readily amenable to being handled by computers. Collectively these are known as numerical methods, with the computer programs that apply them being known as solvers.

Over the years, a variety of such numerical methods have been devised, each with strengths and weaknesses, in terms of relative efficiency, accuracy and stability. For instance, so-called fixed-step solvers involve numerical methods that calculate the overall system state at fixed time intervals. Such methods are relatively straightforward to implement but often suffer from the fact that the time interval chosen may be too long when parts of the system are changing rapidly and too short when nothing much is changing at all. The result is that too much information may be lost in the first case and too much unnecessary information produced in the second.

Alternative numeric methods have been devised that use variable time steps, whose length is determined by the underlying variability of the data. These tend to be faster or more accurate (or both) than fixed-step solvers, depending on overall system variability. As a rule they require the user to specify a maximum time interval (or "step size") and two different tolerance values for estimated (local) error, one relative and one absolute. The relative tolerance specifies the maximum estimated error value tolerated for any calculation of a system state variable, as a proportion of its calculated value. The absolute tolerance specifies the maximum estimated error tolerated in absolute terms, and is mainly intended for situations where the variable in question has been calculated as having a value of zero, in which case the relative tolerance is not calculable, or close to zero, in which case it may not be appropriate.

Put simply, having calculated the system state at some time point, the solver will then calculate the state at a time point in the future, according to the maximum time interval allowed. If, during the subsequent calculations, the value for any system variable is estimated to have an error in excess of both of the two specified tolerances, any further calculations will be abandoned and an earlier time point in the future will be chosen at which to restart the calculations. The underlying logic is that shorter time intervals after the last system state calculation will tend to be associated with less change in state variables, meaning that values can be calculated more confidently (i.e. with less likelihood of error).

Only once the calculation has determined that none of the associated local errors exceed both of the tolerance can the solver proceed with calculating the next time point in the future, using the maximum permitted time step from the current time point, as described earlier. Otherwise it must keep shortening the step size until either all such local errors are estimated to be below the required limits, or else the next calculated time step falls below some specified system minimum.

Alternatively, it may exceed some system-designated time limit without reaching a solution. In either of the latter cases the solver will typically report an error and abandon any further calculations, a scenario often referred to as a convergence failure. Essentially this means that the general error (the accumulated local errors), is either rising with time or failing to decrease.

Some systems (known as “stiff systems”) can be particularly problematic in this last respect. Whilst there is some disagreement as to what this term means, exactly, it is frequently encountered in systems where, at certain time points, some system variables are displaying large, rapid changes in value, whereas others are barely changing value at all. Situations like these are often encountered in systems of chemical reactions, such as might be seen within cells.

Fortunately, there are numerical methods that deal well with these situations and, in practice, the standard solvers offered by most modelling tools are chosen to be able to operate well over a wide range of conditions, including all but the stiffest of systems.

1.2.6.2 Partial differential equations: their use in modelling reaction diffusion systems

in the previous section it was shown how a compartmental model could be developed using only ODEs. As explained earlier, this is possible because in compartmental models one is generally only interested in changes of a dependent variable (concentration) with respect to one independent variable (time) for each species. (This implies that each species is uniformly distributed within its respective compartment.) However, with spatial models of reaction-diffusion systems this assumption no longer holds. Indeed, the whole point of such models is to see how species concentration varies at each point in space, as well as with time. This means that the model must now take into account changes in concentration with respect to more than one independent variable, time plus however many spatial dimensions we are interested in, meaning four independent variables in the case of a 3D spatial model. This requires the model to be based on partial differential equations (PDEs) in which the derivative with respect to one changing independent variable is defined, while holding the other independent variables constant.

In a spatial model one is interested at each moment in time in the local species concentrations at each of a set of discrete points within each compartment. Given that these species concentrations will not be the same at each point, one must therefore take into account diffusion, in which species tend to move from areas of higher to lower concentration at a characteristic rate. Thus, in calculating the new local concentration of a species at a given point in space and time, one must take into account its current concentration, any changes resulting from reactions in which it is involved in (either as a product or reactant, with the rate of such reactions depending on the local concentration of reactants), plus any changes resulting from diffusion.

For each species diffusion can be represented mathematically by the equation:

$$\partial \frac{[Species]}{\partial \mathrm{t}} = \mathrm{D}.\Delta\,[\mathrm{Species}]$$

in which [Species] represents the concentration of the particular species in question, D is the diffusion coefficient for that species and Δ, the Laplace operator (sometimes written ∇^2) denotes the sum of partial second derivatives of [Species) with respect to each spatial dimension.

This diffusion equation is added to the combined equation for all the reactions this species is involved with, to give the combined equation for that species. So, to take the same example used in Chapter 2, where species A is involved in the following two irreversible mass action reactions (with respective forward rate constants of k1 and k2):

B + C-> A

A + D -> E

when put together with the relevant diffusion equation, gives the combined equation for species A:

$$\partial \frac{[A]}{\partial t} = k_1.[B].[\mathrm{C}] - k_2.[\mathrm{A}].[D] + \mathrm{D}.\Delta\,[\mathrm{A}]$$

Here, the difference with ODE/compartmental models is that this equation is applied to each discrete point in the problem space, whereas in ODE models it is applied to the whole compartment in which the species resides. But, in common with ODE models, the set of such combined equations for each of the model species typically forms the mathematical core of PDE-based spatial models. The difference, of course, is that, in addition to these equations, geometric specification of the various model compartments adds a great deal more complexity to the model, which is likely to be reflected in its underlying mathematical representation, as compared to the simple volumetric representation of compartments in ODE models.

Again, as with ODE models, exact, continuous solutions for these sets of PDEs by symbolic means are not practical. Instead, approximate, discretised solutions are derived by the use of numerical

methods. These methods fall into four general categories: finite element, finite difference, finite volume and mesh-free methods. The finite element method breaks down the continuous problem space into smaller discrete elements and finds solutions for these in terms of algebraic and ordinary differential equations. The finite difference method does something similar but finds approximate solutions in terms of finite difference equations, whilst the finite volume method breaks a problem space down into a set of discrete cells or volumes and then finds approximate solutions for each volume in terms of a set of algebraic equations representing flows (from adjacent volumes) through each of its volume surfaces. Mesh-free methods are used where discretising the continuous problem space into a mesh of discrete elements with a predefined set of neighbours (as in the methods described above) may be inappropriate, such as in moving, deformable bodies.

1.2.6.3 Previous attempts to model FAs, and how they compare with the proposed model

Judging from the literature, there have not been any previous attempts to exclusively model whole FAs, except at the level of one of two FA proteins, and certainly nothing of the same scale and ambition. Most such modelling has focused on various aspects of actin dynamics [153,154,155,1], using deterministic mathematical models, mainly ODE based. There have, however, been other modelling approaches used, looking at other FA-related areas, although mostly not with FAs as their primary focus.

Amongst these are an ABM-based study of integrin clustering [157], two Boolean models of growth-factor receptor signaling that include a number of FA-associated proteins, such as Src, integrins, FAK, DOCK1, Rac1, and Cdc42[142,158], and another Boolean model, this time of platelet signaling that includes integrins and talin[159]. There has also been a novel rule-based model of the roles of

In 742.3 [1156]

talin, Dok1, and PIPKI in integrin signaling dynamics [160]. Other than these, there have been a number of FA-related mathematical models, including a hybrid deterministic/stochastic model of integrin-mediated signalling and adhesion dynamics[161], a model of Rho GTPase cycling [162], an algebraic model of FA-related mechanical stress [163], and a model of cell migration that incorporates various aspects of FA dynamics [164].

On the whole, though, it is fair to say that there have been comparatively few attempts to model FAs in their entirety.

1.3. Overall aims of the project

1.3.1 Primary aims

The primary aim of this Ph.D. project is to develop a model of FAs in stages, so as to develop insights into the functioning of FAs in the wider context of lamellipodia formation (with which they are primarily associated, along with filopodia formation), as well as in their isolated form. Important questions to be answered include what are the primary drivers of FA formation, in terms of determining both the rate at which they form and the overall levels achieved, what are the other major factors helping to determine these formation rates and levels, how are these different drivers and factors regulated, and how do they interact with each other. The overall impression gained from the literature is that FA regulation is a highly complex process. Hopefully, this model will, by the end, have revealed much more about this overall process, but also have indicated what the key regulatory processes are, what more needs to be discovered about such regulation, and what areas of investigation should be prioritised.

In terms of the model itself, this will start with a working submodel of the early stages of inside-out signalling, as described above (Figure 1.3), in this case involving the chemokine SDF1, acting on its cognate G protein-coupled receptor CXCR4 [165]. This will result in a signalling cascade leading to Rac1-mediated lamellipodia formation, as described briefly above. The aim at the first stage is to develop this lamellipodia submodel incrementally, to the point where Rac1 activation is generating actin elongation and cross-linking. This part of the project will be described in Chapter 2.

The next stage will involve the development of a working submodel of FA dynamics, including FA assembly and disassembly, and the many aspects that regulate these two processes. It is intended that eventually it will take outputs from the first inside-out lamellipodia model, including PIP2 and actin as critical inputs, relieving the submodel of some unnecessary complexity. But

initially these will be supplied at a suitable rate, as determined by the lamellipodia model and by evidence from the literature. This stage will be described in Chapter 3.

Finally, the two sub-models will be linked, so that the outputs from the lamellipodia model directly inform the FA model, forming a complete modular model of inside-out signalling. This process will be described in Chapter4.

1.3.2 Modelling software used

During the course of the project, two different differential equation modelling tools will be employed, COPASI and Virtual Cell. Whilst both can be used to develop compartmental/ODE-based models, Virtual Cell, unlike COPASI, also allows spatial/PDE modelling (in 1D, 2D and 3D form), something that will be illustrated briefly in Chapter 3. This Virtual Cell model might offer additional insights into aspects of FA dynamics not revealed by compartmental modelling approaches.

It is intended that this dual use of compartmental and spatial modelling will be continued into the third and final combined model to obtain the maximum benefits of both, whilst avoiding the excessive computational overhead of attempting to model the whole in spatial form. In other words, only the specific aspects that will clearly benefit from spatial modelling will adopt this approach. However, this will involve establishing suitable criteria for determining where these benefits are most likely to be realised.

Hopefully, by the time the reader has finished reading this report, they will have a much better appreciation of the challenges and benefits of using mathematical modelling approaches in the study of FAs and of cell biology in general. Hopefully, and more importantly too, they will have learnt much more about the many different aspects of FA behaviour.

Chapter 2. Phase 1 - The SDF1-CXCR4 chemotaxis lamellipodia model

2.1 Introduction

Chemotaxis, the induced movement of a motile cell (or larger organism) along a concentration gradient of an attractant substance, commonly involves a class of receptors called G protein coupled-receptors (GPCRs) [166]. These are transmembrane receptors that respond to various signals outside the cell (including neurotransmitters, hormones and other external stimulants) by acting as guanine nucleotide exchange factors (GEFs) for an associated trimeric G protein. Such G proteins are made up of three molecules (named α, β and γ), with the last two collectively forming a functional subunit. The α subunit can bind either a GTP or GDP nucleotide [166,167].

Whilst bound to the GDP, the α subunit remains tightly bound to the βγ subunit, in which form they are all inactive [166]. However, activation of an associated GPCR causes it to facilitate the exchange of this GDP for a GTP nucleotide, leading to dissociation of the attached α subunit from its βγ subunit partner [166]. In this dissociated form both subunits can then activate, or otherwise affect, other cellular proteins. Two of the principal downstream pathways which GPCRs affect are the cAMP [166,167,168] and the phosphatidylinositol signal pathways [169,166]. Which of these two (or other) pathways are affected is principally determined by the properties of the α subunit that has been activated [169].

For instance, in the case of the pathway being modelled here, a chemotactic protein SDF1 (stromal cell-derived factor 1, also known as CXCL12) principally binds to a GPCR called CXCR4 [165]. Such binding activates CXCR4, resulting in the exchange of GDP for GTP (principally) in two different classes of α subunit , one of which, the αi2 subunit, appears to be primarily involved in SDF1-induced chemotaxis [170]. However, in some cases, such as in this instance, the dissociated βγ subunit (Gβγ) can also be involved in downstream activity [171,169], including Gβγ-mediated activation of PI3Ks [171,169].

In this first part of the project we are modelling the Gβγ-associated role in SDF1-mediated chemotaxis, as a typical example of inside-out FA signalling. More specifically, this particular pathway involves PI3K-mediated phosphoinositide signalling, involving the phosphorylation of PIP2 to PIP3, with consequent activation of Rac1, resulting in the lamellipodia formation characteristic of chemotaxis, and of inside out signalling generally. As stated in the previous chapter, the overall object here is to have a working lamellipodia submodel, including appropriate phosphoinositide signalling, whose outputs can be used to inform a separate FA submodel, to be implemented in the next phase. SDF1-CXCR4 chemotactic signalling has been chosen for this reason, abbreviated to exclude those parts of the pathway not immediately relevant to the intended submodel, and to keep the submodel relatively manageable.

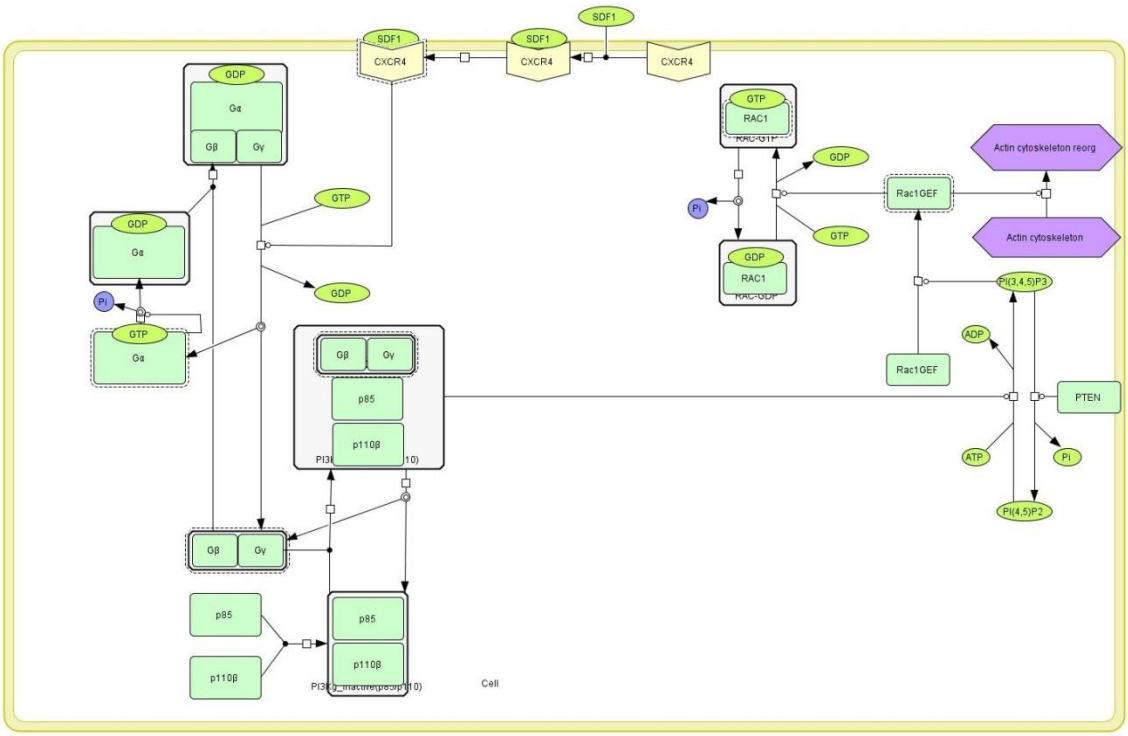

Figure 2.1 Schematic depiction of SDF1-CXCR4 chemotaxis mechanism, as used in this lamellipodia submodel. (See below for detailed explanation. Rac1-mediated actin remodelling steps not included in this diagram, for the sake of conciseness.) Arrows represent changes in state of model species. (For instance, the change in state from PI(3,4,5)P3 to PI(4,5)P2 on the right-hand side of the diagram.) Lines ending in circles denote that a change in state is modulated by another model species, e.g. the enzymatic modulation of the above reaction by PTEN.

This pathway is summarised in Figure 2.1 above, in somewhat simplified form.

Here, SDF1 binding to the CXCR4 receptor leads to CXCR4 activation, resulting in GTP exchange for GDP within its associated G protein and subsequent dissociation of the latter. The dissociated Gβγ subunit then binds to a Class IA phosphatidylinositol kinase, PI3Kβ, composed of the p85 and p110β subunits [172], activating it [173]. This in turn phosphorylates PIP2 to PIP3, with the resulting PIP3 binding to and activating GEFs (Rac1GEFs) specific to the monomeric GTPase Rac1 [174]. As with GPCRs, these GEFs exchange GTP for the GDP nucleotide associated with the inactive form of Rac1, activating the latter. Rac1 is well-documented in the literature as a potent activator of lamellipodia formation (a key process not only in chemotaxis and other forms of cell migration, but also in cell spreading [175,176]). As explained earlier, It does this by activating the WAVE regulatory complex [177], thereby driving Arp2/3-mediated actin elongation and, possibly, branching [178]. The resulting cytoskeletal rearrangement [176,112] provides the necessary force for lamellipodia extension.

Finally, just as PI3K phosphorylates PIP2 to PIP3, another phosphatase, PTEN, dephosphorylates PIP3 back to PIP2, thus tending to down-regulate lamellipodia formation and, therefore, also chemotaxis [179].

2.2 Methods

2.2.1 Introduction

In the previous chapter it was shown how a reaction-diffusion system, such as the one described here, can be modelled as a set of differential equations, model species (in this context all the proteins and other molecules, including lipids, in the system being modelled), and additional parameters. These equations specify all the various reactions possible between model species. It was then shown how various simulations of such models can be performed (or "run") by integrating all these reaction equations using an appropriate solver.

2.2.1.1 Plots and reports and parameter scans

It then remains to analyse the results. One of the key tools for this are plots (i.e. graphs), which all modelling tools will normally provide. Thus, for instance, COPASI can generate a number of predefined plots as standard, including plots of particle (i.e. molecular) numbers, concentrations, particle rates and concentration rates (i.e. the rates at which these are changing over time) and reaction fluxes (the rate at which product is being created from reactants over time).

Such plots can be very useful in analysing the behaviour of the model, particularly over time, and many modelling tools (including COPASI and Matlab) allow the user to tailor their own plots from the underlying output data. Other alternatives include the ability to export the data in formats suitable for re-import into spreadsheet and graphing applications, and the ability to generate reports from it. Such reports may come in predefined formats (as is the case with COPASI, which allows particle numbers, reaction fluxes and the like to be output in report form rather than as plots) or it may be possible for the user to define their own report formats.

However, plots and reports are often not enough on their own if one is to fully understand how a model behaves. Other forms of analysis can offer further insights, two of the most important

being parameter scans and sensitivity analysis. Parameter scans allow the user to explore how system outputs vary with changes in system inputs. So, for instance, the user might choose to set up a scan in which the initial concentration of some model species is increased in set increments along a range. Alternatively, some model parameter value may be increased logarithmically or sampled randomly within a set range or, in the case of a nested scan, two or more different parameter values might be altered concurrently. In this respect, parameter scans only automate what the user could also achieve by manually altering such parameter values and then generating plots or reports to see the results, albeit with more time and effort required.

Sensitivity analysis, as its name implies, measures the sensitivity of some user-specified system output to changes in various system inputs. Such inputs may also be user-specified, and may include initial concentrations of species, specific reaction rate constants, compartment volumes or other system variables and parameters. Output values may include steady-state or peak levels of some reaction product, or overall levels of some other species over a given time course.

Various mathematical techniques can be used to produce a sensitivity value for each input parameter, describing how much the chosen output value changes in response to a change in the input value. So, for instance, if the concentration of a product falls substantially in response to a slight increase in one of its associated reaction rate constants, one would expect sensitivity analysis to assign a high negative value to this particular rate constant. (Here the negative sign is only denoting that the output value is changing in the opposite direction to the input value.) By contrast, if the same product concentration increased only very slightly in response to a large increase in the initial concentration of one of its reactants, one would expect sensitivity analysis to assign a low positive value to this parameter.

2.2.1.2 Sensitivity analysis and parameter estimation

Sensitivity analysis can be very useful in analysing complex models, where parameter scanning and the study of plots can lead to confusion and where it can be hard to determine which of the many potential input values to focus ones attention on. It can also be very important when it comes to optimising the model. For instance, with complex models it is usually very hard to find reliable values for various parameters, including initial concentrations and reaction rate constants in the literature and in many cases the modeller is forced to rely on educated estimates. Even where such values are available, their inclusion may lead to the model behaving differently to how the system behaves physiologically.

It is common in such circumstances to resort to parameter fitting, which is to say automatically adjusting input values until the model's overall behaviour is deemed acceptable. (I.e., until, overall, the model reliably produces the desired outputs.) Sensitivity analysis can be critical here because it can identify a small subset of parameter values that may require only modest changes to steer the model's behaviour in the desired direction. This enables the modeller to constrain parameter fitting to this subset, which can greatly increase the chances of success.

Ideally, though, such artificial fitting of parameters should be kept to a minimum and, if possible, avoided altogether. Again, sensitivity analysis can be very useful in this regard, because, by indicating the parameters to which the model is most sensitive, it allows the modeller to focus most of their time and attention on finding reliable published values for these, whilst avoiding excessive time spent trying to find the most reliable values for other parameters, when in fact these values barely effect overall model behaviour. Since, as mentioned earlier, it often turns out that only a rather small subset of input parameters has a major effect on the principal model outputs, there can be large savings in time. An additional benefit is the increased confidence that, if one has found reliable values for these inputs, any differences seen between the model's behaviour and that of the system it is modelling, are likely to be worth investigating.

2.2.1.3 Other analysis methods

In most cases plots, reports, parameter scans and sensitivity analysis provide sufficient tools for analysing most aspects of model behaviour. However, modelling software environments may provide other tools, such as steady-state and stability analysis, metabolic control analysis and similar. As well as providing different methods for analysis, methods for improving the model may be provided. For instance, COPASI, provides optimisation, which is similar to parameter estimation, if rather more specific in the outputs it focuses on.

2.2.2 Analysis methods used

A broad range of standard COPASI time course plots were employed, namely particle number, concentration, particle number rate, concentration rate and reaction flux. For the sake of brevity, not all of these are shown in the Results section (2.4). (No user-defined plots were required.)

Various parameter scans were performed (all using the Time Course task), with settings described in the Results section. Again, and for the same reasons, not all the results of these are shown in the Results section. Sensitivity analysis was also performed, based on the Time Series subtask, with Effect set to Single Object and Cause set to All Parameter Values. The Delta factor value was 0.01, and the Delta minimum was 1e-12.

Parameter estimation was not attempted, for reasons explained later.

2.2.3 Modelling application version number

COPASI: Version 4.22 (Build 170)

2.2.4 System units used

Time units:	Seconds (s)
Quantity units:	Femtomoles (fmol)
Volume units:	Femtolitres (fl)

2.2.5 Solver settings

The Deterministic (LSODA) solver was used for all Time Course simulations, with a relative tolerance value of 1e-06 and an absolute tolerance of 1e-12. Maximum internal steps were set at 10,000. Time Course durations were typically of 20,000 seconds, with 2000 intervals, and therefore an interval size of 10 seconds.

The next section describes how the SDF1-CXCR4 chemotaxis model was implemented, using the compartmental/ODE modelling tool COPASI, as described briefly in section 1.2.4.2 of Chapter 1.

2.3 Model setup

2.3.1 Introduction

The SDF1-CXCR4 lamellipodia submodel just described is modelled as a set of 31 ordinary differential equations, each defining how the concentration of one or more products is determined by the concentrations of reactants. For the CXCR4-G protein interaction, this actually results in a total of 18 different reactions, covering all of the various permutations, in accordance with a reaction schema first proposed by Kinzer-Ursem and Linderman[180], in many cases using parameter values specified in a model of platelet phosphoinositide and calcium regulation [181]. Whilst this platelet model was specifically modelling the G protein-coupled responses of an ADP-binding GPCR rather than CXCR4 (and acting on a Gαq- rather than a Gαi-family subunit), this seemed like the best option in the absence of values specific to CXCR4. (Specific rate values for Gαi association and dissociation from Gβγ and for CXCR4-SDF1 binding, taken from the literature, were incorporated in place of the values in the platelet model.)

Of the remaining 13 reactions, 6 relate to activation and binding of PI3K, 2 to activation, binding and degradation of PTEN, and one each to phosphorylation of PIP2 (by activated PI3K) and dephosphorylation of PIP3 (by activated PTEN). The final 3 relate to activation of Rac1GEFs, and activation (by activated Rac1GEFs) and degradation of Rac1.

The model also incorporates 32 different species, of which nearly two-thirds (20) are involved in CXCR4-SDF1 binding and associated G protein events. The other 12 species are involved in phosphoinositide modification or in Rac1 activation. Non-zero copy numbers of all proteins, other than SDF1, were taken from a published HeLa cell quantitative proteome [182], with Rac1GEF numbers taken as the sum of copy numbers of TIAM1, Vav2, Vav3, ECT2, ARHGEF4, Trio and PRex1, all reported in the literature to be GEFs for Rac1. SDF1 concentrations were specified

according to a published value for an optimal chemoattractant gradient for SDF1 (300 ng/ml, or 2.2e-08 M), with this value held fixed to ensure that the model is continuously driven by SDF1.

Whilst the HeLa proteome lists G protein subunits individually, this model starts with 340,000 molecules of the whole GDP-bound trimeric protein, with all other G protein-associated model species starting with zero copy numbers. This is to ensure, as much as possible, that the model is driven by Gβγ solely resulting from CXCR4-SDF1 ligation. (This copy number, i.e. 340,000, is taken from the value reported for Gαi subunit proteins in the proteome.)

Copy numbers of PI(4,5)P2 were given a large value of 100,000 molecules fixed to ensure a ready supply of PIP2 for phosphorylation by PI3K to PIP3, and to ensure, as much as possible, that PIP2 levels do not influence the model in anyway. For similar reasons, GTP and GDP nucleotide concentrations (1 mM) were specified to be in excess. Where copy numbers were initially specified, concentrations were automatically calculated by COPASI (i.e. by dividing the specified copy number by the relevant cell compartment volume). Similarly, where concentrations were initially specified (i.e. SDF1, GTP and GDP), copy numbers were automatically calculated by COPASI from these values.

The resulting model species, reactions and other parameters are shown below.

2.3.2 Model species, their copy numbers and concentrations

These species, with their copy numbers and concentrations, are listed in Table 2.1 below, with all model reactions listed in Table 2.2.

Species	Compartment	Copy number	Concentration (M)
CXCR4	PM	100,000 *	2.6e-05
CXCR4*	PM	0	0
CXCR4*.Gi	PM	0	0
CXCR4*.Gi.GDP	PM	0	0
CXCR4*.Gi.GTP	PM	0	0
CXCR4*.SDF	PM	0	0

CXCR4*.SDF.Gi	PM	0	0
CXCR4*.SDF.Gi. GDP	PM	0	0
CXCR4*.SDF.Gi. GTP	PM	0	0
CXCR4. Gi. GDP	PM	0	0
CXCR4.SDF	PM	0	0
CXCR4.SDF.Gi. GDP	PM	0	0
Gα.GDP	Cytosol	0	0
Gα.GTP	Cytosol	0	0
Gβγ	Cytosol	0	0
Gi.GDP	Cytosol	340,000 * †	3.76e-07
Gi.GTP	Cytosol	0	0
GDP	Cytosol	~9.0e+08 **	0.001
GTP	Cytosol	~9.0e+08 **	0.001
PI(3,4,5)P3 (PIP3)	PM	0	0
PI(4,5)P2 [PIP2]	PM	100,000 (fixed) **	2.6e-05 Fixed
PI3K	Cytosol	2800 *	3.1e-09
PI3K*.Gβγ	Cytosol	0	0
PI3K*.Gβγ.PM	PM	0	0
PI3K.PM	PM	0	0
PTEN	Cytosol	3200 *	3.54e-09
PTEN*.PM	PM	0	0
Rac1	Cytosol	330,000 *	8.56e-05
Rac1*	PM	0	0
Rac1GEF	Cytosol	39,000 * ‡	4.32e-08
Rac1GEF*.PIP3	PM	0	0
SDF1	Extracellular	~660,000 (fixed)	2.2e-08 (fixed) ***

Table 2.1 – Table of species used in the COPASI SDF1 chemotaxis model

Key:

* = Protein copy numbers taken from a published HeLa cell quantitative proteome. Source: Nagaraj, N., Wisniewski, J.R., Geiger, T., Cox, J., Kircher, M., Kelso, J., Pääbo, S., Mann, M., 2011. Deep proteome and transcriptome mapping of a human cancer cell line. Mol. Syst. Biol. 7, 548. https://doi.org/10.1038/msb.2011.81.

** = Arbitrary copy number, chosen to ensure that the species concerned is constantly in excess.

*** = Calculated from a published value for an optimal chemoattractant gradient for SDF1 (300 ng/ml), based on a molecular weight of 13.7 kDa. In this model this concentration value is held fixed to ensure that the model is continuously driven by SDF1 in excess. Source: [183]

† = Copy number assumed to be same as that for Gαi subunit protein, as taken from the HeLa quantitative proteome above.

‡ = Sum of copy numbers for the Rac1-associated GEFs TIAM1, Vav2, Vav3, ECT2, ARHGEF4, Trio and PRex1, as taken from the same published HeLa quantitative proteome mentioned above.

2.3.3 Model reactions, their rate laws and rate constant values

Here an * after the species name indicates that it is in activated form

Reaction	Description	Rate law	Parameter values
CXCR4 + SDF1 = CXCR4.SDF1	CXCR4 binding SDF1	Mass action (reversible)	420,000 M-1 s-1 0. 00824 s-1 *
CXCR4* + SDF1 = CXCR4*.SDF1	CXCR4* binding SDF1	Mass action (reversible)	1.407e+06 M-1 s-1 0. 00155 s-1 *
CXCR4.Gi.GDP + SDF1 = CXCR4.SDF1.Gi.GDP	CXCR4.Gi.GDP binding SDF1	Mass action (reversible)	420,000 M-1 s-1 0. 000158 s-1 *
CXCR4*.Gi.GDP + SDF1 = CXCR4*.SDF1.Gi.GDP	CXCR4*.Gi.GDP binding SDF1	Mass action (reversible)	1.407e+06 M-1 s-1 0. 000158 s-1 *
CXCR4.SDF1 = CXCR4*.SDF1	Activation of CXCR4.SDF1	Mass action (reversible)	26.465 s-1 79,000 s-1 **
CXCR4.SDF1.Gi.GDP = CXCR4*.SDF1.Gi.GDP	Activation of CXCR4.SDF1.Gi.GDP	Mass action (reversible)	52.298 s-1 79,000 s-1 **
CXCR4.Gi.GDP = CXCR4*.Gi.GDP	Activation of CXCR4.Gi.GDP	Mass action (reversible)	1725.7 s-1 79,000 s-1 **
CXCR4* + Gi.GDP = CXCR4*.Gi.GDP	CXCR4* binding Gi.GDP	Mass action (reversible)	3.9058e+06 M-1 s-1 18.4375 s-1 **
CXCR4*.SDF1 + Gi.GDP = CXCR4*.SDF1.Gi.GDP	CXCR4*.SDF1 binding Gi.GDP	Mass action (reversible)	3.9058e+06 M-1 s-1 0.199 s-1 **
CXCR4*.Gi + GTP = CXCR4*.Gi.GTP	CXCR4*.Gi binding GTP	Mass action (reversible)	100,000 M-1 s-1 8 s-1 **
CXCR4*.SDF1.Gi + GTP = CXCR4*.SDF1.Gi.GTP	CXCR4*.SDF1.Gi binding GTP	Mass action (reversible)	100,000 M-1 s-1 8 s-1 **
CXCR4*.Gi.GDP = CXCR4*.Gi + GDP	CXCR4*.Gi.GDP releasing GDP	Mass action (reversible)	17.8 s-1 1e+06 M-1 s-1 **

CXCR4*.SDF1.Gi.GDP = CXCR4*.SDF1.Gi + GDP	CXCR4*.SDF1.Gi.GDP releases GDP	Mass action (reversible)	17.8 s-1 1e+06 M-1 s-1 **
CXCR4*.Gi.GTP = CXCR4* + Gi.GTP	CXCR4*.Gi.GTP releasing Gi.GTP	Mass action (reversible)	850 s-1 1e+07 M-1 s-1 **
CXCR4*.SDF1.Gi.GTP = CXCR4*.SDF1 + Gi.GTP	CXCR4*.SDF1.Gi.GTP releases Gi.GTP	Mass action (reversible)	850 s-1 1e+07 M-1 s-1 **
Ga.GTP + Gbg = Gi.GTP	Activated Gi subunit association (to Gi.GTP)	Mass action (reversible)	700,000 M-1 s-1 0.002 s-1 ***
Ga.GDP + Gbg = Gi.GDP	Unactivated Gi subunit association(to Gi.GDP)	Mass action (reversible)	700,000 M-1 s-1 0.002 s-1 ***
Gi.GTP -> Gα.GDP + Gbg	Gi.GTP hydrolysis	Mass action (irreversible)	0.0008 s-1 **
PI(4,5)P2 -> PI(3,4,5)P3; PI3K*.Gbg.PM	Phosphorylation of PIP2 to PIP3 by active PI3K	Michaelis-Menten (irreversible)	kcat = 0.03 s-1 † KM = 7.2e-05 M ††
PI(3,4,5)P3 -> PI(4,5)P2; PTEN*.PM	Hydrolysis of PIP3 to PIP2 by activated PTEN	Michaelis-Menten (irreversible)	kcat = 1.2 s-1 ††† KM = 8.96e-06 M ††††
PI3K + Gbg = PI3K*.Gbg	PI3K binding to and being activated by Gbg	Mass action (reversible)	10,000 M-1 s-1 ‡ 0. 1 s-1
PI3K. PM + Gbg = PI3K*.Gbg	Membrane-bound PI3K activated by Gbg	Mass action (reversible)	10,000 M-1 s-1 ‡‡ 0.1 s-1
PI3K + Rac1 = PI3K.PM + Rac1	Inactive PI3K recruited to PM by inactive Rac1	Mass action (reversible)	100,000 M-1 s-1 ‡‡ 0.15 M-1 s-1
PI3K + Rac1* = PI3K.PM + Rac1*	Inactive PI3K recruited to PM by active Rac1	Mass action (reversible)	100,000 M-1 s-1 ‡‡ 0.15 M-1 s-1
PI3K*.Gbg + Rac1 = PI3K*.Gbg.PM + Rac1	Active PI3K recruited to PM by inactive Rac1	Mass action (reversible)	100,000 M-1 s-1 ‡‡ 0.15 M-1 s-1

PI3K*.Gbg + Rac1* = PI3K*.Gbg.PM + Rac1*	Active PI3K recruited to PM by active Rac1	Mass action (reversible)	100,000 M-1 s-1 0.15 M-1 s-1	‡‡
PTEN + PI(3,4,5)P3 = PTEN*.PM + PI(3,4,5)P3	Binding and activation of PTEN by PIP3	Mass action (reversible)	1.2e+06 M-1 s-1 ‡‡‡ 0.0015 M-1 s-1	
PTEN*.PM ->	Spontaneous PTEN degradation	Mass action (irreversible)	0.15 s-1	‡
Rac1GEF + PI(3,4,5)P3 = Rac1GEF*.PIP3	Binding and activation of Rac1GEFs by PIP3	Mass action (reversible)	10,000 M-1 s-1 0.1 s-1	‡
Rac1 + Rac1GEF*.PIP3 -> Rac1* + Rac1GEF*.PIP3	Activation of Rac1 by activated Rac1GEFs	Mass action (irreversible)	10,000 M-1 s-1	‡
Rac1* -> Rac1	Spontaneous inactivation of Rac1	Mass action (irreversible)	0.001 s-1	‡

Table 2.2. List of all 31 reactions featuring in the COPASI SDF1 chemotaxis model.

Key:

* = On- and off-rates based on published value from surface plasmon resonance study of CXCR4-SDF1 binding. Adjusted as appropriate in some cases to allow for positive effects of prior receptor activation and/or G protein coupling, with these additional coefficient values taken from Purvis Platelet model described in more detail below. Sources: Vega, B., Munoz, L.M., Holgado, B.L., Lucas, P., Rodriguez-Frade, J.M., Calle, A., Rodriguez-Fernandez, J.L., Lechuga, L.M., Rodriguez, J.F., Gutierrez-Gallego, R., Mellado, M., 2011. Technical Advance: Surface plasmon resonance-based analysis of CXCL12 binding using immobilized lentiviral particles. J. Leukoc. Biol. https://doi.org/10.1189/jlb.1010-565. Also, Purvis, J.E., Chatterjee, M.S., Brass, L.F., Diamond, S.L., 2008. A molecular signaling model of platelet phosphoinositide and calcium regulation during homeostasis and P2Y1 activation. Blood 112, 4069–4079. https://doi.org/10.1182/blood-2008-05-157883

** = Parameter values directly taken from a platelet model by Purvis and colleagues, itself based on a generic GPCR equilibrium cubic ternary complex activation model proposed by Kinzer-Ursem and Linderman. Sources: Kinzer-Ursem, T.L., Linderman, J.J., 2007. Both Ligand- and Cell-Specific Parameters Control Ligand Agonism in a Kinetic Model of G Protein–Coupled Receptor Signaling. PLOS Comput. Biol. 3, e6. https://doi.org/10.1371/journal.pcbi.0030006. Also, Purvis, J.E., Chatterjee, M.S., Brass, L.F., Diamond, S.L., 2008. A molecular signaling model of platelet phosphoinositide and calcium regulation during homeostasis and P2Y1 activation. Blood 112, 4069–4079. https://doi.org/10.1182/blood-2008-05-157883

*** = Off-rate value inferred from published Kd and on-rates. Source: Sarvazyan, N.A., Remmers, A.E., Neubig, R.R., 1998. Determinants of Gi1α and βγ Binding measuring high affinity interactions

in a lipid environment using flow cytometry. J. Biol. Chem. 273, 7934–7940. ttps://doi.org/10.1074/jbc.
273.14.7934.

† = Converted from specific activity value for p110ß/p85a of 16 nmol/min/mg and molecular weight of 110 kDa. Source: PI3K (p110 beta/p85 alpha), active (I1036) - Datasheet - i1036-061m0726dat.pdf [WWW Document], n.d. URL http://www.sigmaaldrich.com/content/dam/sigma-aldrich/docs/Sigma/Datasheet/9/i1036-061m0726dat.pdf (accessed 8.29.15).

†† =Woscholski, R., Dhand, R., Fry, M.J., Waterfield, M.D., Parker, P.J., 1994. Biochemical characterization of the free catalytic p110 alpha and the complexed heterodimeric p110 alpha.p85 alpha forms of the mammalian phosphatidylinositol 3-kinase. J. Biol. Chem. 269, 25067–25072.

††† = McConnachie, G., Pass, I., Walker, S., Downes, C., 2003. Interfacial kinetic analysis of the tumour suppressor phosphatase, PTEN: evidence for activation by anionic phospholipids. Biochem J 371, 947–955.

†††† = Wu, C.-W., Bell, R., Storey, K.B., 2013. Regulation of PTEN function and structural stability in hibernating thirteen-lined ground squirrels. Cryobiology 66, 355–356. https://doi.org/10.1016/j.cryobiol.2013.02.055.

‡ = Arbitrary values. Chosen as relatively conservative and in the absence of appropriate values in the literature. PTEN degradation value based on various reports in the literature (including in HeLa cells) that this degradation is rapid after PTEN is recruited to the cell membrane, e.g. Das, S., Dixon, J.E., Cho, W., 2003. Membrane-binding and activation mechanism of PTEN. Proc. Natl. Acad. Sci. 100, 7491–7496. https://doi.org/10.1073/pnas.0932835100.

‡‡ = Arbitrary kon and koff values, chosen to yield a Kd value for binding of p110ß to Rac1 of approximately 1.5e µM, as measured in the article below.. The assumption is that, having been recruited to the PM by Rac1, the two then separate. (This is so as to reduce the number of subsequent reaction equations for Rac1 and PI3K.)Source: Fritsch, R., de Krijger, I., Fritsch, K., George, R., Reason, B., Kumar, M.S., Diefenbacher, M., Stamp, G., Downward, J., 2013. RAS and RHO Families of GTPases Directly Regulate Distinct Phosphoinositide 3-Kinase Isoforms. Cell 153, 1050–1063. https://doi.org/10.1016/j.cell.2013.04.031.

‡‡‡ = Parameters for membrane-PTEN binding, determined from SPR analysis. Source: Das, S., Dixon, J.E., Cho, W., 2003. Membrane-binding and activation mechanism of PTEN. Proc. Natl. Acad. Sci. 100, 7491–7496. https://doi.org/10.1073/pnas.0932835100. The additional assumption used in this model is that that PIP3 is, in some way, then activating PTEN. Source: McConnachie, G., Pass, I., Walker, S., Downes, C., 2003. Interfacial kinetic analysis of the tumour suppressor phosphatase, PTEN: evidence for activation by anionic phospholipids. Biochem J 371, 947–955.

2.3.4 Model compartments and their volumes

The model also has 3 compartments, representing the HeLa cell cytoplasm, the PM and the extracellular region. Based on the literature, typical volumes for HeLa cells range from between 500-4000 fl [184](having the same value in μm^3). For the sake of convenience, a value of 3000 fl, from the higher end of this range, was chosen. This choice was largely informed by a published value of 374 fl for the HeLa cell nucleus [2], an organelle that is commonly described as accounting for around 10% of the volume of eukaryotic cells [186]. Given that HeLa cells are a cancer cell line, and therefore likely to contain more DNA than noncancerous cell lines, such a nuclear volume seemed compatible with a cell of 3000 fl.

Having made further estimates concerning the overall volume of mitochondria, ER and ribosomes, based on typical values in the literature, and then subtracting the total of these from the cell volume to obtain the cytosolic volume, the following values were obtained:

Organelle /cell compartment	**Volume (fl)**
Nucleus	374
Mitochondria	600
ER (rough and smooth) & Golgi	480
Peroxisomes, lysosomes & endosomes	90
Cytosol	~1500
Total:	~3000

Table 2.3. Volumes of major organelles, as estimated for a HeLa cell with a hypothetical volume of 3000 fl. See above for how HeLa cell volume and nucleus were arrived at. Mitochondria, ER and Golgi estimates based on a report from the literature giving volumes of 20%, 10% and 6%, respectively, of the overall cell volume [187], with peroxisomes, lysosomes and endosomes reported as 1% each. Cytosol volume was derived by subtracting the volumes for the other organelles from the overall cell volume.

The extracellular compartment was given an arbitrary volume of 50,000 fl. This was chosen purely to allow levels of SDF1 to be massively in excess of CXCR4, whilst keeping its

[2]G. G. Maul and L. Deaven, 'Quantitative Determination of Nuclear Pore Complexes in Cycling Cells with Differing DNA Content', *The Journal of Cell Biology* 73, no. 3 (1 June 1977): 748–60, https://doi.org: for/10.1083/jcb.73.3.748.

concentrations at a level reported to be optimal for SDF1-mediated chemotaxis [183] (2.2e -08 M). In any case, actual effective volumes for the extracellular compartment are likely to vary substantially from case to case, making a reliable estimate unfeasible.

Lastly, the PM was given a volume of 6.4 fl. This estimate was based on a reported value of 1600 μm^2 in the literature [188]. This was then multiplied by another reported value of 4 nm for the cell membrane [189], to arrive at the PM volume just mentioned.

Taken together, the three compartment volumes used were as follows:

Compartment	Volume (fl)
Extracellular	50,000
PM	6.4
Cytosol	1500

Table 2.4. Summary of compartment volumes used in the SDF1-CXCL chemotaxis lamellipodia model. (See above for explanation of how these estimates were arrived.)

2.3.5 Model global variables

Finally, 28 global variables were used in this COPASI model. Fourteen of these had fixed values, being used to avoid having to enter the same local variable value repeatedly for the same reactions, with the increased risk of errors that might result. However, two were kcat values (for the PI3K and PTEN reactions needed for the Vinax calculation described below), with the rest being rate constant values for various CXCR4 G protein-related reactions, as taken from the Purvis model [181](using variables suggested by the Kinzer-Ursem and Linderman published GPCR model [180]). The remaining fourteen were assignment variables, meaning that their values were arithmetically derived from other values, in this case the Purvis global fixed variables just mentioned. Of these, all but two were used as rate constant values for CXCR4/G protein-related reactions, with the remaining two acting as functions that returned Vmax values for the PI3K and

PTEN reactions. (Where Vmax is derived by multiplying the relevant enzyme concentration by its corresponding kcat value, as described in the previous chapter).

These global variables are summarised below

Global variable name	Type	Value / Expression
Alpha	Fixed	3.35
Beta	Fixed	6.62
Gamma	Fixed	9.39
Delta	Fixed	9.85
Ka	Fixed	0 *
Kact	Fixed	0.0001
k1	Fixed	7.9
k3	Fixed	420000
k 11	Fixed	590000
alpha_k1	Assignment	alpha * k1
beta_k1	Assignment	beta * k1
alpha_k3	Assignment	alpha * k3
beta_k11	Assignment	beta * k11
k1/Kact	Assignment	k1 / Kact
alpha_beta_delta_k1	Assignment	alpha * beta * delta * k1
Kg	Fixed	32000
k11/ Kg	Assignment	k11 / Kg
delta_gamma_Kg	Assignment	delta * gamma * Kg
kGTP	Fixed	0.0008
k11/delta_gamma_Kg	Assignment	k11 / (delta * gamma * Kg)
Kd	Fixed	3.47e-08
Kd_k3	Assignment	Kd * k3
Kd_k3/gamma	Assignment	Kd * k3 / gamma
Kd_k3/delta_gamma	Assignment	Kd * k3 / (delta* gamma)
PI3K_Vmax (M s^{-1})	Assignment	PI3K_kcat * [PI3K*.Gbg.PM]
PI3K_kcat	Fixed	0.03 s^{-1}
PTEN_kcat	Fixed	1.2 s^{-1}
PTEN_Vmax (M s^{-1})	Assignment	PTEN_kcat * [PTEN_PM]

Table 2.5. Summary of Global Variables, together with their types and assigned values or expressions. Key: * = Typo, not used.

This completes the model setup specification at the commencement of model development. Any changes or additions to the model, made in the light of later results, are described in the next section.

2.4 Results

2.4.1 Results from Version 1 of the COPASI CXCR4-SDF1 chemotaxis lamellipodia model

This first version of the CXCR4-SDF1 chemotaxis model produces initially promising results, as shown in Figure 2.2, the plot of particle numbers. Promising in the sense ofCXCR4 SDF1 binding occurring and Gβγ being produced.

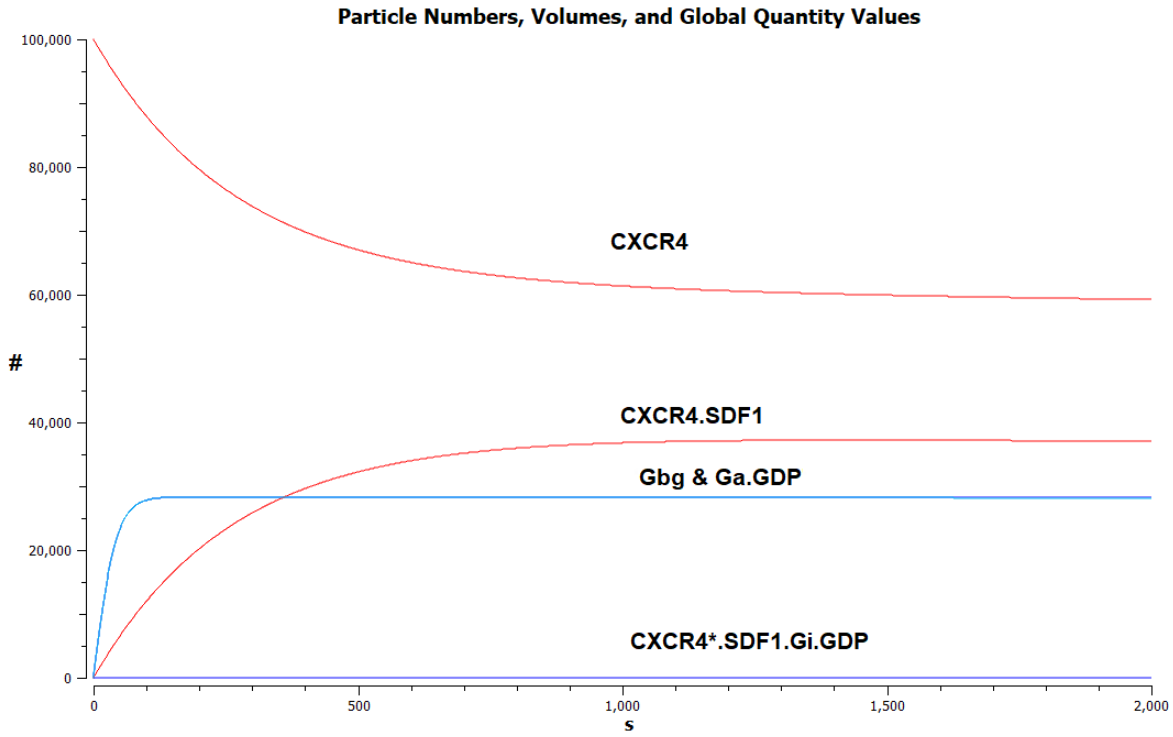

Figure 2.2. Particle numbers forCXCR4, CXCR4.SDF1, CXCR4*.SDF1.Gi.GDP, Ga.GDP and Gβγ. Here, and in future graphs, # means number of molecules/particles

Within around 20 minutes (i.e. around 1200 seconds) of first exposure to SDF1, almost 40% of CXCR4 has become bound to it. Although only a small fraction of this bound CXCR4 is activated at any given time (as represented by the mauve curve for CXCR4*.SDF1.Gi.GDP), large amounts of Gα.GDP and Gβγ are rapidly produced (around 27,000 molecules in less than 2 minutes). This would seem to imply that SDF1-bound CXCR4 is being activated rapidly, as Gα.GDP and Gβγ are the breakdown products of Gi.GTP, whose formation in this model is intended to be principally driven by activated CXCR4 complexes, acting as a GEF to anyGi.GDP to which it is bound.

However, despite so much Gβγ being available, only a small amount of PI3K becomes activated by it, as shown in Figure 2.3, with PI3K*.Gβγ peaking at only 5 molecules within the first 2 minutes and PI3K*.Gβγ.PM peaking at only 9 molecules, after 20 minutes have elapsed.

The main reason for this can be seen in Figure 2.4, which shows that the vast majority of available PI3K (2800 molecules) is converted into membrane-bound PI3K.PM, without subsequently becoming activated by binding to Gβγ. The problem here is that, despite the large number of molecules of Gβγ in the cytosol, such is the volume of the cytosol (at 1500 fl) that it represents only a very small concentration, at around 31 nM.

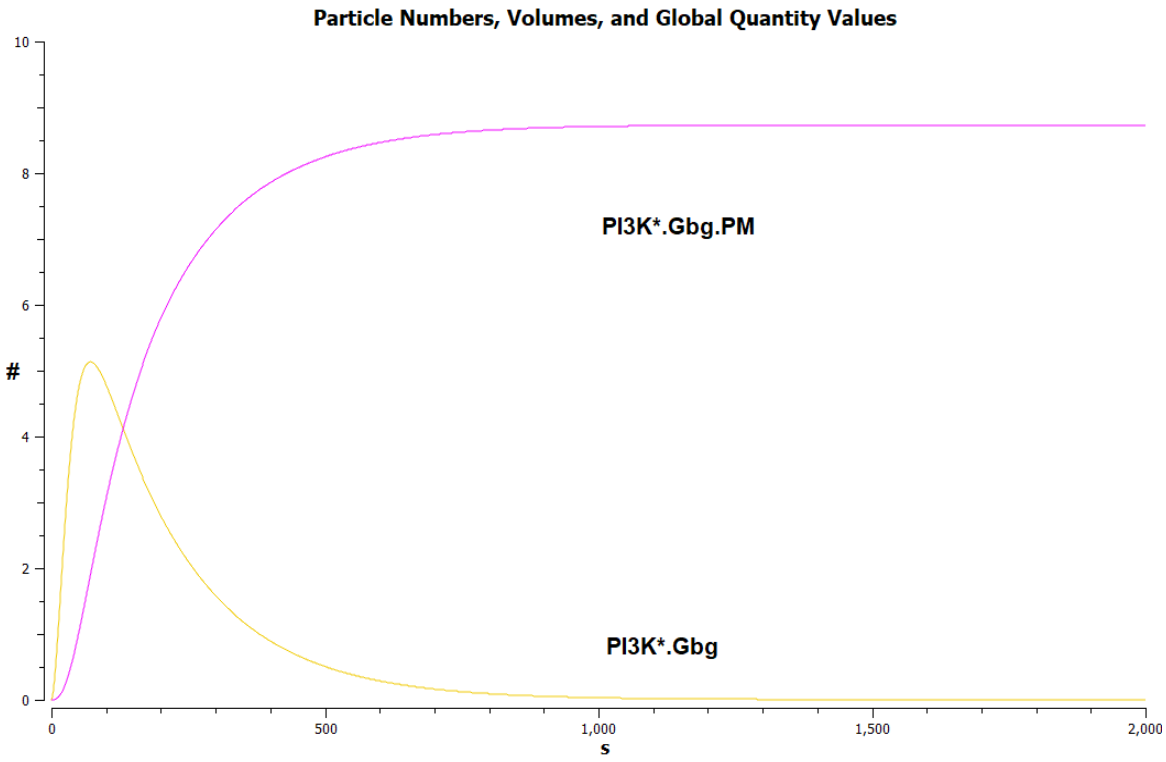

Figure 2.3. Particle numbers of membrane-bound PI3K*.Gβγ and activated PI3K*.Gβγ.PM.

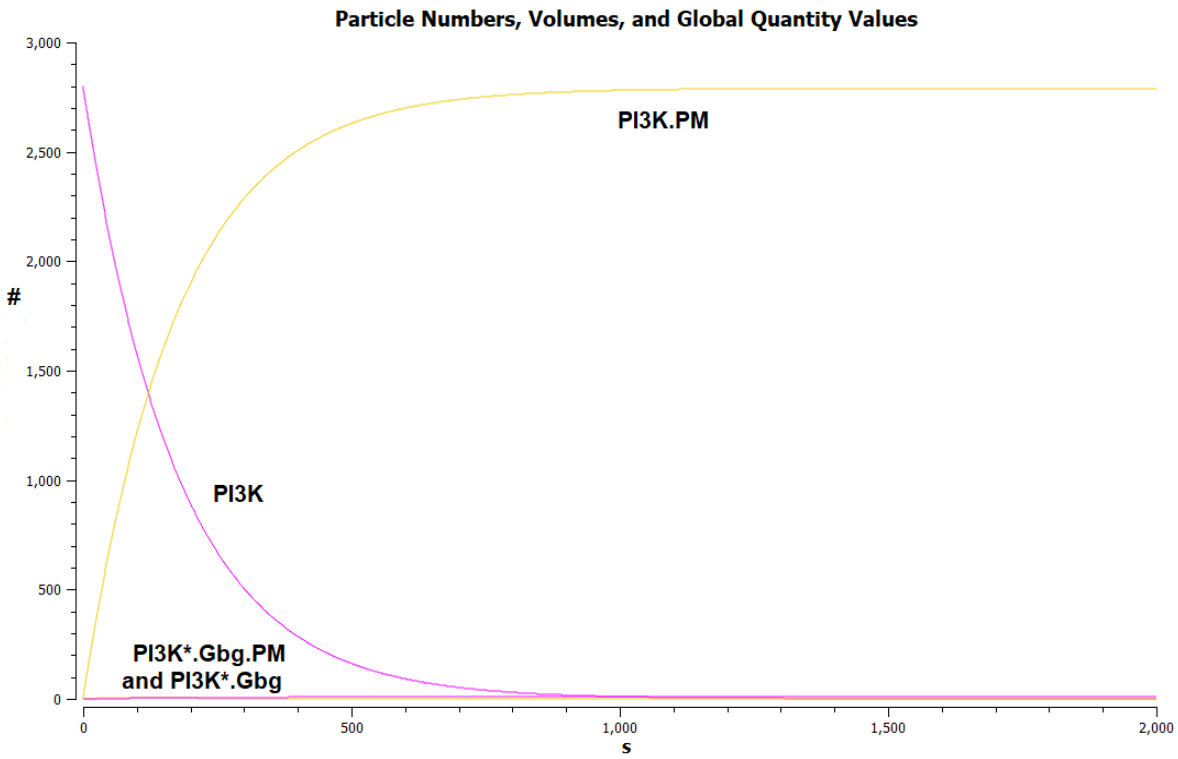

Figure 2.4. Particle numbers of all PI3K species.

In comparison, membrane-resident Rac1, which, in activated and inactivated form, recruits PI3K to the cell membrane, exists in much higher numbers than Gβγ (330,000 as against around 27,000 molecules) and represents an even higher concentration, proportionately, (around 860 µM as against 31 nM) by virtue of residing in the much smaller PM compartment (6.4 as against 1500 fl).

Since the rate of mass action reactions are determined by the concentrations of reactants, rather than particle numbers, this inevitably means that membrane recruitment of PI3K will proceed at a much faster rate than Gβγ binding, unless compensated for by a much slower (second-order) rate constant for the first reaction. In fact, the opposite is the case, with the former reaction having a forward rate constant of 100,000 M^{-1} s^{-1}as against 10,000 M^{-1} s^{-1} for Gβγ binding. (The reverse rate constants are not really comparable, being second-order for Rac1 recruitment and first-order for Gβγ binding, by virtue of the fact that the model assumes that, having recruited PI3K, Rac1 and PI3K then separate, whereas PI3K is assumed to stay bound to Gβγ after activation.)

As a result, the rate of membrane recruitment of unbound PI3K greatly exceeds the rate of Gβγ binding to PI3K, as can be seen from a comparison of the relevant reaction fluxes or particle-formation rates, as shown in Figure 2.5. As a consequence, the rate of formation of activated PI3K*. Gβγ.PM decreases.

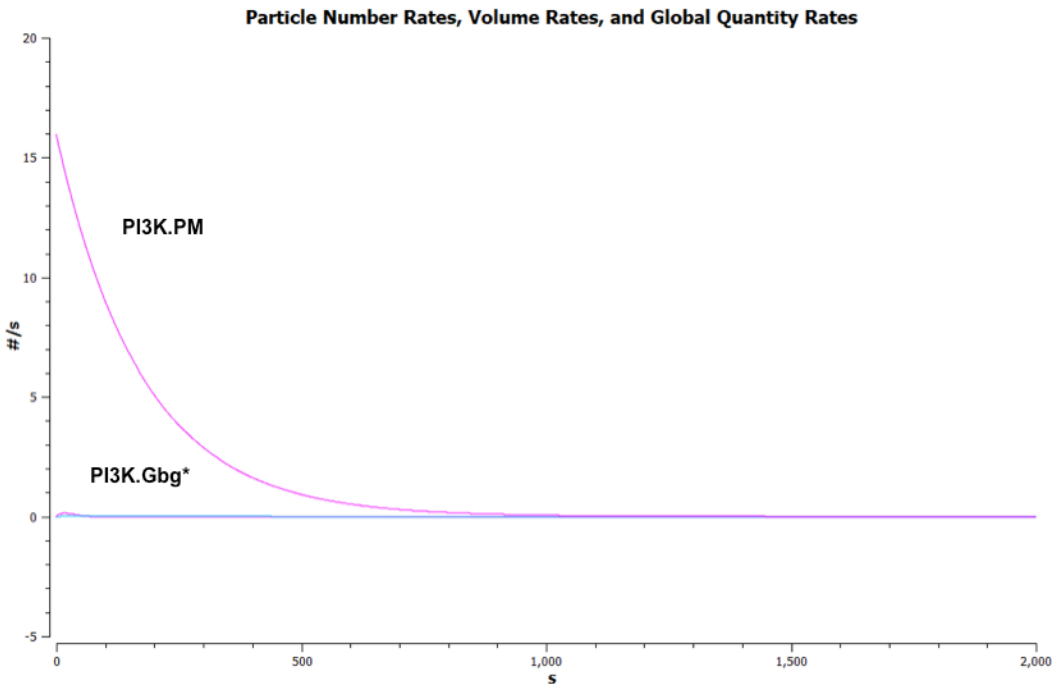

Figure 2.5. Particle number rates of formation of PI3K.PM and PI3K*.Gβγ.

The reason for this decrease is that, of the two ways of forming it, either by membrane recruitment of PI3K*.Gβγ (by Rac1) or through PI3K.PM binding Gβγ, the second route is much slower, owing to the very low concentrations of cytosolic Gβγ, compared to the very high concentrations of PM-resident Rac1, as explained earlier. But, as increasing proportions of PI3K become membrane-bound, this becomes the only route available.

Another possible reason, as shown in Figure 2.6, is that whereas, in concentration terms, Gβγ is a little less than an order of magnitude in excess of cytosolic PI3K (at approximately 31 nM as against 3.2 nM), this same concentration is substantially lower proportionally in relation to that of

membrane-resident PI3K.PM (which, at 720 nM, is roughly one-and-a-half orders of magnitude higher).So, it could be argued that, given the 1:1 stoichiometry in both cases, Gβγ is acting as more of a drag on the second reaction than PI3K is acting as a drag on the first one.

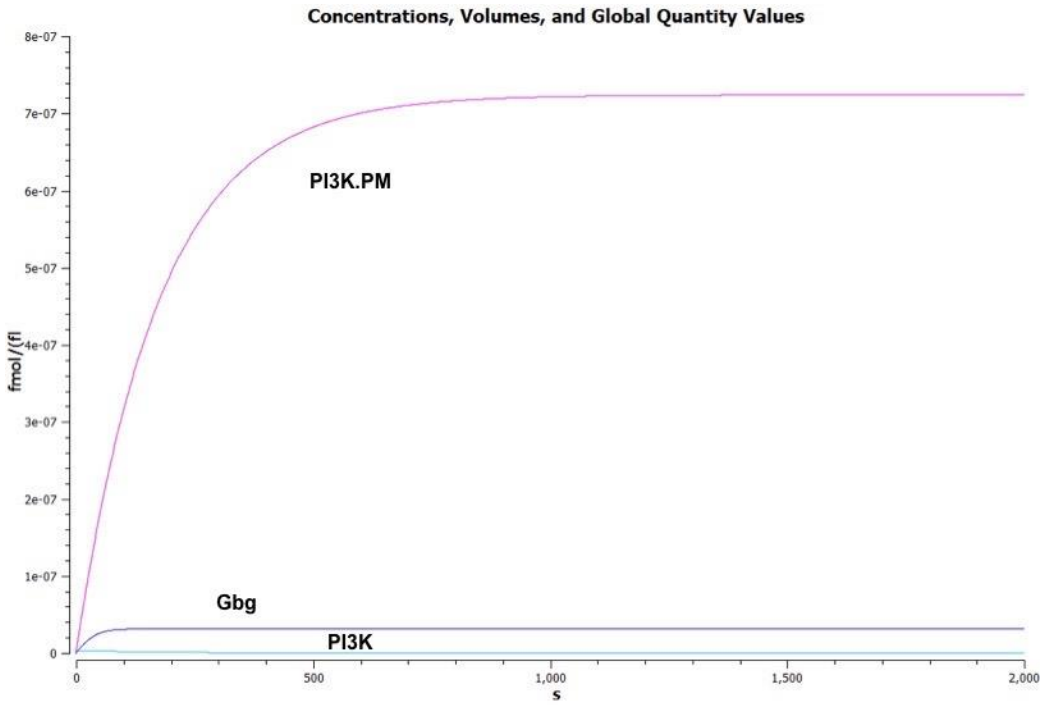

Figure 2.6. Concentrations of PI3K, PI3K*.Gβγ and PI3K.PM.

Because this model, not unreasonably, allows only activated, membrane-bound PI3K to convert PIP2 to PIP3, and only around nine molecules, on average, are available in this form at any given time, levels of PIP3 are correspondingly low. As shown in Figure 2.7, levels of PIP3 formation peak at only 0.017 particles per second in the first few hundred seconds, before falling again, so that with 10,000 seconds elapsed, levels of PIP3 have barely reached 90 molecules (Figure 2.8].

Since Rac1GEFs require PIP3 binding to become activated, levels of activated Rac1GEFs are inevitably very low as a consequence of the very low levels of PIP3, as shown in Figure 2.9, where average levels are still well below one molecule with 50,000 seconds elapsed. Because each activated Rac1GEF can activate more than one Rac1, of which there are many (330,000 molecules), and because both activated Rac1GEFs and inactive Rac1 are both membrane-bound

(meaning both are in relatively high concentrations, owing to the very low PM volume), even this low number of Rac1GEFs is able to activate a much larger number of Rac1 molecules (over 500, with 50,000 seconds elapsed, Figure 2.10).

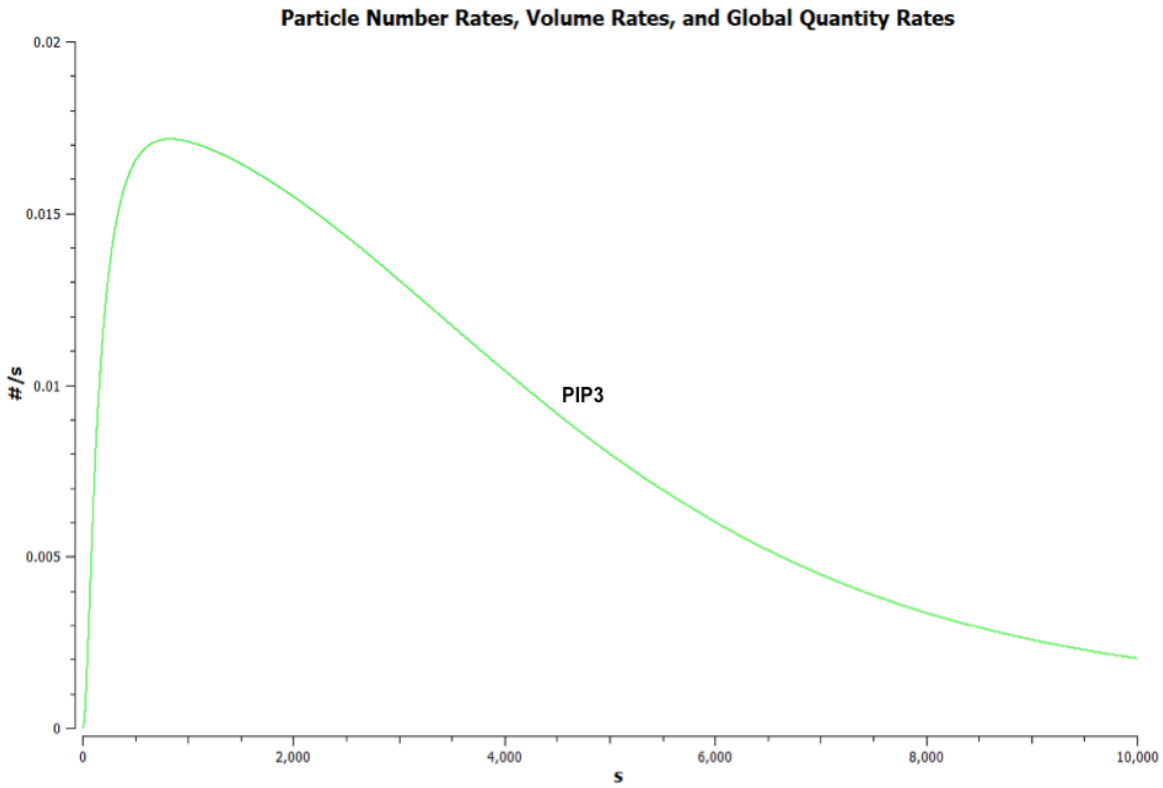

Figure 2.7. Particle formation rates for PIP3.

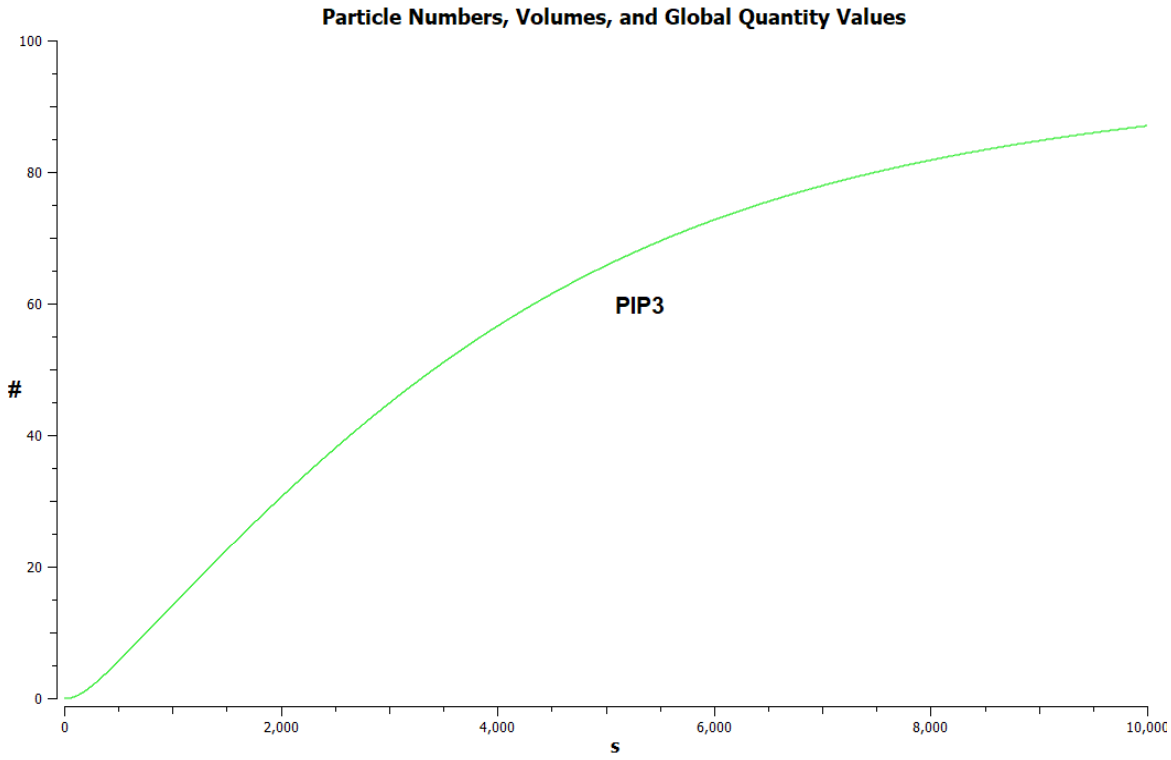

Figure.2.8. Particle numbers of PIP3.

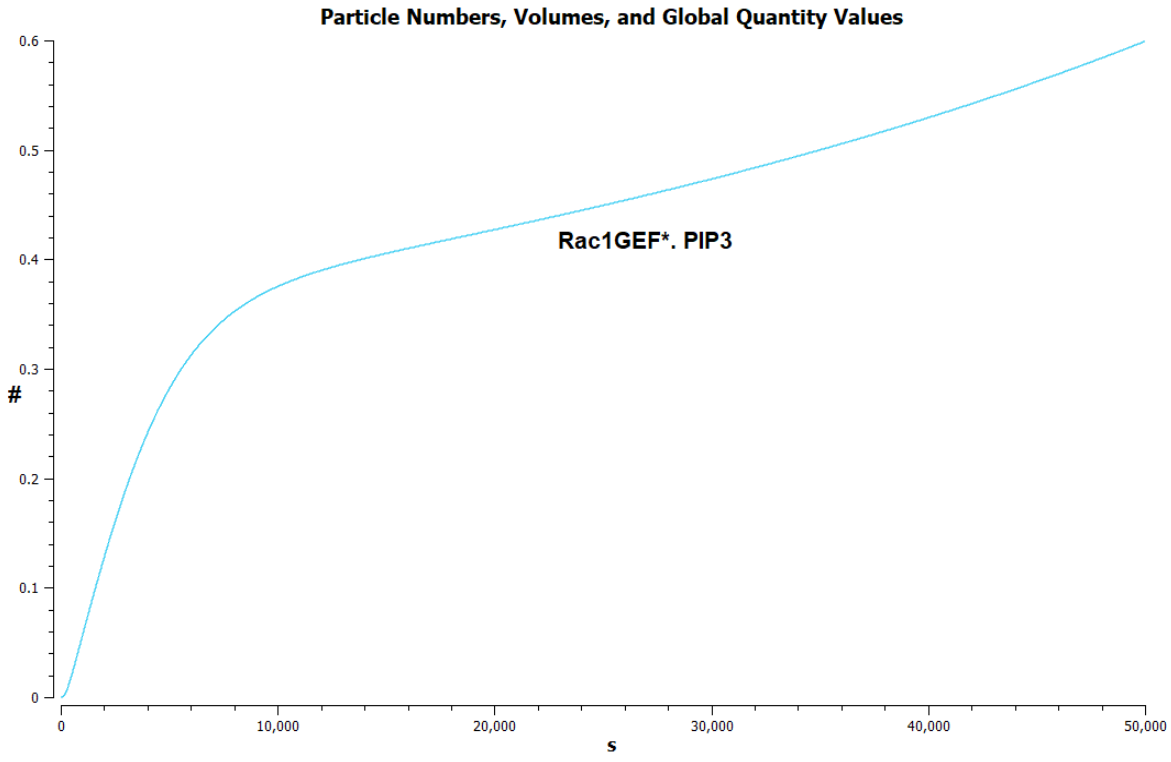

Figure 2.9. Particle numbers of activated Rac1GEF*. PIP3.

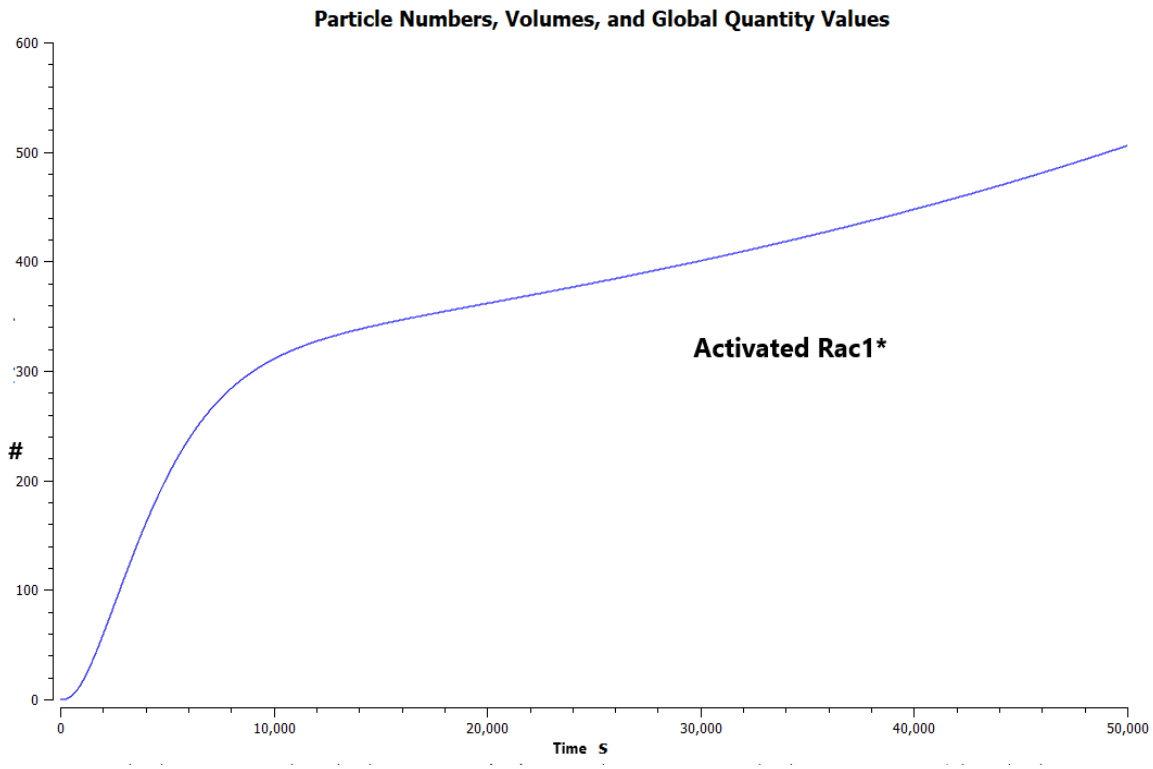

Figure 2.10. Particle numbers of activated Rac1*.

However, given the time taken and the large numbers of available Rac1 molecules (330,000), as well as the fact that SDF1 in this model is provided in continual excess, these actually represent very low numbers of activated Rac1. As mentioned earlier, a major reason for this is that the large levels of Gβγ (around 27,000 molecules) created downstream of this SDF1-CXCR4 binding, only result in around 90 molecules of PIP3 being generated within 10,000 seconds, by which time the rate of PIP3 formation is steadily falling. And, as explained, the main cause of this deficit is the very low rate at which membrane-bound inactive PI3K is activated by Gβγ.

This suggests that one way to substantially increase the overall PIP3 yield would be to prevent PI3K recruitment to the cell membrane until it has been activated. There is some support for this assumption in the literature, including evidence suggesting that prior binding by Gβγ is required for membrane recruitment of the p110γ catalytic subunit [190], but clearly this is a different enzyme with a different regulatory subunit, so may not be relevant. However, whilst examining this evidence, it was realised that the existing model contains an incorrect assumption, namely that

trimeric G proteins and their subunits are mainly cytosolic. In fact, for much of the time, whether in associated or dissociated states, they remain bound to the plasma membrane[191],either attached to their associated receptor or diffusing laterally within the membrane.

Adjusting the model to incorporate this new assumption to some extent obviates the need for prior Gβγ binding, as it will tend to reduce the disadvantage inactive membrane-resident PI3K has against its cytosolic counterparts, in terms of Gβγ-mediated activation. This is confirmed by the results from this new version of the model, as can be seen in Figure 2.11.

2.4.2 Results from Version 2 of the COPASI model

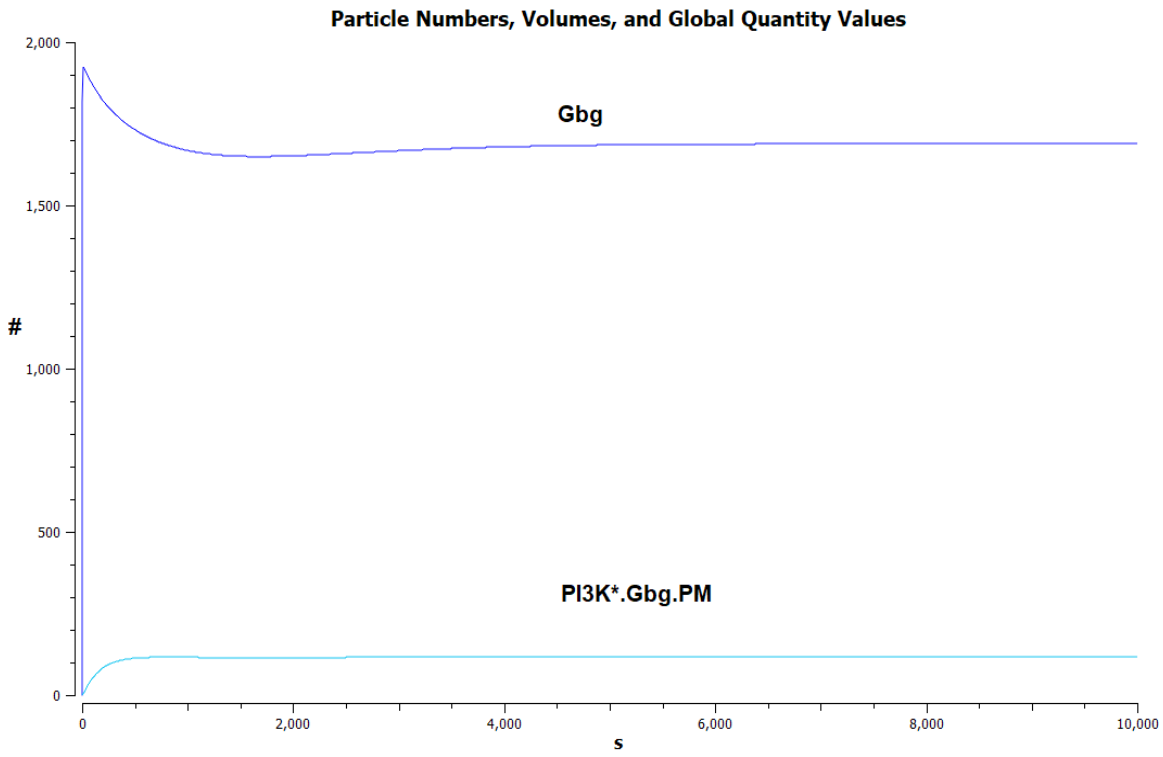

Figure 2.11. Particle numbers of Gβγ and activated, membrane-boundPI3K*.Gβγ.PM, in the second version of the model.

Now, despite much less Gβγ being created (with steady-state levels of around 1700 molecules as opposed to 27,000 previously), PI3K is being converted into activated, membrane-bound

PI3K*.Gβγ.PM in much larger quantities. Thus, after 10,000 seconds, around 120 molecules of the latter have been created, whereas, after the same time had elapsed in the original model, less than 10 molecules had been created. Consequently, much more PIP3 is produced in this new version of the model (around 440 molecules as opposed to around 90 molecules formally (Figure 2.8). In other words, nearly 5 times as much PIP3 is produced (after 10,000 seconds), as seen in Figure 2.12.

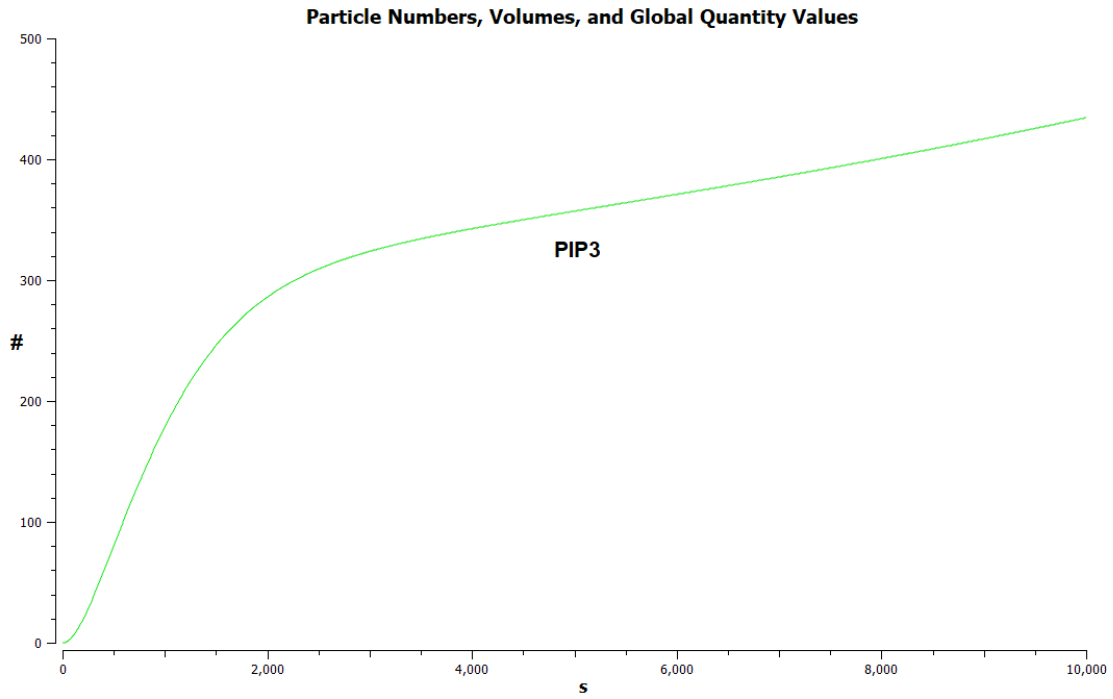

Figure 2.12. Particle levels of PIP3 in the second version of the model.

Levels of activated Rac1GEFs and Rac1 are similarly higher compared with the previous model (with average particle numbers of around 25 molecules of Rac1GEF*.PIP3 as against 0.6 molecules, and nearly 20,000 molecules of Rac1*as against nearly 500 molecules, with 50,000 seconds elapsed), as can be seen by comparing Figures 2.13 and 2.14with Figures 2.9 and 2.10.

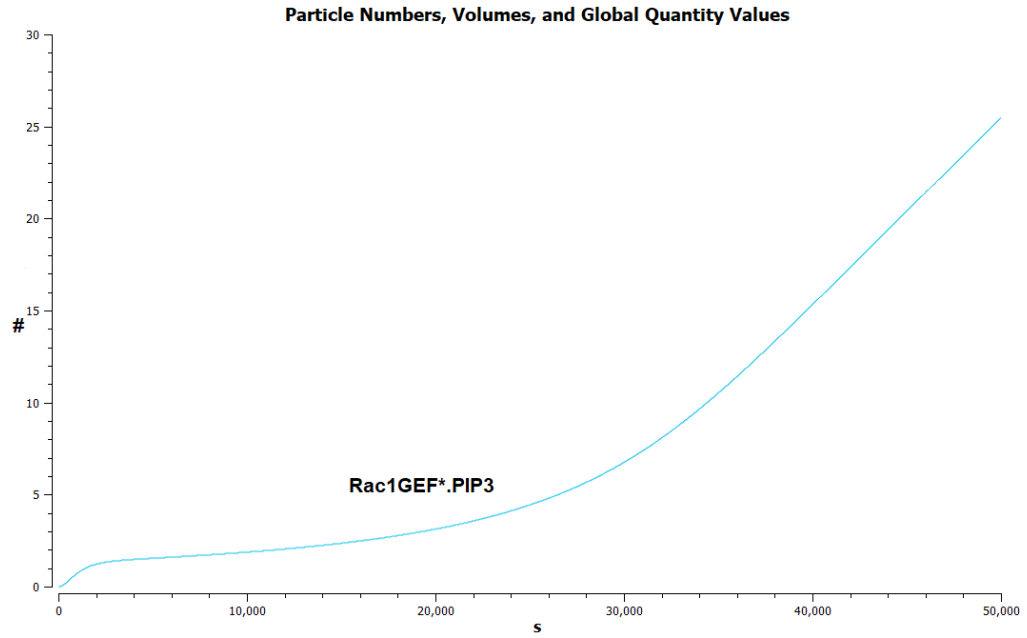

Figure. 2.13. Particle levels of activated Rac1GEF*.PIP3 in the second version of the model.

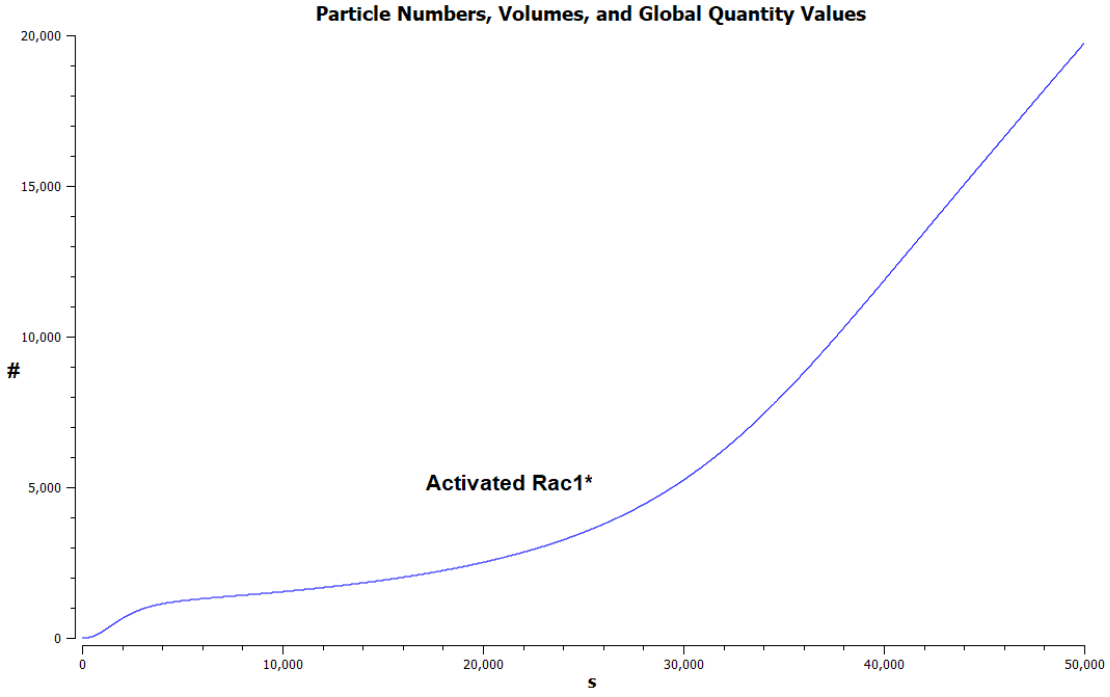

Figure 2.14. Particle levels of activated Rac1* in the second version of the model.

Thus this apparently minor change in assumptions (i.e. making G protein subunits membrane-resident) has resulted in a 40-fold increase in levels of activated Rac1, with a similar increase for Rac1GEF*.PIP3.

However, given the highly favourable conditions mentioned earlier, this still demonstrates a limited response, even allowing for the fact that HeLa cells are relatively slow-moving in response to external stimulation [192], and thus may be constitutively refractory in this respect.

Another change that could be considered is to make the further assumption that inactive Rac1GEFs are membrane-resident rather than cytosolic.

2.4.3 Results from Version 3 of the COPASI model

When this change is made, whilst leaving G proteins and their subunits cytosolic, this results in substantial increases in levels of activated Rac1GEF*.PIP3 and activated Rac1*, compared to the original model. This is despite levels of Gβγ and PI3K*. Gβγ.PM and PIP3 (once one allows for increased take-up by Rac1GEFs) remaining largely unchanged. Specifically, in the course of 50,000 seconds, levels of activated Rac1GEF*.PIP3 rise to around 127 molecules (as against less than one molecule in the original model), while levels of Rac1* rise to roughly 80,000 molecules, as against 500, over the same time period.

This can be seen by comparing Figures 2.15, 2.16 and 2.17 with Figures 2.8, 2.9 and 2.10. As these Figures show, in this latest version of the model more than half of available Rac1GEFs are activated before 50,000 seconds have elapsed, and more than three times as much Rac1 (over 70,000 molecules as compared to 20,000 previously).

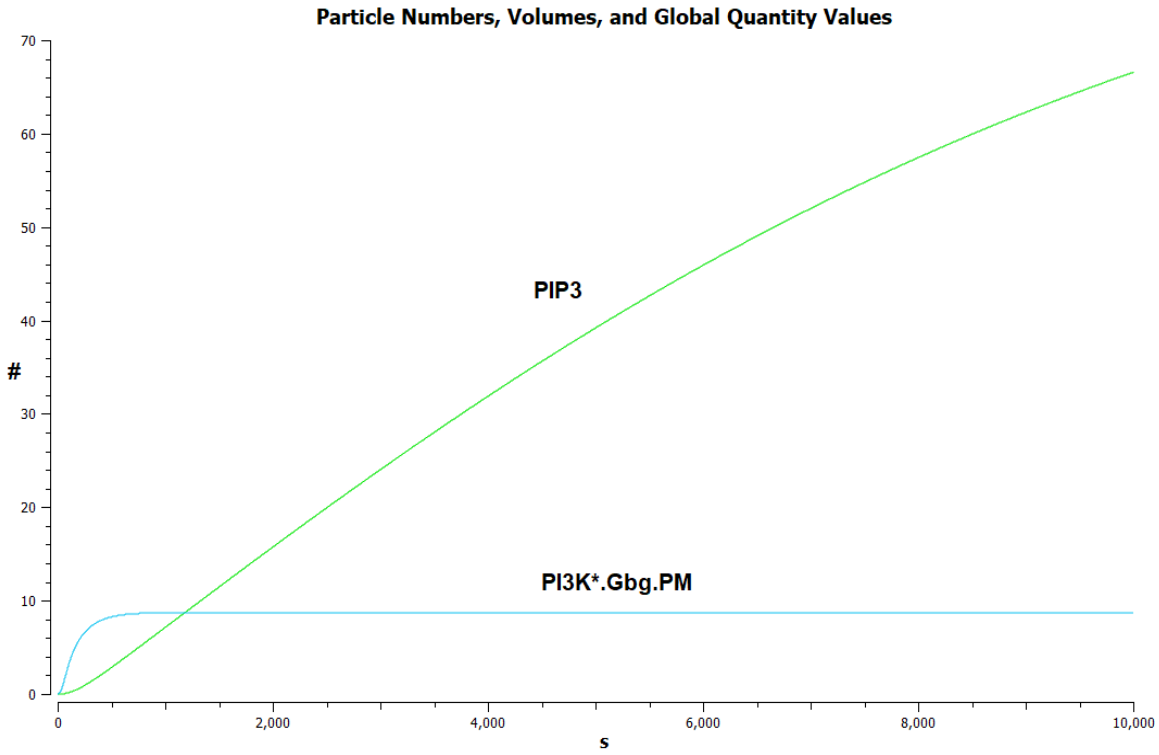

Figure 2.15. Particle numbers of PI3K*.Gβγ.PM and PIP3 in the third version of the model.

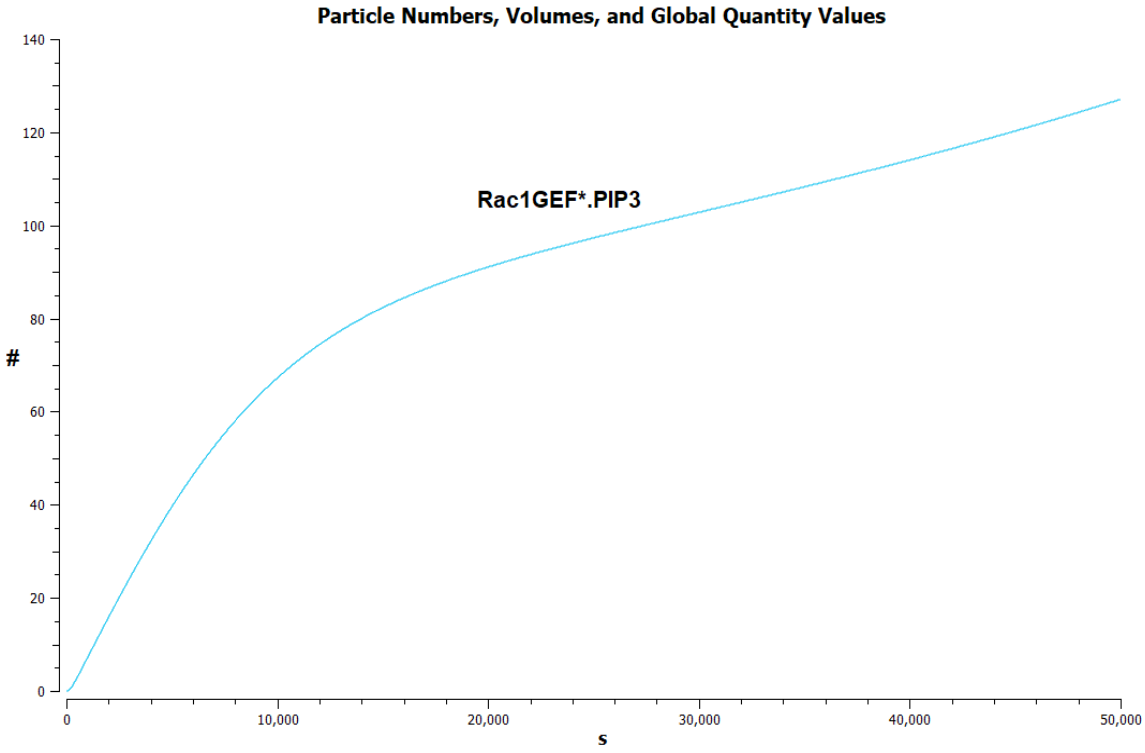

Figure 2.16. Particle levels of activated Rac1GEF*.PIP3 in the third version of the model.

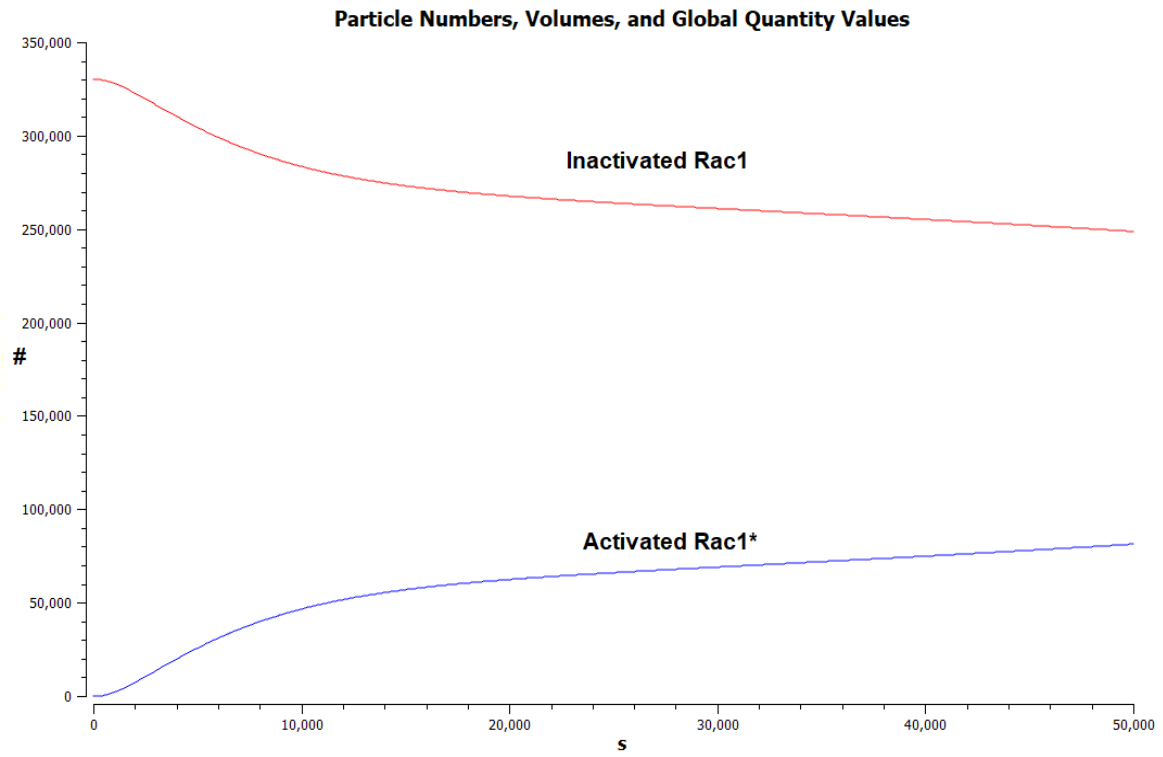

Figure 2.17. Particle levels of activated vs activated Rac1* in the third version of the model.

Clearly this is a very substantial increase in all the key outputs of the model, and can be explained by increased uptake of PIP3 by membrane-resident Rac1GEFs, whose concentration is higher than when it is resident in the cytosol. As explained earlier, this higher relative concentration results in a higher reaction rate, which can be understood non-mathematically as the result of a much-increased likelihood of PIP3 and Rac1GEFs encountering each other when both are membrane-resident. This is confirmed by comparing reaction fluxes for Rac1GEF-PIP3 binding in this new version of the model, as compared with the original, as shown in Figures 2.18 and 2.19.

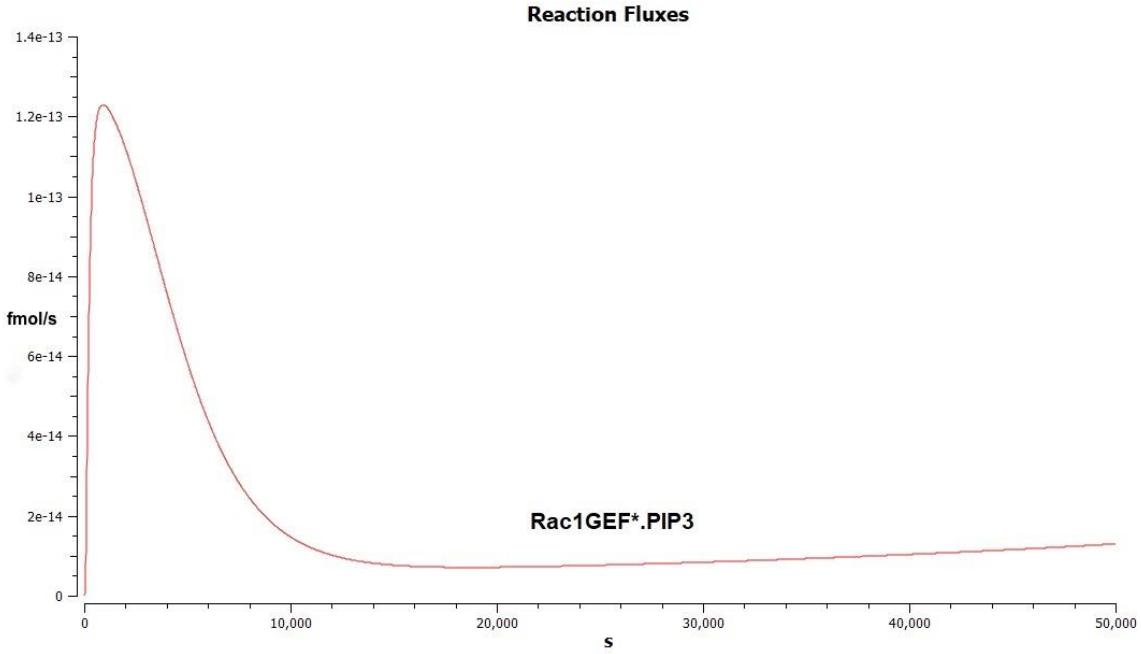

Figure 2.18. Reaction flux for Rac1GEF*.PIP3 formation from cytosolic Rac1GEF and PIP3. (First version of model.)

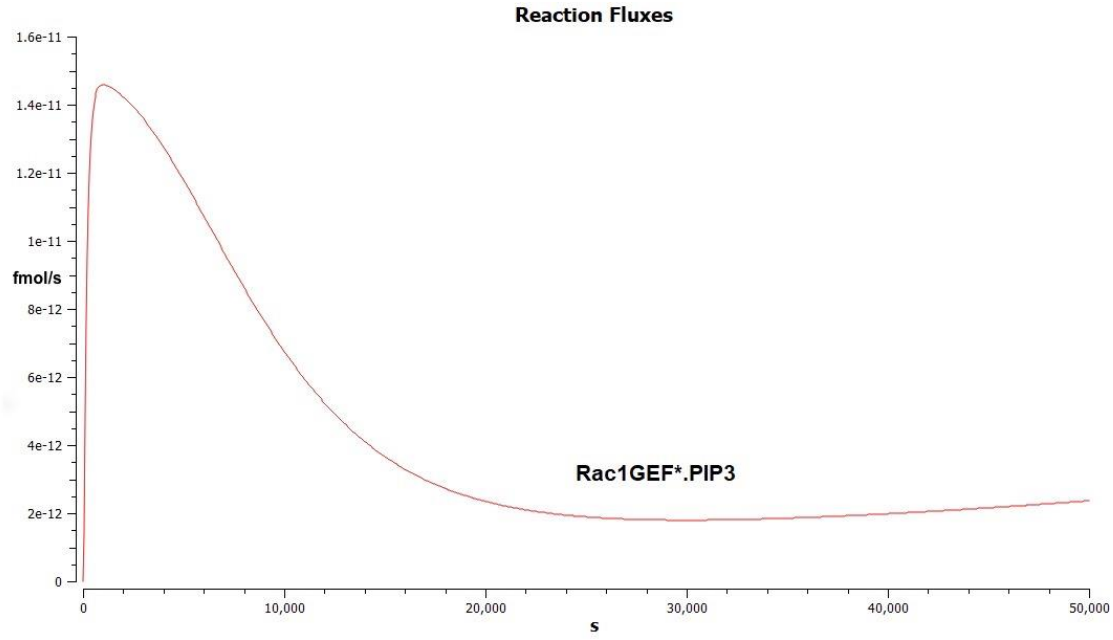

Figure 2.19. Reaction flux for Rac1GEF*.PIP3 formation from membrane-resident Rac1GEF and PIP3. (Third version of model.)

Thus, assuming that inactive Rac1GEFs are primarily membrane-resident leads to a nearly 100-fold increase in the reaction rate for Rac1GEF-PIP3 binding. However, it has to be acknowledged that the evidence in support of Rac1GEF being predominantly membrane-resident, is only limited.

Certainly, a large proportion of Rac1GEFs contain PH domains (often in combination with DH domains), which typically bind to phosphoinositide-rich membranes[193,194]. However, such binding appears to be absent in the case of many Rac1GEFs[195]. For this reason, despite the obvious increase in Rac1 activation, this assumption is not included in further models, whereas G protein membrane-residency is retained, as a closer reflection of the underlying physiological reality.

At this point there are still a number of unrealistic assumptions in the model that have been included for the sake of simplifying the model. Amongst the most critical of these are the twin assumptions that both SDF1 and PIP2 are held at fixed levels. This means that they are effectively inexhaustible. Replacing these assumptions should help to make the model more realistic, in terms of better reflecting the underlying physiological reality that it is supposed to model.

Given that SDF1 levels (2.2e-08 M) in this model are not considered unreasonable (being held at a concentrations deemed physiologically optimal in a published chemotaxis assay result [183]), a straightforward way of making their availability more realistic is to use events. Events allow the user to introduce changes to the model during runtime. This allows SDF1 to be introduced sometime after the model has run (to see how it runs without SDF1) and then withdrawn at a later stage, in order to see how the model responds to this withdrawal. Other than this, all that is required is to set the initial concentration of SDF1 to zero and make its levels reaction-dependent rather than fixed.

Unfortunately this same approach cannot be used for PIP2 because, being a precursor for PIP3, it is important to see how variability in its availability affects levels of PIP3, as well as other downstream aspects of the model. For this reason it is better to model PIP2 formation in ways that more faithfully reflect how it is generated in vivo, especially in response to chemotactic signals. This requires the introduction of four new species (PI, PIP, PI4K and PI5K) and two new reactions (one for phosphorylation of PI to PIP, by PI4K, and the other for phosphorylation of PIP to PIP2 by PI5K).

Rather than introducing both of these new changes together in the same model, it is usually better to introduce major changes like this in stages, that is to say, one model at a time. In this case, PIP2 generation is introduced first. Including the changes to G protein residency introduced earlier, the changes to the original model, up to this point, can be summarised in Tables2.6and 2.7.

In common with the second version of the model (which just had membrane-resident G proteins), levels of Gβγ are relatively low, at 1700 molecules, after 10,000 seconds elapsed (Figure 2.20). Over the same time period, levels of PI3K*.Gβγ.PM are about the same as the second version (at 117 molecules, as opposed to 120, as shown in Figure 2.21).

Species	Compartment	Particle#	Concentration (μM)	Notes
Ga.GDP	PM	0	0	Moved from cytosol to PM
Ga.GDP	PM	0	0	ditto
Gβγ	PM	0	0	ditto
Gi.GDP	PM	340,000	88.2	ditto
Gi.GTP	PM	0	0	ditto
PI	PM	100,000	26	New species, Estimate, based on relative levels reported by Xu and colleagues[196].
PI(4,5)P2	PM	2800	0.73	Quantity no longer fixed. Estimate, based on relative levels reported by Xu and colleagues [196].
PI4K	PM	24,000	6.2	New species. Copy numbers taken from HeLa cell quantitative proteome [182].
PI5K	PM	15,000	3.9	New species. Copy numbers taken from HeLa cell quantitative proteome [182].
PIP	PM	2000	0.52	New species, Estimate, based on relative levels reported by Xu and colleagues [196].

Table 2.6. Summary of species, and changes to compartment of existing species in models after the base model.

Reaction description	Reaction equation	Rate law	Rate constant values
PI phosphorylation to PIP	PI -> PIP; PI4K	Michaelis-Menten	kcat = 2.77 s^{-1} Km = 1.6e-05 M[181]
PIP phosphorylation to PIP2	PIP ->PI(4,5)P2; PI5K	Michaelis-Menten	kcat = 1.02 s^{-1} Km = 1.0e-05 M[181]

Table 2.7. Reactions added to the fourth version of the model.

2.4.4 Results from Version 4 of the COPASI model

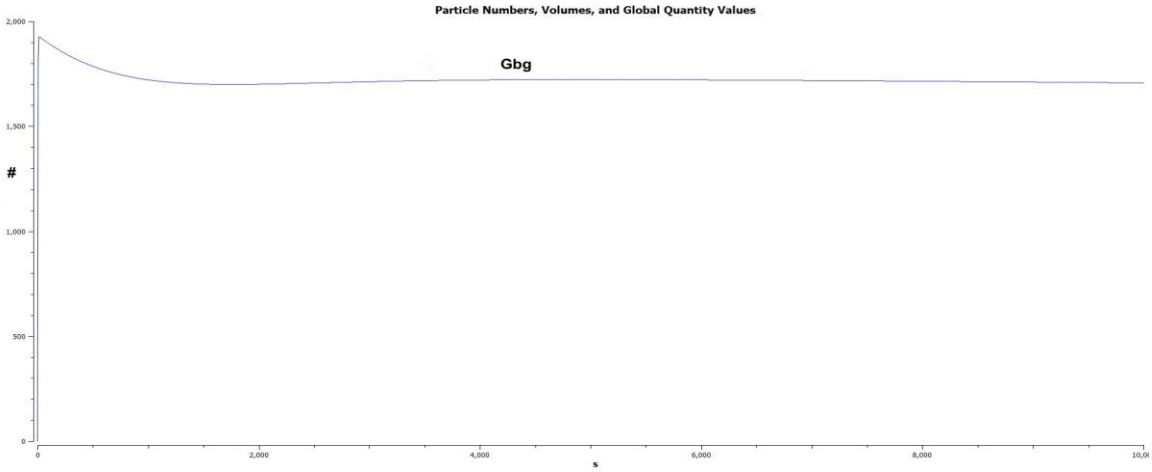

Figure 2.20. Particle levels of Gβγ in the fourth version of the model. Peak particles numbers of Gβγ ~ 1900. Time course 10,000 s.

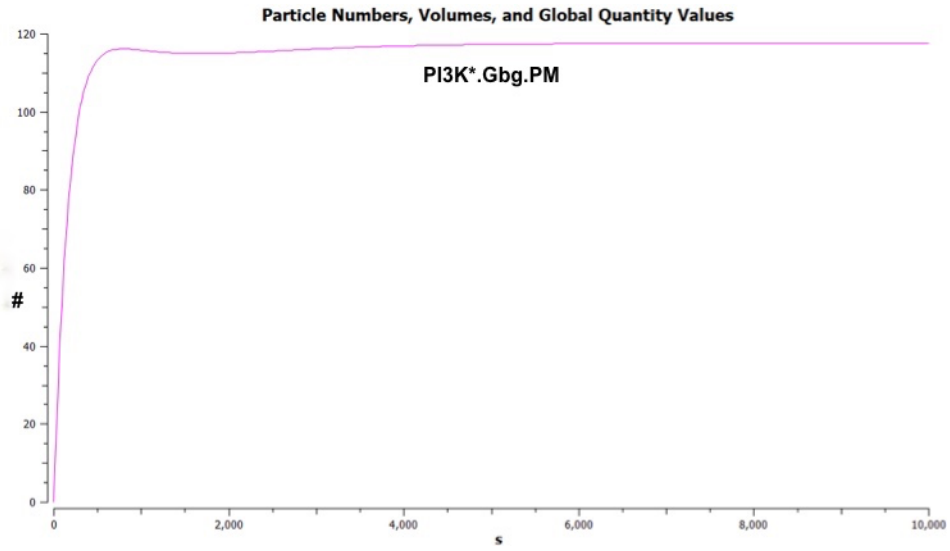

Figure 2.21. Particle levels of PI3K*.Gβγ.PM in the fourth version of the model. Peak particle levels = ~ 117. Time course 10,000 s.

However, because PIP2 levels are so much higher from early on (at around 105,000 molecules, after 30 seconds elapsed, Figure 2.22), as compared to a fixed level of 20,000 molecules in the second model, these 117 molecules of PI3K*.Gβγ.PM are able to generate substantially higher peak levels of PIP3 (8500 molecules, Figure 2.23) as against 440 molecules in the second model (Figure 2.12). This is confirmed by a comparison of the reaction fluxes for PIP2 phosphorylation (Figures 2.24 and 2.25), for the two models, which shows that within 500 seconds, the flux rate for this fourth model has reached a peak level of approximately 1.6e-09 fmol/s (Figures 2.25), whereas, over the same time period in the second version, this rate has plateaued at around 4e-10 fmol/s (Figure 2.24).

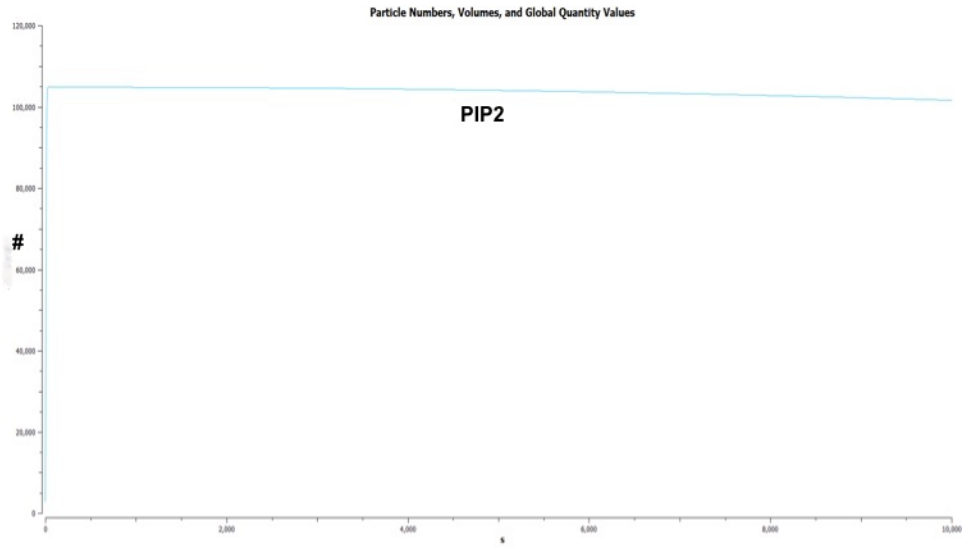

Figure 2.22. Particle levels of PIP2 in the fourth model. Peak particles numbers ~ 105000. Time course 10,000 s.

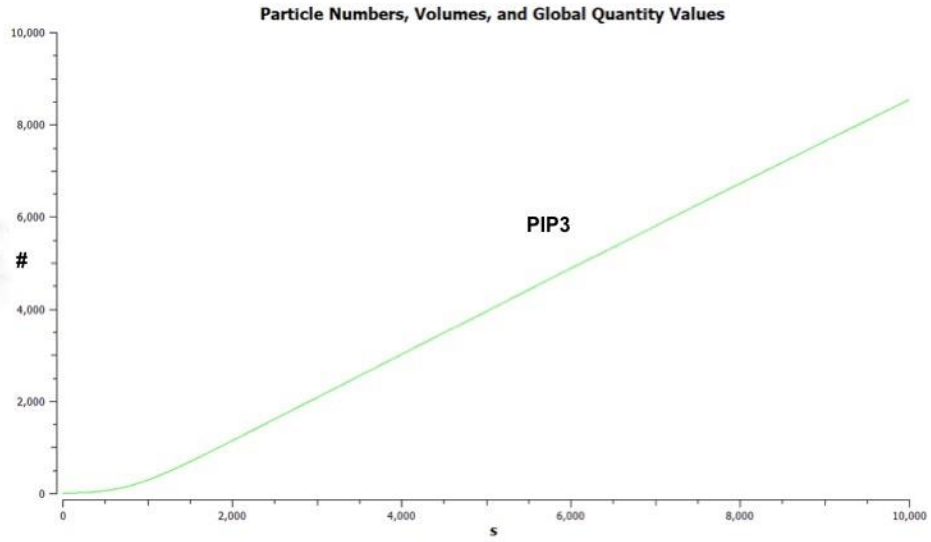

Figure 2.23. Particle levels of PIP3 in the fourth model.

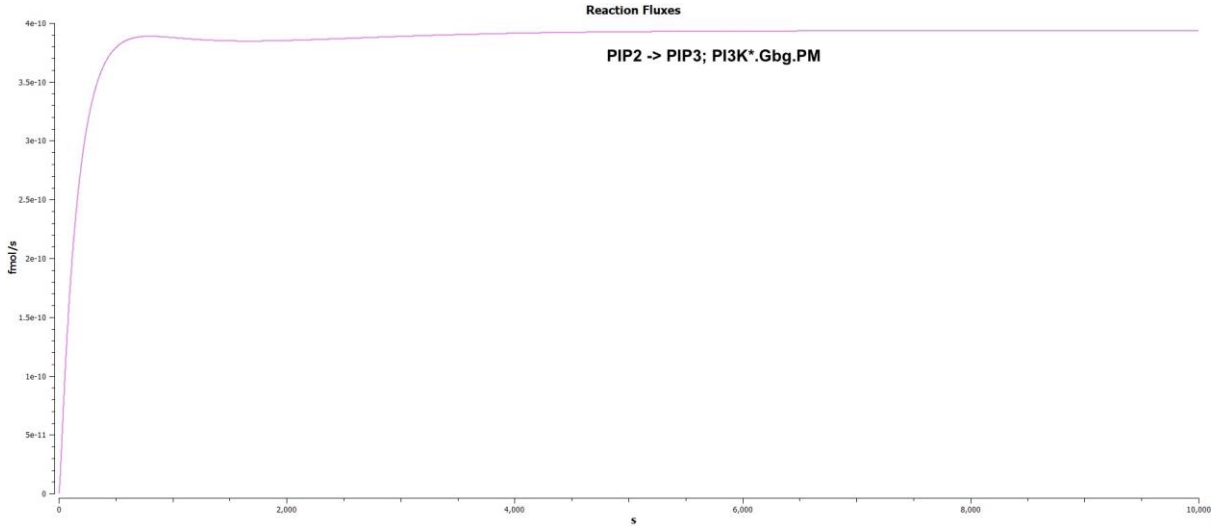

Figure 2.24. Reaction flux for PIP2 phosphorylation to PIP3 by PI3K, in second version of model.

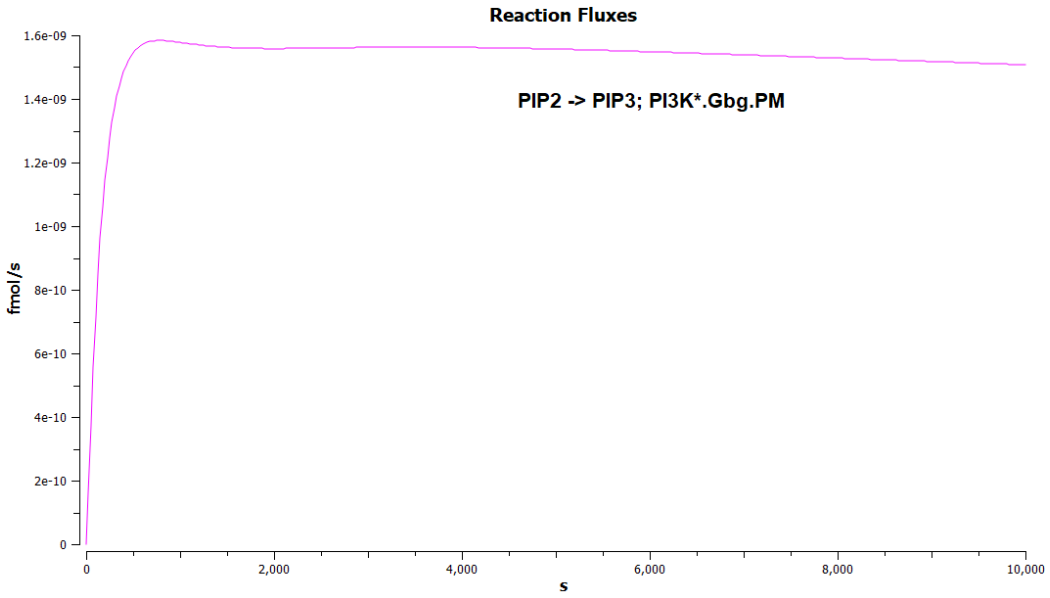

Figure2.25. Reaction flux for PIP2 phosphorylation to PIP3 by PI3K, in fourth version of model.

Over 50,000 seconds, this results in a nearly sevenfold increase in activated Rac1GEF*.PIP3 (170 molecules as against 25) and a fivefold increase in activated Rac1*(around 100,000 molecules as against 20,000 in the second model). These can be seen by comparing Figures 2.26 and 2.27, with Figures 2.13 and 2.14. In the second version of the model (which just had membrane-resident G proteins), levels of Gβγ are relatively low, at 1700 molecules, after 10,000 seconds elapsed (Figure 2.11). Over the same time period, levels of PI3K*.Gβγ.PM are about the same as the second version at 117 molecules (Figure 2.21), as opposed to 120, as shown in Figure 2.11.

Adding in SDF1 as a bolus, rather than being continuously available (and at a fixed level), can most conveniently be done with an event, as was explained earlier. However, including just SDF1 addition in this event would mean that PIP2 generation (which now includes PIP generation from PI) is still occurring from the outset. Given how much PIP2 is now being generated, and how quickly it reaches these peak levels, this could seriously distort the model results.

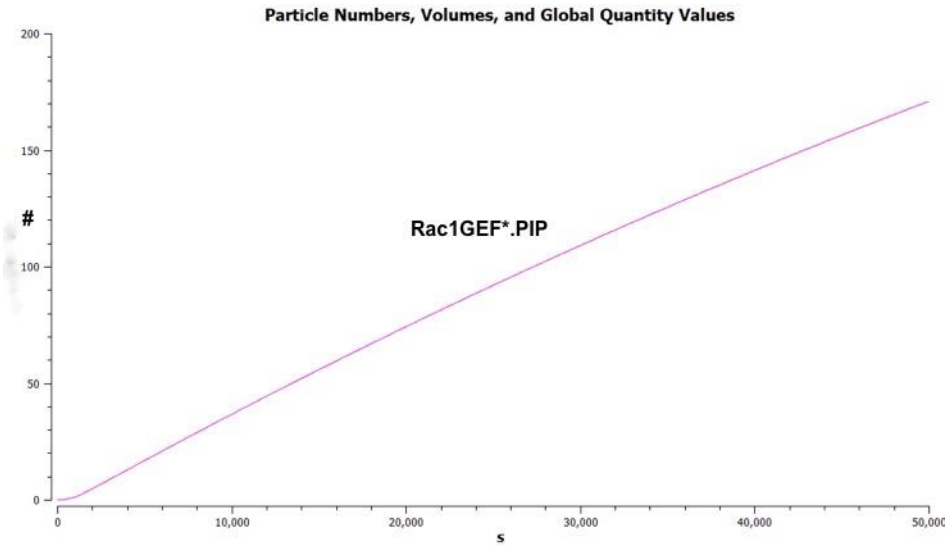

Figure 2.26. Particle numbers of activated Rac1GEF*.PIP3 in fourth version of the model. Time course 10,000 s

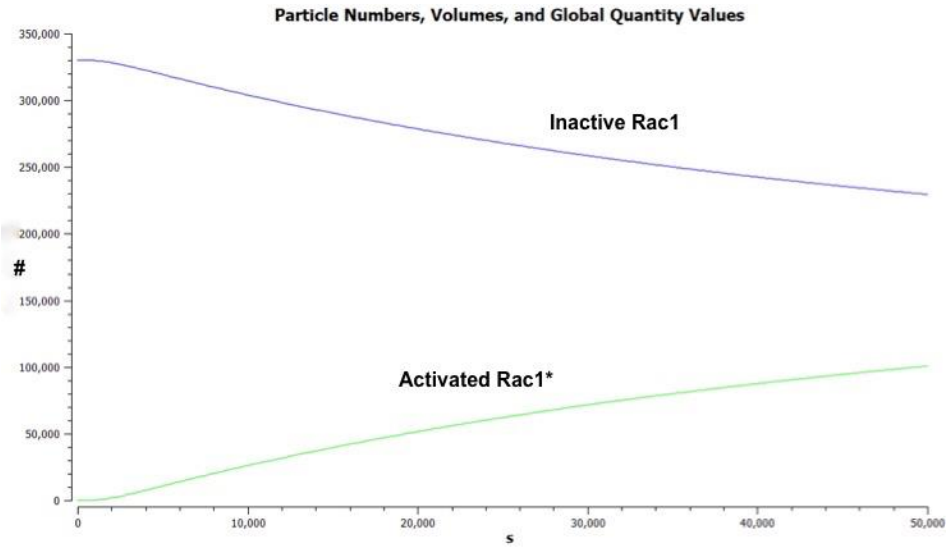

Figure 2.27. Particle numbers for inactive Rac1 and activated Rac1*in fourth version of the model. Time course 10,000 s

This is because there is likely to be some free Gβγ available, even in the total absence of SDF1, simply as a result of normal G protein metabolism. This in turn makes it likely that there will be some activated PI3K, ready to turn this large amount of PIP2 into PIP3, with downstream consequences for Rac1GEF and Rac1 activation.

In any case, the underlying assumptions in the previous version of the model, that PI4K and PI5K are active (and in earlier versions, that PIP2 is constitutively available in large amounts), is only valid if we assume that SDF1 stimulation also occurs from the outset. (Such large amounts of PIP2 are not normally seen in resting cells, being typically seen as the downstream of major stimulation events, such as by external chemotactic agents.)

For this reason It makes sense to include PI4K and PI5K activation in the same event as SDF1 addition and, in fact, because of the way COPASI implements events, the most straightforward way of doing this is to assume that, like SDF1, both kinases only come into existence at the same time as SDF1. This means also that, to keep the model consistent, both kinases are removed at the same time as SDF1, all of which requires a separate event.

Taken together, this requires further modifications to the original model, as shown in Tables 2.8 and 2.9.

Species	Compartment	Particle #	Concentration (M)	Notes
SDF1	Extracellular	0 or …	0 or 2.2e-08	Starts at zero levels, then is introduced and removed as two events.
PI4K	PM	0 or 24,000	0 or 6.2e-06	ditto
PI5K	PM	0 or 15,000	0 or 3.9e-06	ditto

Table 2.8. Changes in species quantities for fifth version of model, compared to previous versions.

Event name	Time triggered (Model time)	Results
SDF1, PI4K, PI5K On	1500 seconds	SDF1, PI4K and PI5K increased from zero levels to their working levels.
SDF1, PI4K, PI5K Off	8500 seconds	SDF1, PI4K and PI5K decreased from their working levels to zero levels.

Table 2.9. New events added to this fifth version of the model.

2.4.5 Results from Version 5 of the COPASI model

However, running this new model reveals a serious defect in all the models, which was not apparent until now. This can be seen from Figures2.28, 2.29, 2.30 and 2.31.

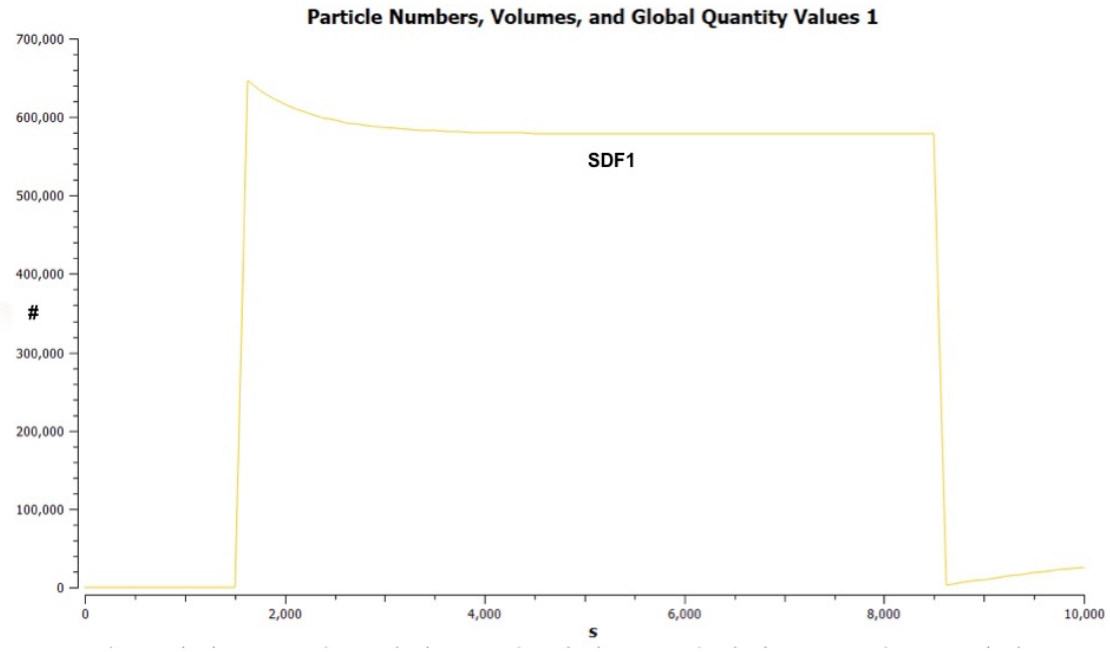

Figure 2.28.SDF1 particle numbers in fifth version of the model.
Peak particle numbers ~ 1900. Time course 10,000 s

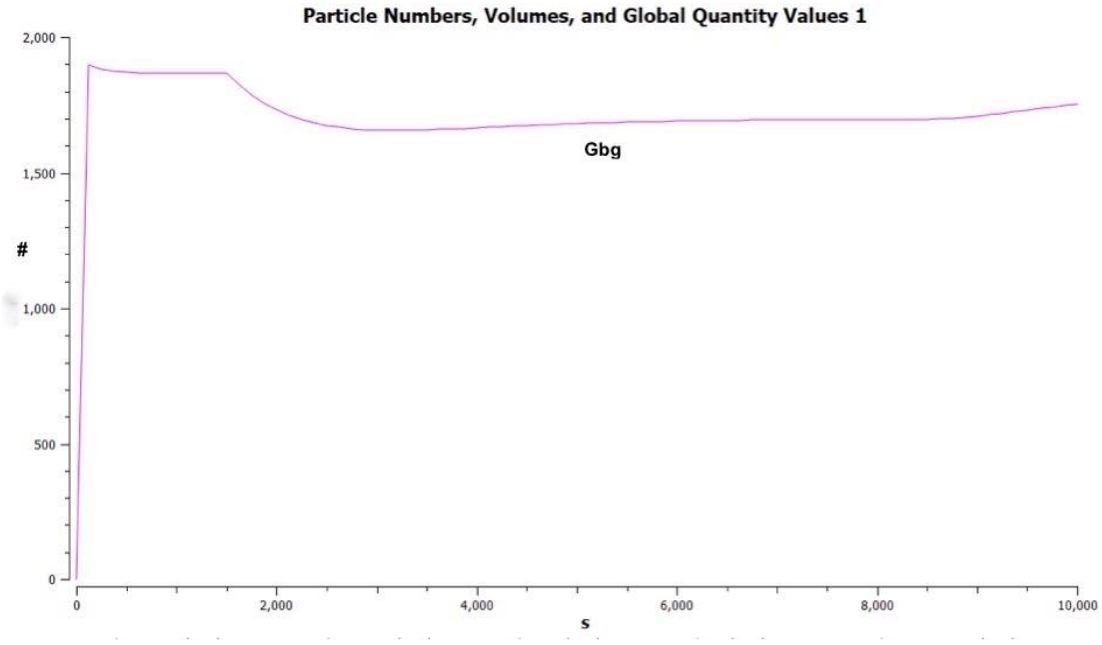

Figure 2.29. Gβγ particle numbers in fifth version of the model.

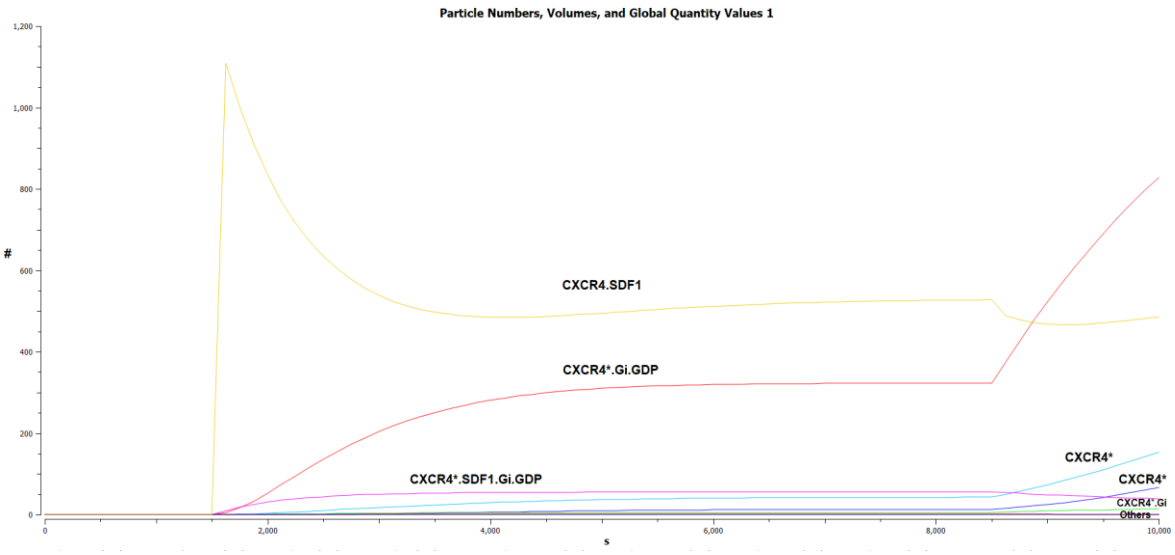

Figure 2.30. Particle numbers for all activated and/or SDF1-bound CXCR4 species in fifth version of model. (Except CXCR4.SDF1.Gi.GDP, for which peak levels are too high to fit in the same graph.) Peak particle levels of CXCR4.SDF1 ~ 1100 for reference. Time course 10,000 s.

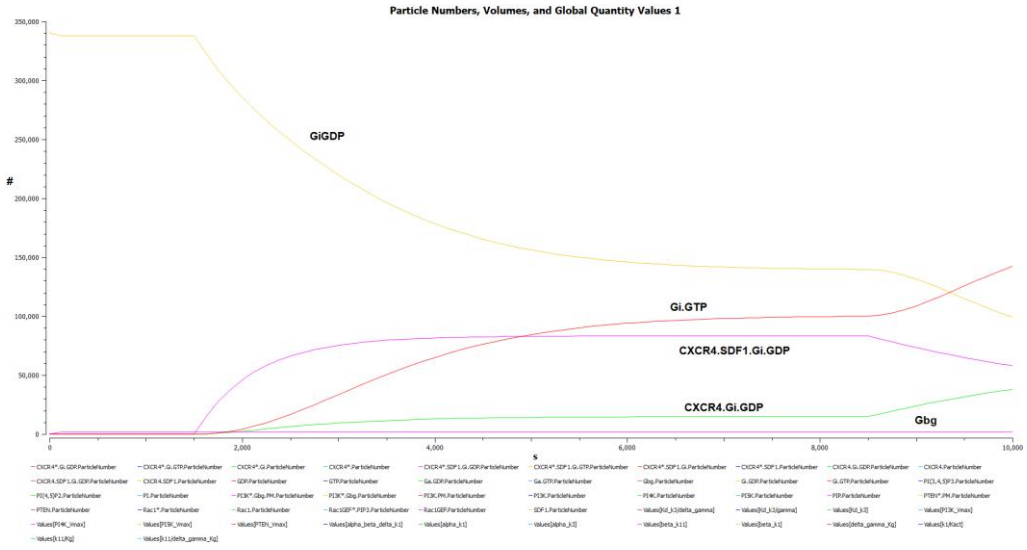

Figure 2.31.Particle numbers for CXCR4.Gi.GDP, CXCR4. SDF1.Gi.GDP, Gi.GTP, Gi.GGP and Gβγ in fifth version of model.

Despite SDF1 not being introduced until 1500 seconds have elapsed (Figure 2.28), levels of Gβγ have almost immediately risen to a steady state level of 1700 molecules (Figure 2.29), i.e. the same steady-state level as seen in the second and fourth versions of the model.

In other words, the Gβγ stimulus in this, and these previous models, is not generated as a result of SDF1 binding to CXCR4 but simply as a breakdown product of unstimulated G protein metabolism. This is confirmed by Figure 2.30, showing that no SDF1-bound or otherwise activated CXCR4 species are seen in any number until 1500 seconds have elapsed, i.e. as a result of SDF1 stimulation.

The reason for the lack of Gβγ can be seen in Figure 2.31, with all Gβγ, other than the initial 1700 molecules, locked up in CXCR4.Gi.GDP,CXCR4.SDF1.Gi.GDP and Gi.GTP. On further looking into this, it seems clear that the primary reason is that the G protein reaction schema adapted from Purvis and colleagues [181] (based on the Kinzer-Ursem and Linderman ternary complex model[180]) is faulty. An unusual feature of it is that, instead of the attached Gi.GTP separating into its component Gαi.GTP and Gβγ parts, while still attached to the activated CXCR4*receptor, the Gi.GTP is first released before being hydrolysed. In fact, the hydrolysis rate constant kGTP has been copied incorrectly from the Purvis model (as 0.0008 s^{-1}, instead of 0.013 s^{-1}) but correcting this does not resolve the problem of Gβγ effectively being locked up. (Although it does alter relative levels of Gα.GDP and Gα.GTP).Another unusual feature is that the reaction schema has a reaction step for activated Gα subunit re-association (Gα.GTP + Gβγ = Gi.GTP), which in this model is deemed to occur at the same rate as inactivated Gα subunit re-association (Gα.GDP + Gβγ = Gi.GDP). (In this regard, the model follows the same practice as the Purvis model, albeit substituting reaction rates specific for Gi G proteins.)

As a starting point, we should try replacing these reactions with a set of reactions that better reflect a more conventional understanding of GPCR dynamics. At the same time, now that it is more directly associated with GTP hydrolysis, the erroneous kGTP value can be replaced with a value based on a published GTP hydrolysis rate for Gαi subunits[197]. This results in the changes seen in Tables 2.10 and 2.11.

Reaction	Reaction equation	Action	Reason /Reference
Activated Gab subunit association (to Gi.GTP)	Ga.GTP + Gbg = Gi.GTP	Removed	Not considered conventional
CXCR4*.Gi.GTP releasing Gi.GTP (to CXCR4* + Gi.GTP)	CXCR4*.Gi.GTP = CXCR4* + Gi.GTP	Altered to CXCR4*.Gi.GTP = CXCR4* + Ga.GTP + Gbg	Replacement with more conventional interpretation of GPCR dynamics
CXCR4*.SDF1.Gi.GTP releases Gi.GTP (to CXCR4*.SDF1 + Gi.GTP)	CXCR4*.SDF1.Gi.GTP = CXCR4*.SDF1 + Gi.GTP	Altered to CXCR4*. SDF1.Gi.GTP = CXCR4*. SDF1 + Ga.GTP + Gbg	Ditto
Gi.GTP hydrolysis (to Ga.GDP + Gbg)	Gi.GTP ->Ga.GDP + Gbg	Altered to Ga.GTP ->Ga.GDP	Ditto

Table 2.10. Changes to reactions in this model, compared to previous models.

Rate constant value	Function	Old value	New value	Reference
kGTP	Rate of hydrolysis of Gα-associated GTP (to Gα.GDP)	0.008 s^{-1}	0.05 s^{-1}	Replacement with Gαi.GTP appropriate value[197].

Table 2.11. Change to rate constant value (Global variable), compared to previous models.

2.4.6 Results from Version 6 of the COPASI model

Running this new version of the model leads to some changes in G protein-related behaviour, as shown in Figures 2.32and 2.33below.

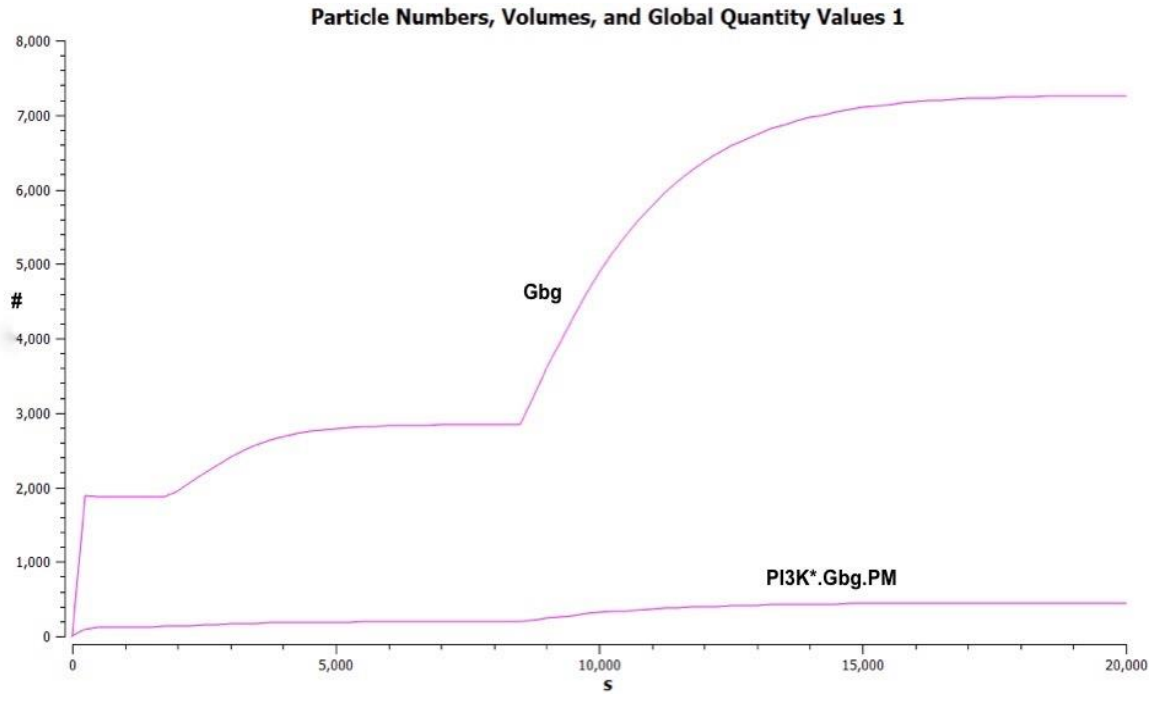

Figure 2.32. Particle numbers for Gβγ and PI3K*.Gβγ.PM in sixth version of model.

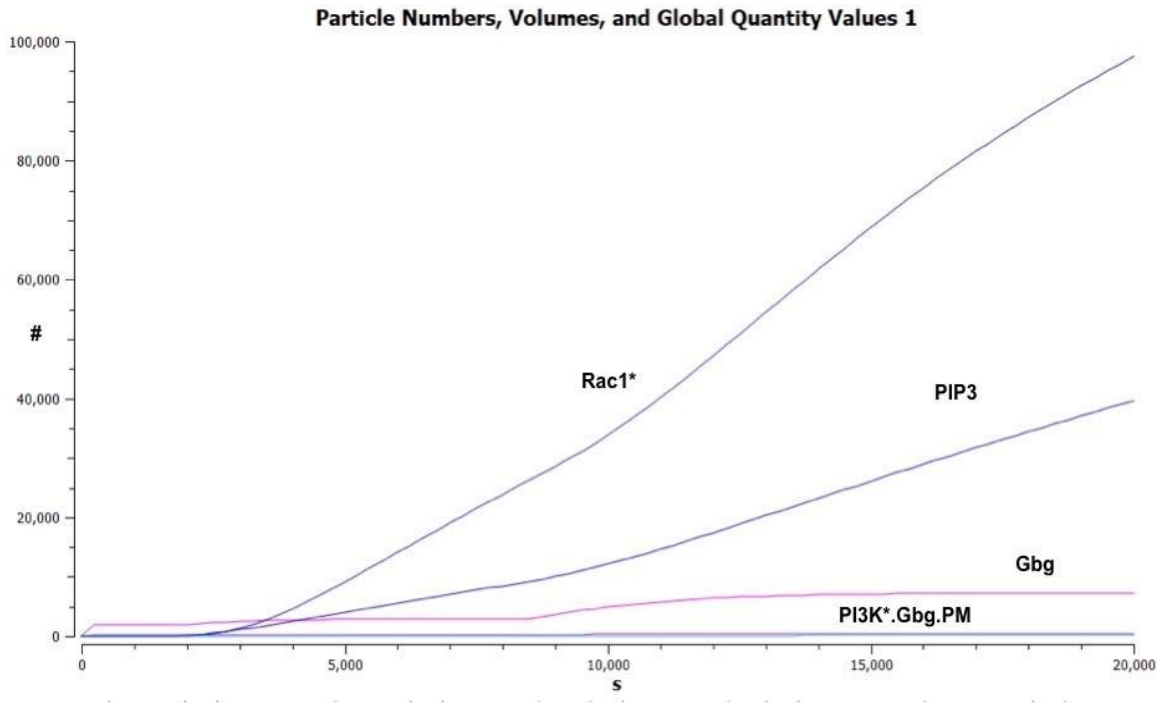

Figure 2.33. Particle numbers for Gβγ, PI3K*.Gβγ.PM, Rac1GEF*.PIP3 and Rac1*in sixth version of model.

These show a distinct increase in levels of Gβγ and (less obviously) in levels of activated PI3K*.Gβγ.PM (Figure 2.32) with concomitant increases in levels of Rac1GEF*.PIP3 and activated Rac1*(Figure 2.33). Furthermore, It seems clear that these elevated levels of Rac1GEF*.PIP3 and Rac1*are the direct result of SDF1 stimulation in this new version of the model, given that they both only start to rise after the point where SDF1 stimulation first starts. What is very puzzling in both these Figures is that, instead of falling at the 8500 second mark, when SDF1 is removed, there is actually a further increase in both. Indeed, this increase is clearly even larger than when SDF1 is introduced. This is further supported by the greater rate of increase in Rac1GEF*.PIP3and Rac1*seen shortly after this time.

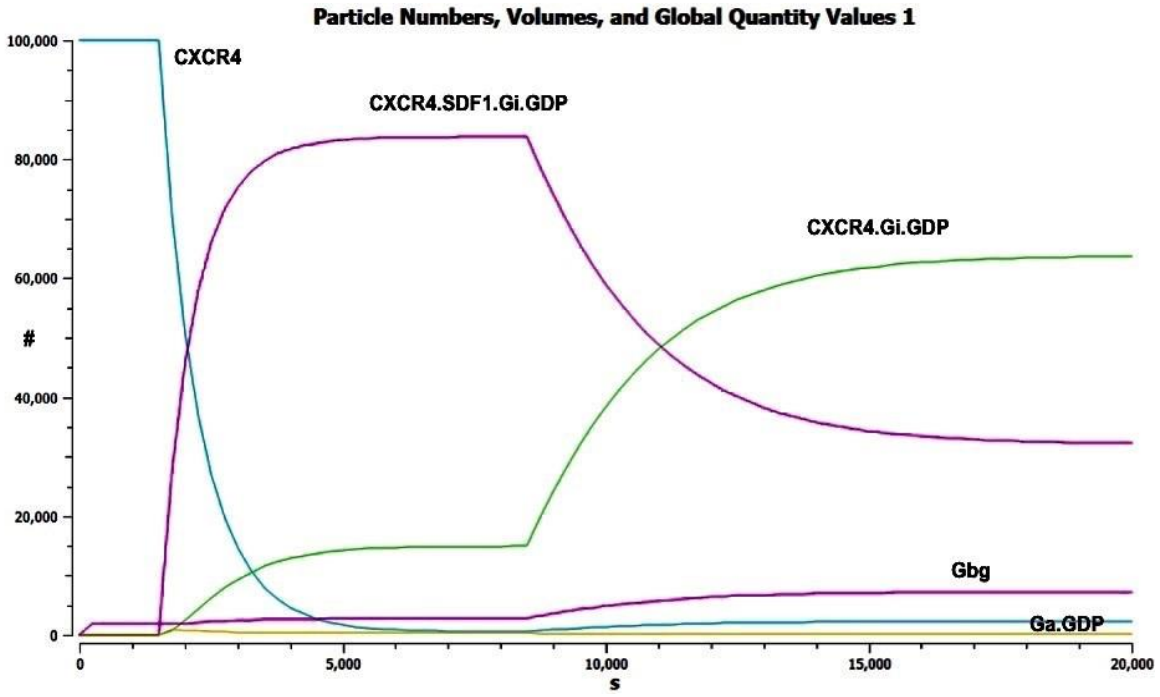

Figure 2.34. CXCR4.Gi.GDP, CXCR4. SDF1.Gi.GDP, Gi.GTP, Gi.GGP and Gβγ in sixth version of model.

Looking at Figure 2.34, it seems clear that, as SDF1 is added, the vast majority of CXCR4 is replaced by SDF1-bound but inactive CXCR4.SDF1.Gi.GDP. As this SDF1 is then removed and levels of CXCR4.SDF1.Gi.GDP fall away, only around three-quarters is replaced by CXCR4.Gi.GDP, implying that the remainder must be filled by other non-SDF1-bearing CXCR4 complexes. Surprisingly, as Figure2.35shows, several of these are activated CXCR4 forms. Since the vast majority of this is in the form ofCXCR4*.Gi.GDP, it is only to be expected that the bound Gi.GDP will then be activated (to Gi.GTP) and then hydrolysed to Gα.GDP and Gβγ.

Furthermore, since it is an assumption of the model (in common with the Purvis [181] and Kinzer-Ursem and Linderman[180]models) that release of G protein components by activated CXCR4*doesn't immediately result in its inactivation, there may well be time for further cycles of Gi.GDP activation, release and hydrolysis while this unbound CXCR4*is still active. Given that this underlying assumption is not unreasonable, it may be that this "double bounce" effect can be viewed as a credible prediction of the model, to be tested against real-world behaviour.

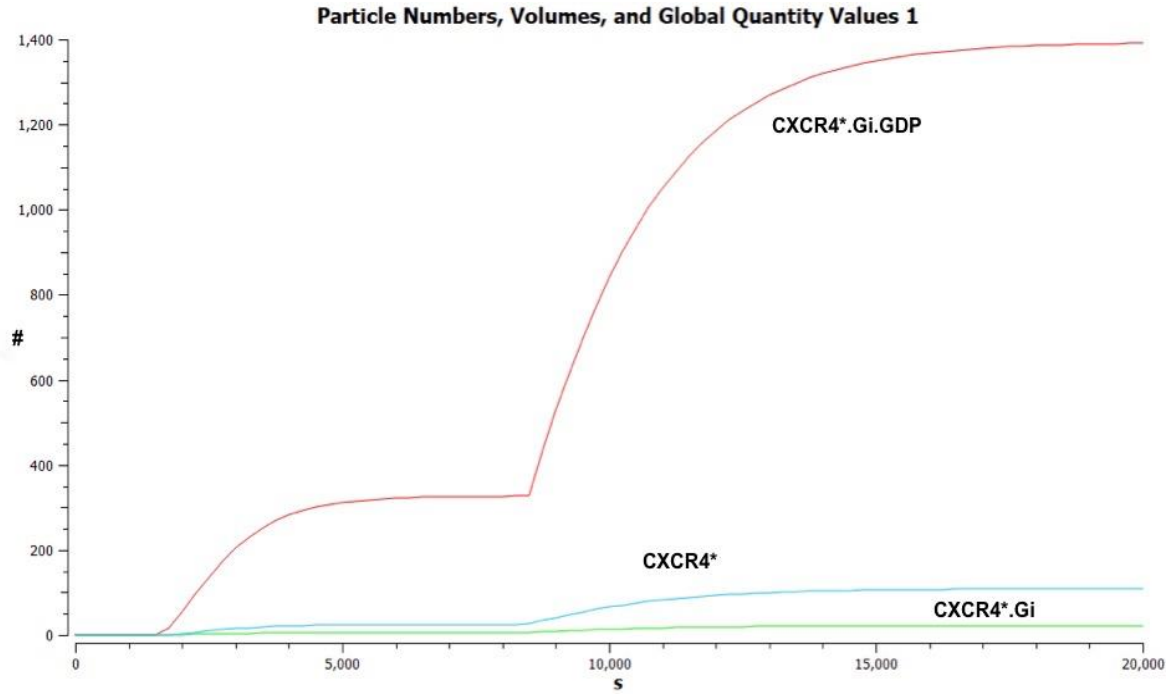

Figure 2.35. CXCR4*.Gi.GDP,CXCR4*.Gi and CXCR4*in sixth version of model.

However, it is unlikely that the same can be said for the fact that Gβγ levels fail to fall away as SDF1 stimulation is removed. Rather, this appears to be the result of further limitations of the model as it currently stands. One possible cause, not previously addressed, is that intracellular levels of GTP and GDP are assumed to be available in excess and in equal amounts (specifically, at concentrations of one millimolar each). In fact, according to some reports, under physiological conditions, intracellular levels of GTP are typically around 10 times more numerous than GDP[195]. Taking an average value from the literature for GTP [198], this leads to further changes to the existing model, as shown in Table 2.12.

Species	Compartment	Particle #	Concentration (μM)	Notes
GTP	Cytosol	~450,000,000	500	Average of two reported values [198]
GDP	Cytosol	~45,000,000	50	One-tenth GTP value [195]

Table 2.12. Changes to species in seventh version of model.

2.4.7 Results from Version 7 of the COPASI model

Implementing these changes leads to Gβγ levels falling after SDF1 removal, although not by much, as can be seen in Figure 2.36.

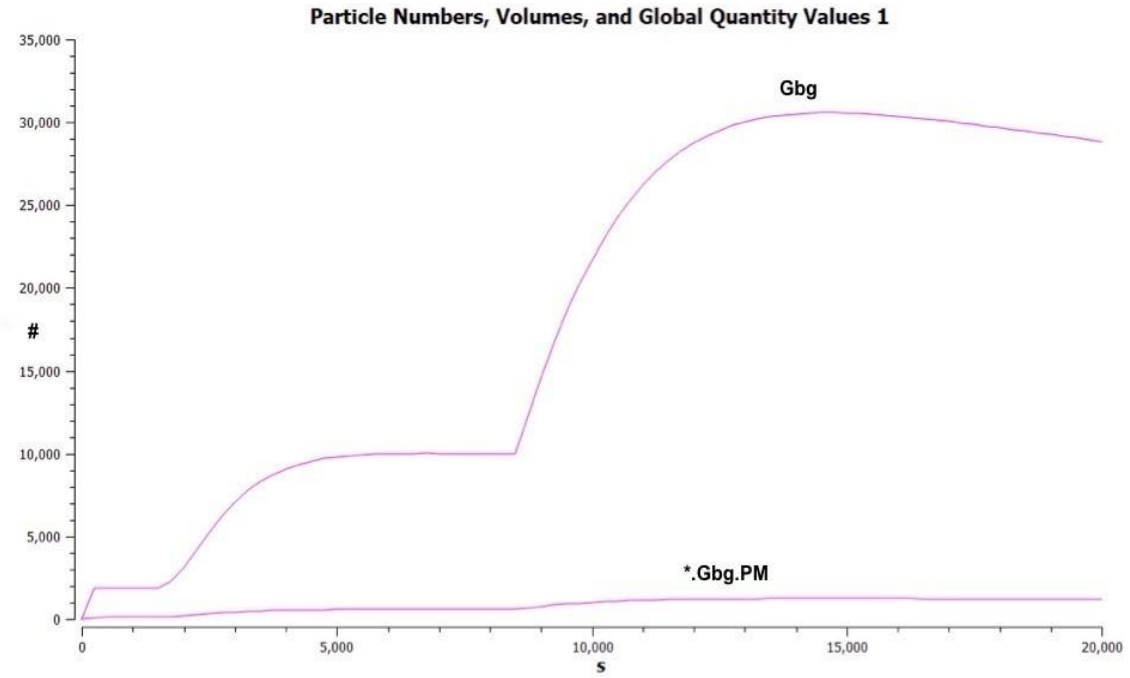

Figure 2.36. Particle levels for Gβγ and activated PI3K*.Gβγ.PM in seventh version of model.

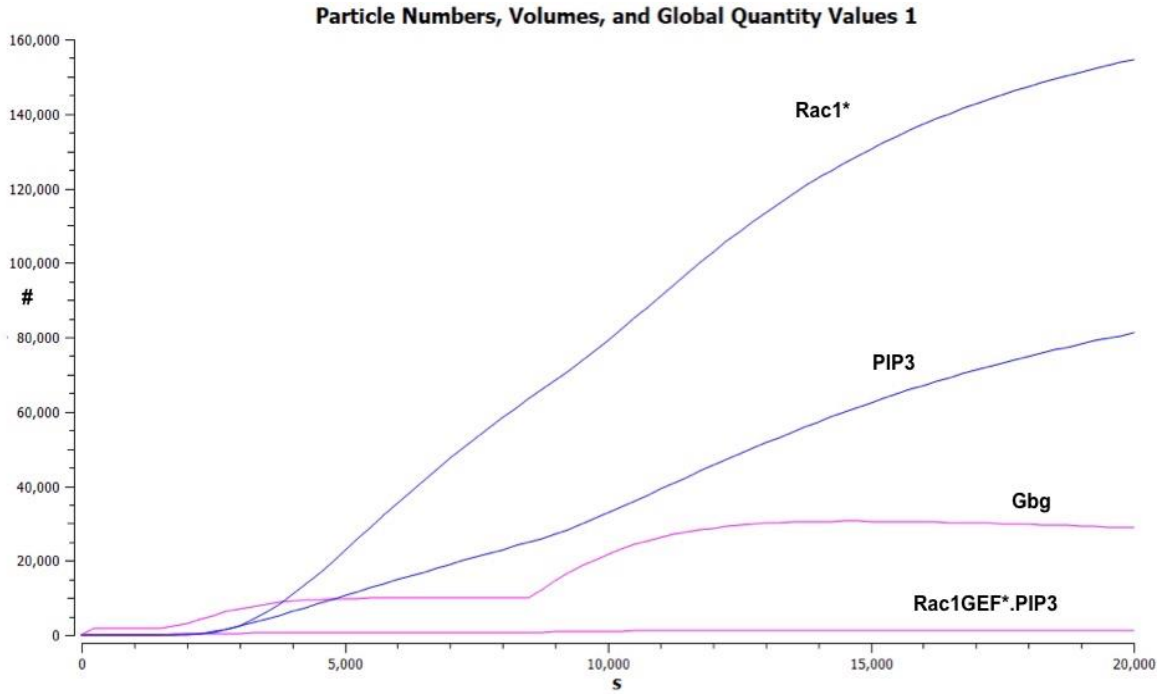

Figure 2.37. Particle levels for Gβγ, PIP3, Rac1GEF*.PIP3 and Rac1* in seventh version of model.

More surprisingly, reducing both GTP and GDP levels and altering the balance between them, has led to a nearly fourfold increase in peak Gβγ levels (from around 7500 molecules to over 30,000, over 20,000 seconds), a near-doubling of PIP3 levels (from around 40,000 molecules to 80,000) and an approximately 60% increase in activated Rac1*levels (from around 100,000 to around 160,000 molecules), as shown in Figure 2.37. In fact, as the results of parameter scans on GTP and GDP concentrations show, in Figures 2.38 and 2.39, decreasing GDP levels results in a substantial increase in Gβγ levels, whereas decreasing GTP levels has the opposite result.

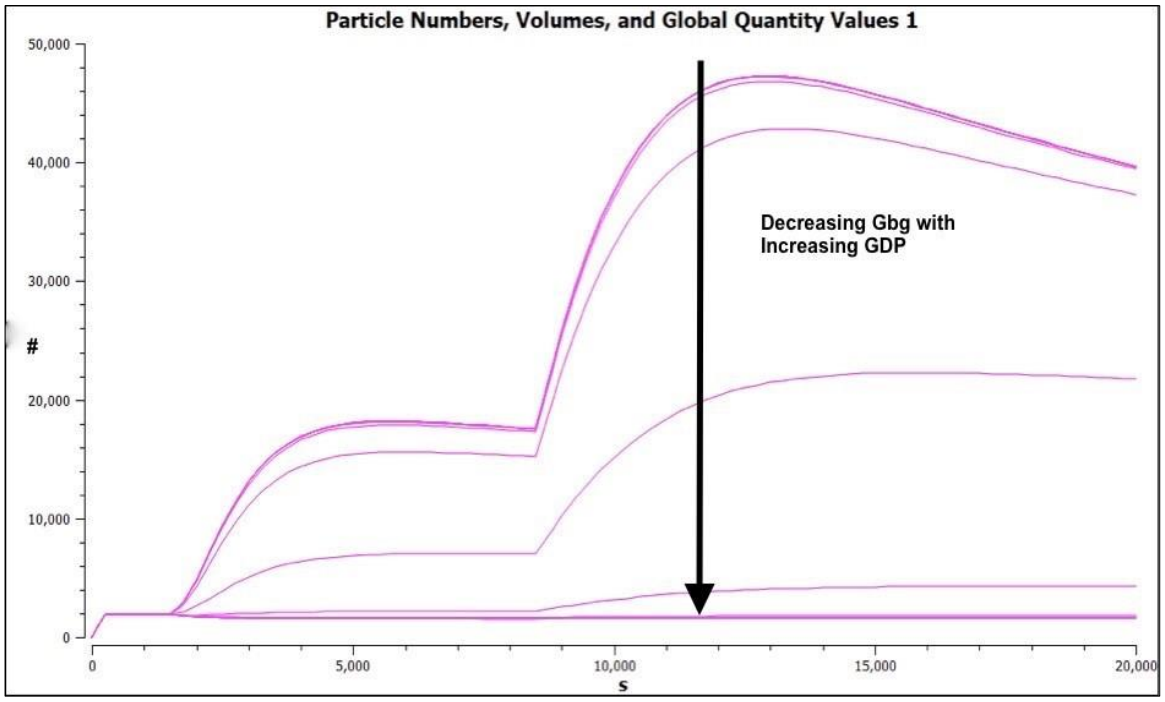

Figure 2.38. Results of parameter scan on GDP concentrations as affecting Gβγ levels in seventh version of model.

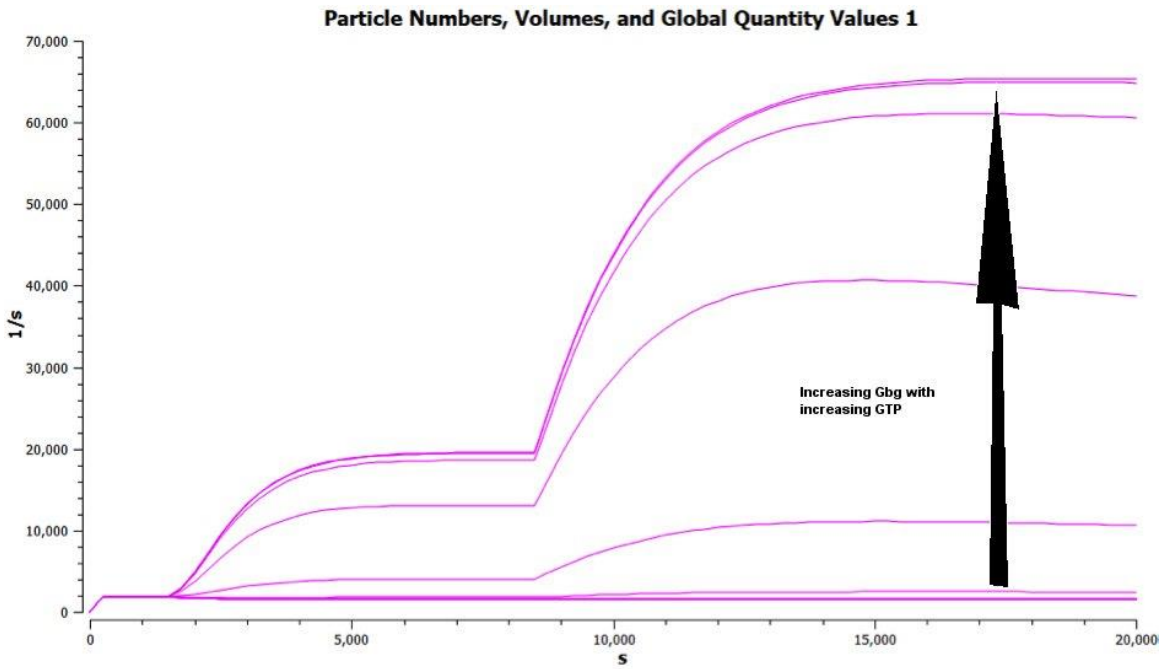

Figure 2.39. Results of parameter scan on GTP concentrations as affecting Gβγ levels in seventh version of model.

It is noticeable also, that the lowest GDP concentrations in Figures 2.38 are associated with a decline in Gβγ levels, after the initial two increases, something not seen at higher GDP concentrations.

At this point, despite some of the obvious problems with the submodel, as just discussed, and in terms of its overall runtime, its development has begun to stabilise to the point where various forms of sensitivity analysis can be justified. As explained earlier, such analysis helps to determine what parameters are critical to a model's behaviour. Performing a time series(as opposed to a steady state) analysis for Gβγ, PIP3 and activated Rac1* levels over 20,000 seconds (and then scaling the result), produces the results seen in Figures 2.40-42.

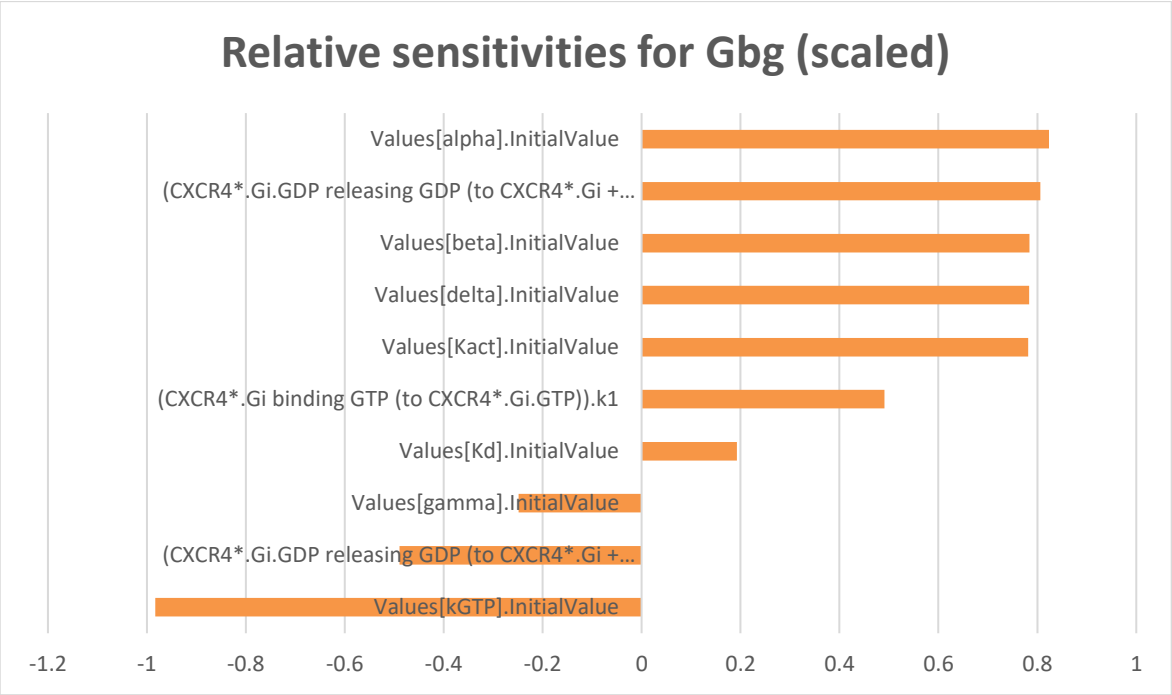

Figure 2.40. Selected results of time series sensitivity analysis for model Gβγ particle levels over 20,000 seconds. Scaled output plotted in MS Excel.

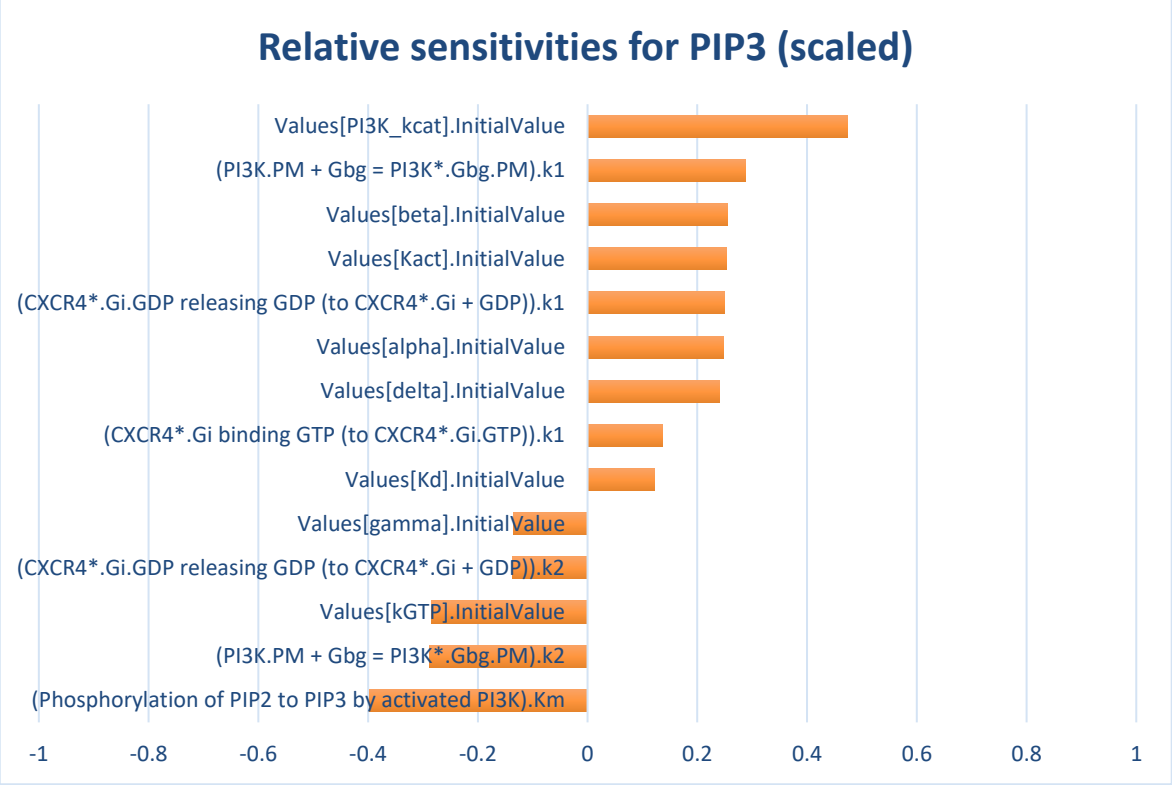

Figure 2.41. Selected results of time series sensitivity analysis for model PIP3 particle levels over 20,000 seconds. Scaled output plotted in MS Excel.

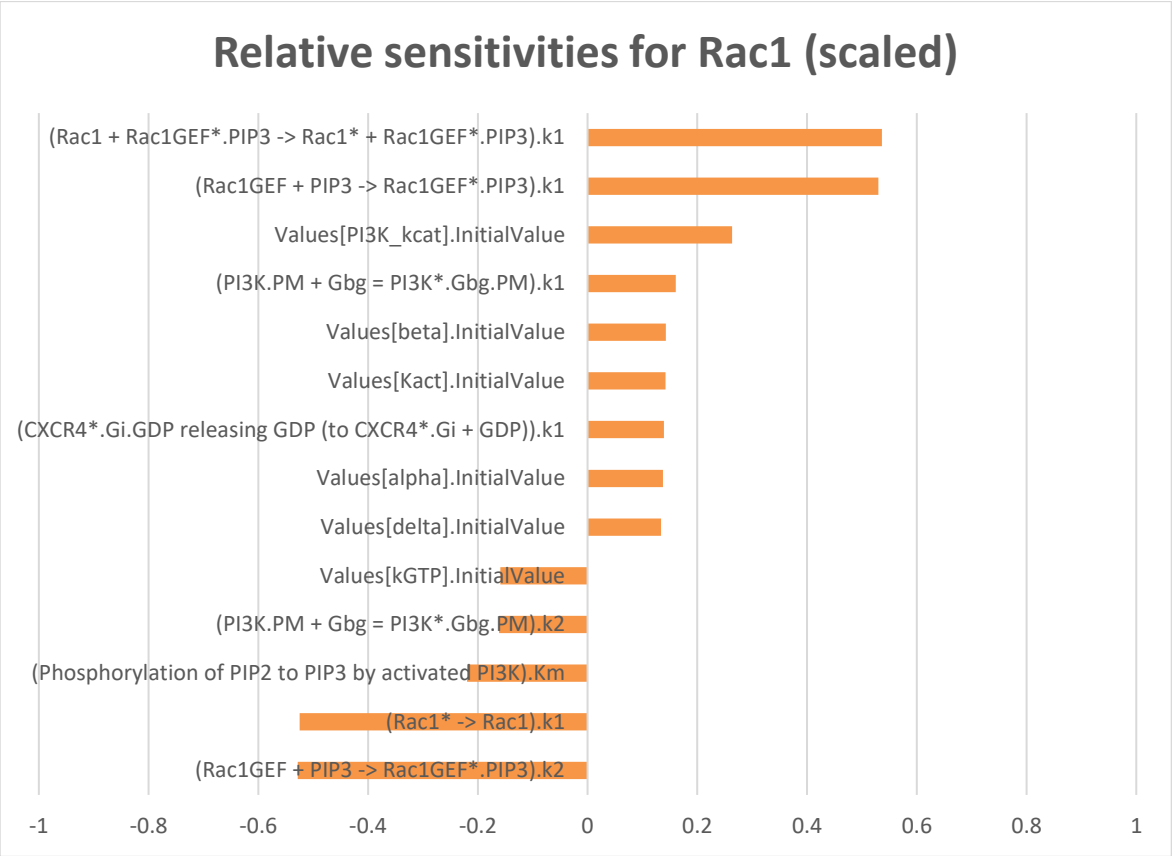

Figure 2.42. Selected results of time series sensitivity analysis for model Rac1* particle levels over 20,000 seconds. Scaled output plotted in MS Excel.

An interesting trend noticeable here is how various constants that determine CXCR4-SDF1 binding and associated G protein changes, show a decreasing influence as the sensitivity analysis moves further downstream from CXCR4-SDF1 binding. So, for instance, the constant Alpha, a measure of the extent to which SDF1 binding facilitates CXCR4 activation, has a relative sensitivity value of 0.82 (i.e. 82% of the maximum) for Gβγ levels, but only a value of 0.14 for levels of activated Rac1*. Similarly, the constant kGTP (the GTP hydrolysis rate constant) decreases substantially from a very high value of -0.98, as a negative influence on Gβγ levels, to a more modest value of -0.16 for activated Rac1* levels.

Of course, there is no surprise in this, given that, the further downstream the analysis, the more factors are inevitably involved. However, it illustrates a common finding: that increased reaction

path length in signalling cascades is often associated with increased robustness, in terms of the final response to the initial stimulus.

Certainly, a number of systems biology models have been shown to have a very low sensitivity to changes in kinetic parameters[199], a finding that extends to signalling networks[200,201,202]. This may help to explain the notable insensitivity of the more downstream species, such as Rac1*, to changes in upstream reaction rate constants, observed in the HeLa model described above.

Unfortunately, these latest sensitivity analysis results serve only to reinforce an overall impression that this lamellipodia submodel is likely to remain a poor model of the underlying reality, even after further substantial changes are made. What this analysis seems to be suggesting is that the scope for adjusting the model, to produce a much more rapid Rac1*response, maybe limited to only a few very downstream factors. Given how much the model needs to be speeded up (as explained in the next section), it is not at all clear that such changes could be justified by reference to the literature or by other arguments based on our current understanding of the relevant biology.

Even if this is not the case, other considerations suggest the model does not justify further development. One of these is the number of parameter values that have had to be estimated, in the absence of reliable data from the literature. Given the large number of such unknowns, together with the continuing lack of confidence in the submodel's behaviour, and lack of understanding as to why it is not behaving much better than it is, it is hard to justify the application of parameter estimation in order to bring its behaviour within more acceptable limits. It would, in any case, be very challenging to specify desirable constraints for such a parameter fitting process, so that the results achieved what was intended. Put simply, the model is too large, there are too many unknowns and the outputs are too disparate for such an exercise to be likely to prove beneficial.

For all these reasons, further development of the model is not considered worthwhile beyond this point, with no further results taken. This decision is explained in more detail in the next section.

2.5 Discussion

2.5.1 Reasons for abandoning thisSDF1-CXCR4 lamellipodia submodel

The fact that after a number of changes have been made to it, G protein dynamics in this submodel still continue to raise doubts, while the overall timescale required for critical outputs such as PIP3 and Rac1* to reach levels of several thousand molecules still remains in the order of hours rather than minutes, suggest that the model does not justify further development. By this is meant not just addition of the various additional species and reactions necessary for Rac1-mediated cytoskeletal rearrangement, referred to in the introduction to this chapter (and set out as a principal aim of the project in the previous chapter), but also any further attempts to speed up the model to a more realistic time frame.

What should such a time frame be? Various reports suggest that the timescale of a single cycle of extension and retraction for an individual lamellipodia is in the order of 3-10 minutes [203,204,205],consistent with other reports that the timescale of cell spreading is of the same order [206]. This timescale also seems consistent with various reports concerning average lifetimes of lamellipodia-associated FAs, ranging from seconds to minutes depending on size and maturation levels [207,208,209]. Given these FA lifetimes, it seems highly unlikely that this submodel can be altered to provide outputs such as PIP3 and actin-related data, which can usefully inform the intended FA submodel to be developed in the next phase.

If the development of this lamellipodia model had been the primary aim of the project, a different decision might have been reached. However, the fact that the main aim is to investigate FA dynamics, and (as will be shown later) there are alternative ways of achieving this, means that it

makes most sense to abandon this submodel at this point, and to progress to the next phase, albeit in a rather different form than originally intended.

2.5.2 Revised project aims

As was explained in the previous chapter, as well as FAs being activated by so-called inside-out signalling (modelled in the form of SDF1-stimulation here), there is another form of activation, outside-in signalling (Figure 1.3). This happens when integrin-ligation by appropriate ECM proteins (such as collagens, fibronectin and laminin) stimulates formation of FAs, via a similar PIP2-driven mechanism by which inside-out signalling drives such formation.

Therefore, the primary aim of this project will now be to develop such an outside-in model. This has the additional benefit that it does not require integration with another submodel, as originally intended, and of generally being much less ambitious. Given the problems encountered with this submodel, such a scale-back in ambition appears the wisest course of action in terms of this project.

In light of the lessons learnt here, it would seem to make sense to develop this model more gradually than was attempted in this case. Therefore, only a relatively simple FA model, consisting of integrin, a generic ECM Ligand, talin, and a very basic phosphoinositide mechanism (even simpler than the one used here) will be developed in the next phase, described in Chapter 3. If this demonstrates overall behaviour consistent with what is seen physiologically, this will then be developed in the final phase into a much more developed model of FA assembly and disassembly. This phase will be described in Chapter 4.

However, there are always lessons to be learnt, even from failed models, and these will be discussed in the next section.

2.5.3 Insights and lessons to be drawn from the SDF1-CXCR4 lamellipodia submodel

The SDF1-CXCR4 chemotaxis model described in this chapter illustrates many of the challenges of modelling cellular biology, many of which challenges are not just encountered with differential equation models. Some of the more important ones are listed below.

2.5.3.1 Lack of relevant data in the literature to inform this submodel

In particular, reliable information relating to species initial concentrations and reaction rate constants can be extremely elusive, if it is available at all. This was seen in the case of Rac1GEF activation by PIP3, where arbitrary values were chosen for binding kinetics, after an exhaustive search in the literature failed to find consensus values for Rac1GEFs. Similarly, the decision to base CXCR4-SDF1 binding and related G protein dynamics on a published model of a different GPCR system was heavily influenced by considerations of the time likely to be required in finding rate constants for all the different reactions involved. As it turned out, this decision had unexpected consequences, with the model not behaving as expected, especially concerning Gβγ dynamics. This may be because, in the original model, there was no downstream role for Gβγ, whereas, of course, in this model it plays a critical role.

2.5.3.2 Choices of which proteins and isoforms to include in this submodel

There may be additional challenges determining which of a number of different isoforms of a particular protein to include in a model, and how best to model certain interactions between different species. Thus, for PI3K, there are five different forms of the p85 regulatory subunit (p85α, p85ß, p55α, p55ß, p50α) and three different isoforms of the p110 catalytic subunit that they regulate (p110α, p110ß, p110δ) [Geering et al, 2007].Only one of these three catalytic subunits (p110ß) is known to be influenced by Gβγ, with such simulation being particularly implicated in GPCR-mediated cell migration [Kurosuet al, 1997; Dbouk et al, 2012]. However, in addition to the three p110 catalytic mentioned, there is a fourth one, p110γ, that is regulated by another class of regulatory subunit, p101,, and is also known to be simulated by Gβγ and involved in chemotaxis [Brock et al, 2003].

Given that PIP3 is thought to be the principal intermediate link between such PI3K activation and pro-migratory activities such as lamellipodia formation, there was a strong case for including both the p110β and p110γ isoforms of the model from the outset. Certainly. including the latter would clearly have made the model more complete. However, it would also have required additional reactions for its activation and PIP2 phosphorylation activity, adding to model complexity and, most likely, requiring considerable time to obtain the necessary reaction rate constant values. Alternatively, both isoforms could have been batched together. However, this would have required estimating suitable reaction rates for their joint activation and PIP2-phosphorylation activity, which might well have been very challenging to determine. For these reasons it had been intended to leave the addition of the PI3Kγ isoform to a later version of the model, once the model had begun to show generally acceptable behaviour, and when the effects of adding it could be more clearly assessed.

As another example of the often difficult choices to be made concerning what species to include, a number of Rac1 GEFs (including TIAM1, TRIO, PREX1 and two VAV species) have been grouped together as a single species (named Rac1GEFs). These have been selected from a larger group of reported Rho-family GEFs, based on their reported Rac1 activity. However, this is by no means a clear-cut decision. For instance, a third member of the VAV family (VAV3) has been excluded, purely on the basis that it is reported to show a preference for RhoA and RhoG over Rac1 [210], meaning that the copy number of Rac1GEFs is 4000 less than it might otherwise be [182].

2.5.3.3 Decisions over reaction aggregation and simplification

Where, as in the case here, a system is comprised of a relatively large number of reactions, it may be necessary and beneficial to simplify and abbreviate some of these reactions, at least at the beginning of model development. For instance, in this model Rac1GEF activation of Rac1 has been abbreviated to a single reaction with irreversible mass action kinetics, whereas, in fact, like other GEFs, they accelerate the exchange of GDP for GTP (bound by the Rac1 monomeric GTPase) in an essentially enzymatic reaction. A fuller modelling of such activation would require three reactions at the very least, one for Rac1GEF-Rac1 binding, another (employing enzyme kinetics) for GDP-GTP exchange and a final one for Rac1-GTP activation. However, this would involve adding two further reactions to the twenty-eight in the initial model. For a model in the early stages of development, it is unlikely that such a modification would add any real benefit, and is something that can usually be deferred to a later stage.

2.5.3.4 Decisions over what functionality to include in this submodel

The same considerations apply on a larger scale, when considering what aspects of system behaviour to include in the model. For instance, it is known that tyrosine phosphorylation also plays a large role in activating PI3Ks, with such phosphorylation typically mediated by receptor tyrosine kinases, such as growth factor receptors[211].It has also been shown that the Gαq G protein subunit plays a role in SDF1-mediated chemotaxis, apparently exerting an inhibitory role [170]. However, adding in both of these aspects would greatly add to the complexity of the model, without necessarily making it any more useful as a predictive or explanatory tool.

Before expanding any model in this way, it is essential to have it working reliably in the first place. Ideally, in most cases, this means starting with the simplest model that can produce predictable results, then using experimental data to validate it. Only once this model has shown itself to reliably produce valid results should one then proceed to expand it to include other behaviours, and only then by small increments, further validating it with each step. Clearly by this standard, the model described here as been over-ambitious from the outset.

2.5.3.5 Significance of choices concerning which compartments species are resident in

Turning to more specific aspects of the model, it became clear very early on that determining which model compartment a particular species should be resident in can be critical to the subsequent behaviour of the model. This is because the rate of most reactions is heavily determined by the concentrations of the various reactants, with reactants in higher concentrations being more likely to encounter each other than those in lower concentrations (simplifying it somewhat).

For example, changes to the second version of the model to make G protein subunits (including, most importantly, Gβγ) membrane-resident rather than cytosolic (as in the first version) resulted in much higher levels of PI3K.Gβγ .PM, PIP3, Rac1GEF*.PIP3 and Rac1*(Figures 2.11-14) than in the first version (Figures 2.2-3, 2.8-10, respectively). Similarly, changing inactive Rac1GEF from cytosolic to membrane-resident in the third version resulted in much higher levels ofRac1GEF*.PIP3 and Rac1*(Figures 2.16 and 2.17) than in the second version (Figures 2.13 and 2.14, respectively). As the case of Gβγ shows, such decisions can be particularly important in the case of species early on in a signal cascade, or similar, where there are successive amplifications of other downstream species. Obviously, though, the modeller must be guided by the evidence in deciding which compartments allocate a particular species. In this case the published literature conclusively pointed to G protein subunits being membrane-bound; however, there was insufficient evidence to support such residency for inactive Rac1GEF.

2.5.3.6 Differences resulting from changing modes of availability of key species

Changing PIP2 from being available *ab initio* as a fixed amount (20000 molecules) to being generated by PI5K (from PIP, itself generated from PI by PI4K) resulted in much more PIP2 in the fourth version of the model as compared to earlier versions (Figure 2.22) and therefore in increased levels of PIP3 (despite no increase in levels of activated PI3K), Rac1GEF*.PIP3 and Rac1* (seen by comparing Figures 2.23, 2.26 and 2.27 with 2.12-14, respectively). This is to be expected, given that within 30 seconds peak levels of PIP2 had reached 105,000 molecules (Figure 2.22), more than five times the fixed level of earlier versions, leading to similar increases in Rac1GEF*.PIP3 and Rac1*. However, it seems remarkable that so much PIP2 can be generated so quickly (using enzymatic rates based on the literature), giving some idea of how PIP2 is able to drive lamellipodia and accompanying FA formation at the rate seen physiologically.

Introducing SDF1 as a delayed bolus in version 5, rather than making it available immediately, as in earlier models, revealed a problem with the way G protein dynamics was modelled, not apparent previously. The early rapid increase in Gβγ levels that had been assumed to be a response to SDF1 stimulation, turned out to be entirely the product of normal G protein metabolism, as based on a reaction scheme used in a published model[181], itself based on a published reaction schema [180]. Since the G protein reaction scheme in question appeared to work without problems in that model, it was assumed that the modified version used here would also work acceptably. However, in that particular model Gβγ did not perform any useful role so any issues, if they arose, might not have been noticed.

Having removed, amended or replaced some of the more unconventional G protein-associated reactions from the model resulted in increases in Gβγ that could be attributed entirely to SDF1 stimulation, as initially intended. Moreover, this sixth version of the model also shows further increases in PI3K.Gβγ .PM, PIP3, Rac1GEF*.PIP3 and Rac1*(Figures 2.32-33) over and above those already seen. However, there is also some puzzling behaviour by the model in relation to G protein dynamics. Instead of levels of Gβγ, PI3K.Gβγ .PM, PIP3, Rac1GEF*.PIP3 and Rac1*falling back after the removal of SDF1 stimulus, this removal is accompanied by another further rapid increase of these species (Figure 2.32-33). It is possible (if unlikely) that this sustained "double bounce" behaviour is in agreement with what is seen physiologically, and could be considered a valid testable prediction of the model. However, the failure of Gβγ levels to fall after SDF1 removal, suggests that there is still a problem with the model, in terms of G protein-related behaviour, that needs to be resolved. If, after such resolution the same "double bounce" behaviour continues then experiments to see if this is replicated in reality might then be considered.

2.5.3.7 Considerations when supplying key species in excess

The last version of the model altered levels of GTP and GDP both absolutely and relative to each other, so that there was 10 times as much GTP as GDP, as opposed to being in equal numbers in earlier versions. Despite the fact that both had been greatly in excess of all other cellular species (at one millimolar concentrations), and remained so after the change (at 450,000,000 and 45,000,000 molecules respectively), these alterations in relative levels led to a fourfold increase in peak Gβγ levels (as seen by comparing Figure 2.36 with 2.32) and substantial increases in activated downstream species such as Rac1GEF*.PIP3 and Rac1*(Figures 2.38 and 2.39). Subsequent parameter scans revealed that decreasing GDP levels and increasing GTP levels both result in increasing peak Gβγ levels, explaining this difference. This result serves as a reminder that relative levels of species can still be important in determining model behaviour, even when vastly in excess.

2.5.3.8 Path length and system robustness

Finally, sensitivity analysis of the model reveals that sensitivities to upstream factors in the model decrease as one moves downstream, so that, for instance, Rac1* levels are less sensitive to changes in factors affecting G protein dynamics than are Gβγ levels, even though there is some amplification observable through the system. Thus the downstream behaviour can be said to be relatively robust in this respect, in the sense of being influenced by a large number of different factors, both upstream and downstream (Figures 2.40-42).

2.5.3.9 Overall slowness of the model

It is clear that the model would need a number of changes to the overall timescale of events, if it is to reflect the physiological reality. The timings for lamellipodia and FA lifetimes mentioned in section 2.4.1 suggests that this would need to be in the order of tens of minutes, at most, rather than taking more than 10 hours, as it currently does. However, as was seen in the case of moving G proteins and inactive Rac1GEFs, reducing the size of compartments can lead to very substantial reductions in the rate of the reactions in which they are involved. This, in turn, can result in substantial reductions in the overall runtime of model simulations.

In reality, the actual (or effective) volume of both the cytosol and PM are much smaller than they might appear from a simple calculation. After all, as well as being full of various organelles, such as the nucleus and ER, already discounted here (section 2.2), the cytosol is also full of proteins and other macromolecules, as well as fibres, as in microtubules and the actin cytoskeleton. Given that both of these are associated with lamellipodia formation, which we are modelling here, one could reasonably assume that the effective reaction volume in the vicinity of lamellipodia maybe substantially reduced, compared to the rest of the cytosol. Similar considerations apply to the PM, where, in the immediate vicinity of lamellipodia formation, the local resident protein population will almost certainly increase considerably, as a result of large-scale recruitment of cytoskeleton-associated protein such as profilin, CapZ and gelsolin by PIP2 and other phosphoinositides[212,213].

However, this finding does not just apply to the major cellular compartments, such as the PM and cytosol. In reality, the actual (or effective) volume in which a reaction occurs may be much smaller than this. The obvious example is membrane species confined to lipid rafts, or similar membrane microdomains. Here, clearly, any reaction between species resident in such microdomains will involve a reaction volume far smaller than the volume of the PM itself. Similar

considerations apply where reactions are confined in close proximity as a result of being bound together on scaffold proteins. A good example here is the protein IQGAP1, which acts as a scaffold protein for PI4K, PI5K and PI3K, feeding the product of one into the other, and thereby considerably speeding up the production of PIP3 [214]. Confinement of reactants within lipid compartments, such as endosomes, is another case where one would expect reactions to proceed far more quickly (all other things being equal) than when the same number of reactants were resident in the cytosol. However, as stated earlier, there are other problems with the submodel that do not justify going down this route.

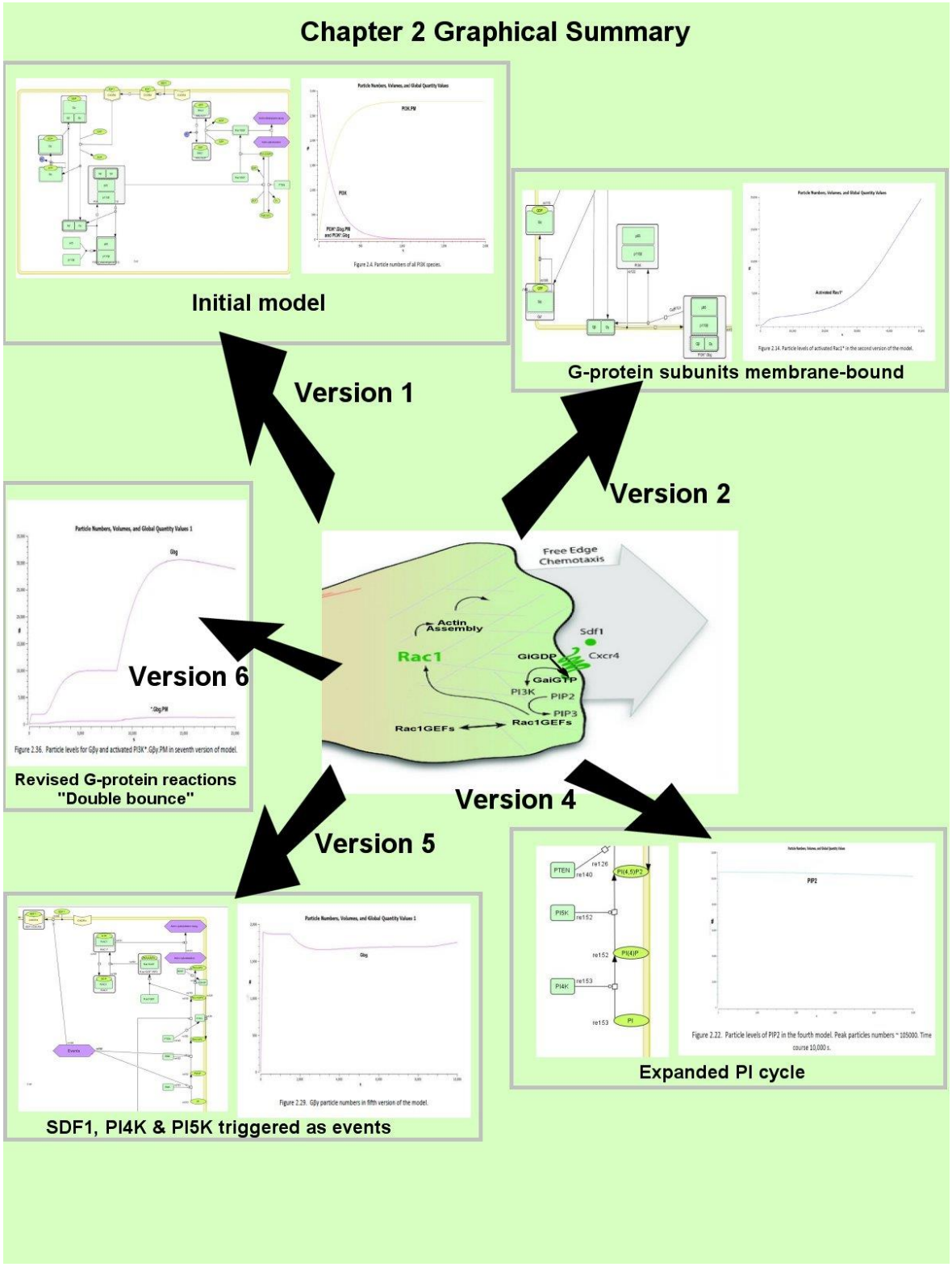

Figure 2.43. Graphical summary of Chapter 2 results.

Chapter 3 - A simple model of FA assembly

3.1 Introduction

As briefly alluded to in Section 1.1.3, and shown in Figure 1.3, outside-in signalling refers to a mechanism in which FA formation, within protruding lamellipodia, is driven by ECM-ligation of integrins [215]. This is in contrast to inside-out signalling, where such events are driven by growth factors, chemokines and other signalling molecules, all typically soluble and acting via their cognate receptors, rather than via integrin [216,215]. As stated in section 1.3, the original primary aim of this project had been to model inside-out signalling, beginning with a lamellipodia submodel, driven by CXCR4-mediated SDF1 signalling, as described in the previous chapter. The intention had then been to develop a separate submodel of FA dynamics in this chapter, integrating the two sub-models in Chapter 4.

Following the decision to stop development of this lamellipodia submodel, it has been decided to proceed instead with an-outside-in signalling model. This has the advantage in many ways of being less complex than inside-out signalling, at least in the early stages. There is also no need to implement it as separate sub-models. A single model, incrementally developed, will suffice. As a result of these design decisions, the aims of the project have been revised as follows:

3.1.1 Aims of the project

The primary aim now of this project is to develop a single model of outside in signalling-mediated FA dynamics. This will include a basic model of FA assembly, together with a limited phosphoinositide mechanism (incorporating PI, PIP, PIP2, PIP3, IP3 and DAG, with the last two produced from PIP2 by PLCγ), a simple IP3-driven calcium release mechanism, and calcium-driven calpain cleavage of FAs. This outside in-model will be developed in two stages. A simple model of

early FA assembly will be produced in this chapter (Chapter 3). This will include a generic integrin, binding a generic ECM ligand, PI, PIP2 and a generic PI5K kinase, producing PIP2 from PIP. This will then be developed much further in Chapter 4, resulting in the full model described above, and extensively analysed to see what insights can be gained from it.

3.1.2 Modelling software to be employed

This remains unchanged from those stated in section 1.3.2.Before outlining the details of the proposed model to be implemented in this chapter, it is necessary to outline the relevant cell biology underpinning it.

3.1.3 Biological background to this model

Talin activates integrins by holding the cytoplasmic tails of the integrin heterodimer apart, forcing it to adopt a conformation with a higher affinity for its natural extracellular ligand [217,218]. It also promotes integrin clustering via cross-linking [219,3,4]. Together, these are believed to play a major role in both inside-out and outside-in signalling, although, in the latter case, the full mechanism for this is not fully understood [215,220].

Rather than include somewhat speculative mechanisms for triggering initial outside-in-associated integrin activation In this model, the assumption is made that a few (outside-in-mediated) integrin-ECM ligation events have already occurred. In turn, this has triggered some PIP2 production, sufficient for outside-in binding to continue throughout the rest of the cell. (Any

[3]Caroline Cluzel et al., 'The Mechanisms and Dynamics of Avβ3 Integrin Clustering in Living Cells', *The Journal of Cell Biology* 171, no. 2 (24 October 2005): 383–92, https://doi.org/10.1083/jcb.200503017.
[4]Frederic Saltel et al., 'New PI(4,5)P2- and Membrane Proximal Integrin-Binding Motifs in the Talin Head Control ?3-Integrin Clustering', *The Journal of Cell Biology* 187, no. 5 (30 November 2009): 715–31, https://doi.org/10.1083/jcb.200908134.

other species that would result from these few prior events have not been included in the model, to reduce complexity).

The importance of PIP2 here, is that it is required to activate talin, by releasing an autoinhibitory association between the talin head and rod regions [221], allowing the talin FERM domain first to bind to PIP2 and other anionic lipids in the membrane, and then to the integrin tail region [222]. However, in order for outside in-signalling to be sustained, PI5K must first be activated, so as to maintain production of PIP2 from PIP, just as it did in the inside-out signalling model in the previous chapter.

However, in this model membrane-bound PI5K is activated by ECM-ligated integrins, rather than being spontaneously activated, as in the SDF1-CXCR4 lamellipodia model. Such activation is known to occur, by a mechanism that is not fully understood but may involve phosphorylation by Src-family kinases (SFKs) [223], which are known to be activated by ECM-integrin ligation [224]. However, here this PI5K activation is simply modelled as being performed by the species representing activated, ECM-ligated integrin (which are all assumed to be SFK containing.

3.1.4 Outline of proposed early FA model

This suggests a simple starting model in which cytosolic talin is recruited to the membrane by PIP2. This leads to the now-activated talin binding to integrin. Having been activated by talin-association, integrins are free to bind ECMs, with the resulting integrin-ECM complex then activating (membrane-resident) PI5K. This activated PI5K then phosphorylates PIP to PIP2. This is what will be initially implemented in the Virtual Cell model, firstly as a compartmental/ODE application. This application will be developed further to more accurately reflect the underlying physiology. If it is judged to be behaving acceptably by this criterion, the latest version of the compartmental application will then be used as a basis for a further 2D spatial application.

3.2 Methods

3.2.1 Analysis methods used

Much of what was written in the corresponding Methods (2.2.1) section in Chapter 2, concerning the analytical tools available in COPASI, also applies to the ODE/compartmental and PDE/spatial modelling used here in Virtual Cell. This includes the various analytical methods, including plots and parameter scans. However, some of the plotting options available, as standard, in COPASI (such as plots of particle numbers and concentration rates), are not available in Virtual Cell. Sensitivity analysis, while available for the compartmental/ODE (but not the spatial) application, has not been applied in this case, as the relative simplicity of this model did not appear to merit it. For similar reasons, parameter estimation was not used either.

3.2.2 Modelling application version number

Virtual Cell: Version 7.0.0 (Build 9)

3.2.3 System units used

Time units:	Seconds (s)
Length units:	µm
Area units:	μm^2
Volume units:	μm^3
Substance units:	$\mu M.\mu m^3$
Membrane substance units:	molecules
Volume concentration units:	µM
Membrane conc. units:	$molecules.\mu m^{-2}$
Diffusion units:	$\mu m^2.s^{-1}$

3.2.4 Solver settings

In terms of Solver settings, these were as follows.

3.2.4.1 ODE/compartmental application

For the ODE/compartmental application the default Combined Stiff Solver (IDA/CVODE) was used, with both absolute and relative error tolerances set to values of 1.0E-9. The maximum step size was set to 1.0 seconds, with simulations run for either 1000 or 10,000 seconds, depending on circumstances. Output settings varied, according to requirements.

3.2.4.2 PDE/2D spatial application

For the PDE/2D spatial application, the Fully-Implicit Finite Volume, Regular Grid (Variable Time Step) integrator was used, with an absolute error tolerance of 1.0E-9 and a relative error tolerance of 1.0E-7. The maximum step size was set to 0.1 seconds, with simulations run for 1000 seconds in all cases. The output interval was either 1.0 or 5.0 seconds, according to requirements. The Mesh Size was 101 elements for both X and Y dimensions, giving a total size of 10201 elements and a spatial step of 1.016 μm in both the X and Y planes.

3.3 Model Setup

Whereas, in the COPASI submodel of the previous chapter, the various reactions, species and other parameters were listed as a set of tables, the relative simplicity of the present model allows an alternative approach. Here, each stage of the model specification process is illustrated, step-by-step, until the entire model has been implemented. This will hopefully give some more insight into how reaction-diffusion models can be implemented as a set of species, differential equations (representing reactions between those species) and various parameters, using appropriate tools, such as Virtual Cell, specifically designed for cell biologists, and requiring no mathematical background.

3.3.1 Compartment (Structure) setup (Version 1)

A good starting point with Virtual Cell (and with modelling in general) is to determine what compartments will be needed in the new model. In the case of this model, at least as a starting point, three will suffice, namely the cytosol, the plasma membrane and the extracellular space. These can then be added in the Reaction Diagram section of the tool as shown in Figure 3.1.

Virtual Cell calls compartments "structures" and these can be seen listed separately in the Structures section (or added there, Instead of in the Reaction Diagram section). At this point, unlike in COPASI, no volume is specified for these compartments and, in fact, for the most part, the early design stages in Virtual Cell modelling involve specifying the logical model, which Virtual Cell calls the Physiology. The next stage typically involves populating each model compartment with its resident species, which can also be done in the Reaction Diagram. This is shown in Figure 3.2 for the FA model.

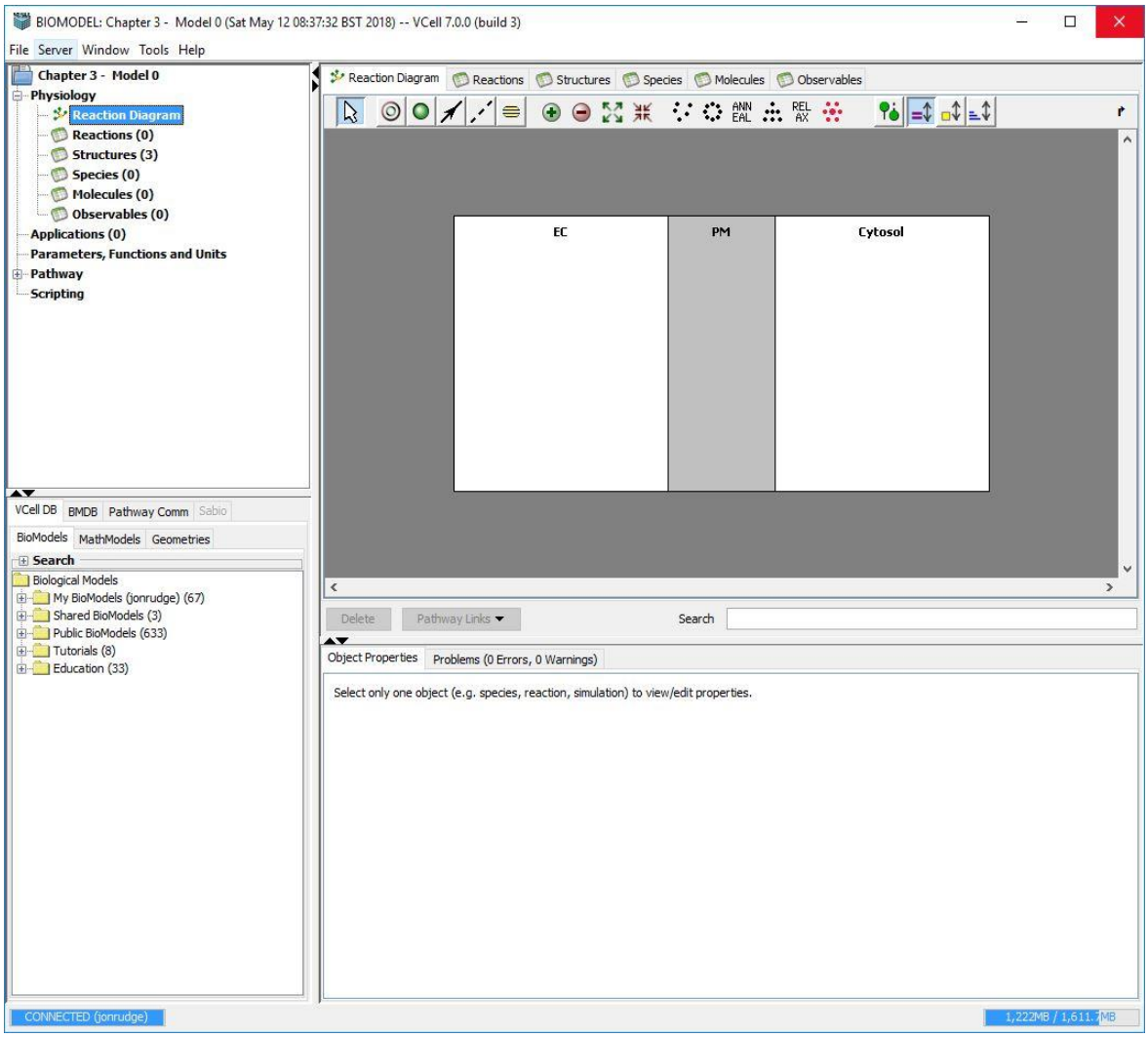

Figure 3.1. Compartment specification in Virtual Cell.

3.3.2 Species specification (Version 1)

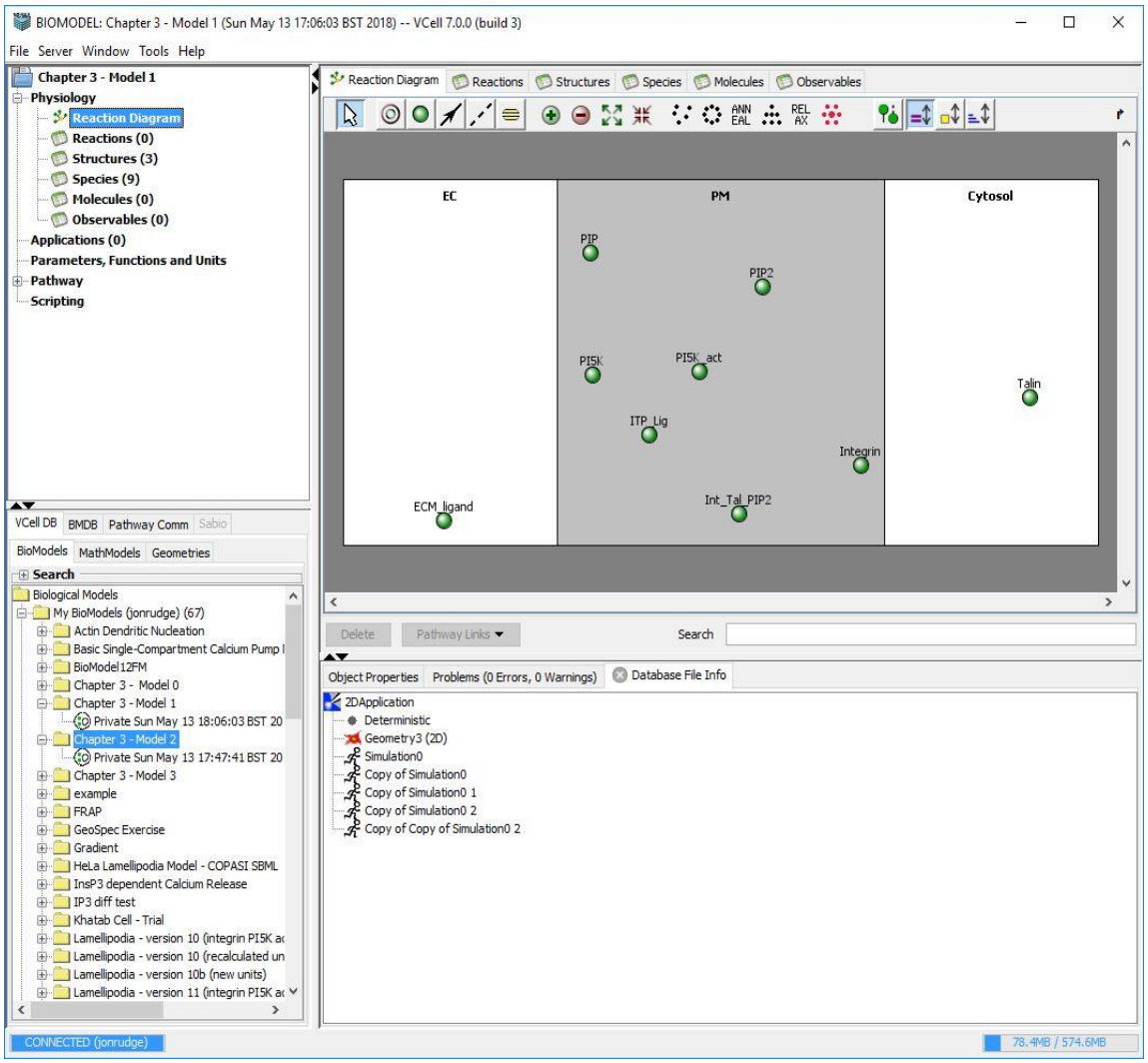

Figure 3.2. Species specification in Virtual Cell.

As with compartments/structures, these model species can also be seen listed in the separate Species section, and added there instead of in the Reaction Diagram. Again, unlike in COPASI, no initial concentrations (or particle numbers) are specified, with the species, at this point, merely forming another part of the logical model. The next step then is to specify reactions for all the species. In fact, it is not necessary to specify all reaction species in one go, prior to specifying reactions and, in practice, additional species are often found to be necessary, typically as intermediate products.

3.3.3 Reaction specification (Version 1)

In the same way that structures and species can be added in their respective sections, reactions can also be added in the Reactions section. However, it is often more convenient to use the GUI tools of the Reaction Diagram for this purpose, as in this case. The four reactions are shown in Figure 3.3.

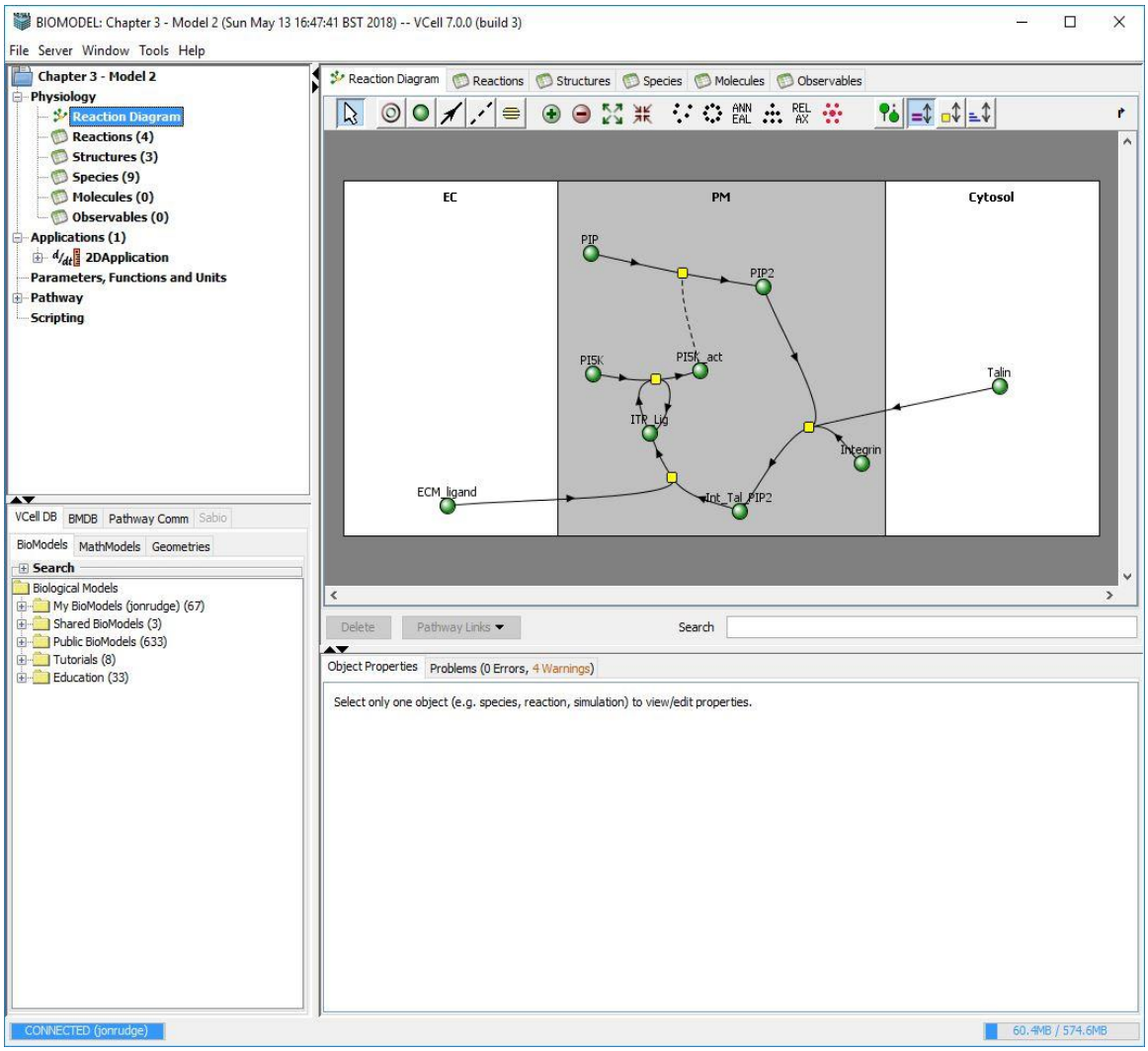

Figure 3.3. Reaction network specification in Virtual Cell.

Unlike the case with compartments and species, specifying the reaction network alone isn't sufficient. For each reaction the rate law and the relevant rate constant values must be specified, as part of the model Physiology. This must be undertaken in the Reaction section. The four

different reactions are as follows. One involves integrin, talin and PIP2 binding to form the activated complex Int_Tal_PIP2. This goes on to bind an extracellular matrix ligand (ECM_Ligand), forming the complex ITP_Lig. This complex then activates PI5K, which converts PIP to PIP2. Since PIP2 is itself involved in the reaction in which integrin is activated, the overall system can be considered to have an element of positive feedback.

The attempt to find rate constant values for these reactions raises some issues. Firstly, only a Kd value could be found relevant to the integrin talin-PIP2 reaction proposed here. Moreover, it is clear from the journal article from which it was obtained[225]that it isn't appropriate for the third-order reaction (in which integrin, talin and PIP2 simultaneously bind to each other) specified here. Secondly, activation of PI5K by ligated integrins appears to involve phosphorylation by integrin-associated Src, meaning that it is an enzymatic, not mass-action reaction, as specified here. Fortunately there is good binding data for various integrins with extracellular matrix proteins, for instance for integrin α5β1-fibronectin binding [226], which will be used in this model.

From Figure 3.3, it is clear that the third-order integrin-talin-PIP2 reaction needs breaking up into two reactions. The question is which binds first -whether it is to be integrin and talin, talin and PIP2 or integrin and PIP2. A search of the literature suggests that talin binds to PIP2 first [33], which makes sense, given that talin is normally resident in the cytosol and must first be "activated", requiring a conformational change, before it can bind to the membrane. This requires the presence of acidic phospholipids such as PIP2, without which talin is repulsed from the membrane surface[33]. So, without prior PIP2 binding, talin is unlikely to bind to membrane-embedded proteins such as integrin.

Taking the above into account the model reactions need to be altered as shown in Figure 3.4.

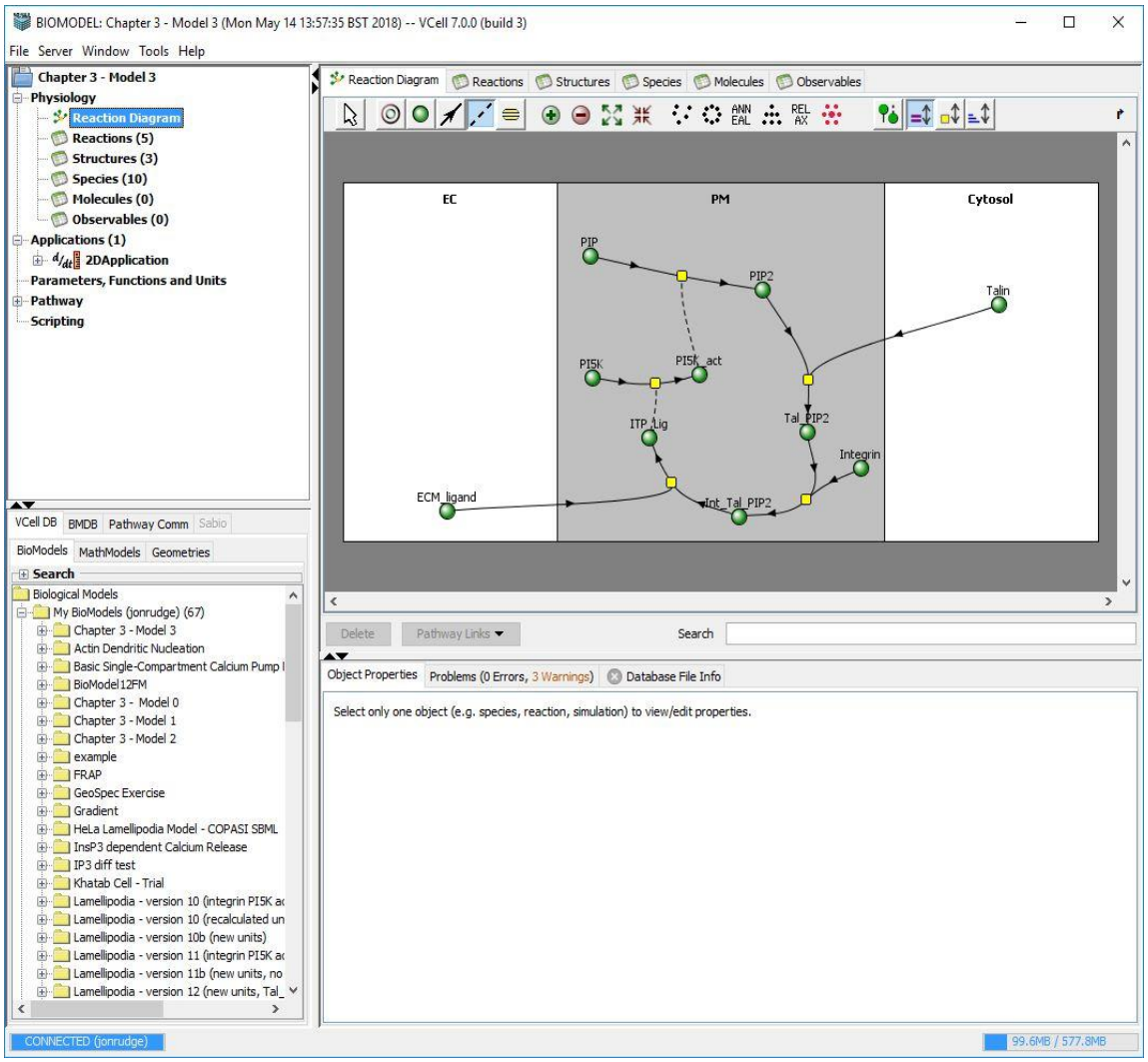

Figure 3.4. Revised reaction network specification in Virtual Cell.

Now the talin-PIP2-integrin reaction has been broken down into two mass-action reactions (firstly a talin-PIP2 binding reaction, followed by a reaction for integrin binding to the talin-PIP2 complex). PI5K activation has now been turned into an enzymatic reaction. Having revised this reaction network to better reflect the reported reality, it remains to specify the reaction rate laws and rate constants for the five reactions of which the model is now comprised. These are listed in Table 3.1, together with an explanation of how they were derived.

Reaction	Description	Rate law	Parameter values
Talin + PIP2 = Tal_PIP2	Talin binding PIP2 to form Tal_PIP2 complex	Mass action (reversible)	kf = 1.4e+05 M-1 s-1 kr = 0.35 s-1 *
Integrin + Tal_PIP2 = Int_Tal_PIP2	Integrin binding to talin-PIP2 complex	Mass action (reversible)	kf = 1.85e+04 M-1 s-1 kr = 3.6e-03 s-1**
Int_Tal_PIP2 + ECM_Ligand = ITP_Lig	Int_Tal_PIP2 ligation by an ECM ligand	Mass action (reversible)	kf = 7.7e+05 M-1 s-1 kr = 4.9e+01 s-1***
PI5K -> PI5K_act; ITP_Lig	PI5K activated by (Src-containing) ITP_Lig	Michaelis-Menten (irreversible)	kcat = 10 s-1 KM = 3.0e-05 M****
PIP -> PIP2; PI5K_act	PIP phosphorylated to PIP2 by activated PI5K	Michaelis-Menten (irreversible)	kcat = 1.02 s-1 Km = 1.0e-05 M*****

Table 3.1. List of model reactions, together with specified rate law and parameter values. Key:* = kr value taken from[160].Kf value derived by dividing kr value by mean Kd value of 2.5 µM, derived from two published reports [225,33]. ** = [227]- Mean average values from two reported for F2 and F3 binding to β3 tails. (Off-rate agrees well with value of 9e-03 s-1 from published FRAP study [228].)*** = Based on reported kinetics for α5ß1 binding to fibronectin[226]. **** = Based on value taken from published model, in itself derived from the literature [160,229]. ***** = Taken from [181] (Table 1).

As can be seen from the Key, ligation of the ECM_Ligand to the (activated) integrin complex Int_Tal_PIP2 is based on a report on fibronectin binding kinetics to the α5β1 integrin dimer. Kinetic rate constant values for integrin binding to talin-PIP2 complex were based on a report for binding kinetics for various talin head groups to integrin β3, as no data for talin-β1 binding could be found. However, given that talin-β1 and talin-β3 binding affinities, as indicated by reported Kd values, appear to be reasonably similar[215,230,227], this seemed a reasonable estimate.

This illustrates a general problem when attempting to find reliable rate constant values for binding reactions. Kd values are often relatively easy to obtain from the literature. However, in themselves they are essentially unusable as they represent the ratio of the binding off-rate divided by the on-rate, neither of which can be inferred from this Kd value without additional information. Unfortunately, arbitrarily assigning on- and off-rate values that yield the required Kd value is not generally acceptable, as a quick test with a range of permutations will often show that model behaviour varies substantially between permutations. And, where this is not the case,

there is always the risk that such differences will emerge later on in model development but will not be picked up at the time.

This issue arises again with talin-PIP2 binding, where a number of studies [231,225,33]report similar Kd values for PIP2 binding (in the range 0.18-0.98 µM for the F2F3 head groups) with various talin head groups, but provide no on- or off-rates. Fortunately, in this case, an off-rate has been found [160], which can be divided by a consensus Kd value to yield the missing on-rate. However, given that the Kd values and the off-rate were obtained from different experiments, this is not an ideal method.

Fortunately, obtaining relevant parameter values for the remaining two reactions (PIP phosphorylation by PI5K and PI5K activation by ITP_Lig) has proved more straightforward, as is often the case with enzymatic reactions of wide interest to the research community.

3.3.4 Compartmental/ODE application setup (Version 1)

Having entered the reaction data, the physiological model is now complete. The next stage is to create applications that work on this model. In other words, it is only at this stage (unlike the case with COPASI) that one specifies initial concentrations for species and sizes for compartments. The underlying reason for this is that applications can be compartmental (i.e. ODE-based) or spatial-based, or spatial-based models being either one-, two- or three-dimensional. For this reason, compartment dimensions will differ between applications and this may mean that species initial concentrations will vary also.

In the first instance, it usually makes sense for the first application to be compartmental, as these are generally easier to set up. Also, having specified the application, the next task is to specify and run simulations, in order to study model behaviour, and compartmental simulations tend to run many times faster than spatial models (typically in the order of a few seconds, as against a few minutes for 2D spatial models). As a number of simulations are typically required to establish that the model is correctly specified, it makes little sense to spend more time on this than is necessary.

3.3.4.1 Compartment size determination (Version 1)

Having created a new application, Virtual Cell first requires the user to specify whether it is to be deterministic, stochastic or network-free (meaning rule-based). In this case, we specify deterministic. Immediately a new application is created, with a compartmental (i.e. non-spatial) geometry, which is the default, together with default values for the EC, PM and cytosolic compartments. In this case, the EC and cytosolic compartments can be given the same values as in the COPASI model. By coincidence, the EC default value is the same as the COPASI model, i.e. 50,000fl (or 50,000 μm^3, as here), while the Cytosol volume has to be altered to 1500 fl (1500 μm^3).

However, a major difference between Virtual Cell and COPASI is that the former models membranes as two-dimensional surfaces, whereas COPASI models them as three-dimensional volumes. For this reason, the PM volume from the COPASI model (6.4 fl) cannot be used. Fortunately, the surface area for HeLa cells is available from the literature [1600 μm^2, from BioNumbers[188,189]] and, in fact, was used to derive the PM volume for the COPASI model. (And, in the absence of such information, could, of course, be derived from a membrane volume by dividing it by a suitable value for the membrane thickness.)

This yields the following compartment volumes for the ODE-based application for this model:

Compartment	Compartment size (volume or area)
Extracellular (EC)	50,000 fl (50,000 μm^3)
Plasma membrane (PM)	1600 μm^2
Cytosolic	1500 fl (1500 μm^3)

Table 3.2. Compartment sizes for first, compartmental/ODE-based, application of this model.

3.3.4.2 Species quantification (Version 1)

Having entered these into the geometry section for the compartmental application, the next task is to specify species concentrations and densities for the liquid and membrane compartments respectively. These are shown in Table 3.3:

Species	Compartment	Quantity	
		Copy # per HeLa cell	Density/ Concentration †
ECM_Ligand	EC	~450,000	0.5 µM *
Integrin (α5β1)	PM	30,000 *	18.75molecules/μm^2 **
Int_Tal_ PIP2	PM	0	0 molecules/μm^2
ITP_Lig	PM	0	0 molecules/μm^2
PI5K	PM	15,000 *	9.375 molecules/μm^2 **
PI5K_act	PM	0	0 molecules/μm^2
PIP	PM	200,000 **	125 molecules/μm^2 ***
PIP2	PM	1000 **	0.625 molecules/μm^2 ***
Talin	Cytosol	98,000*	0.11µM **
Talin_PIP2	PM	0	0molecules/μm^2

Table 3.3. Model species initial quantities. Key: † = Liquid compartment species specified as concentrations in µM, membrane compartment species as densities (molecules per μm^2). * = Ligand specified so as to be in excess of integrins. ** = Copy number value taken from a published HeLa cell quantitative proteome [182]. Densities were calculated by dividing these copy numbers by the PM surface area (1600 μm^2), concentrations by converting copy number value to µmoles and then dividing by cytosolic volume in litres (1.5e-12). *** = Assumption, with PIP initial copy numbers designed to eventually allow sufficient activation of talin by PIP2, and with initial PIP2 levels sufficient to allow some integrin activation in order to bind fibronectin.

As can be seen in the Table Key, many of these initial values have been taken from the same quantitative proteome for HeLa cells used in the COPASI model in the previous chapter.

Species that are the products of reactions are assumed to have zero levels at the outset, whilst fibronectin and 'PIP' (which feature only as reactants) are assumed to be in excess, so as to prevent them from being rate-limiting for the system under investigation (which might distort the results and their interpretation). Finally, PIP2 levels have been chosen so as to allow some initial talin activation in the form of Tal_PIP2, and therefore subsequent integrin activation as Int_Tal_PIP2, but not so much that, subsequent PI5K activation (which is the result of integrin-fibronectin ligation, and provides the forward feedback control mentioned earlier) is essentially redundant.

3.3.4.3 Running a simulation (Version 1)

Having specified these initial levels it is now possible to run a simulation of the model, which is done in the Simulations section for this application. This offers the possibility of overriding various model parameters (in the Parameters section), including previously specified reaction rate constants and the species levels just entered. However, in this instance we keep these unaltered, also accepting all the default solver settings, except for the Ending value, which we change to 1000 seconds, in order to allow the model to run for this length of time. The results of this are shown in the next section.

3.4 Results

3.4.1 First version

The results from running the simulation for this first version of the model are shown below, in the form of time series plots of concentrations/densities for all model species:

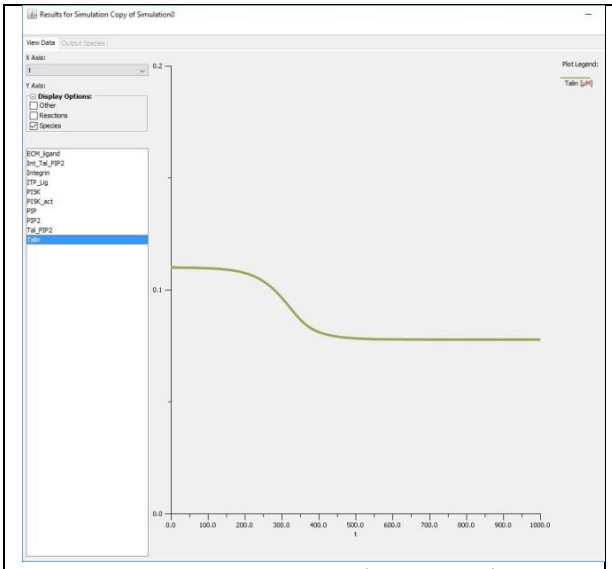

Figure 3.5. Talin levels (µM, 1000 s)
Minimum particle numbers: ~70,200 molecules

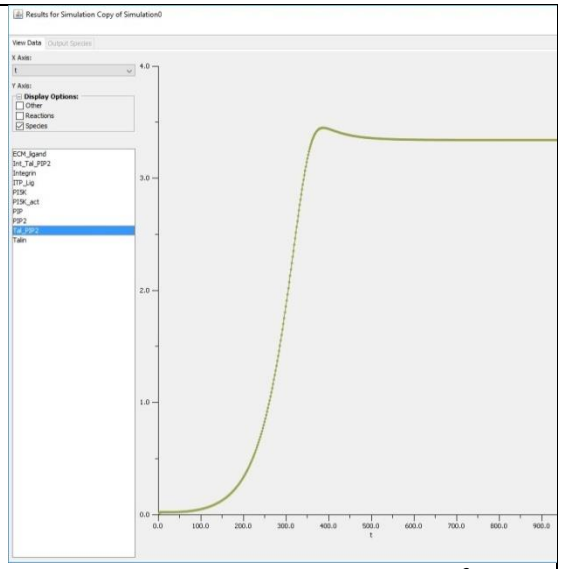

Figure 3.6. Tal_PIP2 levels (molecules.µm^2, 1000 s)
Peak particle numbers: ~5517 molecules

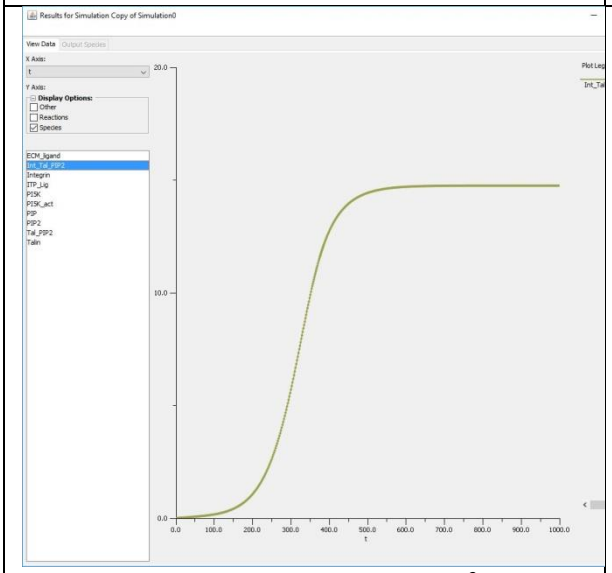

Figure 3.7. Int_Tal_PIP2 (molecules.µm^2, 1000 s)
Peak particle numbers: 23,600 molecules

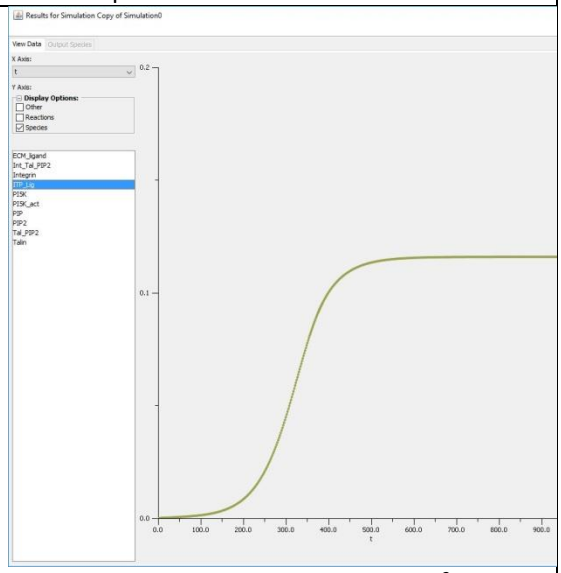

Figure 3.8. ITP_Lig levels (molecules.µm^2, 1000 s)
Peak particle numbers: ~185 molecules

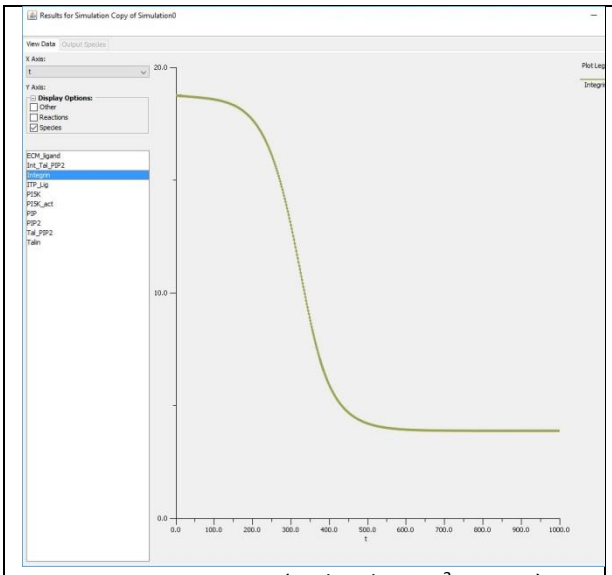

Figure 3.9. Integrin (molecules.µm^2, 1000 s)
Minimum particle levels: 6235 molecules

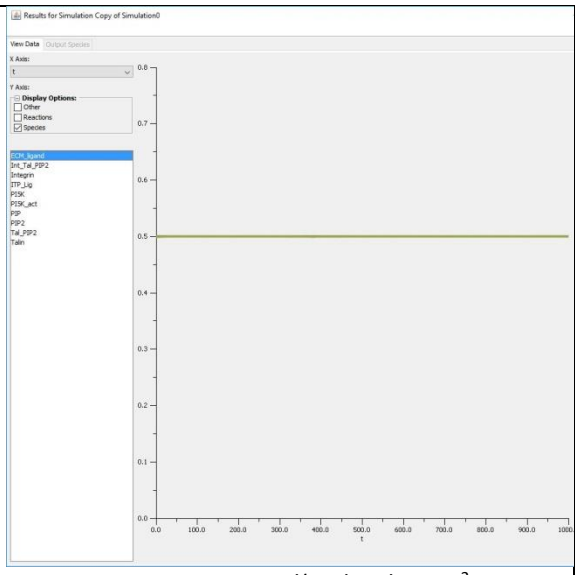

Figure 3.10. ECM_Ligand(molecules.µm^2, 1000 s)Particle levels: ~15,000,000 molecules

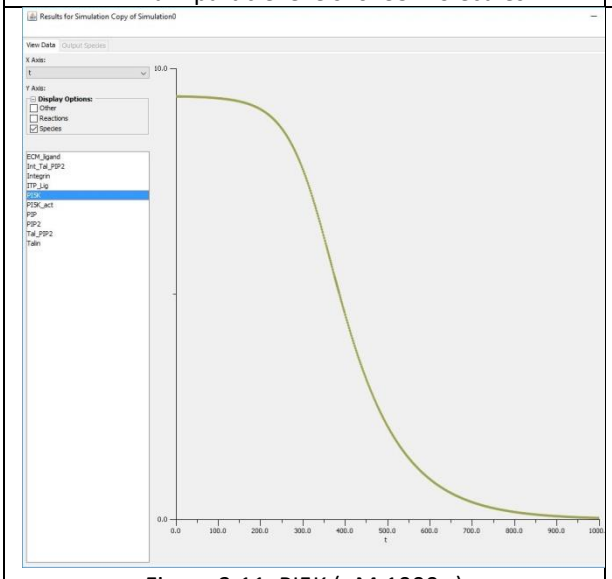

Figure 3.11. PI5K (µM,1000 s)
Minimum particle levels: ~46 molecules

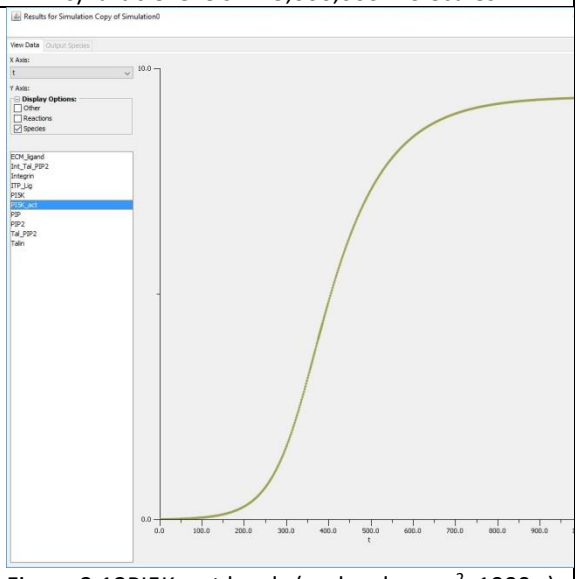

Figure 3.12PI5K_act levels (molecules.µm^2, 1000 s)
Peak particle levels: ~14,954 molecules

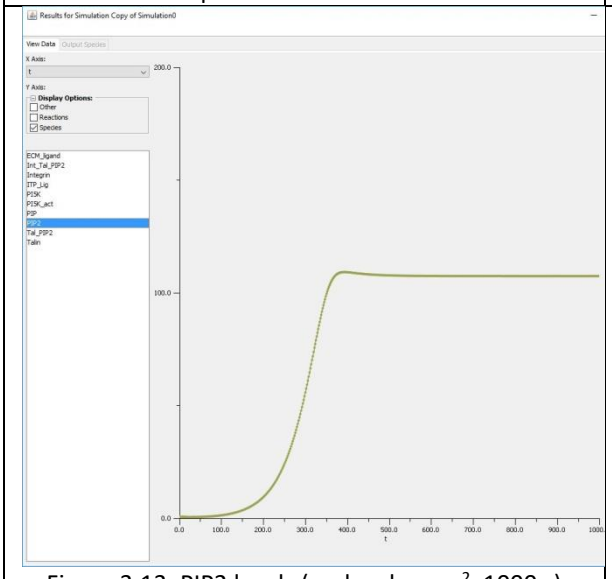

Figure 3.13. PIP2 levels (molecules.µm^2, 1000 s)
Peak particle levels: 174,400 molecules

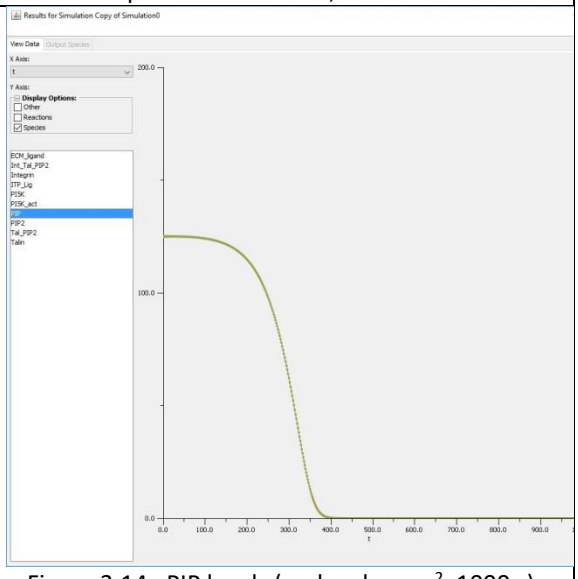

Figure 3.14. PIP levels (molecules.µm^2, 1000 s)
Minimum particle levels: ~0 molecules

These show that the model is working broadly as one might predict. That is to say, there is an initial lag period of around two minutes (120 seconds), when only small amounts of PIP2 are available, because the absence of activated PI5K means there is little, if any, additional PIP2 being converted from PIP. However, after that, there is a period of exponential increase from around the two-minute mark to around the five-minute (300 seconds) mark in which levels of model species such as PIP2, Tal_PIP2, Int_Tal_PIP2, ITP_Lig and PI5K_act increase rapidly. This is as a result of the positive feedback loop mentioned earlier, whereby increased PIP2 levels lead to increased integrin activation and ligation and therefore increased PI5K activation.

This period of exponential increase comes to an end as levels of available PIP become exhausted, as a source of further PIP2 (Figure 3.14). Although Figure 3.11 suggests that exhaustion of (unactivated) PI5K may also be contributing to this decline in the rate of increase, this appears to be happening 100 or more seconds after the event, and to be coincidental rather than contributory. Reasons for believing this include the fact that levels of PIP are close to zero at the 400 second mark (Figure 3.14), at a time when nearly half of the originally available PI5K (approximately 4.5 out of 9.375 molecules/μm^2) is still unactivated (Figure 3.11). Another reason is that plots for Tal_PIP2, Int_Tal_PIP2, ITP_Lig and PIP2 show a closer relationship with the PIP plot than with the PI5K plot, all reaching peak levels well before PI5K is exhausted (Figures 3.6, 3.7, 3.8 and 3.13).

It also seems clear that integrin dynamics are being driven by PIP2, rather than the reverse. Support for this view comes from the fact that peak levels of both Int_Tal_PIP2 and ITP_Lig occur after peak levels of PIP2 (at around 550 seconds and 600 seconds, respectively, as against around 400 seconds). Also from the fact that, at the time that they reach these peak levels, neither integrin or talin are used up at this point (with around 64% and 21% still remaining, as shown in Figures 3.9 and 3.5, respectively), despite levels of PIP2 at this point being substantially in excess

of either of them. This is an agreement with the general pattern in which particle level curves for model species are seen to lag the particle level curves for PIP2 and PIP, as can be seen by, for example, comparing the time points at which such curves are seen to transition from broadly linear to logarithmic rates of change, and vice versa. Typically such lags are in the order of 100 seconds or so (Figures 3.5-3.14).

However, it also has to be acknowledged that too small a proportion of activated integrins (Int_Tal_PIP2) are subsequently becoming ECM-litigated ITP_Lig. Specifically, although there are approximately 23,600 molecules of Int_Tal_PIP2 available, as a result of all the PIP2 creation described above, only 185 molecules of these subsequently become ITP_Lig. Given that ITP_Lig is required to activate PI5K, this suggests that the positive feedback loop described is not quite as all-encompassing as it might first appear.

Aside from this, an interesting general result of this first version of the model is that the overall timescale of talin and integrin activation, ECM-integrin ligation, PIP2 creation, and other lamellipodia-associated events, are happening within 400 seconds or so (i.e. less than 10 minutes). This agrees reasonably well with what is seen in reality[232,233,234]. However, it has to be acknowledged that this is only a very simplified version of FA dynamics, lacking other major FA-associated proteins, such as Src, FAK and vinculin and based on a number of physiologically invalid assumptions, such as excess levels of PIP and with inactive PI5K in the model being PM-resident, when, in reality, much of it is cytosolic [235].

Making this model more realistic in this respect, therefore, require some mechanism for cytosolic PI5K recruitment to the PM. Fortunately such a mechanism appears to exist in the form of talin, which has a well-attested binding site for at least one of the PI5K isoforms[236], PIPKIγ. Moreover, this isoforms is thought to be important to FA assembly[237,238], by being responsible for producing

PIP2 in a spatially and temporally favourable fashion [35,238]. It is also known that talin and PIPKIγ translocate to the PM when complexed together[239,35,238].In fact, it also appears that talin1 (the predominant talin isoform in HeLa cells[182], and a normally cytosolic protein) binds preferentially to PIPKIγ[239], except in the presence of negatively-charged (acidic) phospholipids, especially PIP2, where it binds preferentially to the cytoplasmic tail of certain integrins[239,225].

This suggests a simple model in which cytosolic talin binds to cytosolic PI5K, with this complex then being recruited to the membrane by PIP2[225]. This leads to the now-activated talin to separate from PI5K and bind to integrin, leaving PI5K membrane-resident and subject to activation by any ECM-ligated integrins.

Since the model assumes that some outside-in signalling has already occurred, we will also assume that some PI5K has already been recruited to the PM. This means there will be two inactive PI5K populations, one membrane-resident and the other cytosolic, with the majority in the latter compartment, given its larger volume. Therefore, implementing this modification should be relatively straightforward, requiring two changes to species and two new reactions, as detailed in Tables 3.4 and 3.5.

3.4.2 Second version

New species	Compart-ment	Copy #	Density /concentration	Notes
PI5K_PM	PM	5000	3.125 molecules/μm^2	Renamed existing species, with one third of the original levels
PI5K_cyto	Cytosol	10,000	0.01107 μM	New species, two-thirds of original

Table 3.4. Changes to model species for the second version of the model.

New reaction	Description	Rate law	Parameter values	Notes
Talin + PI5K_cyto =Tal_PI5K	Binding of cytosolic talin and PI5K	Mass action (reversible)	Kf = 4.4 μM^{-1} s^{-1} Kr = 27.0 s^{-1}	*
Tal_PI5K + PIP2 = Tal_PIP2 + PI5K_PM	Recruitment of talin-PI5K complex to PM by PIP2	Mass action (reversible)	Kf = 0.14 μM^{-1} s^{-1} Kr = 0.1 $\mu m^2.molecules^{-1}.s^{-1}$	**

Table 3.5. Additional reactions required for the second version of the model. Key * = Based on experimental data for talin F2-F3 FERM domain binding to PIPK Iγ[239]. ** = Copied from existing Talin-PIP2 binding reaction, with Kr value altered to allow for it being a second-order reaction here, as opposed to a first-order reaction in the original.

However, as can be seen from Table 3.5, adding in talin-mediated PI5K recruitment to the cell membrane has caused a complication, in that, after the talin-PI5K complex is recruited by PIP2, PI5K is then released. Unfortunately, no kinetic data could be found for this reaction and so basing it on the talin-PIP2 reaction seems the best solution (given that, in both reactions, talin ends up being bound to PIP2 alone). However, since this new version also includes the release of PI5K (in the form of PI5K_PM), this means that the reverse reaction is second-order, rather than first-order, as in the original reaction. It is not possible to convert first-order rate constants to second-order, because, as a quick comparison will show, the relationship is non-linear. (Or rather, linear changes in first-order rate constant values require non-linear changes in second-order values if the reaction rate is to remain the same.) Here the reverse rate constant has been altered from 0.35 s^{-1} to 0.1 $\mu m^2.molecules^{-1}.s^{-1}$ to make it slower than the forward rate, but it is not a very satisfactory solution.

In fact, the results from running this new version of the model show it to be considerably slower than its predecessor, as can be seen from the timescales in Figures 3.15-3.26.

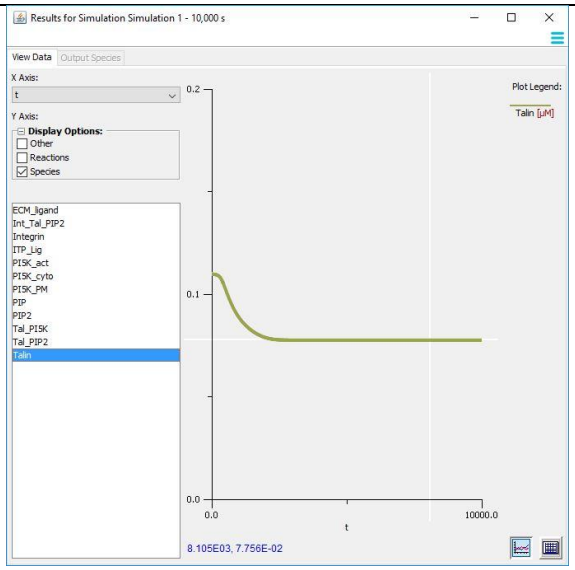

Figure 3.15. Talin (µM, 10,000s)
Minimum particle numbers: ~70,000 molecules

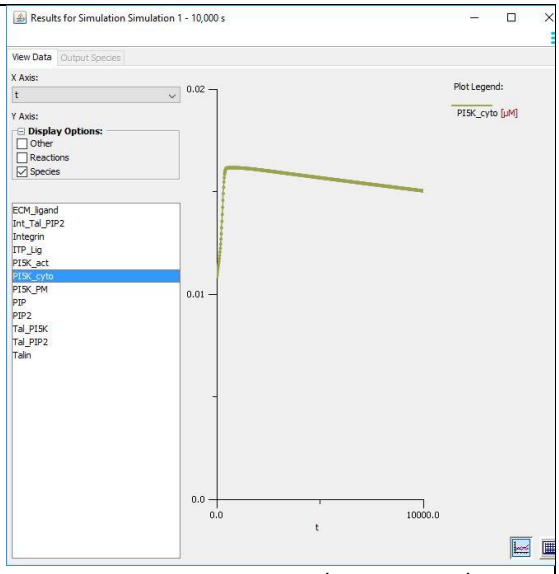

Figure 3.16. PI5K_cyto (µM, 10,000s)
Peak particle numbers: ~15,000 molecules

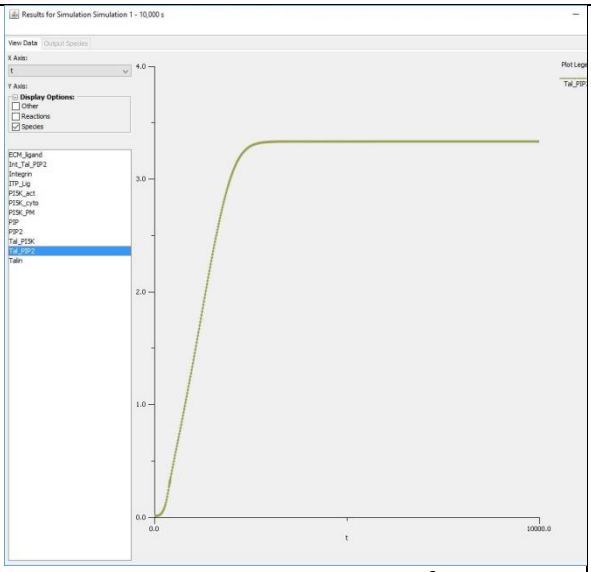

Figure 3.17. Tal_PIP2 (molecules.µm^2, 10,000s)
Peak particle numbers: ~5333 molecules

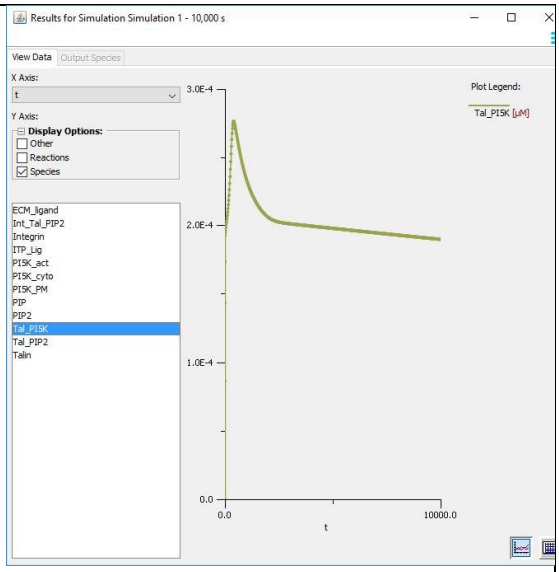

Figure 3.18. Tal_PI5K (µM, 10,000s)
Minimum particle numbers: ~172 molecules

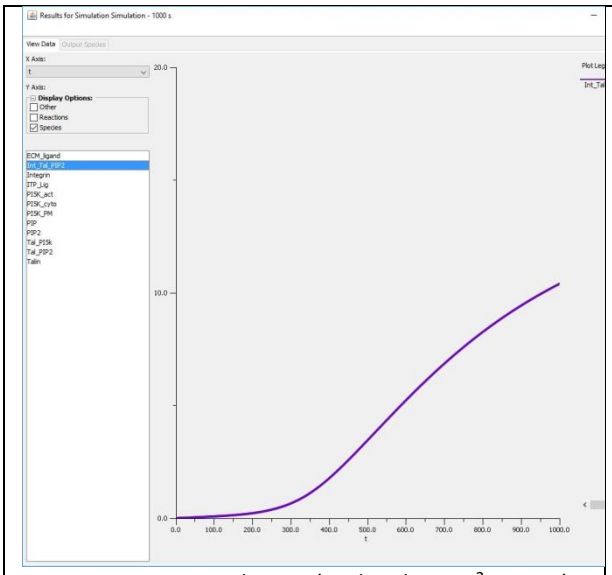

Figure 3.19. Int_Tal_PIP2 (molecules.µm^2, 1000s)
Peak particle levels: 16,640 molecules

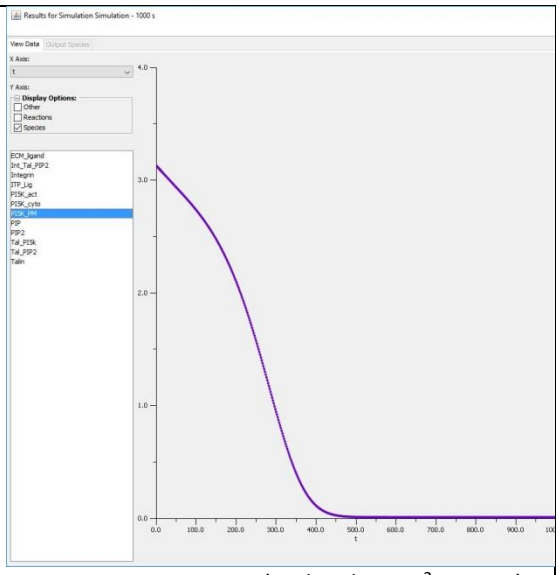

Figure 3.20. PI5K_PM (molecules.µm^2, 1000 s)
Minimum particle levels: ~14 molecules

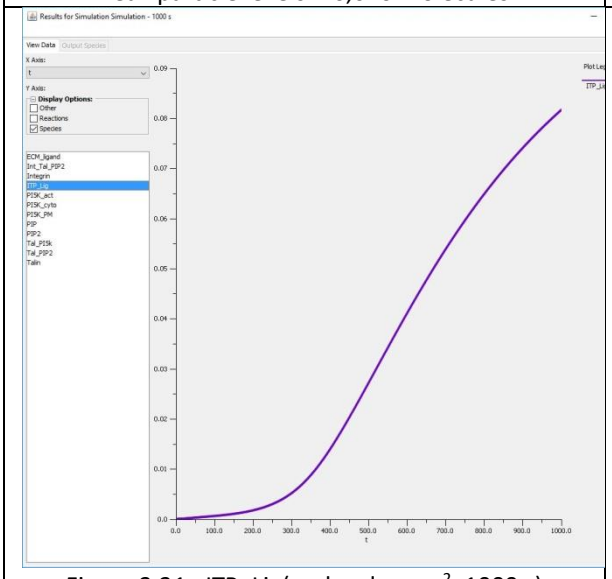

Figure 3.21. ITP_Lig(molecules.µm^2, 1000 s)
Peak particle numbers: ~131 molecules

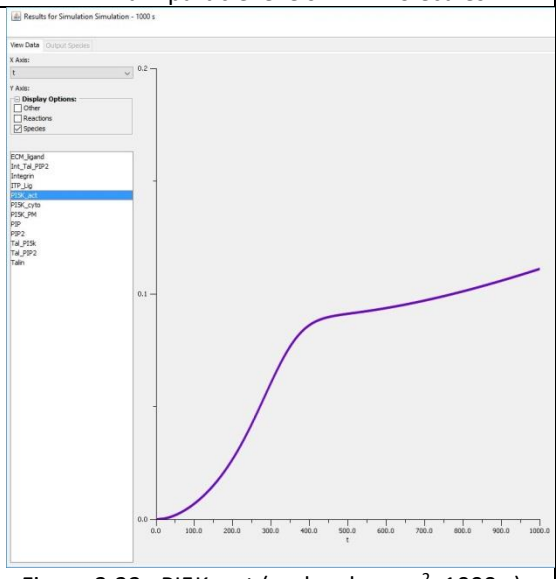

Figure 3.22. PI5K_act (molecules.µm^2, 1000 s)
Peak particle levels: 178 molecules

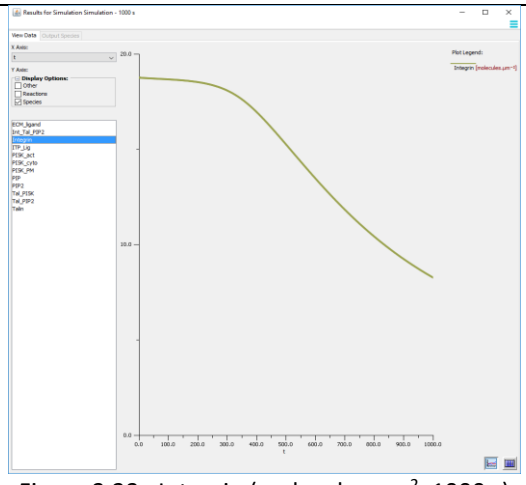

Figure 3.23. Integrin (molecules.µm^2, 1000 s)
Minimum particle numbers: 13,286 molecules

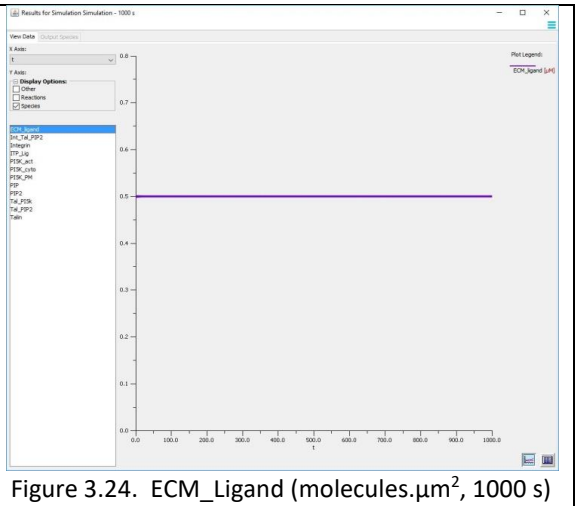

Figure 3.24. ECM_Ligand (molecules.µm^2, 1000 s)
Particle numbers: ~15,000,000 molecules

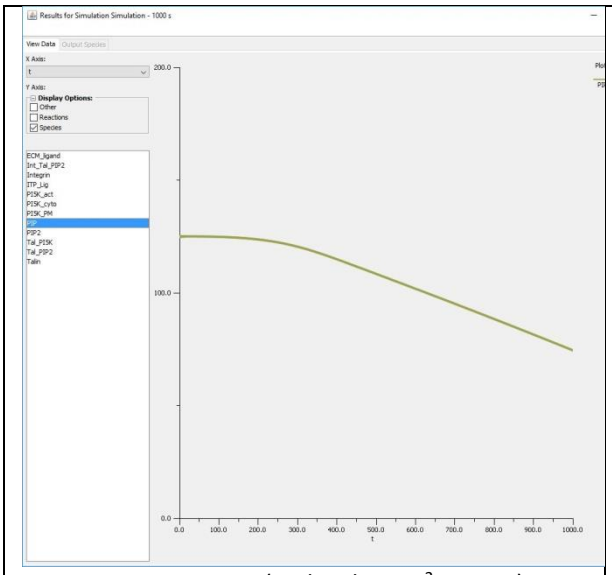

Figure 3.25. PIP (molecules.µm^2, 1000 s)
Minimum particle numbers: 119,200 molecules

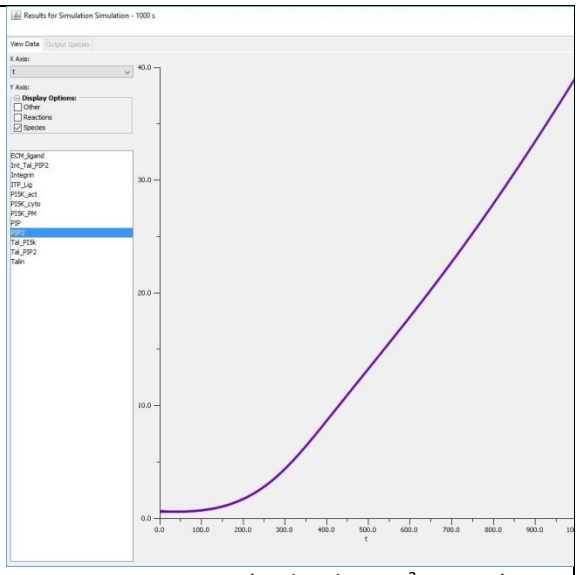

Figure 3.26. PIP2 (molecules.µm^2, 1000 s)
Peak particle numbers: ~62,850 molecules

As a result of the changes made in the second model, it is clear that the behaviours seen in the previous version of the model, such as the decline in PIP (Figure 3.25) and its replacement by PIP2 (Figure 3.26), and the rise of Tal_PIP2, Int_Tal_PIP2, ITP_Lig, as well as a decline in talin and integrin, are still all occurring, but over a much longer time scale. Given the previous discussion, an obvious reason for this might be that at the reverse rate constant (or off-rate) for the recruitment of talin-PI5K complex to PM by PIP2 might, at 0.1 µm^2.molecules^{-1}.s^{-1}, maybe too high.

This is confirmed by a logarithmic parameter scan on this value (over 10,000 seconds, with Kr values varied from 0.00001 to 0.1 µm^2.molecules^{-1}.s^{-1}), with substantial reductions in the time required for PIP levels to decline to zero (Figure 3.27). However, even at a level of 0.00001 µm^2.molecules^{-1}.s^{-1} this has only reduced to 600 seconds, whereas in the previous model this required only 400 seconds, and it is arguable whether such a relatively low off-rate can be justified in the absence of any support in the literature.

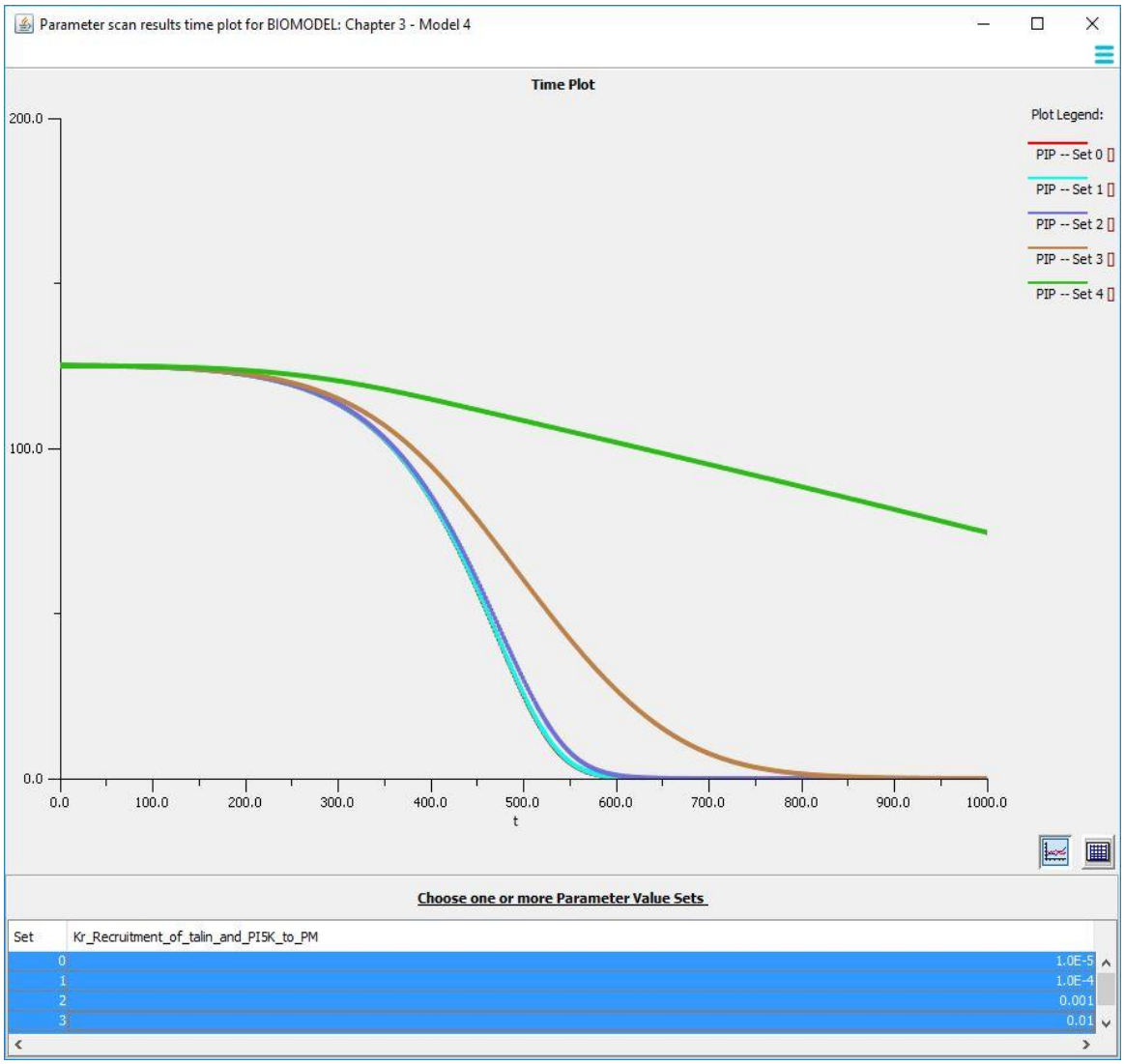

Figure 3.27. Logarithmic parameter scan, showing effect of reducing off-rate (Kr value) for talin-PI5K dissociation on levels of PIP in the second version of the model. (PIP levels shown as densities in molecules.µm^2, time shown in seconds. 100molecules.µm^2 represents 160,000 molecules.)

Another approach, which can be more easily justified, is to break down this reaction into two parts, with one reaction for PIP2 binding to Tal_PI5K (to create Tal_PI5K_PIP2 complex), followed with another reaction where PI5K dissociates from this complex, leaving Tal_PIP2 and PI5K_PM. This has the added advantage that the reverse reaction for PIP2 binding to Tal_PI5K is first-order again, meaning that the same first-order rate constant value used for talin-PIP2 binding in the third version of the model (i.e. 0.35 s^{-1}) can also be used here. However, the disadvantage is that it requires an additional reaction (for dissociation of the Tal_PI5K_PIP2 complex), moreover, one

for which there appears to be no available kinetic data in the literature. That aside, we assume here, in the absence of any more convincing alternatives, that this is a simple reversible mass action reaction with modest forward and reverse rates (Kf = 1.0 s^{-1}, Kr = 0.0001µm^2.$molecules^{-1}$.s^{-1}), necessitating the changes shown in Tables 3.6 and 3.7.

3.4.3 Third version

New species	Compartment	Copy #	Density /concentration	Notes
Tal_PI5K_PIP2	PM	0	0 molecules/µm^2	

Table 3.6. New species required for Version 3of the model.

New reaction	Description	Rate law	Parameter values	Notes
Tal_PI5K + PIP2 = Tal_PI5K_PIP2	Binding of cytosolic talin_PI5K with PIP2	Mass action (reversible)	Kf = 0.14 µM^{-1} s^{-1} Kr = 0.35 s^{-1}	*
Tal_PI5K_PIP2 -> Tal_PIP2 + PI5K_PM	Dissociation of Tal_PI5K_PIP2 to form Tal_PIP2 and PI5K_PM	Mass action (irreversible)	Kf = 1.0 s^{-1} Kr = 0.0001 µm^2.$molecules^{-1}$.s^{-1}	**

Table 3.7. New reactions required for Version 3 of the model. Key: * = Same parameter values as used for talin-PIP2 binding reaction. See relevant Key description in Table 3.1, Methods section. ** = Arbitrary value, chosen to be well within normal range for such reactions.

The Reaction Diagram for Version 3 the model now looks like this:

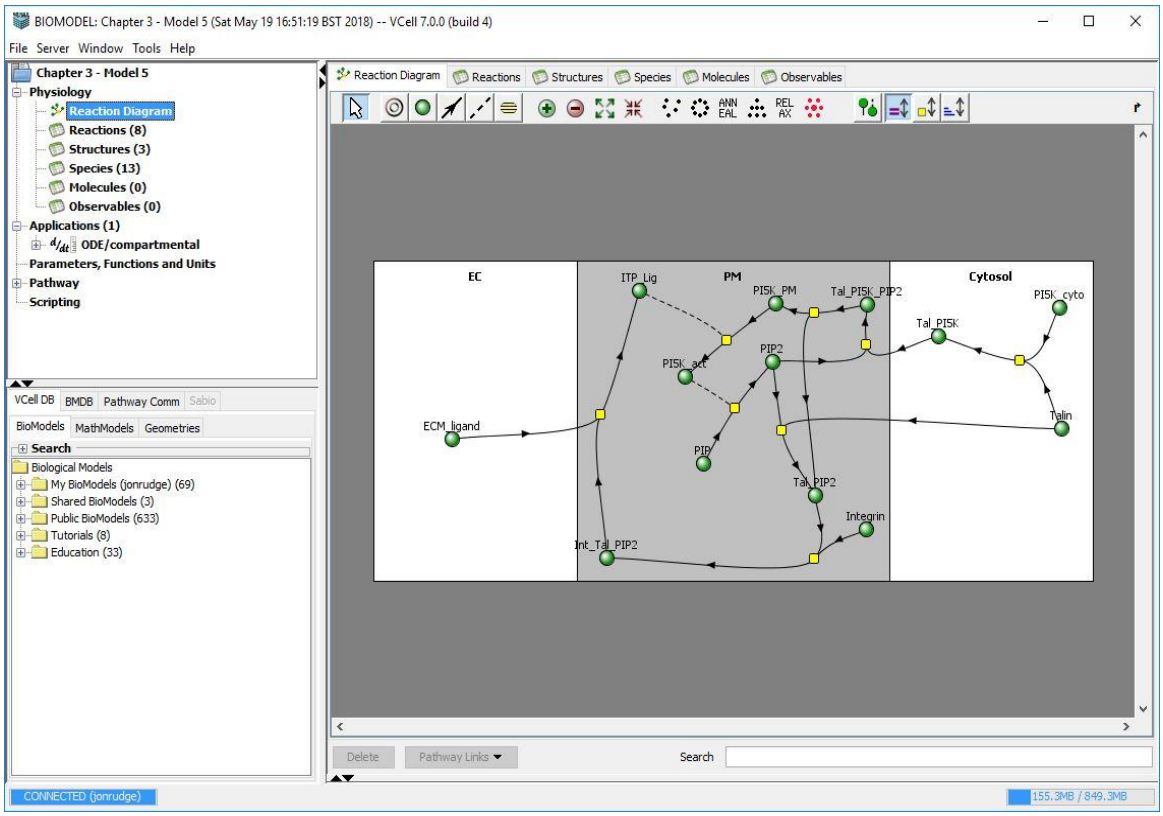

Figure 3.28. Reaction Diagram for Version 3 of the model.

Running this version of the model produces the results shown below.

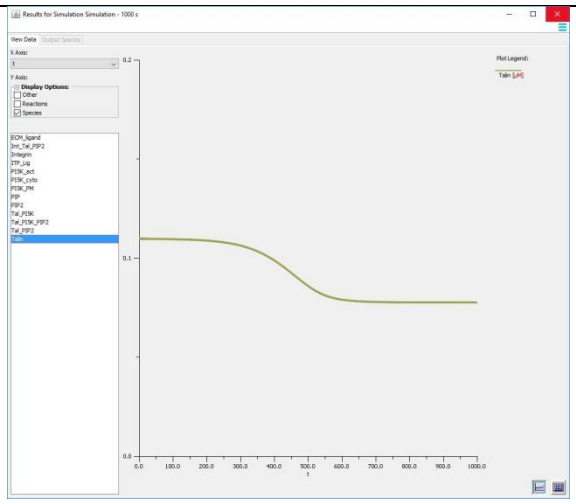

Figure 3.29. Talin (µM, 1000 s)
Minimum particle numbers: ~70,144 molecules

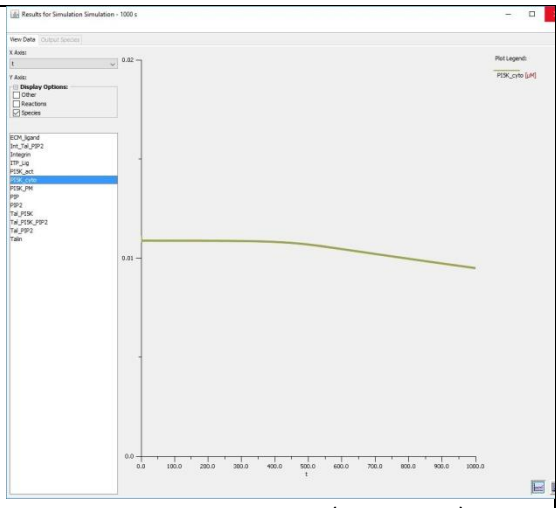

Figure 3.30. PI5K_cyto (µM, 1000 s)
Minimum particle numbers: 8515 molecules

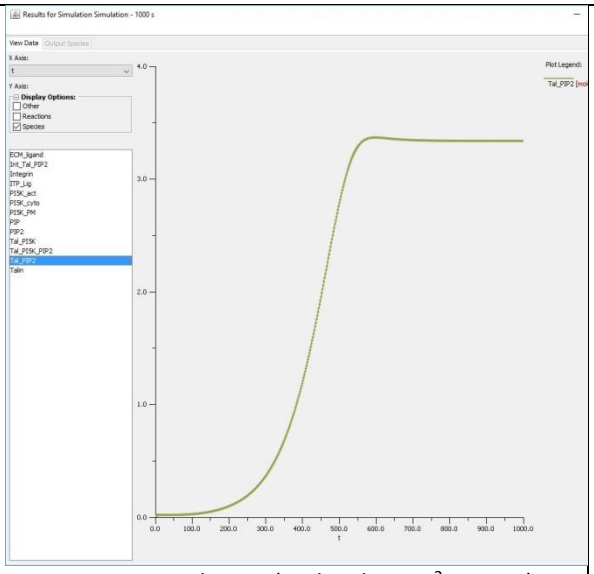

Figure 3.31. Tal_PIP2 (molecules.µm^2, 1000 s)
Peak particle numbers:5387 molecules

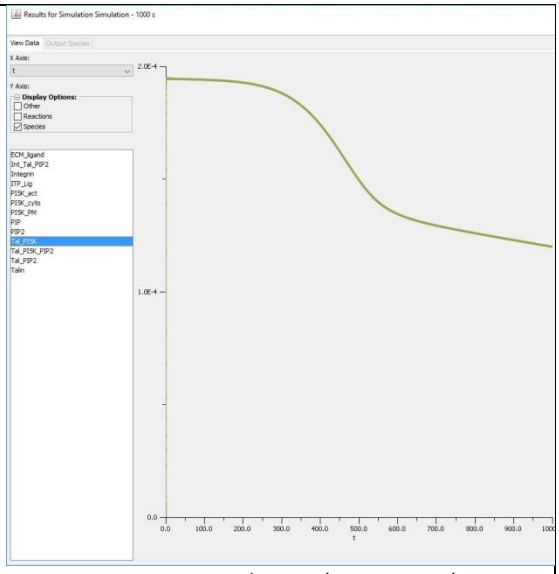

Figure 3.32. Tal_PI5K (µM, 1000 s)
Minimum particle numbers: 3589 molecules

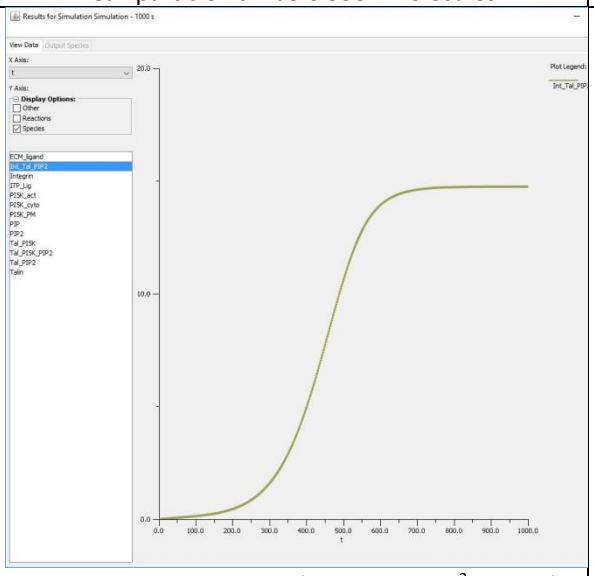

Figure 3.33. Int_Tal_PIP2 (molecules.µm^2, 1000 s)
Peak particle numbers: 23,600 molecules

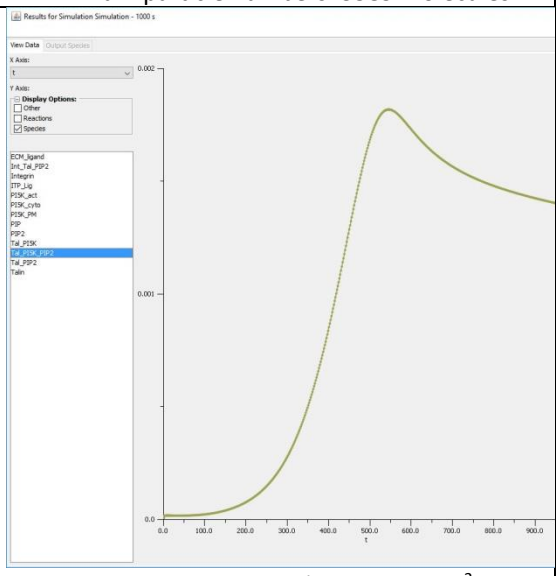

Figure 3.34. Tal_PI5K_PIP2 (molecules.µm^2, 1000 s)
Peak particle numbers: ~2 molecules

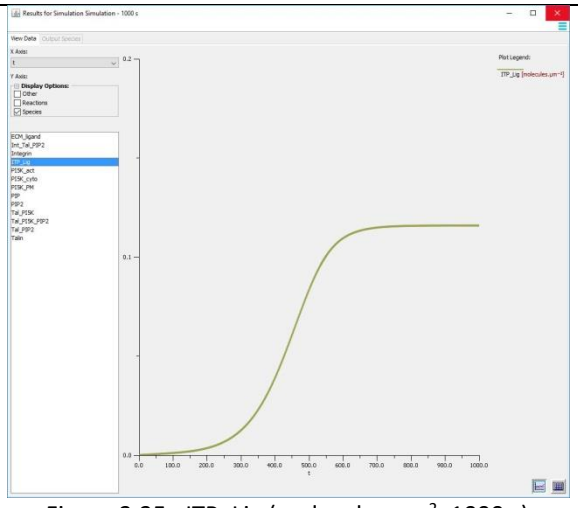

Figure 3.35. ITP_Lig (molecules.µm^2, 1000 s)

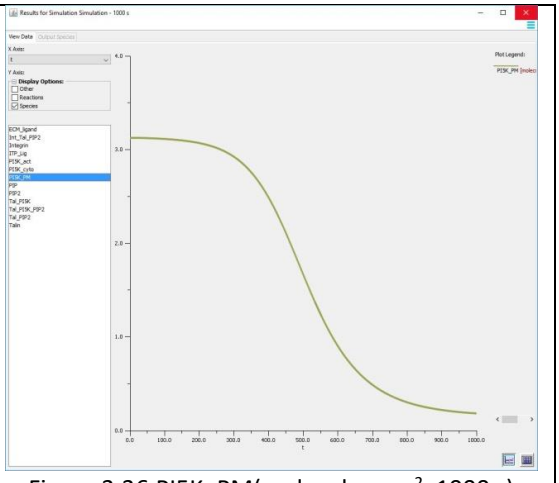

Figure 3.36.PI5K_PM(molecules.µm^2, 1000 s)
Minimum particle numbers: ~295 molecules

Peak particle numbers: 185 molecules

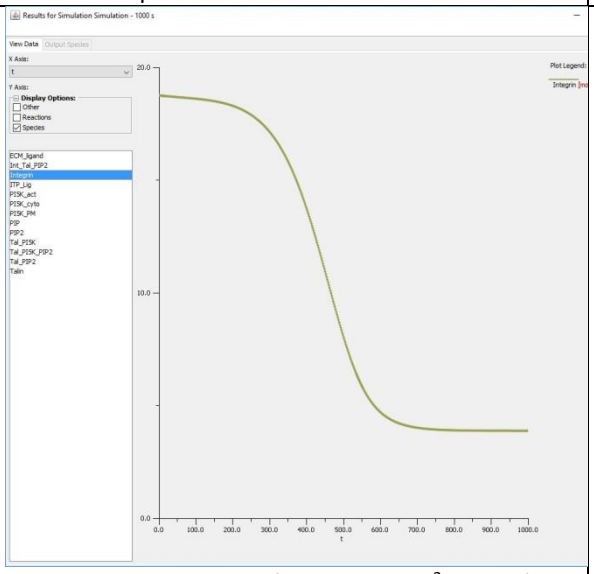

Figure 3.37. Integrin (molecules.µm^2, 1000 s)
Minimum particle numbers: ~6210 molecules

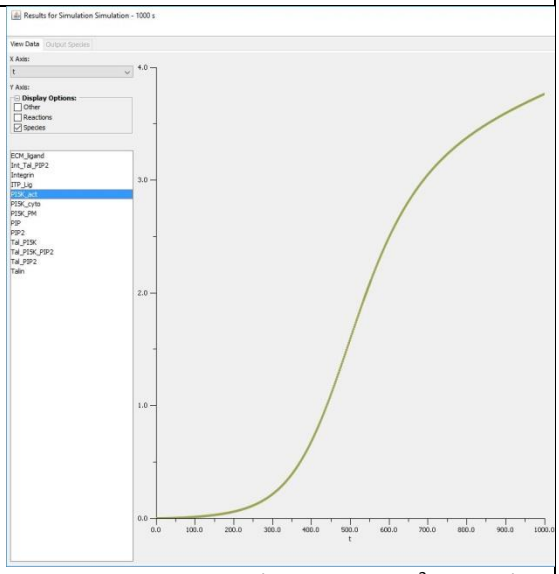

Figure 3.38. PI5K_act(molecules.µm^2, 1000 s)
Peak particle numbers: ~6029 molecules

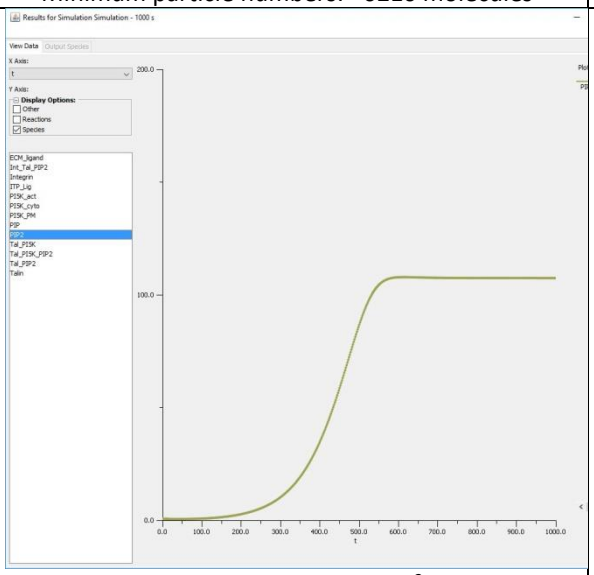

Figure 3.39. PIP2 (molecules.µm^2, 1000 s)
Peak particle numbers: 172,320 molecules

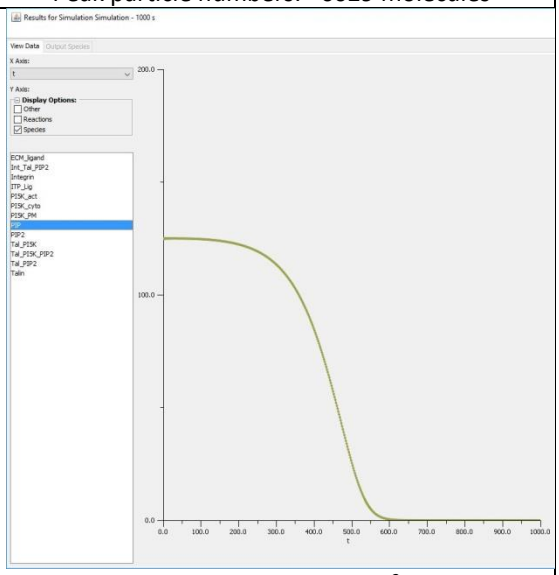

Figure 3.40. PIP (molecules.µm^2, 1000 s)
Minimum particle numbers: ~0 molecules

Overall, comparing Figures 3.29 to 3.40 to their corresponding plots in, it is clear that Version 3has reversed much of the speed decrease seen in Version 2 (Figures 3.15-3.26), restoring the rates closer to that seen in Version 1 (Figures 3.5 to 3.14), albeit running a little bit slower. Specifically, PIP levels require around 600 seconds to fall to zero in this version, as opposed to around 400 seconds, with a similar delay (i.e. around 200 seconds) seen in the time required for other model species to peak or fall to their minimum levels.

In other words, it takes around 10 minutes for maximal integrin activation and ECM ligation, PIP2 conversion, as well as for other related processes to reach terminal steady-state levels. Given that, at this point, around 24,000 out of 30,000 (i.e. around 80%) of the original available integrins have now bound to other model species, and 200,000 molecules of PIP have been converted to PIP2, this seems a physiologically realistic timescale. For instance, cell spreading has variously been reported to take tens of minutes [240,241,242]. Since this involves lamellipodia formation all around a cell, one might reasonably expect, after this was all complete that PIP2 levels would be very high, and that the majority of integrins would be activated and a maximal steady-state to have been reached.

However, there is still a major problem in that only 185 molecules of integrin is ligated to ECM ligands in the form of ITP Lig, or around 0.6% of the various integrin-associated species. This despite the fact that, at the same time, 23,000 molecules (around 14.7 molecules per μm^2, or approximately 78.4% of all integrin species) are available for such ligation in the form of Int_Tal_PIP2, and despite ECM_Ligand being available in excess.

Looking again at the kinetic parameters for the Int_Tal_PIP2-ECM_Lig binding reaction in Table 3.1 of the Methods section, and comparing it with similar reactions in the literature, it is noticeable that the off-rate value of 49 s^{-1} is surprisingly high for such a reaction. For instance, a study of integrin α5β1 binding to various fibronectin binding-site fragments, using surface plasmon resonance, calculated values of 4.9 (±4.9) x 10^5 M^{-1} $s^{-1}$for the on-rate and 6.5 (±0.1) x 10^{-3} s^{-1} for the off-rate, in wild-type fragments[243]. In other words, whereas the reported on-rate is very similar to the existing on-rate, the off-rate is nearly 5 orders of magnitude smaller. This, and other reports[244], suggest that a substantial reduction in this off-rate value would be justifiable.

For this reason, we choose a fairly conservative value of 0.1 s^{-1} for the next version of the model, leading to the changes seen in Table 3.8.

3.4.4 Fourth version

New reaction	Description	Rate law	Parameter values	Notes
Int_Tal_PIP2 + ECM_Ligand =ITP_Lig	Ligation of Int_Tal_PIP2 by ECM ligand	Mass action (reversible)	Kf = 0.77 μM^{-1} s^{-1} Kr = 0.1 s^{-1}	Adjustment to Kr value from 49 s^{-1}

Table 3.8. Change to reaction for Version 4 of the model.

Typical results, when running this version, are shown in Figures 3.41-3.46. Comparing these results with those of the third version of the model (Figures 3.5-3.14), it is immediately noticeable that, as well as substantially increasing the amount of ECM-ligated integrins (from approximately 185 to 22,480 molecules, or 0.125 to 15 molecules.μm^{-2}), PIP depletion takes around half the time (200 as opposed to 400 seconds) to reach zero levels, with a similar halving of the time taken for PIP2 to reach peak levels. There are also substantial reductions in the time taken for Tal_PIP2 and Int_Tal_PIP2 to reach peak levels, with the latter showing a much lower peak (around 5250 molecules as opposed to 22,480 molecules, or 3.5 as opposed to 15 molecules.μm^{-2}), as much more of it is ligated to ITP_Lig. Overall, far more integrins are being ligated by ECM protein (over 24,000 molecules more than in the first version) and the model is generally running substantially faster.

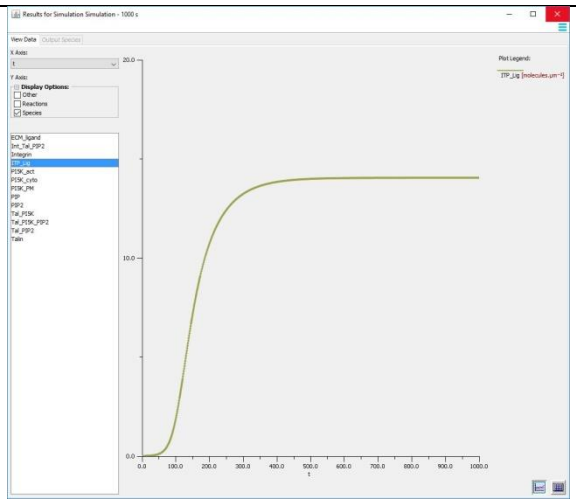

Figure 3.41. ITP_Lig (molecules.µm^2, 1000 s)
Peak particle numbers: 22,480 molecules

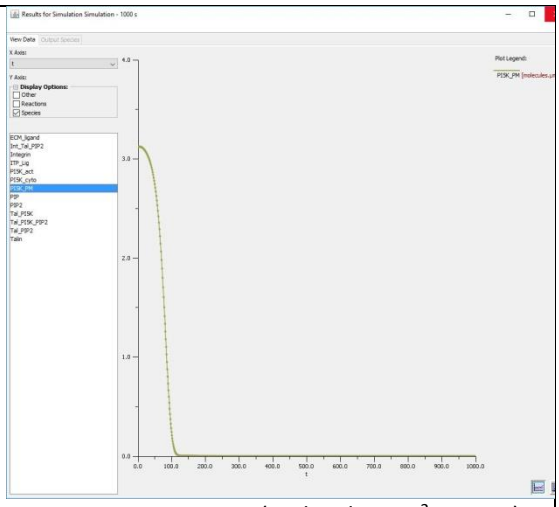

Figure 3.42. PI5K_PM(molecules.µm^2, 1000 s)
Minimum particle numbers: ~0 molecules

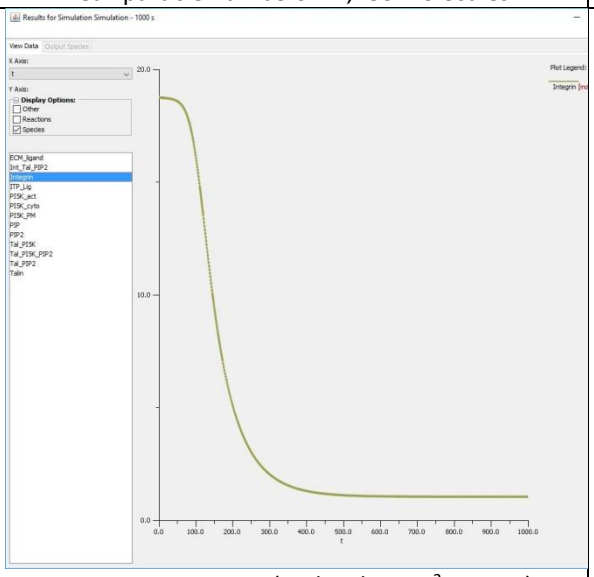

Figure 3.43. Integrin (molecules.µm^2, 1000 s)
Minimum particle numbers: 1674 molecules

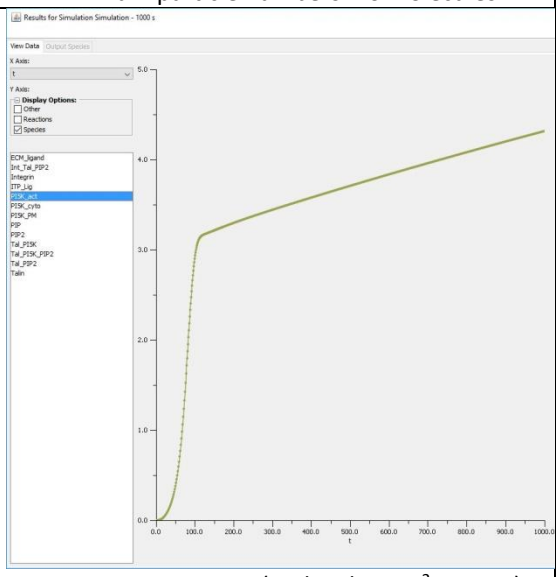

Figure 3.44. PI5K_act (molecules.µm^2, 1000 s)
Peak particle numbers: ~6906 molecules

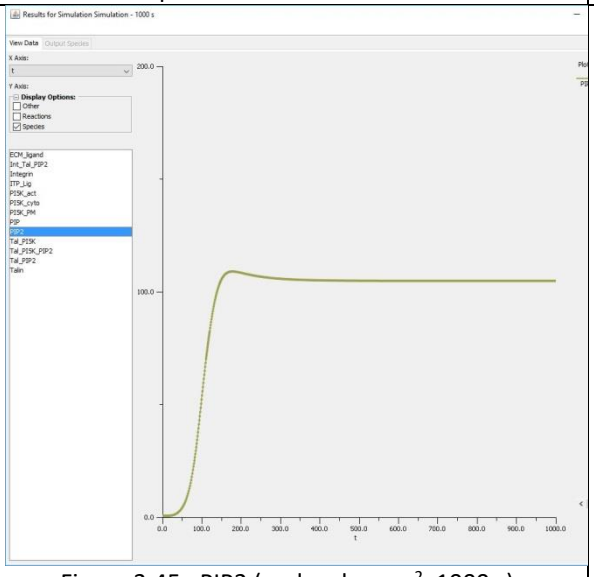

Figure 3.45. PIP2 (molecules.µm^2, 1000 s)
Peak particle numbers: 174,400 molecules

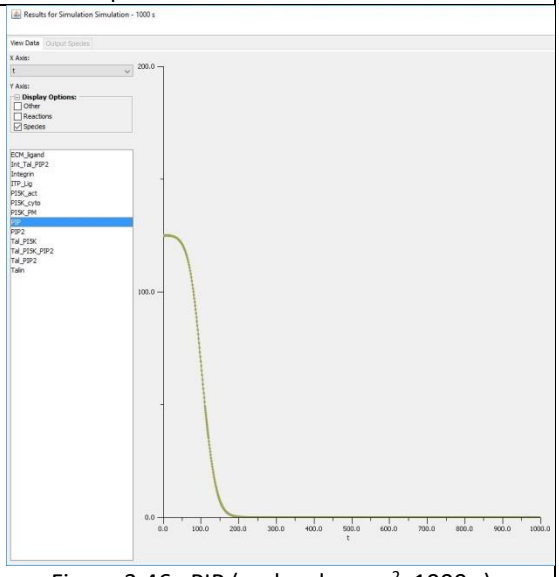

Figure 3.46. PIP (molecules.µm^2, 1000 s)
Minimum particle numbers: ~0 molecules

At this point, with a version of the model that is consistently showing behaviour that appears to be reasonably in accordance with experimental evidence, it makes sense to progress to a spatial model. In this case a 2D spatial model will be constructed for efficiency reasons, as such models run much faster than 3D equivalents (in minutes rather than hours), as a result of having far this calculations to perform. Before doing this, however, some thought has to be given as to shape and dimensions to give this model.

3.4.5 Fifth version (2D spatial)

Evidence from the literature suggests that a flattened disk of approximately 20 microns diameter would be an acceptable approximation of a HeLa cell, representing a typical cell grown to confluence [245]. If we assume, for the sake of simplicity, that the interior of this cell is all cytosol (i.e. it has no nucleus, other organelles, microtubules, etc), this implies a cytosolic area of just over 310 μm^2 (π *(10 $\mu m)^2$, where 10 µm represents the cell radius). Given that, in the compartmental/ODE application, we assumed that the cytosol had a volume of 1500 μm^3, this implies a height of 5 µm.

If we divide the extracellular volume (50,000 μm^3) from the ODE model by this same height we get an EC area of 10,000 μm^2, ensuring the same ratio of EC to cytosol sizes in both forms of the model.

Finally, the PM will be the circumference of this disk, meaning it will have a length of around 63 µm (π * 20 µm, where 20 µm is the diameter of the cell). This is considerably less than the area of the PM in the compartmental model (1600 μm^2) divided by the height, representing the fact that, modelling the HeLa cell in 2D plan view, like this, loses the vast majority of PM, most of which

would be on the top and bottom surfaces in a 3D representation of the same disk. The question is whether this disproportionate reduction of PM size will substantially alter system behaviour.

Having established the shape and dimensions of the 2D model geometry, it remains to implement it. This requires creating a new deterministic application. In the application Geometry Definition section, we add a new geometry, selecting the geometry type "Analytic Equations (2D)". Having renamed the opening subdomain0 to "EC", while accepting its default value of 1.0, we then edit the domain, increasing the X and Y sizes to 100 in both cases, whilst leaving the X and Y origin at (0, 0). A further (analytic) subdomain is added for the HeLa cytosol, replacing the default name by "Cytosol" and the default value by the following expression:

$$((((x - 50.0) \wedge 2.0) + ((y - 50.0) \wedge 2.0)) < (10.0 \wedge 2.0))$$

This will create a circle with a radius of 10 μm, centred on the coordinates (50.0, 50.0). However, before it can be visualised or used in the model these two newly-created geometric subdomains must be mapped to their corresponding logical structures, as defined in the Physiology. This is done in the Structure Mapping section, using the line tool (Figure 3.47). Returning to the Geometric Definition section, the 2D model of a HeLa cell can now be visualised in the Surface View subsection (Figure 3.48), while the respective compartment sizes can be seen in the Geometric Region Details subsections. This shows that the cytosol has approximately the right area (~300 μm^2) but the EC area is too small. This can be rectified by editing the domain again and increasing the X and Y sizes to 101.6 in both cases, yielding the compartment sizes shown in Table 3.9 below. These are now sufficiently close to the desired values to accept this geometrical specification for the application.

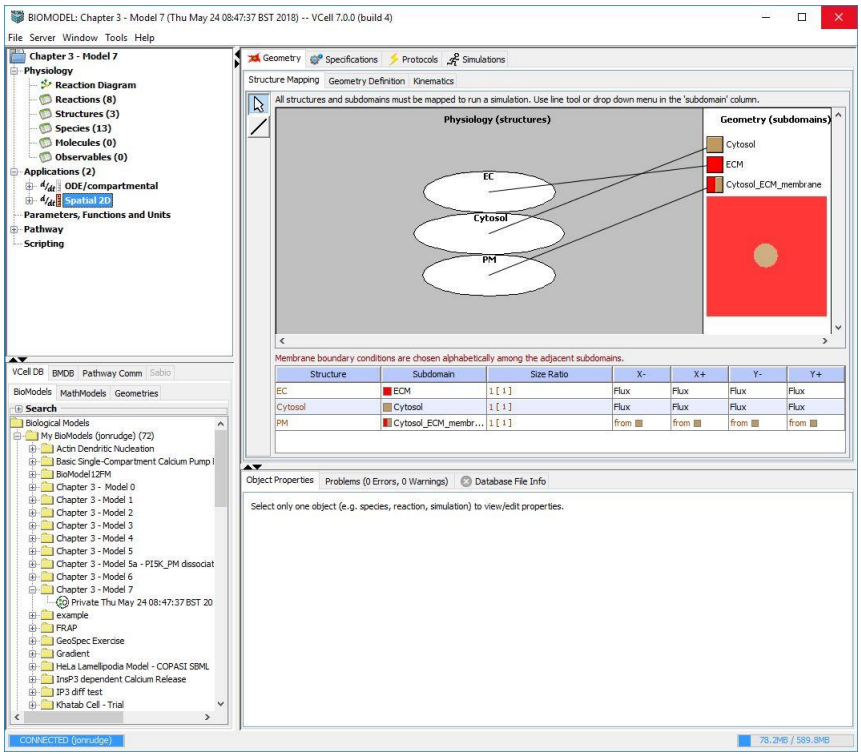

Figure 3.47. Structure Mapping for 2D spatial model (version 5).

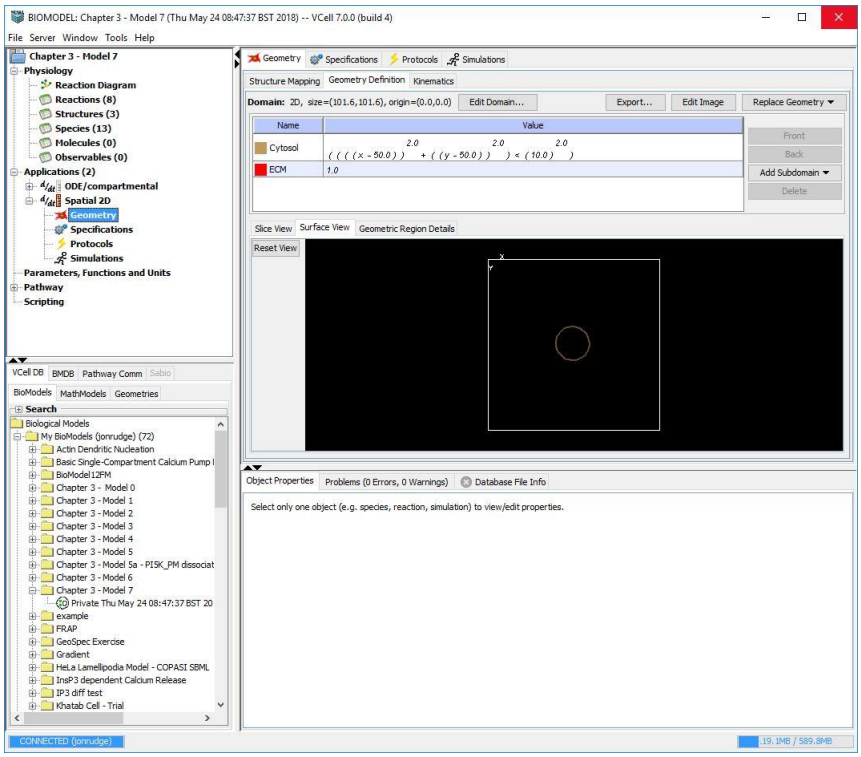

Figure 3.48. Geometric View of 2D spatial model (Version 5).

Compartment name	Compartment size
EC	1008 .75 μm^2
Cytosol	313.8 μm^2
PM	62.6 µm

Table 3.9. Compartment names and sizes for 2D spatial model (Version 5).

Having specified the geometry for this 2D spatial application, the next task is to specify one or more simulations in order to run it.

Again, as with the compartmental application, this is relatively straightforward in this case. Firstly one checks the Specifications section for the application, to see if any of the default settings need changing:

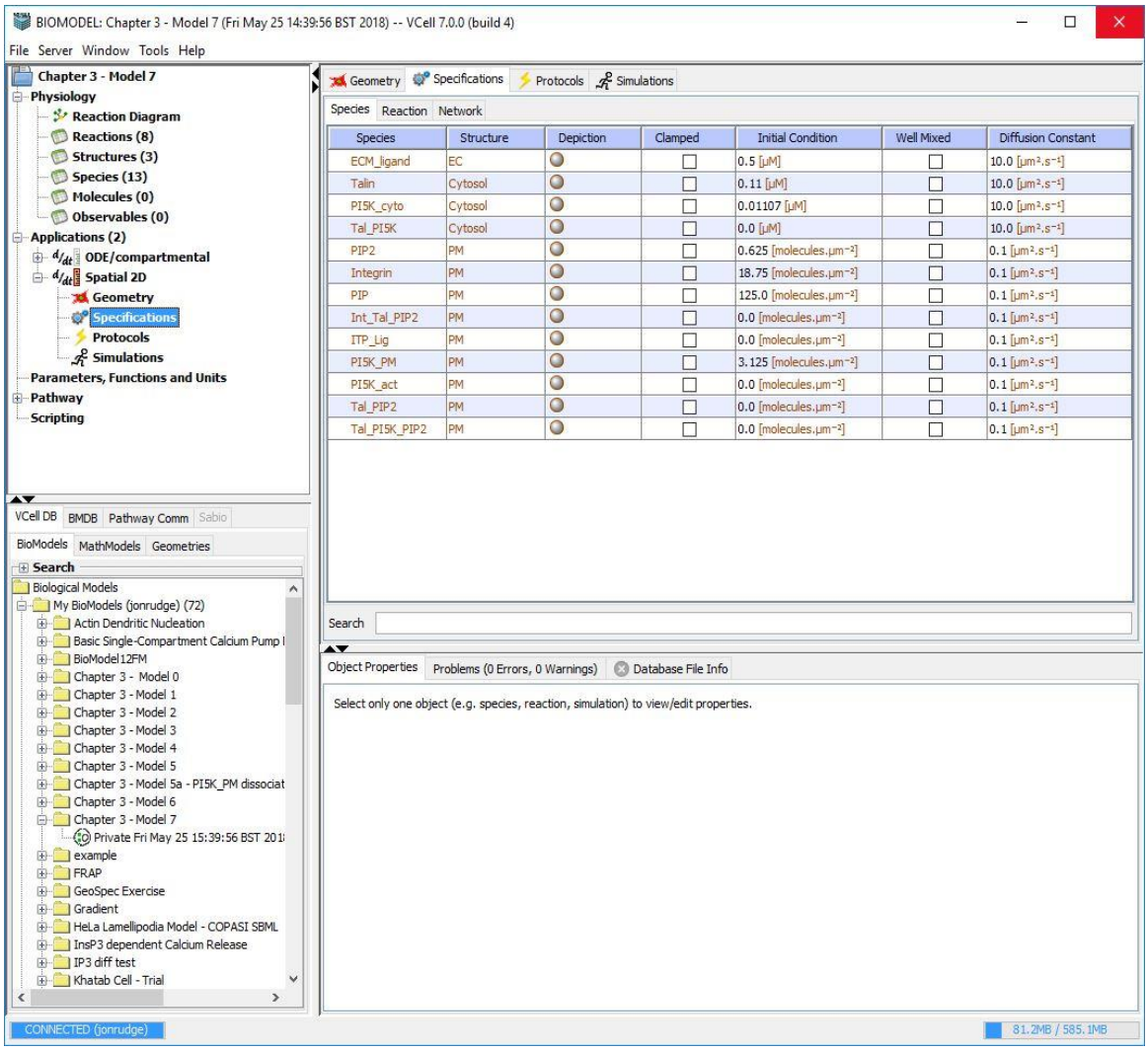

Figure 3.49. Specifications section for 2D spatial application (Version 5)

The Specification section is similar to the same section for the compartmental application, except for the addition of two columns at the end. The first one specifies whether the species in question is assumed to be well-mixed or is allowed to have different concentrations at different points. Since the whole point of such spatial models is to see how local concentrations of the different species vary with time, we accept the default state (i.e. unmixed). The diffusion coefficient (which determines the rate at which each species diffuses within their compartment) justifiably specifies much lower diffusion rates for membrane species (0.1 $\mu m^2.s^{-1}$) as opposed to liquid compartment species (10.0 $\mu m^2.s^{-1}$). Again, in this case, we accept these default values. The only change that is necessary is to add the concentrations for nonzero species, such as integrin. These are identical to the values entered for the compartmental application (shown in Table 3.3), despite the fact that the dimensions of all compartments are different between the two applications. In this case Virtual Cell converts the concentrations and densities into the appropriate values, as is also the case for the PM diffusion coefficient.

Having accepted the application specifications, a simulation can be added, following much the same procedure as for compartmental models. Again the user has the opportunity to override the application specifications, reaction rate constants and some other model parameters in the Parameters section. This time there is a Mesh section, where the user can specify the spatial resolution for the model run, which will affect both the accuracy and the time taken for the model simulation to run. Here, as for the Parameters section, we accept the default values. In the Solver section, we again accept the default integrator (I.e. Fully-Implicit Finite Volume, Regular Grid (Variable Time Step)) but override the Ending and maximum time step fields to 1000 and 1 (seconds) respectively. Finally, the output interval is increased to 1.0 seconds, chosen to give smoother, more accurate plots. (This will generate a warning, when running the simulation, and lead to plot results being very slow to appear. Normally a longer output interval of, say, five seconds would be more appropriate.)

Finally, running this model produces the following results:

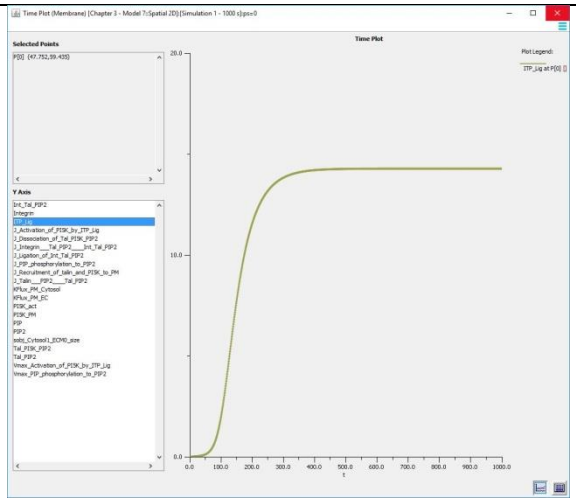

Figure 3.50. ITP_Lig (molecules.µm^2, 1000 s)
Peak particle numbers: ~893 molecules

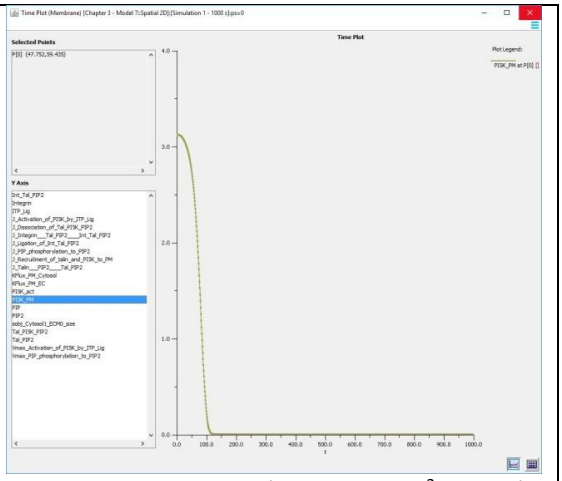

Figure 3.51. PI5K_PM (molecules.µm^2, 1000 s)
Minimum particle numbers: ~0 molecules

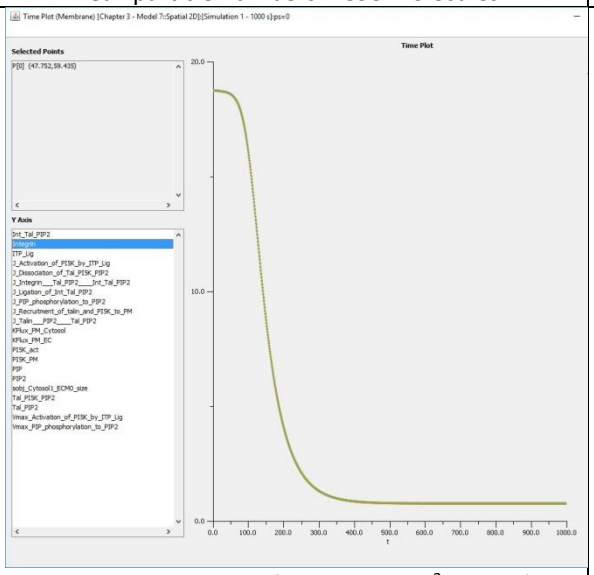

Figure 3.52. Integrin (molecules.µm^2, 1000 s)
Minimum particle numbers:~48 molecules

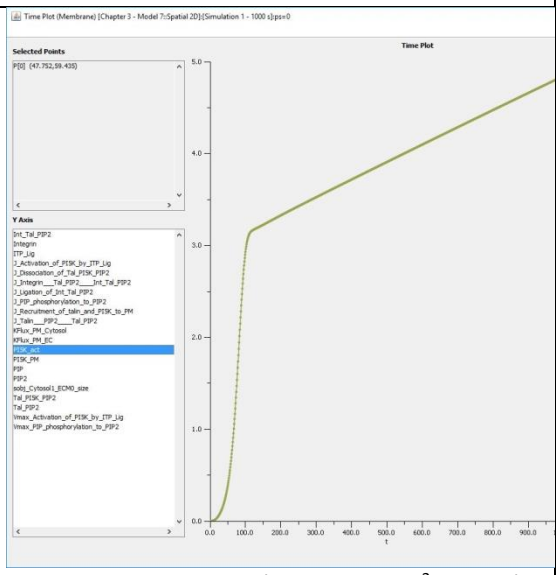

Figure 3.53. PI5K_act (molecules.µm^2, 1000 s)
Peak particle numbers:~303 molecules

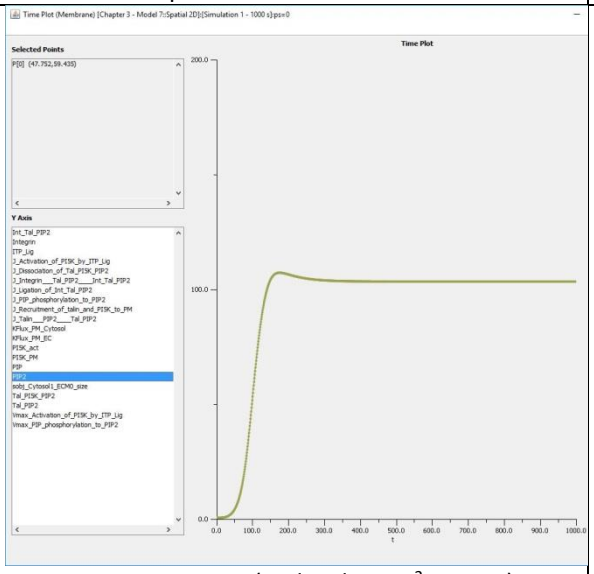

Figure 3.54. PIP2 (molecules.µm^2, 1000 s)
Peak particle numbers:~6712 molecules

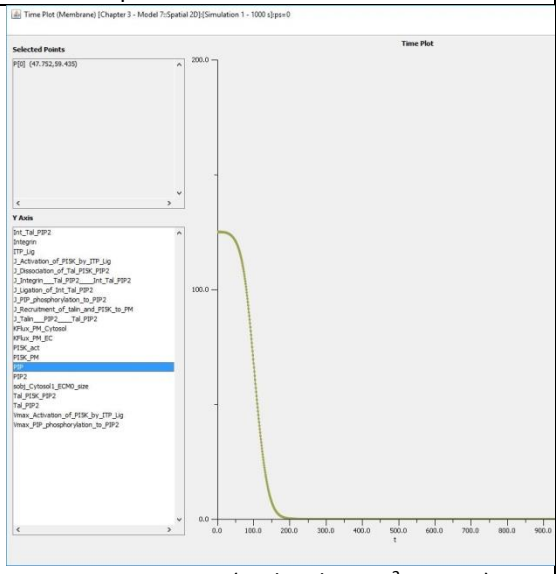

Figure 3.55. PIP (molecules.µm^2, 1000 s)
Minimum particle numbers: ~0 molecules

At first sight these results seem very similar to those of the previous compartmental model (i.e. version 4, shown in Figures 3.41-3.46), after allowing for differences in the size of the chart area in their respective plotting dialogs. However, it has to be remembered that these show densities or concentrations, not particle numbers, and, given that the PM is proportionally around one-fifth the size it is in compartmental model, this means its particle numbers are generally around one fifth the level they should be, also. Moreover, some species show obvious differences in density as well as number between the compartmental and spatial models, as can be seen from Figures 3.56 to 3.61.

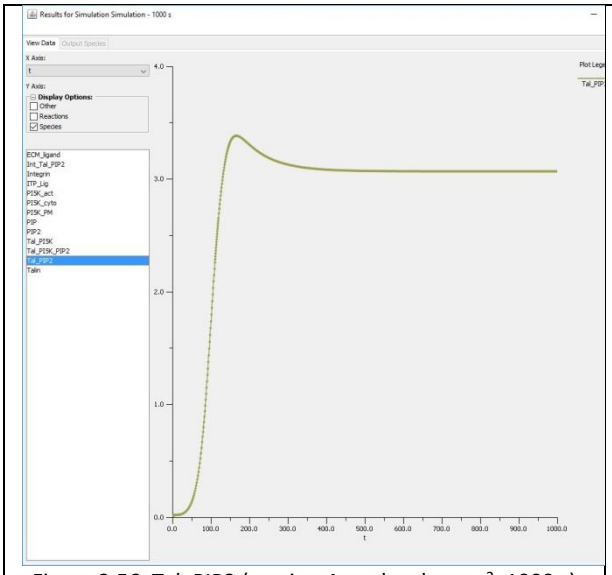

Figure 3.56. Tal_PIP2 (version 4, molecules.µm^2, 1000 s)
Peak particle numbers:5408 molecules

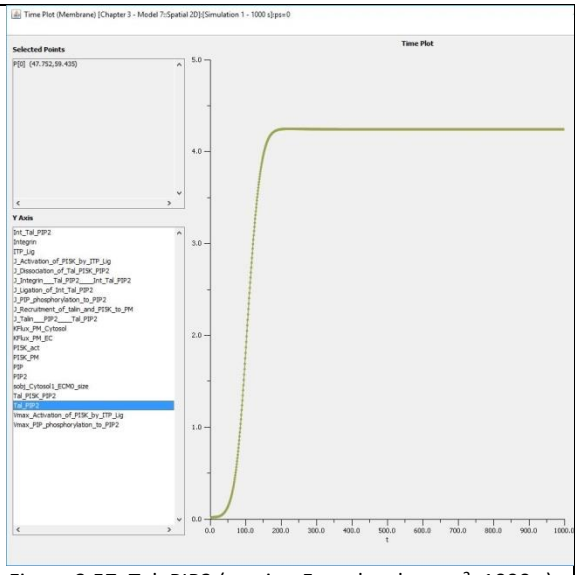

Figure 3.57. Tal_PIP2 (version 5, molecules.µm^2, 1000 s)
Peak particle numbers: ~266 molecules

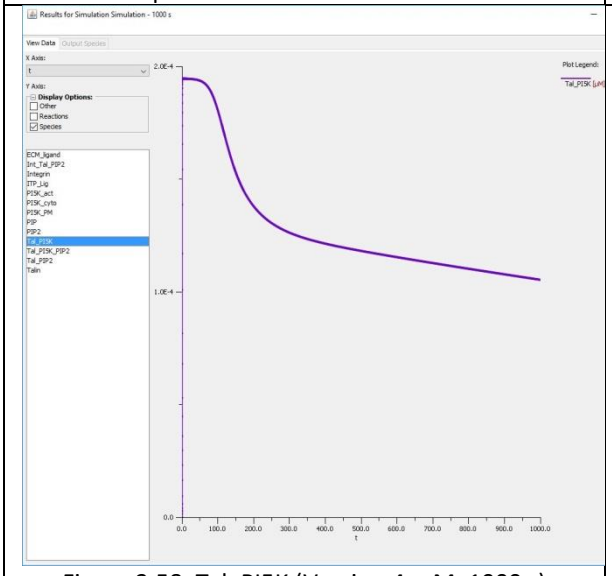

Figure 3.58. Tal_PI5K (Version 4, µM, 1000 s)
Minimum particle numbers: ~95 molecules

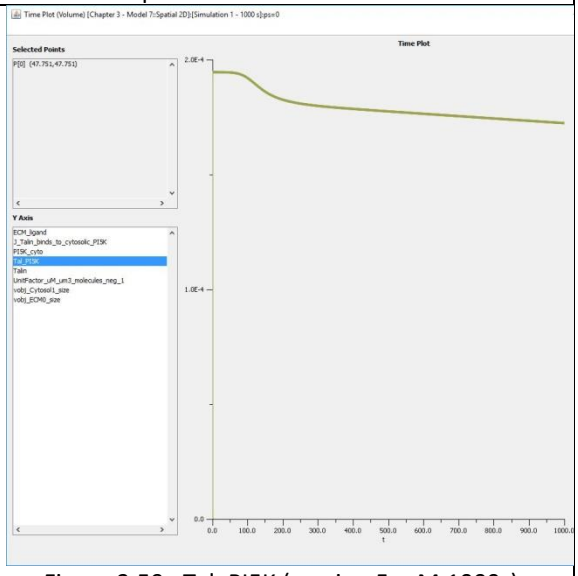

Figure 3.59. Tal_PI5K (version 5, µM,1000s)
Minimum particle numbers

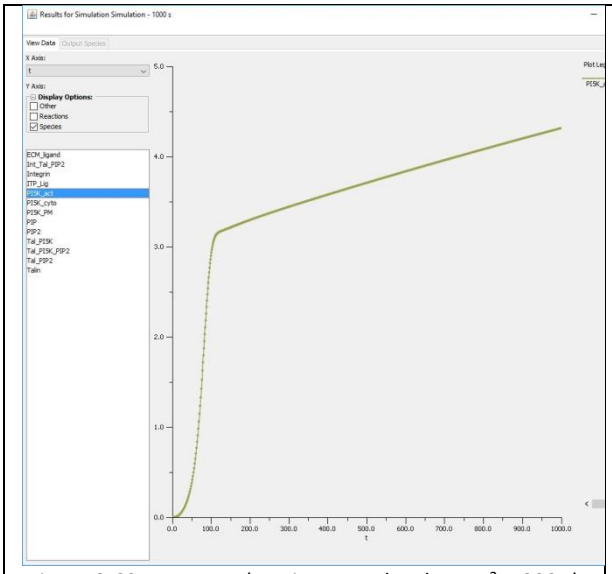

Figure 3.60. PI5K_act (version 4, molecules.µm^2, 1000 s)
Peak particle numbers: ~6906 molecules

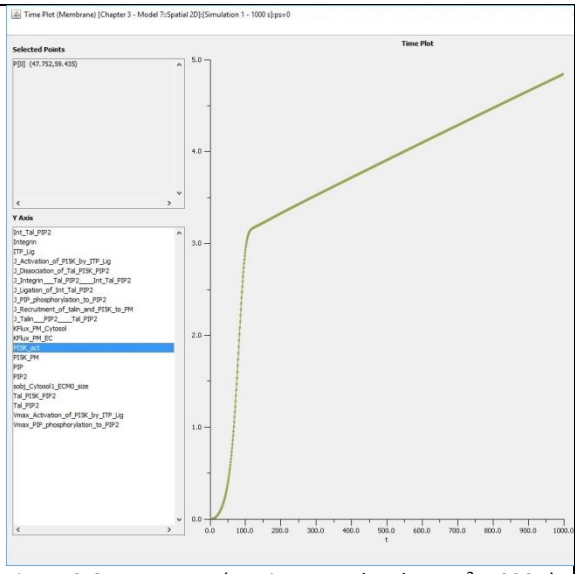

Figure 3.61. PI5K_act (version 5, molecules.µm^2, 1000 s)
Peak particle numbers: ~303 molecules

To see whether these discrepancies arise because of the compartmental differences, proportionally, between the two models, or for some other reason, we need to run another simulation. This time we will take advantage of an option in the simulation Parameters section that allows rescaling of compartments relative to their original geometric specification. Given that the compartmental model has a surface area of 1600 µm^2 and the other two compartments were divided by an assumed height of 5 µm, then keeping the PM length in proportion to these compartments implies dividing the PM surface area by the same amounts, which gives a desired length of 320 µm. Dividing this desired length by its actual length (i.e. 320 µm / 62.61 µm) gives the required scaling factor, i.e. 5.11. This can now be entered in the appropriate AreaPerUnitArea field in the simulation Parameters section, as shown in Figure 3.62.

3.4.6 Adjusted fifth version (2D Spatial)

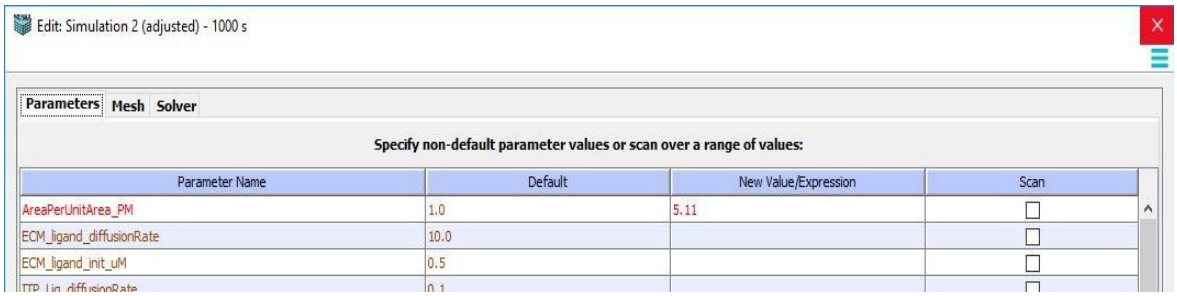

Figure 3.62. PM unit area rescaling, in this case multiplying PM unit area by 5.11.

This produces the following results:

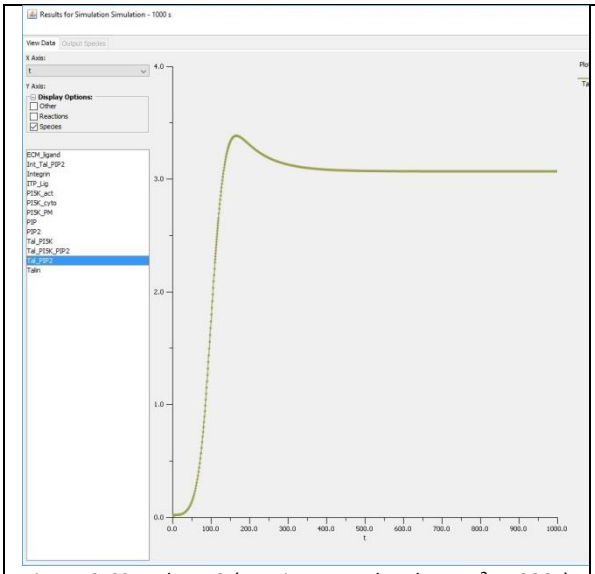

Figure 3.63. Tal_PIP2 (Version 4, molecules.µm^2, 1000s) Peak particle numbers: ~5408 molecules

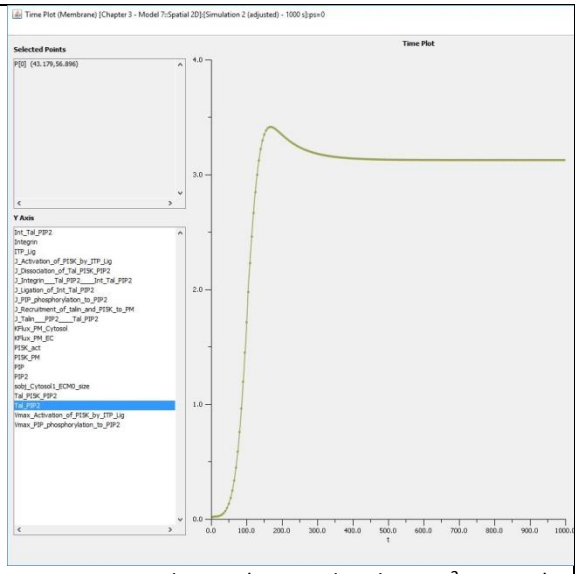

Figure 3.64. Tal_PIP2 (v.5, molecules.µm^2, 1000 s) Equivalent peak particle numbers: ~5461 molecules

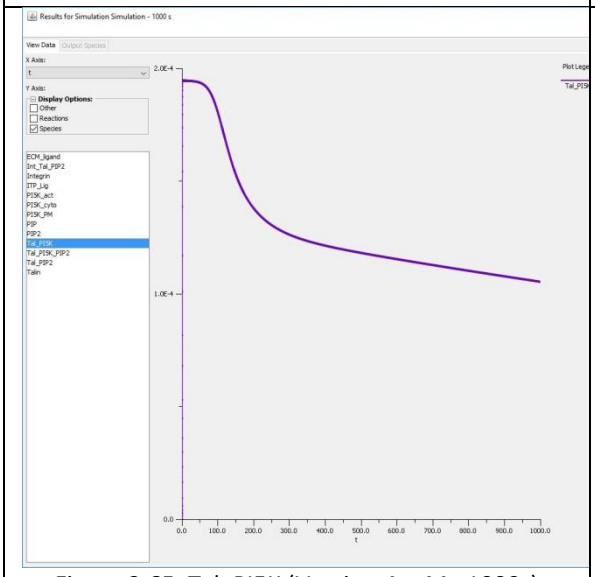

Figure 3.65. Tal_PI5K (Version 4, µM, 1000s)

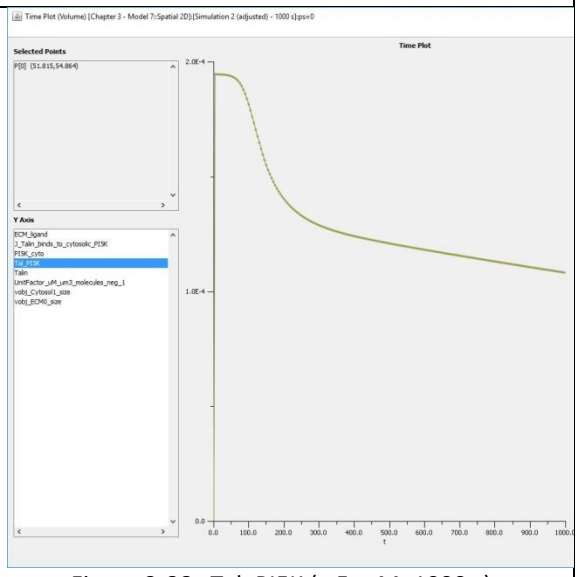

Figure 3.66. Tal_PI5K (v.5, µM, 1000 s)

Minimum particle numbers: ~176 molecules	Equivalent minimum particle nos.: ~176 molecules
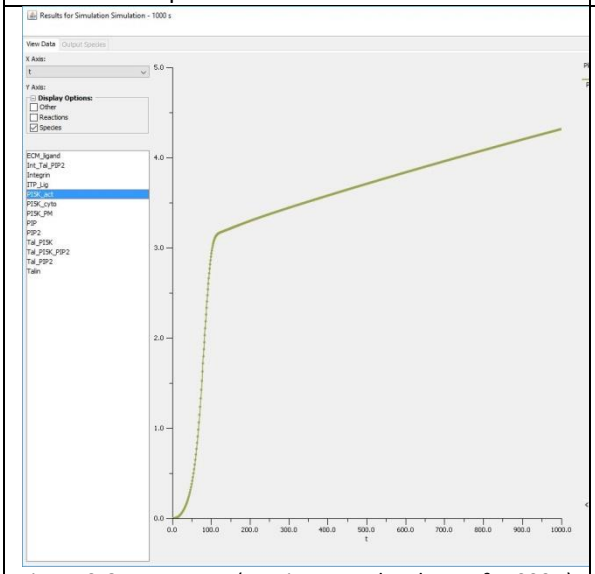	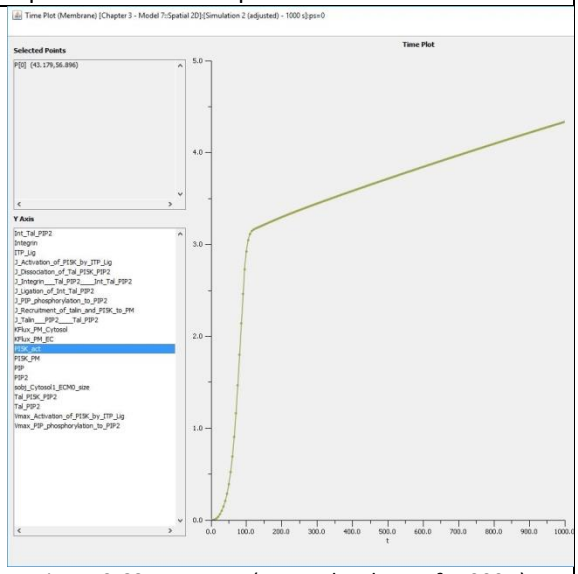
Figure 3.67. PI5K_act (Version 4, molecules.µm^2, 1000 s) Peak particle numbers: ~6906 molecules	Figure 3.68. PI5K_act (v.5, molecules.µm^2, 1000 s) Equivalent peak particle numbers: ~6936 molecules

Again, allowing for differences in the display format between the two models, these results are now essentially identical both in density/concentration and in equivalent particle numbers. (For PM-resident species equivalent values are obtained multiplying the adjusted Version 5PM species densities by the new PM length of 320 µm, and then multiplying these particle numbers by the assumed height of 5 µm. For cytosolic species equivalent values were obtained by multiplying the molar value by an assumed cytosolic volume of 1500 fl, i.e. the same cytosolic volume using the Version 4 model.)

This can be more clearly seen by comparing columns 2 and 4 in Table 3.10, and columns 2 and 3 in Table 3.11. Allowing for possible measurement error, this strongly suggests that the proportional differences in compartment sizes between the PM and the other two compartments are the reason for the discrepancies in results seen between the previous compartmental application and this spatial version.

PM Species	Version 4 peak/minimum particle numbers	Adjusted Version 5 peak/minimum particle numbers	Equivalent Version 5 3D peak/minimum particle numbers *
Integrin	1674	~329	~1645
ITP_Lig	22,480	~4500	22,496
Int_Tal_PIP2	5848	~1171	~5853
Tal_PIP2	5408	~1092	~5461
Tal_PI5K_PIP2	~3	~1	~3
PI5K_PM	~0	~0	~0
PI5K_act	~6906	~1387	6936
PIP2	174,400	34816	174,080
PIP2	~0	~0	~0

Table 3.10. Comparison of particle number counts for PM-resident species in Version 4 (compartmental/ODE) and Version 5 (2D-spatial/PDE) model, after PM length has been adjusted in Version 5 to match that of Version 4. (Values rounded to nearest integer.) Equivalent Version 5 values in the third column derived by multiplying (unrounded) values in the second column by 5 μm, the assumed height of the HeLa cell in this model.

Cytosolic species	Version 4 peak/minimum particle numbers	Adjusted Version 5 equivalent peak/minimum particle numbers
Talin	~67,304	~67,305
PI5K_cyto	~8010	~8056
Tal_PI5K	~176	~176

Table 3.11. Comparison of particle number counts for cytosolic species in version 4 (compartmental/ODE) and Version 5 (2D-spatial/PDE) model, where PM length has been adjusted in Version 5 to match that of Version 4. (Particle numbers calculated from reported molar values, assuming a cytosol with a volume of 1500 fl in both cases, then rounded to nearest integer.)

Overall, then, it can be seen that, as long as suitable adjustments are made to ensure that all compartment sizes are near-identical, spatial models (or at least 2D spatial can be a good match for compartmental models and vice versa. All these results will be discussed in the next section.

3.5 Discussion

3.5.1 Discussion of these early FA model results

On the whole, these results, from the early FA model, generally show that overall, integrin dynamics (including integrin activation by talin and PI5K activation by ECM-ligated integrin) are driven by levels of PIP2, rather than the reverse being the case. This was the conclusion reached from the results for the Version 1 model in which particle level curves for all key species, including ITP_Lig and ITP Lig-activated PI5K_act, are seen to lag particle level curves for PIP2 and PIP, typically by around 100 seconds or more. Specifically, comparisons of the times at which various species reach peak levels (or reach minimum levels, including exhaustion), and when their particle level curves transition from broadly linear to logarithmic rates of change (or vice versa) all tend to confirm that PIP2 is primarily driving model events.

Nevertheless, the rate at which integrin binds its respective ECM ligand is still critical. As the Version 4 model results show, not only does this determine how much activated integrin goes on to bind such ligands, but also the rate at which PIP2 is formed from PIP (as well as associated behaviour such as Tal_PIP2 formation) and overall system behaviour.

These two results point to a very complex interdependence between critical system species. It also implies, of course, that getting physiologically accurate rate constants for these two reactions (i.e. integrin-ECM ligation and PI5K-mediated PIP phosphorylation to PIP2) is particularly important if the model is to sufficiently reflect reality to be a useful predictive tool.

However, as in the previous chapter, such considerations come up against one of the main challenges in model development in cell biology, namely how to obtain reliable rate constant values for system reactions. Assuming one can locate the relevant rate constant in the literature

in the first place, how does one determine its reliability? Taking the example just mentioned, it was only at a later stage of model development that it became clear that a critical rate constant (the off-rate for integrin-ECM protein ligation) was likely to be incorrect by several orders of magnitude, despite being published in a well-respected journal [226], and involving a reaction of great interest to researchers . Had there been a substantial number of such values readily available in the literature, it seems likely that this off-rate would have featured as an outlier and been discounted as such.

This is not to question the competence of those conducting the experiment that produced this value, nor even the experimental protocol used. It is rather to point out the intrinsic difficulties of measuring reaction rates, given the large number of different proteins (and isoforms thereof), their great structural variation, and widespread differences in the ways they all interact with each other. Even without such variability, various technical challenges and very incomplete knowledge of the conditions (chemical and physical) under which reactions take place, can lead to large differences between in vivo and in vitro measurements, not helped by a general absence of agreed standards by which such reactions should be measured[246].

This means that the overall lack of reaction-specific data (at least in a form that allows such reactions to be reliably modelled) is compounded by reservations about the reliability of such data as there is. It may turn out that the rate constant value in question is, in fact, a better guide to the actual rate than has been supposed in this chapter. In which case, other explanations for the model's apparently aberrant behaviour must be found. But in the absence of other specific data to support a judgement either way, one is forced to base such judgements on more tangential evidence. Such evidence may include data from similar reactions, estimates based on other pertinent data, or any general principle thought applicable to the case.

A similar rate constant-related problem arose when trying to add talin-mediated PI5K recruitment into the model, as Version 2. In the absence of any kinetic data specifically relating to this reaction in the literature, it was hoped to model this using an existing reaction (talin-PIP2 binding) for which the data was already available. However, adding in PI5K release meant that the reverse reaction was now second-order rather than first-order. Unfortunately, as explained, there is no method for converting a first-order rate constant into a second-order constant, or indeed for converting any mass action reaction of one order to another. Adding in an additional reactant can completely alter the nature of a reaction, meaning that the existing reaction rate (and its associated rate constant) may be a very poor guide as to the rate of the new reaction. Removing reactants from a reaction will likely have similar results.

Talin-mediated PI5K recruitment to the PM, as introduced in version 4, also created a further problem, in terms of considerably slowing down overall model behaviour. Initially this appeared to be because the ad hoc reverse rate constant had been set too high. However, the results of a logarithmic parameter scan showed that only extremely low values for this rate constant would reverse the changes described, indicating that the fault might lie elsewhere. The problem, in fact, turned out to lie in the way the reaction (for talin-mediated PI5K recruitment) had been specified as a single reaction. Breaking this down into two separate reactions, one for Tal_PI5K-PIP2 binding and the other for dissociation of the resulting Tal_PI5K_PIP2 complex into two separate components (Tal_PIP2 and PI5K_PM) greatly speeded up overall model behaviour, largely restoring the timescale of events to that seen in the previous version.

In the writer's experience this problem is often encountered, typically when an attempt to simplify a model leads to multistep reactions (such as the one described) being compressed into a single one. Or it may, as in this instance, be driven by a (generally misplaced) desire to have a particular reaction, for which no kinetic data is known, in a form sufficiently similar to another

reaction (for which such data is available), so that one can justify to oneself using the same rate constant values.

As the example here shows, it is usually best to resist this temptation, even if it means having more reactions, each using ad hoc rate constant values in the absence of any better alternatives. Whilst, as just demonstrated, the difficulty of finding reliable reaction rate constant values is one of the great challenges of modelling reaction-diffusion systems in cell biology, it is not the only one.

The lack of published quantitative proteomes means that one is restricted to a small number of cell types for which quantitative proteomes have been published, if one is to justify the initial concentrations of species used in ones model. Furthermore, it has to be remembered that such quantitative proteomes, as have been published, are normally based on lysates from a number of cells, not just one [182,247]. If these cells have been cultured for long enough they may be at different stages in the cell cycle. Thus any quantitative proteome will typically represent an average value for many forms of the cell in question. At the same time, the lysates themselves represent a snapshot in time, before and after which protein copy numbers may vary considerably.

Overall, though, the results of this chapter suggest that the limitations and challenges discussed here need not get in the way of a model producing useful insights and predictions. In the previous chapter a compartmental/ODE model, implemented in COPASI, showed how rapidly increasing levels of PIP2 can play a prominent role in driving Rac1-mediated lamellipodia formation. Here, another compartmental/ODE model has been developed, this time in Virtual Cell, and showing how similar levels of PIP2 can simultaneously promote formation of FAs in numbers typically associated with lamellipodia. Moreover, whereas in the first example we saw

an example of inside-out signalling (in the form of SDF1 signalling via CXCR4 GPCRs), here we have seen an example of outside-in signalling in response to ECM-integrin binding.

In the process, this model has revealed the complex interaction between the FA-associated proteins integrin, talin and PI5K on the one hand and phosphoinositides such as PIP and PIP2, on the other. However, its comparative simplicity means that no very novel insights have arisen from its behaviour. Nor was it realistically expected that any would.

In the next chapter this model will be developed further to incorporate other FA-associated proteins and phosphoinositide species, as discussed in Section 3.1.1. To compensate for this much increased level of complexity, this new version of the model will be reduced in scale to model only a single lamellipodia. As well as reducing the computer load and therefore the simulation runtimes, such a lamellipodia-focused investigation will be more characteristic of the physiological setting in which FAs normally function.

In addition, the new model will be implemented both in COPASI and then in Virtual Cell, using compartmental/ODE modelling. This will allow a direct comparison between these two modelling tools, both in terms of the modelling process and in the results produced.

However, the principal gain will come from the additional complexity. An extensive search of the literature has found no other model that has looked at early FA dynamics, certainly not at this scale or level of detail. It is hoped that, as a result, more insights about FA behaviour may result, some of them entirely novel.

3.5.2 Discussion of spatial modelling approaches

This chapter has looked at some of the differences between modelling in Virtual Cell, a software environment primarily aimed at spatial/PDE modelling, compared with modelling in COPASI (exclusively designed for compartmental/ODE modelling), as described in the previous chapter. In the course of this illustration it has been shown that it frequently makes sense to develop a compartmental version of the model first. This avoids spending time at the outset specifying a spatial geometry for the model, but the chief time-saving comes from the much reduced run-time for simulations, with a 1000 seconds compartmental simulation typically requiring a solver run of the order of 10-30 seconds, where a 2D spatial equivalent might require 3-30 minutes, and a 3D spatial version 1 day or more. Generally, it can be said that there are few major issues concerning model behaviour that might arise in a spatial model, which will not also be apparent in a compartmental version. So it usually pays to initially develop the model as a compartmental application until the model is generally behaving as originally intended, and it has achieved a relatively stable state of development. By this has meant that the model has reached a stage where all glaring errors have been removed and the focus is now on incremental improvements in the model.

Of course, an alternative would be to first develop the compartmental version in COPASI, before porting it into Virtual Cell. However, this is not as straightforward as it might first appear. As has been seen, COPASI compartmental modelling differs from Virtual Cell modelling. Firstly, COPASI models compartments as volumes, whereas Virtual Cell models them as surfaces. Secondly, as a consequence, second-and higher-order mass action reactions involving only membrane-resident reactants must, in Virtual Cell, use rate constant values with units of $\mu m^2.molecules^{-1}.s^{-1}$, as opposed to volumetric units (i.e. per unit concentration per unit time) used by COPASI. (Similar considerations apply to membrane-associated Michaelis-Menten enzymatic and higher-order mass action reactions.)

Thirdly, whilst the facility to export COPASI models in SBML (Systems Biology Markup Language) file format, and to re-import such files into Virtual Cell, should allow seamless porting of such models, the reality is very different. This is because neither software environment is fully SBML-compatible[248,136], with a number of features in both environments (e.g. global variables, as implemented in COPASI, and electrophysiology in Virtual Cell) not currently supported by SBML.

In particular, with multi-compartmental models, COPASI does not require that any spatial relationship is specified between compartments, or that compartments are specified as membrane or non-membrane. There is no obvious reason within COPASI for such specifications, therefore exported COPASI SBML files do not contain this information either. However, Virtual Cell models must specify such information and, in its absence, all imported compartments are assumed to be non-membrane, with a spatial relationship that is essentially random. (For example, a Cytosol compartment may appear sandwiched between an EC and PM compartment, none of them specified as membranes.) Nor does it appear to be possible to change this specification. The result is that, unless a COPASI model is very simple, it is usually easier to specify it *ab initio* in Virtual Cell.

By contrast, as we have seen, progressing from a compartmental to a spatial version of a model is largely seamless within Virtual Cell. For all these reasons, if the intention from the outset is to develop a spatial model, then, in the absence of any overriding considerations (e.g. the need to take advantage of particular COPASI features, such as its more advanced sensitivity analysis and parameter estimation functionality), it will usually make sense to start with a compartmental version, implemented in Virtual Cell.

However, the results from this chapter raise questions about the actual benefits of spatial modelling. In the end, after having corrected the compartment sizes in the spatial model, the results were found to be essentially identical to those of the equivalent compartmental model, in terms of overall levels of model species as they evolve with time. Therefore, if this is all that one is interested in, it generally makes little sense to use spatial modelling, with all the accompanying overhead (including defining the spatial geometry and deciding on suitable diffusion constants), as well as the greatly increased simulation run-time.

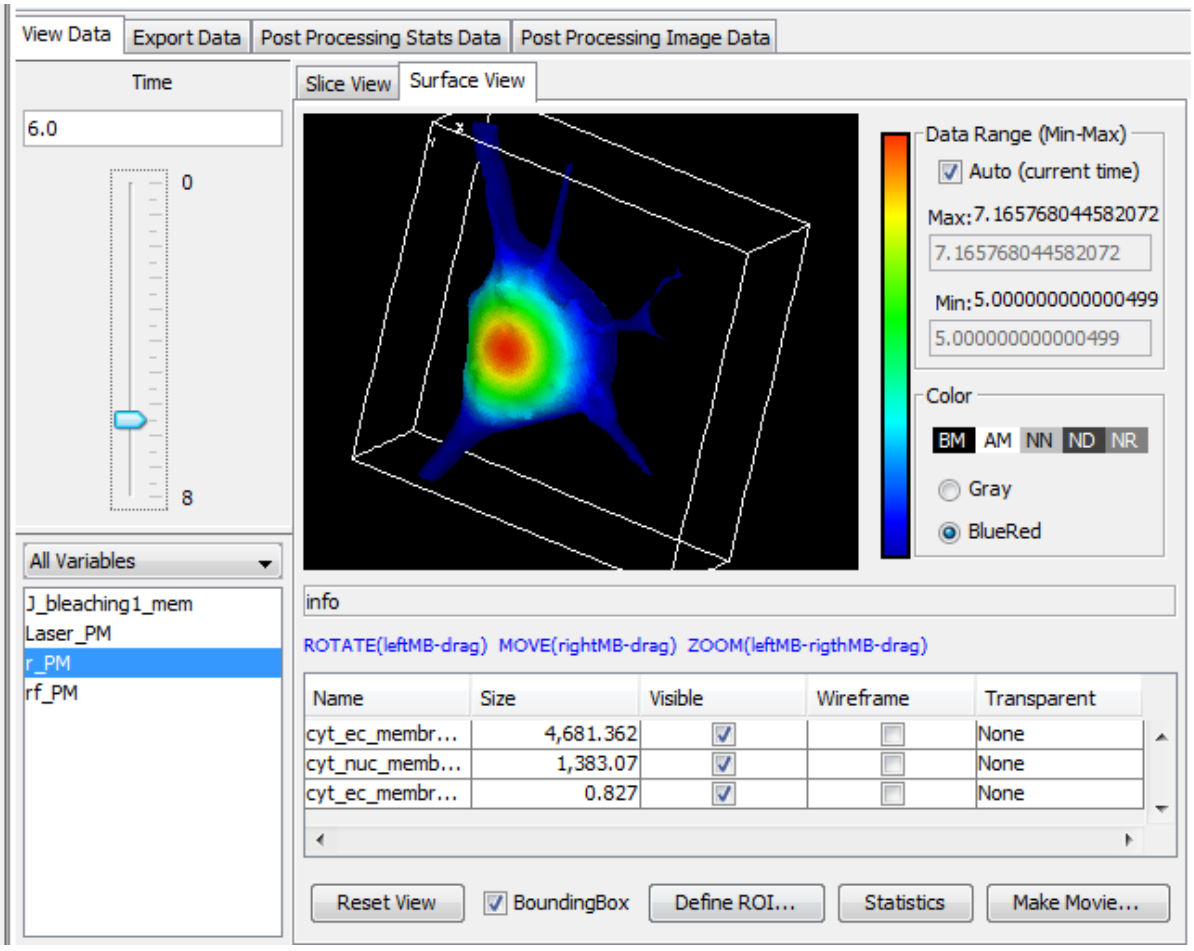

Figure 3.69. Example of selected species (r_PM here) density for a given simulation time point displayed as a colour scale over a user-defined rendered surface.

However, If one is interested in detailed spatiotemporal changes in species concentrations (for instance, how a signalling molecule, such as calcium, propagates through the cytosol), Virtual Cell spatial modelling provides useful analytical options. These include the ability to define various

linear and area-based (as well as single point) regions of interest over which changes in density/concentration of chosen species can be plotted, and (for 3D spatial models) displays of rendered surfaces, over which individual species densities are shown as a relative colour scale for any time point during the simulation run (Figure 3.69). Finally there is a facility to record a chronological sequence of such colour-mapped densities for a species (over a user-defined time period) as a video.

Together, these are certainly very useful features to have, making it much easier to see how, overall, local concentrations of given species vary, both spatially and temporally, over the course of a simulation. Without these one would have to rely on a series of plots of individual species levels, taken at single spatial points.

However, it must also be remembered that differential equation modelling (i.e. ODE/PDE compartmental/spatial modelling, as described in this and the previous chapter) models changes at the population, rather than the particle level. Put another way, tools such as COPASI and Virtual Cell are primarily designed to model reaction-diffusion problems, involving changes in space and time of species concentrations, rather than changes at the molecular level. This means that there are many molecular-focused spatial modelling tasks (for instance, studies of integrin clustering) that usually require alternative approaches, such as agent-based modelling [249] or molecular dynamics[250].

This aside, the spatial complexity of cells suggests that there must be problems in which the results from compartmental modelling are unlikely to be adequate on their own. Signalling gradients would seem to be one obvious example, since ODE models would surely require many separate compartments if it were to give any useful results, even where the gradient is not influenced by complex topography. Where such complex topography is present, such as in

Purkinje cell dendritic spines, the need for some form of spatial modelling seems obvious, as shown by at least one example [251]. But such complex topographies are relatively uncommon in cell biology, meaning that in most cases the modeller must often rely on personal experience when deciding whether to employ spatial modelling, taking account of the overheads previously mentioned.

Fortunately this does not have to be an all-or-nothing decision. In the model of the Purkinje dendritic spines just referred to, only some aspects of the problem were modelled spatially, with other aspects employing compartmental modelling. And even some of the spatial modelling required only one-dimensional models, considerably reducing the overhead incurred. It should be added that this modular approach, in which different aspects of a problem are modelled using the methodology judged most appropriate to that situation (a judgement which will usually include considerations of computational overheads), and with outputs from such modules typically informing other modules, has much to recommend it generally, in other words, not just in scenarios where spatial and compartmental differential equation modelling seem most appropriate.

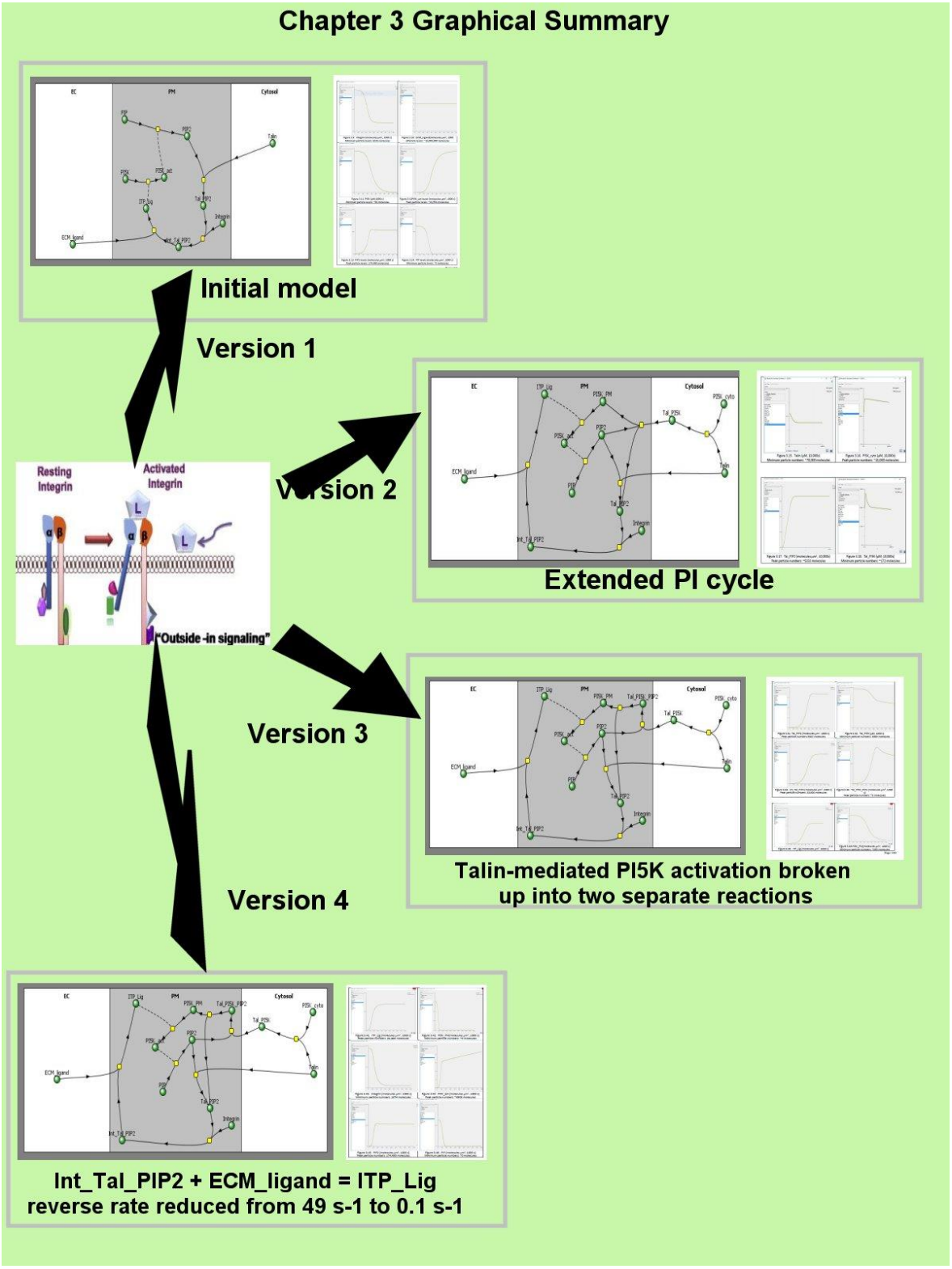

Figure 3.70. Graphical summary of Chapter 3 results.

Chapter 4. A complex model of early FA dynamics in lamellipodia

4.1 Introduction

4.1.1 How phosphoinositides and calpain cleavage combine to regulate FA assembly/disassembly

4.1.1.1 Possible roles for PIP2 in FA assembly

As shown in Chapter 3, the membrane-resident phosphoinositide PIP2 has a critical role in driving FA assembly by promoting and influencing talin (and thereby PI5K) membrane recruitment and activation. In particular, we saw how, by activating PI5K, PIP2 increased its own production (and, therefore, further talin-mediated integrin activation), thus exhibiting positive-feedback. But this role in FA assembly extends much further. Most importantly, PIP2 binding is not only pre-eminent in the cell membrane-recruitment of talin but is also thought to be important in recruiting FAK to the same membrane. As in the case of talin, such recruitment by PIP2 also activates FAK, by releasing both from their autoinhibitory conformation [221,33,52]. Activated talin and FAK in turn promote further FA assembly [252,253,52].

As well as activating integrins by forcing it to adopt a conformation with a higher affinity for its natural extracellular ligand, and promoting integrin clustering via cross-linking, as described in section 3.1,3, talin also provides binding sites for vinculin and actin. In the former case, such binding sites are revealed only after mechanical stretching has revealed them, thus giving talin a role as a force-activated mechanical switch [32,26].

Similarly, FAK, in its role as a scaffold protein, provides binding sites for other proteins, including growth factor receptors, PI3-kinase and Src [37,254,255]. Having been recruited, Src is then activated, leading to further FAK phosphorylation [256,257]and the formation of a FAK-Src complex. This

complex, amongst other things, regulates Rho-family GTPases, either by direct Src-mediated phosphorylation of p190RhoGAP [258], or, less directly, via the Dock180-ELMO and GIT2-β-Pix pathways [259], following phosphorylation of paxillin and p130Cas.

There is also experimental evidence showing that PIP2 binding to vinculin is essential to the latter's critical function in FA dynamics, probably involving a role in vinculin's membrane-recruitment or activation, or both [260].

4.1.1.2 Possible roles for PIP2 in FA disassembly

What of disassembly? As well as driving new FA assembly by promoting recruitment and activation of key FA proteins, as described above, PIP2also appears to promote turnover of existing FAs [261]. A number of ways in which PIP2 might do this have been proposed. For instance, it has been suggested that high levels of PIP2 may lead to lower vinculin residency times in FAs, so preventing vinculin-mediated stabilisation of FAs [262]. Or, conversely, high PIP2 levels may promote FA disassembly by prolonging residency of FAK [261,256].

However, this does not appear to fully explain disassembly, failing to provide a mechanism for spontaneous disaggregation of the remaining FA protein constituents, such as talin, paxillin and FAK. Nor would falling levels of PIP2, given that PIP2 does not appear to perform a critical structural role within FAs, by cross-linking or other means. Declining levels of PIP2 might reduce levels of FA formation, but there is no obvious reason why they should promote disintegration of existing FAs.

Another possible explanation lies in endocytosis, with some reports suggesting a role for PIP2 in driving clathrin-mediated endocytosis. This appears to involve an interaction with dynamin and other clathrin-associated proteins [263,264,265,266], possibly requiring FAK [50,267]. However, there is only

limited evidence for PIP2 as a direct driver of FA-associated endocytosis, and a precise mechanism has yet to be elucidated.

More generally, the idea that high levels of PIP2 can drive both assembly and disassembly is problematic, unless there is some form of switching mechanism to turn off disassembly where not required, either by lowering levels of PIP2 or by preventing PIP2 mediated disassembly. Otherwise FAs could never mature. Possibly FA-stabilising vinculin may provide such a switch, since we can reasonably assume that mechanical force tends to hold vinculin within FAs (by providing vinculin binding sites within talin and actin binding sites within vinculin), thus acting in opposition to the tendency of PIP2 (observed by Chandrasekar and colleagues [262]) to reduce vinculin residency times within FAs. However, precisely how PIP2 acts on vinculin is still very much open to question.

4.1.1.3 Possible more indirect roles for PIP2 in FA disassembly

Recently, however, it has become clear that the phosphoinositides may play a more indirect, but critical, role in FA disassembly. Hydrolysis of PIP2 by PLC isoforms yields the membrane lipid DAG and the cytosolic molecule IP3. IP3 binds to IP3R receptors in the ER, triggering the release of calcium ions (of which the ER is a major reservoir) into the cytosol. This cytosolic calcium then binds to and activates calpain, a calcium-dependent cysteine protease, which, amongst other things, cleaves a number of FA-associated proteins, including talin, FAK, paxillin, vinculin, α-actinin and at least one integrin isoform [268,269,270,271,272,273,274]. FA disassembly and subsequent degradation of FA component proteins appear to be substantially driven by this calpain cleavage [271,269].

However, as before, this account appears contradictory. On the one hand, PIP2 promotes assembly and, on the other, indirectly promotes disassembly, implying that these opposite

processes would tend to cancel each other out for any given level of PIP2. The simplest way this can be resolved is if assembly and disassembly occur at different times, again implying some sort of switching mechanism. If so, an obvious candidate for this switching role is PLCγ, since it converts PIP2 to IP3.

In support of this, evidence suggests that PLCγ is primarily recruited to the cell membrane by PIP3 [275,276], the downstream product of PIP2 phosphorylation. Also, it appears that PI-3 kinase, which performs this phosphorylation, and PLCγ are both activated by FA-associated tyrosine kinases (including Src family kinases, FAK and Integrin Linked Kinase), following integrin ligation and cross-linking (coincident with clustering) by extracellular ligands [277,176,278,279,280,281].

Thus it is very plausible that PLCγ activation and IP3 creation occurs later than the initial burst of PIP2 formation accompanying FA assembly, being driven by integrin ligation and clustering. If so, one might suppose it is the default response to these two events, preventing further FA maturation and instead promoting FA disassembly, unless a further event inhibits this response. Thus it may be that force-related changes (e.g. talin/vinculin elongation) in some way turn off PLCγ activation, or in some other way inhibit calpain cleavage, and so allow FA maturation. In the absence of these, within a limited timeframe, calpain cleavage and consequent FA disassembly and recycling proceed automatically.

Altogether, it seems clear that FA dynamics involves a large number of processes and interdependencies, making up a complex system.

4.1.2 What remains to be discovered

Whilst a general idea is emerging of how PIP2 and calcium-driven calpain cleavage contribute to focal adhesion assembly and disassembly, important aspects are still missing from the picture. For instance, the relative importance of each factor to overall FA dynamics is far from clear. And we still lack a precise understanding of the underlying mechanics and how they are regulated.

Crucially, the precise mechanism leading to rapid PIP2 creation from PM-associated PIs is poorly understood. The recent discovery that talin appears to recruit PIPKIγ to the PM, at the same time as being itself recruited to the PM (and activated) by PIP2, provides an obvious source of PIP2, together with a positive-feedback mechanism for its generation. However, such a mechanism surely requires a separate trigger event, both to initiate it in appropriate circumstances (e.g. after inside-out or outside-in stimulation) but also to prevent it happening inappropriately. At the time of writing this trigger mechanism has yet to be fully determined.

Also, whilst evidence suggests that PIP3 and Src phosphorylation are critical to PM-recruitment and activation of PLCγ, respectively, the underlying mechanism is less well understood. That is to say, what exactly drives Src phosphorylation and PI3K-mediated PIP3 creation, how is it synchronised, and how is that mechanism governed by external events?

Also, how do the Rho-family GTPases, Rac1, Cdc42 and RhoA, in their role as key governors of overall FA dynamics [112,282], fit into the above picture? What role does internalisation and recycling play in FA assembly and disassembly [283]? Do PIP2 and calpain-cleavage have the same relative importance in inside-out and outside-in signalling? How does the calcium signal propagate, and how is that signalling and its FA response affected by cell morphology? Does the calcium signalling mechanism permit individual FAs to generate and receive different calcium signals to their neighbours, or do adjacent FAs receive the same summated signal? If so, what is the typical

size, spatially or numerically, of FA groups receiving such summated signals? Finally, what effect, if any, does the size and shape of lamellipodia (or similar FA-active regions) have? Similarly, how does the size and shape of calcium stores, and the proximity of the latter to the lamellipodial leading-edge, have on the observed FA response?

It is also still far from clear what the precise mechanism is that determines whether a newly-formed FA subsequently disassembles and is recycled or goes on to form a mature adhesion attached to actomyosin fibres. That is to say, although it is fairly well established that mechanical stretching of talin and vinculin play an important role, by recruiting actin fibres to the existing FA complex [284,93], it is far from clear how this relates to disassembly and calpain cleavage. Are disassembly and cleavage default processes that talin and vinculin stretching somehow switch off? If so, how? And how does recycling fit into all this? It may be that vinculin and PIP2 both play an important role here, with a finding that, in the presence of high levels of PIP2, vinculin becomes unbound from FAs, preventing force-induced maturation and promoting cell-spreading, characterised by high levels of FA turnover and recycling [262]. In this respect, then, vinculin is seem to be acting as a kind of PIP2 sensor and switch, with PIP2 acting to recruit vinculin away from FAs, rather than towards them as originally thought.

This would make some sense, effectively setting a timer for force-driven FA maturation. If sufficient mechanical force is present, vinculin remains bound within its FA, become stretched and thereby provides additional binding sites for actin fibres. If no such force is present, the high levels of PIP2 (characteristic of lamellipodia and filopodia) have time to divert it away from its associated FA, preventing maturation. In some way, then, it seems likely that these smaller, vinculin-deficient, immature FAs will be much more prone to calpain cleavage and recycling.

Certainly, this is consistent with the phenomenon of cell-spreading, which is characterised by an absence of traction forces, high levels of PIP2 and of nascent adhesions that fail to mature and are rapidly recycled [285]. One would expect these high levels of PIP2, of course, to be associated with high levels of IP3- and calcium-mediated calpain cleavage However, it should be emphasised that this vinculin-switch hypotheses is solely based on a report from 2005 that has not subsequently received much experimental support in the literature. Also, calpain cleavage itself has been reported to have an important role in inhibiting such cell spreading [286]. So, many aspects of FA dynamics, including the role of PIP2 and calpain cleavage, remain to be determined.

At a simpler level, there is still much confusion as to the precise order of recruitment and activation, as well as the immediate binding partners, of key FA proteins, such as talin, FAK, Src, paxillin and vinculin. Nor is it known how much flexibility there is in this recruitment and activation order.

For instance, it is not known for certain whether talin is recruited to early FAs before FAK, all the other way round [54]. Similarly, it was once suggested that paxillin might be binding directly to β-integrin cytoplasmic tails [287], implying that it might be the earliest recruit to incipient FAs. However, it is now thought that paxillin binds directly to the tail only of α-4 integrins [79]. In the absence of such binding, it is not absolutely certain which constituent FA protein, let alone what part of it, initially recruits paxillin to FAs [288].

4.1.3 Modelling

4.1.3.1 How modelling can further our understanding of FA assembly and disassembly

One reason why a mathematical model can be particularly useful here follows from the interdependence between the two processes described. That is to say, IP3 levels determine levels of calpain cleavage to a large degree, but are themselves in large part determined by levels of

PIP2. Less directly, PIP3 levels also affect calpain cleavage by recruiting PLCγ to the PM [275,276], activating it and encouraging IP3 production. But, of course, these levels of IP3 and PIP3 both come at the expense of PIP2, which drives FA assembly, through talin PM recruitment and subsequent activation. And, as explained, talin recruitment to FAs also results in further PIP2 generation, given that the PI5 kinase PIPKIγ is also recruited and activated at the same time [289,218]. Thus there are elements of both negative and positive feedback at work here, each directly or indirectly involving PIP2.

Processes that incorporate such complex interdependencies tend not to be readily amenable to intuitive analysis, which is why mathematical models can be so useful in studying them. Yet none appear to have been attempted previously, almost certainly none incorporating both PIP2-driven assembly and calpain-driven disassembly.

4.1.3.2 What is new about this PI/calpain cleavage model

The model described here look at the interplay of PIP2 and calpain cleavage, attempting to gain insights into how each affects the overall level of FAs, as well as the timescale over which they might operate in real-world conditions.

For these purposes, two differential-equation modelling tools have been used here, namely COPASI and Virtual Cell. This is primarily because, as we saw, Virtual Cell allows spatial/PDE modelling, where COPASI is confined to compartmental/ODE modelling. However, in addition, as we also saw in the previous chapter, Virtual Cell allows compartmental models. So why not just use Virtual Cell for both compartmental and spatial models, in the same way it was in the previous chapter?

The main reason is because of the much greater complexity of this extended FA model compared to the very simple version in the previous chapter. As stated previously, COPASI provides a much richer set of analysis tools, particularly in terms of plotting options. As will be seen, these will prove particularly helpful in attempting to analyse causative factors of model behaviours, given the complex interdependencies described above.

Given that this model includes IP3-mediated calcium signalling, requiring the ER as an additional, internal calcium-sequestering compartment, there seems a particular benefit to be gained from implementing an additional spatial model. As before, this will require using Virtual Cell and, as seen in Chapter 3, it will make sense to implement at first as a compartmental/ODE model in Virtual Cell, not using the COPASI compartmental model for reasons explained in section 3.5.2.

But also, as we as we have seen, there are differences in the way COPASI and Virtual Cell model membranes and in other ways, as will be explained later. These differences have the potential to provide an interesting direct contrast between the two compartmental models produced by these two modelling environments, as well as providing a further demonstration of the sort of challenges faced in modelling cell biology.

4.2 Methods

4.2.1 Analysis methods used

In terms of analytical methods used (i.e. plots, parameter scans, etc) the same broadly applies here as was written in the corresponding Methods sections in 2.2 and 3.2 of the previous two chapters. Sensitivity analysis was tried on a few occasions and then abandoned as the results in all cases proved contradictory and frequently erroneous. Much of this was undoubtedly due to the highly complex nature of the model, with all its interdependencies, negative feedback effects and similar. Some of it may have been because of the difficulties of specifying the sensitivities required in sufficient detail, again because of the complexity of the model.

For similar reasons parameter estimation was not attempted, as it was felt that similar issues were likely to arise, making any results from the parameter estimation highly unreliable. In any case, without sensitivity analysis working reliably, it would have been difficult to determine which parameter values (including species concentrations) to focus on. This meant that, in such a large model with so many estimated values, there would be so few constraints that any attempt at parameter fitting would be unlikely to prove a useful exercise.

4.2.2 Modelling application version numbers

COPASI: Version 4.22 (Build 170)

Virtual Cell: Version 7.0.0 (Build 9)

4.2.3 System units used

4.2.3.1 COPASI

Time units: Seconds (s)

Quantity units:	Femtomoles (fmol)
Volume units:	Femtolitres (fl)

4.2.3.2 Virtual Cell

Time units:	Seconds (s)
Length units:	µm
Area units:	μm^2
Volume units:	μm^3
Substance units:	$\mu M.\mu m^3$
Membrane Substance units:	molecules
Volume concentration units:	µM
Membrane conc. units:	$molecules.\mu m^{-2}$
Diffusion units:	$\mu m^2.s^{-1}$

4.2.4 Solver settings

4.2.4.1 COPASI

The Deterministic (LSODA) solver was used for all Time Course simulations, with a relative tolerance value of 1e-06 and an absolute tolerance of 1e-12. Maximum internal steps were set at 10,000. Time Course durations were typically of 20,000 seconds, with 250 intervals, and therefore an interval size of 80 seconds.

4.2.4.2 Virtual Cell

In terms of Solver settings, these were as follows. For the ODE/compartmental application the default Combined Stiff Solver (IDA/CVODE) was used, with both absolute and relative error tolerances set to values of 1.0E-9. The maximum step size was set to 1.0 seconds, with simulations run for either 1000 or 10,000 seconds, depending on circumstances. Output settings varied, according to requirements.

For the PDE/2D spatial application, the Fully-Implicit Finite Volume, Regular Grid (Variable Time Step) integrator was used, with an absolute error tolerance of 1.0E-9 and a relative error tolerance of 1.0E-7. The maximum step size was set to 0.1 seconds, with simulations run for 1000 seconds in all cases. The output interval was either 1.0 or 5.0 seconds, according to requirements. The Mesh Size was 101 elements for both X and Y dimensions, giving a total size of 10201 elements and a spatial step of 1.016 μm in both the X and Y planes.

4.3 Model Setup

4.3.1 Compartment volumes and surface areas

4.3.1.1 Lamellipodia cytosolic volume

Lamellipodia dimensions were estimated from the literature, giving a typical height of around 0.15 µm [290,291], a typical protrusion width of 60 µm and a typical depth/diameter of 6 µm [292]. This was then envisaged as a simple rectangular box, with dimensions 60 µm x 6 µm x 0.15 µm, yielding a total volume of 54 μm^3. Ignoring the PM and assuming the internal, liquid, volume to be about 90% of this total (to allow for the actin cytoskeleton and internal organelles, such as the ER), this suggested a lamellipodial/cytosolic volume of approximately 50 μm^3.

4.3.1.2 ER volume

Study of lamellipodia micrographs and calculations of the actin skeleton suggested that in the most of this unavailable volume would be ER, given that lamellipodia appear to be rich in this and it appears to be the only organelle that keeps pace with the lamellipodia front [293]. Thus, at 10% of the lamellipodial cytosolic volume, the ER was assigned a volume of 5 μm^3.

4.3.1.3 Lamellipodia cytosolic area (for 2D spatial models)

Using a similar analogy, two-dimensional lamellipodia were assumed to have a total area of 360 μm^2 (60 µm x 6 µm), yielding a cytosolic area of ~325 μm^2 and an ER area of ~35 μm^2. These were the areas used for these compartments in 2D Virtual Cell models.

4.3.1.4 Lamellipodia PM surface area

Virtual Cell models membranes as two-dimensional surfaces, and therefore requires a surface area (rather than a volume) to be specified. For the purposes of compartmental/ODE models, this was derived from the rectangular box model of the lamellipodia described above, adding together the surface area of 5 of the 6 panels, but excluding the one at the rear, i.e. where the lamellipodia protrudes from the rest of the cell body. This calculation gave a lamellipodial surface area value of approximately 730 μm^2.

4.3.1.5 ER membrane surface area

However, given that the ER observed in lamellipodia appears to be predominantly tubular[294,189,295,296], using the same box analogy to calculate the ER surface area didn't appear valid. Rather, since studies suggest that the ER surface area, in general, may be 10-20 times larger than the cell/PM surface area [294,189], for compartmental/ODE models a value of 2000 μm^2 was allocated to the former, i.e. roughly 3 times that of the lamellipodial surface area. This was based on the assumption that only a small proportion of ER could realistically extend into a developing lamellipodia in the time available.

4.3.1.6 PM and ER membrane lengths (for 2D spatial models)

Virtual Cell membranes are modelled as one-dimensional lengths in 2D spatial models. These were derived from the corresponding surface areas used in the Virtual Cell compartmental models, after dividing by the lamellipodial height (0.15 µm), giving lengths of approximately 4750 µm for the PM and 13,333 µm for the ER membrane.

For the 2D spatial/PDE Virtual Cell model it proved impractical to directly match these membrane lengths to the above-mentioned compartment volumes, using the geometry specification tools available. Therefore, to start with, for the 2D model, only the lamellipodial/cytosolic and ER areas were as specified above (i.e. 325 $\mu m^2$and 35 μm^2, respectively). Whatever membrane lengths were subsequently calculated by Virtual Cell, according to the specification method used, were then initially accepted.

4.3.1.7 PM and ER membrane volumes

For COPASI models (which model membranes as volumes) the membrane surface areas shown above were multiplied by 7.5 nm, the average thickness of cell membranes [189], to give membrane volumes of 5.475 μm^3 (5.475 fl) and 15 μm^3 (15 fl) for the lamellipodial PM and ER, respectively.

4.3.1.8 Extracellular volume

The extracellular volume was assumed in all cases, and varied widely between simulations. This can be justified by the fact that the cell was modelled as stationary, and because the EC compartment in these models contained only fibronectin, which was assumed to have a near-zero diffusion coefficient. Furthermore, the extracellular volume is often effectively indeterminate in the context of any specific cell. In any case, given that fibronectin was supplied massively in excess of available membrane Integrins, it was thought highly unlikely that changes in EC compartment volumes would have any significant effect on results. Subsequent experiments showed this to be the case.

4.3.1.9 Summary of compartment dimensions

For a summary of compartment volumes, areas and lengths used in the two modelling tools, see Table 4.1.

Compartment	COPASI	Virtual Cell	
		ODE	2D
Extracellular	100 fl (100 μm^3)	100 fl	~540 μm^2
Plasma membrane	5.475 fl (5.475 μm^3)	730 μm^2	~4725 µm
Cytosolic	50 fl (50 μm^3)	50 fl	325 μm^2
ER membrane	15 fl (15 μm^3)	2000 μm^2	13,333 µm
Endoplasmic reticulum	5 fl (5 μm^3)	5 fl	~35 μm^2

Table 4.1. Volumes, areas and lengths used for the five model compartments in the COPASI and Virtual Cell modelling tools. COPASI uses volumes for all compartments, liquid or membrane. Areas are used for Virtual Cell ODE model membranes and 2D model membranes, whilst lengths are used for 2D model membranes. * Owing to the challenges of specifying 2D model geometries of the desired size, using the available geometry specification methods in Virtual Cell, incorrect sizes had to be adopted and then later mapped to the desired values at the simulation specification stage, as explained in the text.

4.3.2 Initial concentrations

Since the model was based on a HeLa cell, wherever possible initial protein copy numbers were taken from a published HeLa cell quantitative proteome [182]. These copy numbers (per cell) were then divided by the relevant HeLa compartment volume or, for species appearing in Virtual Cell membranes, the relevant HeLa membrane surface area, to obtain the necessary protein concentration.

So, for example, the copy number for vinculin (approximately 98,000 in the quantitative proteome) was divided into a volume of 1500 μm^3 (1500 fl), which was calculated as the effective HeLa cytosolic volume (50% of the calculated gross volume), to give a cytosolic concentration of approximately 0.34 µM. This was the concentration specified for both the Virtual Cell and COPASI lamellipodia models and, since the lamellipodial cytosolic area was assigned a volume of 50 μm^3 (50 fl) in both models, this yielded a copy number of approximately 10,300 in both cases.

A similar method was used for proteins from other cellular compartments. However, differences between the way Virtual Cell and COPASI model membranes meant that a different approach was needed for membrane-resident proteins. In the case of Virtual Cell, this was fairly

straightforward. First, the copy number for the relevant protein was obtained from the quantitative proteome and this was divided into the surface area which had been calculated for the corresponding HeLa cell membrane, to give an area density for that protein. This was then the value specified for the protein in the Virtual Cell model.

So, for instance, a copy number of 30,000 was obtained for integrin from the quantitative proteome. (This was based on the value for the α5 subunit, as the β1 subunit, associated with α5 in the fibronectin receptor, is more numerous). This was then divided into 1600 μm^2, the surface area calculated for HeLa cell PM, to give a density of 18.75 molecules per μm^2, and this was the value specified for the protein in the Virtual Cell model. Multiplied by 730 μm^2, the lamellipodial PM surface area calculated for this model (see below), gives a copy number of approximately 13,700 for the model.

As COPASI models membranes as volumes (like liquid compartments) it was first necessary to calculate a volume for each of these in a typical HeLa cell. This was done by multiplying the surface area calculated for the relevant membrane by the average thickness for lipid bilayer membranes. The copy number obtained for the protein in question was then divided into this volume to get a membrane concentration, which was a value specified for the COPASI lamellipodia model.

So, for integrin in the example above, the copy number of 30,000 was divided by 12 μm^3 (12 fl), the volume calculated for the HeLa cell PM, to give a concentration of approximately 4.15 μM. This is the concentration value for integrin specified in the COPASI model. Multiplying it by 5.475 μm^3, the volume calculated for the COPASI lamellipodia PM (see below), gives a model particle number of approximately 13,700 for the COPASI lamellipodia model, i.e. the same value as in the Virtual Cell model. Similar calculations were performed for other membrane proteins.

Phosphoinositide concentrations were estimated from relative levels reported in a study by Xu and colleagues [Xu et al., 2003]. These were then apportioned to the PM in the Virtual Cell and

COPASI models, using the same methods described above. Fibronectin was given a concentration of 10 µM, so as to make it in excess of available integrins.

A search of the literature for ER calcium concentrations yielded a wide range of values, broadly in line with the 12 µM- 2 mM range suggested by the literature [297]. Consequently an intermediate value of 250 µM was chosen, based on a figure quoted for the ER-like dense tubular system in platelets [298]. Since the literature suggested a very low (nanomolar) value for resting cytosolic calcium, a reported value of 70 nM [299] was chosen as representative. (In fact, subsequent ODE and PDE models, using a zero value, stop showed only moderate changes in their outputs.)

4.3.3 Diffusion rates

As regards diffusion coefficients, in almost all cases the default Virtual Cell rates were accepted. That is to say, for all membrane species a diffusion coefficient of 0.1 $\mu m^2.s^{-1}$ was specified, and for almost all non-membrane species a rate of 10.0 $\mu m^2.s^{-1}$. The exceptions were Fibronectin, which was given a near-zero diffusion coefficient, given its near-stationary properties. Also, calcium ions (both cytosolic and ER) and IP3 (cytosolic) where, for some simulations, published figures of, respectively,530 $\mu m^2.s^{-1}$ and 280 $\mu m^2.s^{-1}$ were used[300,301]. This was to examine what effects these faster rates would have on spatial model behaviour.

4.3.4 Summary of initial concentrations and diffusion rates used in all models

A summary of the initial concentrations of all species used in the Virtual Cell and COPASI models, together with Virtual Cell spatial application diffusion rates, is shown in Tables 4.2 and 4.3 below.

Species	Compartment	Quantity		Diffusion rate (spatial apps only)***
		Copy # per HeLa cell/ lamellipodia	Concentration	
Ca2+_ER	ER	N/A	250 µM †	10 $\mu m^2.s^{-1}$
Ca2+_cyto	Lamellipodia/Cytosol	N/A	0 .07 µM † §	10 $\mu m^2.s^{-1}$****
Calpain	Lamellipodia/Cytosol	10,000 / 333 *	0.011 µM	10 $\mu m^2.s^{-1}$
Calpain_PM	PM	0	0 molecules.μm^{-2} ††	0.1 $\mu m^2.s^{-1}$
Calpain_act	PM	0	0 molecules.μm^{-2}	0.1 $\mu m^2.s^{-1}$
Cleaved_FAs	PM	0	0 molecules.μm^{-2}	0.1 $\mu m^2.s^{-1}$
DAG	PM	0	0 molecules.μm^{-2}	0.1 $\mu m^2.s^{-1}$
FAK	Lamellipodia/Cytosol	14,000 / 467 *	0.015 µM	10 $\mu m^2.s^{-1}$
FAK_PM	PM	0	0 molecules.μm^{-2}	0.1 $\mu m^2.s^{-1}$
Fibronectin	EC	N/A	10 µM †††	10 $\mu m^2.s^{-1}$
Int_Tal_PIP2	PM	0	0 molecules.μm^{-2}	0.1 $\mu m^2.s^{-1}$
Int_Tal_PI5K _PIP2	PM	0	0 molecules.μm^{-2}	0.1 $\mu m^2.s^{-1}$
Integrin (α5β1)	PM	30,000 / 13,688 *	18.8 molecules.μm^{-2}	0.1 $\mu m^2.s^{-1}$
IntegrinSFKs	PM	0	0 molecules.μm^{-2}	0.1 $\mu m^2.s^{-1}$
IP2	Lamellipodia/Cytosol	0	0 µM	10 $\mu m^2.s^{-1}$
IP3	Lamellipodia/Cytosol	0	0 µM	10 $\mu m^2.s^{-1}$ ****
IP3R	ER membrane	2600 / 220 *	0.11 molecules.μm^{-2}	0.1 $\mu m^2.s^{-1}$
IP3R_act	ER membrane	0	0 molecules.μm^{-2}	0.1 $\mu m^2.s^{-1}$
IP3R_Ca2+	ER membrane	0	0 molecules.μm^{-2}	0.1 $\mu m^2.s^{-1}$
ITP_Lig	PM	0	0 molecules.μm^{-2}	0.1 $\mu m^2.s^{-1}$
ITPPax_Lig	PM	0	0 molecules.μm^{-2}	0.1 $\mu m^2.s^{-1}$
ITPPaxFAK_Lig	PM	0	0 molecules.μm^{-2}	0.1 $\mu m^2.s^{-1}$
ITPPaxFAKVinc_Lig	PM	0	0 molecules.μm^{-2}	0.1 $\mu m^2.s^{-1}$
Paxillin	Lamellipodia/Cytosol	98,000 / 3267 *	0.11 µM	10 $\mu m^2.s^{-1}$
Pax_PM	PM	0	0 molecules.μm^{-2}	0.1 $\mu m^2.s^{-1}$
PI	PM	100,000 / 45,625 **	62.5 molecules.μm^{-2}	0.1 $\mu m^2.s^{-1}$

PI3K	Lamellipodia/Cytosol	2,800 / 93 *	0.003 µM	10 $\mu m^2.s^{-1}$
PI3K_PM	PM	0	0 $molecules.\mu m^{-2}$	0.1 $\mu m^2.s^{-1}$
PI3K_act	PM	0	0 $molecules.\mu m^{-2}$	0.1 $\mu m^2.s^{-1}$
PI4K	PM	24,000 / 10,950 *	15 $molecules.\mu m^{-2}$	0.1 $\mu m^2.s^{-1}$
PI5K	Lamellipodia/Cytosol	15,000 / 500 *	0.017 µM	10 $\mu m^2.s^{-1}$
PI5K_PM	PM	0	0 $molecules.\mu m^{-2}$	0.1 $\mu m^2.s^{-1}$
PI5K_act	PM	0	0 $molecules.\mu m^{-2}$	0.1 $\mu m^2.s^{-1}$
PIP	PM	2,400 / 1095 **	1.5 $molecules.\mu m^{-2}$	0.1 $\mu m^2.s^{-1}$
PIP2	PM	2,800 / 1278 **	1.75 $molecules.\mu m^{-2}$	0.1 $\mu m^2.s^{-1}$
PIP3	PM	60 / 27 **	0.04 $molecules.\mu m^{-2}$	0.1 $\mu m^2.s^{-1}$
PLCg	Lamellipodia/Cytosol	26,000 / 867 *	0.03 µM	10 $\mu m^2.s^{-1}$
PLCg_act	PM	0	0 $molecules.\mu m^{-2}$	0.1 $\mu m^2.s^{-1}$
PLCg_PM	PM	0	0 $molecules.\mu m^{-2}$	0.1 $\mu m^2.s^{-1}$
SERCA	ER membrane	110,000 / 9196 *	4.6 $molecules.\mu m^{-2}$	0.1 $\mu m^2.s^{-1}$
SERCA_Ca2+	ER membrane	0	0 $molecules.\mu m^{-2}$	0.1 $\mu m^2.s^{-1}$
SFKs	PM	43,000 / 19618 *	26.9 $molecules.\mu m^{-2}$	0.1 $\mu m^2.s^{-1}$
Talin	Lamellipodia/Cytosol	98,000 / 3267 *	0.11 µM	10 $\mu m^2.s^{-1}$
Talin_PI5K	Lamellipodia/Cytosol	0	0 µM	10 $\mu m^2.s^{-1}$
Talin_PI5K _PIP2	PM	0	0 $molecules.\mu m^{-2}$	0.1 $\mu m^2.s^{-1}$
Talin_PIP2	PM	0	0 $molecules.\mu m^{-2}$	0.1 $\mu m^2.s^{-1}$
Vinculin	Lamellipodia/Cytosol	310,000 / 10,333 *	0.34 µM	10 $\mu m^2.s^{-1}$
Vinc_PM	PM	0	0 µM	0.1 $\mu m^2.s^{-1}$

Table 4.2. Initial concentrations for Virtual Cell Lamellipodia model. Particle numbers show two values (except when zero or for calcium ions,), with the first denoting the copy number seen in HeLa cells and the second denoting the particle number specified in the Virtual cell model. Key: * = Based on copy numbers seen in a HeLa cell quantitative proteome: N. Nagaraj, J. R. Wisniewski, T. Geiger, J. Cox, M. Kircher, J. Kelso, S. Pääbo, and M. Mann, 'Deep proteome and transcriptome

mapping of a human cancer cell line', *Mol. Syst. Biol.*, vol. 7, p. 548, 2011.Corresponding lamellipodia particle numbers were derived by dividing whole-cell copy numbers by the appropriate HeLa cell compartment volume or membrane surface area (1500 μm^3 for the cytosol, 1600 μm^2 for the PM and 24000 μm^2 for the ER membrane). These were then re-multiplied by the calculated volume and surface areas for the relevant compartment/membranes in the modelled lamellipodia. ** = Assumption. Whole-cell copy number values are based on relative levels seen in cells (where the proportions of PI, PIP and PIP2 are 100, 2.4 and 2.8, respectively.) See Xu, Chang, James Watras, and Leslie M. Loew. 'Kinetic Analysis of Receptor-Activated Phosphoinositide Turnover.' The Journal of Cell Biology 161, no. 4 (26 May 2003): 779–91. doi:10.1083/jcb.200301. Corresponding lamellipodia particle numbers calculated as explained above for lamellipodia PM. *** = In almost all Simulations the default diffusion coefficients supplied by Virtual Cell (i.e. 0.1 µm2.s-1 for membrane compartments and 10 µm2.s-1 for aqueous compartments) have been accepted. **** = In some Simulations the following real-world diffusion coefficients have been used, in place of the default values, as obtained from the literature: Calcium_cyto: 530 µm2.s-1, IP3: 280 µm2.s-1. See Dickinson, G.D., Ellefsen, K.L., Dawson, S.P., Pearson, J.E., Parker, I., 2016. Hindered cytoplasmic diffusion of inositol trisphosphate restricts its cellular range of action. Sci. Signal. 9, ra108.https://doi.org/10.1126/scisignal.aag1625; Donahue, B., Abercrombie, R., 1988. Free diffusion coefficient of ionic calcium in cytoplasm. Cell Calcium 8, 437–48. https://doi.org/10.1016/0143-4160(87)90027-3. † = Calcium concentrations derived from literature for platelets: Rink, T.J., 1988. Cytosolic calcium in platelet activation. Experientia44, 97–100. Sage, S.O., Pugh, N., Mason, M.J., Harper, A.G.S., 2011. Monitoring the intracellular store Ca2+ concentration in agonist-stimulated, intact human platelets by using Fluo-5N. J. Thromb. Haemost. JTH 9, 540–551. †† = All membrane-resident species have concentrations in molecules per unit area, in accordance with Virtual Cell's 2D modelling of membranes. These were calculated as explained above. ††† = Fibronectin concentration assumed, so as to be in excess. §= 0 µM used in some Applications/Simulations.

Species	Compartment	Quantity	
		Copy # per HeLa cell / Lamellipodia	Concentration †
Ca2+_ER	ER	N/A	250 µM ††
Ca2+_cyt	Lamellipodia/Cytosol	N/A	0 .07 µM ††
Calpain	Lamellipodia/Cytosol	10,000 / 333 *	0.011 µM
Calpain_PM	PM	0	0 µM
Calpain_act	PM	0	0 µM
Cleaved_FAs	PM	0	0 µM
DAG	PM	0	0 µM
FAK	Lamellipodia/Cytosol	14,000 / 467 *	0.015 µM
FAK_PM	PM	0	0 µM
FAKPax_PM	PM	0	0 µM
Fibronectin	EC	N/A	10 µM †††
Int_Tal_PI5K_PIP2	PM	0	0 µM
Integrin (α5β1)	PM	30,000 / 13,669 *	4.15 µM
IP2	Lamellipodia/Cytosol	0	0 µM
IP3	Lamellipodia/Cytosol	0	0 µM
IP3R	ER membrane	2600 / 220 *	0.024 µM
IP3R_act	ER membrane	0	0 µM
IP3R_Ca2+	ER membrane	0	0 µM
ITP_Lig	PM	0	0 µM
ITPPax_Lig	PM	0	0 µM
ITPPaxSrc_Lig	PM	0	0 µM
ITPPaxVinc_Lig	PM	0	0 µM
Paxillin	Lamellipodia/Cytosol	98,000 / 3267 *	0.11 µM
PI	PM	100,000 / 45,625**	13.8 µM
PI3K	Lamellipodia/Cytosol	2,800 / 93 *	0.003 µM
PI3K_PM	PM	0	0 µM
PI3K_act	PM	0	0 µM

PI4K	PM	24,000 / 10,950 *	3.32 µM
PI5K	Lamellipodia/Cytosol	15,000 / 500 *	0.017 µM
PI5K_PM	PM	0	0 µM
PI5K_act	PM	0	0 µM
PIP	PM	2,400 / 1095**	0.33 µM
PIP2	PM	2,800 / 1278**	0.39 µM
PIP3	PM	60 / 27**	0.008 µM
PLCγ	Lamellipodia/Cytosol	26,000 / 867 *	0.029 µM
PLCγ_act	PM	0	0 µM
PLCγ_PM	PM	0	0 µM
SERCA	ER membrane	110,000 / 9196 *	1.02 µM
SERCA_Calc	ER membrane	0	0 µM
SFKs	PM	43,000 / 19618 *	5.95 µM
Talin	Lamellipodia/Cytosol	98,000 / 3267 *	0.11 µM
Talin_PI5K	Lamellipodia/Cytosol	0	0 µM
Talin_PIP2	PM	0	0 µM
Vinculin	Lamellipodia/Cytosol	310,000 / 10,333 *	0.34 µM
Vinc_PM	PM	0	0 µM

Table 4.3. Initial concentrations for COPASI model. Particle numbers show two values, with the first denoting the copy number seen in HeLa cells and the second denoting the particle number specified in the COPASI model, calculated as explained in text. Key: * = Based on copy numbers seen in a HeLa cell quantitative proteome: N. Nagaraj, J. R. Wisniewski, T. Geiger, J. Cox, M. Kircher, J. Kelso, S. Pääbo, and M. Mann, 'Deep proteome and transcriptome mapping of a human cancer cell line', *Mol. Syst. Biol.*, vol. 7, p. 548, 2011.** = Assumption. Whole-cell copy number values are based on relative levels seen in cells (where the proportions of PI, PIP and PIP2 are 100, 2.4 and 2.8, respectively.) See Xu, Chang, James Watras, and Leslie M. Loew. 'Kinetic Analysis of Receptor-Activated Phosphoinositide Turnover.' The Journal of Cell Biology 161, no. 4 (26 May 2003): 779–91. doi:10.1083/jcb.200301. Corresponding lamellipodia particle numbers calculated as explained in text. † = With the exception of calcium ions and Fibronectin, all listed initial concentrations were automatically derived by COPASI from the entered lamellipodia particle

numbers (i.e. by multiplying the particle numbers by the appropriate compartment volume). †† = Calcium concentrations derived from literature for platelets: Rink, T.J., 1988. Cytosolic calcium in platelet activation. Experientia44, 97–100. https://doi.org/10.1007/BF01952188; Sage, S.O., Pugh, N., Mason, M.J., Harper, A.G.S., 2011. Monitoring the intracellular store Ca2+ concentration in agonist-stimulated, intact human platelets by using Fluo-5N. J. Thromb. Haemost. JTH 9, 540–551. https://doi.org/10.1111/j.1538-7836.2010.04159. (Particle numbers automatically derived by COPASI, as explained above.) ††† = Fibronectin concentration assumed, so as to be in excess.

4.3.5 Reactions

4.3.5.1 Reaction summary

The model reactions can be summarised as follows:

- Integrin binds to SFKs to form IntegrinSFKs
- Some cytoplasmic talin binds to PI5K, forming Talin_PI5K complex
- Talin_PI5K is recruited to the PM by PIP2 , to form Talin_PI5K_PIP2 complex
- The remaining talin is recruited to the PM by PIP2 , to form Talin_ PIP2 complex
- Talin_PI5K_PIP2 complex dissociates to form Talin_PIP2 and PI5K_PM
- IntegrinSFKs binds to Talin_PIP2 to form Int_Tall_PIP2 complex, whereby integrin is activated
- Fibronectin binds to Int_Tal_ PIP2 complex to form ITP_Lig complex
- Binding of cytosolic paxillin and PIP2 to form membrane-bound Paxillin_PM
- Binding of cytosolic FAK and PIP2 to form membrane-bound FAK_PM
- Binding of cytosolic vinculin and PIP2 to form membrane-bound Vinc_PM
- Binding of ITP_Lig and Pax_PM complexes to form ITPPax_Lig complex
- Binding of FAK_PM to this ITPPax_Lig complex to form ITPPaxFAK_Lig complex
- Binding of Vinc_PM to ITPPaxFAK_Lig complex to form ITPPaxFAKVinc_Lig complex
- Activation of PI5K_PM by SFK-containing ITP_Lig, ITPPax_Lig , ITPPaxFAK_Lig and ITPPaxFAKVinc_Lig, to form PI5K_act
- Binding of cytosolic calpain to PIP2 to form membrane-bound Calpain_PM
- Binding of cytosolic calcium to Calpain_PM to form activated Calpain_act
- Cleavage by Calpain_act of Int_Tal_PI5K_PIP2, Int_Tal_PIP2, ITP_Lig , ITPPax_Lig , ITPPaxFAK_Lig and ITPPaxFAKVinc_Lig species
- PI3K binding to PIP2 to form PI3K_PM
- Activation of PI3K_PM, by SFK-containing ITP_Lig, ITPPax_Lig , ITPPaxFAK_Lig and ITPPaxFAKVinc_Lig, to form PI3K_act
- Conversion of PI to PIP by membrane-bound PI4K
- Conversion of PIP to PIP2 by PI5K_act
- Conversion of PIP2 to PIP3 by PI3K_act
- Binding of PIP3 to PLCγ to form PLCγ_PM
- Activation of PLCγ_PM, by SFK-containing ITP_Lig, ITPPax_Lig , ITPPaxFAK_Lig and ITPPaxFAKVinc_Lig, to form PLCγ_act
- Hydrolysis of PIP2 to IP3 and DAG by PLCγ_act
- Binding of cytosolic IP3 to ER membrane-bound IP3R to form activated IP3R_act, followed by:
    - Binding of ER calcium to IP3R_act to form IP3R_Calc
    - Release of calcium by IP3R_Calc into the cytosol, yielding Calcium_cyt and IP3R_act

- Binding of cytosolic calcium by ER membrane-bound SERCA to form SERCA_Calc
- Release of calcium by SERCA_Calc into the ER, yielding Calcium_ER and SERCA
- Spontaneous conversion of IP3 to IP2
- Spontaneous conversion of IP2 to PIP

See Figure 4.1, for a diagrammatic depiction of this reaction scheme, and Table 4.4 for a listing of all these reactions, together with their rate laws and corresponding rate constant values.

4.3.5.2 Reaction diagram

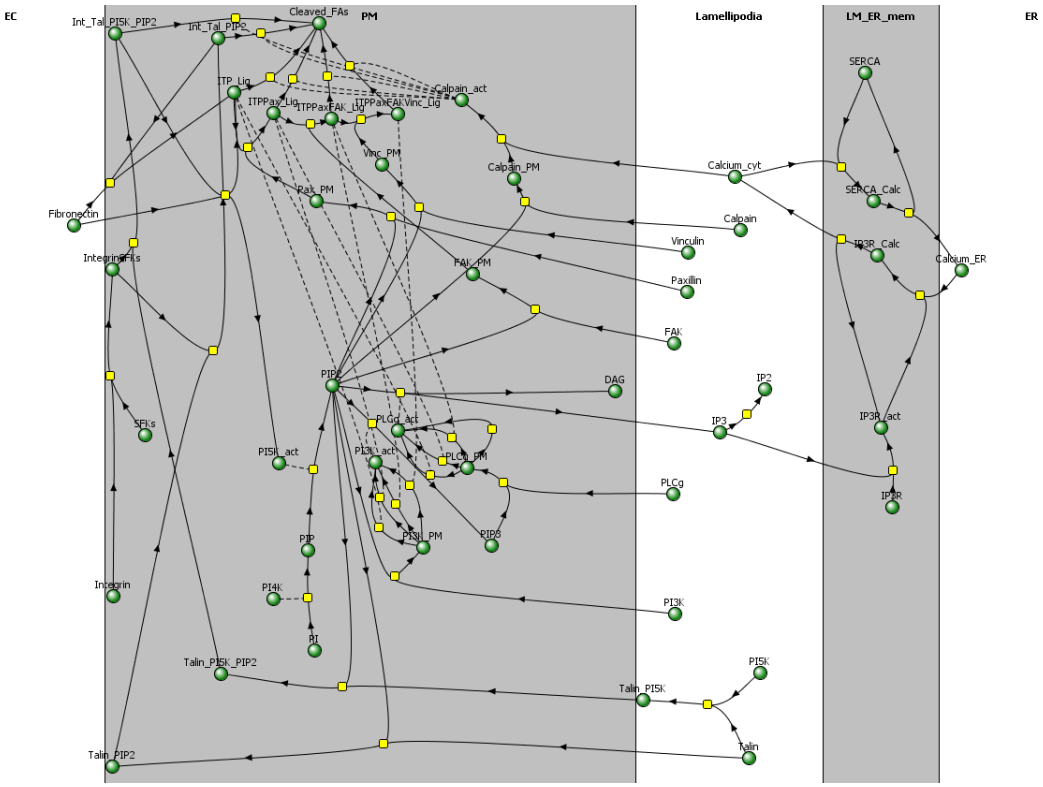

Figure 4.1. Reaction Diagram for COPASI and Virtual Cell models

4.3.5.3 Reaction rate constants

In terms of reaction parameters, all of the enzyme reactions were modelled as irreversible Michaelis-Menten, using reaction parameters taken from the literature, wherever available. The majority of binding reactions were modelled as mass action, in many cases using assumed values, these being 10000 M^{-1} s^{-1} (0 .01 μM^{-1} s^{-1}) for second-order reactions and 0.01 s^{-1} for first-order. (These default values all involved PM-resident, mostly FA-related proteins.)

As in Chapter 3, in order to meet the requirements of Virtual Cell for intra-membrane reactions, it was often required to convert volumetric rate constant values (i.e. values expressed in molar units of measurement) into their areametric equivalents. For instance, the default second-order rate constant value of 10000 M^{-1} s^{-1}, mentioned above, was converted to 2.21E-03 $\mu m^2.s^{-1}.molecules^{-1}$, after first dividing by Avogadro's number, and then by 1000 times the thickness of the membrane.

This same conversion was applied to KM values for membrane-resident enzymatic reactions, as well as to some other second-order mass action reactions rate constants.

The primary reason for using assumed values for many PM/FA-related binding reactions was the general absence in the literature of any reliable data for these reactions. Where any data was available at all this was typically limited to KD values. But, of course, these are simply the quotient of the binding off-rate divided by the on-rate, masking the actual value of either rate. ODE/PDE modelling requires explicit on- and off-rate values to be provided in such instances (i.e. modelling them as reversible reactions). But, more often than not, experiments showed that choosing ad hoc values was not acceptable. Different permutations of on- and off-rates for the same KD value were associated with substantially different model behaviour in many cases. Therefore the use of KD values to constrain binding reaction rates was abandoned in favour of using the same, conservative, default rates.

Another factor in deciding to adopt default binding rates was that some of the binding reactions in the model were simplified one-step abbreviations of more complex multistep reactions, often requiring prior phosphorylation or some other post-translational modification of one or more binding sites. In one or two other instances (e.g. paxillin binding to FAs) the actual binding partner is not known with any confidence, let alone the binding site. It was thought that any value for such a combined reaction rate, even in the unlikely event that it was available, would, in all probability, be of questionable merit.

The actual default binding on- and off-rate values were chosen to fall comfortably within the reported range for binding reactions in general, of 10^{2}-10^{8} M^{-1} s^{-1} for on-rates and 10^{-6}-10^{1} s^{-1} for off-rates [302]. More specifically, the chosen off-rate also falls well within the range of values (0.009-1 s^{-1}) reported from FRAP experiments for protein-protein binding of core FA proteins such as talin, vinculin, FAK and zyxin [228].

For instance, for binding of the vinculin-PIP2 complex (Vinc_PM) to the existing FA (ITPPaxFAK_Lig), the forward rate was 10,000 M^{-1} s^{-1} for the COPASI reaction and 2.21E-03 μm^2 s^{-1} $molecules^{-1}$ for the Virtual Cell reaction (using the conversion method described above). As the reverse rate didn't need converting for the equivalent Virtual Cell reaction, the rate in both cases was 0.01 s^{-1}.

For the sake of consistency it was decided to extend this policy (of adopting default values) to all binding reactions involving PM-resident proteins, even where published alternatives were available. The one noticeable exception was for the binding reaction between PM-resident talin and PI5K, where the rate constant values used in the Chapter 3 model were retained. This was because the reverse reaction was second- rather than first-order, unlike all the other cases mentioned. It was also thought beneficial to retain some consistency with that model, for comparison purposes. For the same reason the rate constant values for the fibronectin-Int_Tal_PIP2 binding reaction were retained from the Chapter 3 model.

Other than the above-mentioned reactions, wherever possible, the relevant rate constants were derived from the literature, as described in the Key description below Table 4.4.

Activation of PI3K by SFK-containing integrin complexes (ITP_Lig, ITPPax_Lig, ITPPaxFAK_Lig and ITPPaxFAKVinc_Lig) was modelled as irreversible Michaelis-Menten, using assumed values. For both PI3K and PLCγ activation the KM value was 1 mM for COPASI, converted to 4520 $molecules.\mu m^{-2}$ for Virtual Cell, while the kcat value was 1 s^{-1} in both cases. These values were chosen as being reasonably conservative. PLCγ activation by the same SFK-containing integrin complexes used the same kcat value, but a relevant KM value of 3 μM was obtained from the literature. Relevant kcat and KM were obtained from the literature for the corresponding PI5K activation reactions.

IP3R reactions used assumed values in all cases, as the intention here was to produce a rapid calcium release from the ER in response to IP3 stimulation, using as simple a model as possible.

This was to avoid adding additional complexity to an already complex model. To this end, it was felt easier to model IP3R as a pump, rather than as an ion channel, which is its true mechanism of action. For these reasons, assumed values, chosen to produce the required response, were felt to be justified. Similar considerations applied to the SERCA pump reactions, which return calcium to the ER.

4.3.5.4 Summary of model reactions and rate constant values

Reaction	Description	Rate law	COPASI Parameter Values	VCell Parameter Values
Ca2+_ER + IP3R_act -> IP3R_ Ca2+	ER calcium loading of activated IP3R	Mass action (irreversible)	Kf = 10000 M^{-1} s^{-1} *	
Calpain_PM + Ca2+_cyto-> Calpain_act	Calpain activation by cytosolic calcium	Mass action (reversible)	Kf = 10000 M^{-1} s^{-1} Kr = 0.01 s^{-1} **	
FAK + PIP2 -> FAK_PM	Recruitment of FAK to PM by PIP2	Mass action (reversible)	Kf = 10000 M^{-1} s^{-1} Kr = 0.01 s^{-1} **	
Fibronectin + Int_Tal_PIP2 -> ITP_Lig + PI5K_act	Fibronectin-FA ligation	Mass action (reversible)	Kf = 770,000 M^{-1} s^{-1} Kr = 0.1 s^{-1}***	
Integrin + SFKs -> IntegrinSFKs	Integrin-SFK binding	Mass action (reversible)	Kf = 100000 M^{-1} s^{-1} Kr = 0.2 s^{-1}****	Kf = 2.21E-02 $\mu m^2.s^{-1}$. $molecules^{-1}$ Kr = 0.2 s^{-1} ****
IP3 -> IP2	IP3 hydrolysis to IP2	Mass action (irreversible)	Kf = 1e-06 s^{-1} *	
IP3R + IP3 -> IP3R_act	Activation of IP3R by IP3	Mass action (reversible)	Kf = 10000 M^{-1} s^{-1} Kr = 0.00001 s^{-1} *	
IP3R_ Ca2+ -> IP3R_act + Ca2+_cyto	Calcium release into cytoplasm by IP3R	Mass action (irreversible)	Kf = 10 s^{-1} *	
ITPPaxFAK_Lig -> Cleaved_FAs; Calpain_act	Calpain cleavage of FA (Paxillin- and FAK-containing)	Michaelis-Menten	kcat = 0.25 s^{-1} Km = 3 µM *****	kcat = 0.25 s^{-1} Km= 13.5 $molecules.\mu m^{-2}$ *****
ITPPaxFAKVinc_Lig -> Cleaved_FAs; Calpain_act	Calpain cleavage of FA (Vinculin -containing)	Michaelis-Menten	kcat = 0.25 s^{-1} Km = 3 µM *****	kcat = 0.25 s^{-1} Km= 13.5 $molecules.\mu m^{-2}$ *****

ITPPaxFAK_Lig + Vinc_PM ->ITPPaxFAKVinc_Lig	Recruitment of membrane-bound Vinculin to FA	Mass action (reversible)	Kf = 10000 M^{-1} s^{-1} Kr = 0.01 s^{-1} **	Kf = 2.21E-03 $\mu m^2.s^{-1}$. $molecules^{-1}$ Kr = 0.01 s^{-1} **
ITP_Lig + Pax_PM -> ITPPax_Lig	Recruitment of membrane-bound Paxillin to FA	Mass action (reversible)	Kf = 10000 M^{-1} s^{-1} Kr = 0.01 s^{-1}**	Kf = 2.21E-03 $\mu m^2.s^{-1}$. $molecules^{-1}$ Kr = 0.01 s^{-1} **
ITPPax_Lig + FAK_PM -> ITPPaxFAK_Lig	Recruitment of membrane-bound FAK to FA	Mass action (reversible)	Kf = 10000 M^{-1} s^{-1} Kr = 0.01 s^{-1}**	Kf = 2.21E-03 $\mu m^2.s^{-1}$. $molecules^{-1}$ Kr = 0.01 s^{-1} **
ITP_Lig -> Cleaved_FAs; Calpain_act	Calpain cleavage of FA (Talin only)	Michaelis-Menten	kcat = 0.25 s^{-1} Km = 3 µM *****	kcat = 0.25 s^{-1} Km= 13.5 $molecules.\mu m^{-2}$ *****
ITPPax_Lig -> Cleaved_FAs; Calpain_act	Calpain cleavage of FA (Paxillin-containing)	Michaelis-Menten	kcat = 0.25 s^{-1} Km = 3 µM *****	kcat = 0.25 s^{-1} Km= 13.5 $molecules.\mu m^{-2}$ *****
Int_Tal_PI5K_PIP2 -> Cleaved_FAs; Calpain_act	Calpain cleavage of Int_Tal_PI5K_PIP2	Michaelis-Menten	kcat = 0.25 s^{-1} Km = 3 µM *****	kcat = 0.25 s^{-1} Km= 13.5 $molecules.\mu m^{-2}$ *****
Int_Tal_PIP2 -> Cleaved_FAs; Calpain_act	Calpain cleavage of Int_Tal_PI5K_PIP2	Michaelis-Menten	kcat = 0.25 s^{-1} Km = 3 µM *****	kcat = 0.25 s^{-1} Km= 13.5 $molecules.\mu m^{-2}$ *****
Paxillin + FAK_PM -> FAKPax_PM	Paxillin binding to FAK and recruitment	Mass action (reversible)	Kf = 10000 M^{-1} s^{-1} Kr = 0.01 s^{-1} **	
PI -> PIP; PI4K	PI phosphorylation to PIP by PI4K	Michaelis-Menten	kcat = 2.77 s^{-1} Km = 1.6E-5 M ******	kcat = 2.77 s^{-1} Km= 72.3 $molecules.\mu m^{-2}$ ******
PI3K_PM ->PI3K_act; ITP_Lig	PI3K activation by FA-associated Src (ITP_Lig)	Michaelis-Menten	kcat = 1 s^{-1} * Km = 1 mM	kcat = 1 s^{-1} Km= 4520 molecules. μm^{-2} *
PI3K_PM ->PI3K_act; ITPPax_Lig	PI3K activation by FA-associated Src (ITPPax_Lig)	Michaelis-Menten	kcat = 1 s^{-1} * Km = 1 mM	kcat = 1 s^{-1} Km= 4520 molecules. μm^{-2} *
PI3K_PM ->PI3K_act; ITPPaxFAK_Lig	PI3K activation by FA-associated Src (ITPPaxFAK_Lig)	Michaelis-Menten	kcat = 1 s^{-1} * Km = 1 mM	kcat = 1 s^{-1} Km= 4520 molecules. μm^{-2} *

PI3K_PM ->PI3K_act; ITPPaxFAKVinc_Lig	PI3K activation by FA-associated Src ITPPaxFAKVinc_Lig	Michaelis-Menten	kcat = 1 s^{-1} * Km = 1 mM	kcat = 1 s^{-1} Km= 4520 molecules. μm^{-2} *
PI5K_PM ->PI5K_act; ITP_Lig	PI5K activation by FA-associated Src (ITP_Lig)	Michaelis-Menten	kcat = 10 s-1 KM = 3.0e-05 M †	kcat = 10 s^{-1} Km= 135 molecules. μm^{-2} †
PI5K_PM ->PI5K_act; ITPPax_Lig	PI5K activation by FA-associated Src (ITPPax_Lig)	Michaelis-Menten	kcat = 10 s-1 KM = 3.0e-05 M †	kcat = 10 s^{-1} Km= 135 molecules. μm^{-2} †
PI5K_PM ->PI5K_act; ITPPaxFAK_Lig	PI5K activation by FA-associated Src (ITPPaxFAK_Lig)	Michaelis-Menten	kcat = 10 s-1 KM = 3.0e-05 M †	kcat = 10 s^{-1} Km= 135 molecules. μm^{-2} †
PI5K_PM ->PI5K_act; ITPPaxFAKVinc_Lig	PI5K activation by FA-associated Src (ITPPaxFAKVinc_Lig)	Michaelis-Menten	kcat = 10 s-1 KM = 3.0e-05 M †	kcat = 10 s^{-1} Km= 135 molecules. μm^{-2} †
PIP ->PIP2; PI5K_act	PIP phosphoryl-ation to PIP2 by activated PI5K	Michaelis-Menten	kcat = 1.02 s^{-1} Km = 1.0E-5 M ******	kcat =1.02 s^{-1} Km= 45.2 molecules.μm^{-2}******
PIP2 + Calpain -> Calpain_PM	Recruitment of Calpain to PM by PIP2	Mass action (reversible	Kf = 10000 M^{-1} s^{-1} Kr = 0.01 s^{-1} **	
PIP2 + PI3K -->PI3K_PM	PI3K recruitment to PM by PIP2	Mass action (reversible)	Kf = 10000 M^{-1} s^{-1}** Kr = 0.01 M^{-1} s^{-1}	
PIP2 + Talin_PI5K -> Talin_PI5K_PIP2	Talin_PI5K binding of PIP2 and PM recruitment	Mass action (reversible)	Kf = 10000 M^{-1} s^{-1}** Kr = 0.01 M^{-1} s^{-1}	
PIP2 + PLCγ_act -->IP3 + DAG; PLCγ_act	PIP2 hydrolysis to IP3 and DAG by PLCγ(_act)	Michaelis-Menten	kcat = 10 s^{-1} KM = 9.9 µM ††	kcat = 10 s^{-1} Km= 44.7 molecules.μm^{-2}††
PIP2 ->PIP3; PI3K_act	PIP2 phosphorylation toPIP3 by PI3K(_act)	Michaelis-Menten	kcat = 0.3 s^{-1} Km = 5.0E-5M †††	kcat = 0.3 s^{-1} Km= 226 molecules.μm^{-2}†††
PLCγ + PIP3 -> PLCγ_PM	Recruitment of PLCγ to PM by PIP3	Mass action (reversible)	Kf = 10000 M^{-1} s^{-1} Kr = 0.01 s^{-1} **	
PLCγ_PM -> PLCγ_act; ITP_Lig	PLCγ activation by FA-associated Src (ITP_Lig)	Michaelis-Menten	kcat = 1 s^{-1} Km = 3e-06 M ††††	kcat = 1 s^{-1} Km= 13.5 molecules. μm^{-2}††††
PLCγ_PM -> PLCγ_act; ITPPax_Lig	PLCγ activation by FA-associated Src (ITPPax_Lig)	Michaelis-Menten	kcat = 1 s^{-1} Km = 3e-06 M ††††	kcat = 1 s^{-1} Km= 13.5 molecules. μm^{-2}††††

PLCγ_PM -> PLCγ_act; ITPPaxFAK_Lig	PLCγ activation by FA-associated Src (ITPPaxFAK_Lig)	Michaelis-Menten	kcat = 1 s^{-1} Km = 3e-06 M ††††	kcat = 1 s^{-1} Km= 13.5 molecules. μm^{-2}††††
PLCγ_PM -> PLCγ_act; ITPPaxFAXVinc_Lig	PLCγ activation by FA-associated Src ITPPaxFAKVinc_Lig	Michaelis-Menten	kcat = 1 s^{-1} Km = 3e-06 M ††††	kcat = 1 s^{-1} Km= 13.5 molecules. μm^{-2}††††
SERCA + Ca2+_cyto->SERCA_ Ca2+	Cytosolic calcium loading of SERCA	Mass action (irreversible)	Kf = 0.1M^{-1} s^{-1}*	
SERCA_Ca2+ ->SERCA + Ca2+_ER	Calcium return to ER by SERCA	Mass action (irreversible)	Kf = 0.003 s^{-1} *	
Talin + PI5K = Talin_PI5K	Talin binding of PI5K	Mass action (reversible)	Kf = 4.4e+06 M^{-1} s^{-1} Kr = 27 s^{-1}†††††	
PIP2 + Talin = Talin_PIP2	Talin binding of PIP2	Mass action (reversible)	Kf = 10000 M^{-1} s^{-1} Kr = 0.01 s^{-1} **	
Talin_ PI5K_PIP2 -> Tal_PIP2 + PI5K_PM	Dissociation of Talin_PI5K_PIP2 to Talin_PIP2 + PI5K_PM	Mass action (reversible)	Kf = 1.0 s^{-1} Kr = 450 M^{-1} s^{-1} ‡	Kf = 1.0 s^{-1} Kr = 0.0001 $\mu m^{2}.s^{-1}$. $molecules^{-1}$ ‡
Talin_PIP2 + IntegrinSFKs -> Int_Tal_ PIP2	Binding of membrane-bound Talin_PIP2 to Src-associated Integrins	Mass action (reversible)	Kf = 10000 M^{-1} s^{-1} Kr = 0.01 s^{-1} **	Kf = 2.21E-03 $\mu m^{2}.s^{-1}$. $molecules^{-1}$ Kr = 0.01 s^{-1} **
Vinculin + PIP2 -> Vinc_PM	Recruitment of Vinculin to PM byPIP2	Mass action (reversible)	Kf = 10000 M^{-1} s^{-1} Kr = 0.01 s^{-1} **	

Table 4.4. Reactions for COPASI compartmental and Virtual Cell compartmental and spatial (2D and 3D) models of lamellipodial FA dynamics in HeLa cells

Key:

* = Assumed values, chosen to give overall model responses similar to physiological.
** = Assumed values, chosen to fall comfortably within reported range for FA-related binding reactions, and for binding reactions generally, as described in the above. (For Virtual Cell models the COPASI forward rate of 10,000 M-1 s-1 has been converted in some instances to the equivalent rate of 2.21E-03 µm2.s-1.molecules-1, as required for intra-membrane reactions, by its 2D representation of membranes.)

*** = Based on reported kinetics for α5ß1 binding to fibronectin, with reverse reaction rate constant value (Kr) revised downwards from 49 to 0.1 s^{-1}. (See Chapter 3 Results Table8 and discussion above, it, for explanation. Corresponding Virtual Cell value calculated as explained above.) Source: Elosegui-Artola et al, 2014. Rigidity sensing and adaptation through regulation of integrin types. Nat. Mater. 13, 631–637. https://doi.org/10.1038/nmat3960.
**** = Rough mean average of binding rates for Lyn and Hck to SH3-binding peptide, similar to protein-rich sequence in integrin tails. Source: Moroco et al., 2014. Differential Sensitivity of Src-

Family Kinases to Activation by SH3 Domain Displacement. PLOS ONE, Volume 9, Issue 8. Pages e105629. (Corresponding Virtual Cell value calculated as explained above.)

***** = Median values taken from range of kcat and KM value for calpain-1 cleavage of proteins published in BRENDA Enzyme database. BRENDA - Information on EC 3.4.22.52 - calpain-1: http://www.brenda-enzymes.org/enzyme.php?ecno=3.4.22.52. (Corresponding Virtual Cell value calculated as explained above.)

****** = Taken from Purvis JE, Chatterjee MS, Brass LF, Diamond SL. A molecular signaling model of platelet phosphoinositide and calcium regulation during homeostasis and P2Y1 activation. Blood. 2008; 112(10): 4069-4079. Doi: 10.1182/blood-2008-05-157883. (Table 1). (Corresponding Virtual Cell value calculated as explained above.)

† = Based on values taken from published model, itself derived from the literature. Sources: Geier et al, 2011. A Computational Analysis of the Dynamic Roles of Talin, Dok1, and PIPKI for Integrin Activation. PLoS ONE 6. https://doi.org/10.1371/journal.pone.0024808. Songyang et al,1995. Catalytic specificity of protein-tyrosine kinases is critical for selective signalling. Nature 373: 536–9. (Corresponding Virtual Cell value calculated as explained above.)

†† = Source: Devika NT, Jaffar Ali BM. Analysing calcium dependent and independent regulation of eNOS in endothelium triggered by extracellular signalling events. Mol Biosyst. 2013 Nov; 9(11): 2653-64. doi: 10.1039/c3mb70258h. PMID: 23954998. (Corresponding Virtual Cell value calculated as explained above.)

††† = kcat value 0.3 s-1 derived from published specific activity value (SignalChem PI3K (p110a/p85a) specific activity assay. http://www.signalchem.com/shared_product_sheets/D241-2.pdf). KM value of 5.0E-5 molecules.µm-2 derived from Huang W, Jiang D, Wang X, et al. Kinetic Analysis of PI3K Reactions with Fluorescent PIP2 Derivatives. Anal Bioanal Chem. 2011; 401(6): 1881-1888.doi: 10.1007/s00216-011-5257-z. (Corresponding Virtual Cell value calculated as explained above.)

†††† = kcat value assumed. KM value taken from Rotin et al. Presence of SH2 domains of phospholipase C gamma 1 enhances substrate phosphorylation by increasing the affinity toward the epidermal growth factor receptor. J Biol. Chem. 1992; 267(14): 9678-9683. (Corresponding Virtual Cell value calculated as explained above.)

†††††= Based on experimental data for talin F2-F3 FERM domain binding to PIPKIγ. Source: Barsukovet al, 2003. Phosphatidylinositol Phosphate Kinase Type 1γ and β1-Integrin Cytoplasmic Domain Bind to the Same Region in the Talin FERM Domain. J. Biol. Chem. 278, 31202–31209. https://doi.org/10.1074/jbc.M303850200

‡ = Arbitrary value, chosen to be well within normal range for such reactions, and to give responses similar to physiological. (See discussion in Results section of Chapter 3, and Table 7 in same section. Corresponding Virtual Cell value calculated as explained above.)

4.4 Results

4.4 .1 COPASI model.

4.4.1.1 Version 1

As indicated in Figure 4.2, a number of model species, including major species such as PIP2 and FAs (here defined as ITP_Lig, ITPPax_Lig, ITPPaxFAK_Lig, ITPPaxFAKVinc_Lig) require around 8000 seconds to reach peak levels. This compares with times of between 200 and 600 seconds for similar events to occur in the Virtual Cell compartmental model in Chapter 3, which, of course, has many differences from this model. Nevertheless, as was noted in the Results section of Chapter 3, the timings seen in that model accorded fairly well with the underlying physiological reality, indicating that this COPASI model is running too slowly, compared to the reality it is attempting to model.

More specifically, Figure 4.2 also suggests that levels of FAs are much lower than one would expect (reaching peak levels of only just over 1300 molecules at around 8000 seconds elapsed), given that nearly 13,700 molecules of unbound integrin are available from the outset. This is confirmed by Figure 4.2, which shows that throughout the model run of 20,000 seconds, the vast majority of integrin remains unbound to fibronectin, in the form of the species Integrin and IntegrinSFKs. Other than these, only ITP_Lig and Cleaved_FAs (representing all integrin species that have been cleaved by calpain) exceed 1000 molecules at any point over that time scale.

Initial sensitivity analysis in this case proves unhelpful, suggesting a variety of different and somewhat contradictory influences on FA levels. (Results not shown.) However, a comparison of the PIP2 and Total FA curves in Figure 4.2 suggests that PIP2 levels may play an important role. Certainly, this makes sense, given the reliance of FA assembly on availability of PIP2 (as described in the introduction to this and the previous chapter, and demonstrated by the results of the model in Chapter 3). In other words, the reason for low levels of FAs might be explained by the

relatively low levels of PIP2 generated in the model, as shown in Figure 4.4. This shows that, of around 48,000 molecules of initial PI, the vast majority is converted into PIP or subsequently cleaved into IP3 and DAG, with PIP2 levels only peaking at around 3000 molecules, i.e. around 6% of the total available.

Since IP3 and DAG can only be hydrolysed from PIP2, this suggests that PIP2 hydrolysis is considerably in excess of PIP2 creation. This is confirmed by Figures 4.5-7 below. Respectively, these show particle formation rates for PIP2, IP3 and DAG (Figure 4.4) and for PI5K_act and PLCγ_act (Figure 4.6), together with particle numbers for these last two species (Figure 4.7). However, a logarithmic parameter scan on the PI5K kcat value (for PIP2 formation from PIP) shows that even substantial increases in PIP2 formation (to the point where 43,000 molecules of PIP2 are formed from the 48,000 maximum available, Figure4.9) is associated with only a maximum peak level of just over 2500 of total FAs (Figure 4.10). In other words, a roughly 14-fold increase in PIP2 (from the peak level of 3000 molecules of the current model to the maximum of 43,000 shown in the parameter scan) is associated with only a doubling in total peak FA levels (from around 1300 molecules in the current model to the maximum of around 2600 molecules shown by the parameter scan).

Clearly there must be other reasons for the shortfall in FAs described above. One obvious limitation is that, whereas, as stated earlier, there are nearly 13,700 molecules of unbound integrin available from the start of each model run, there are only around 3700 molecules of unbound talin available.

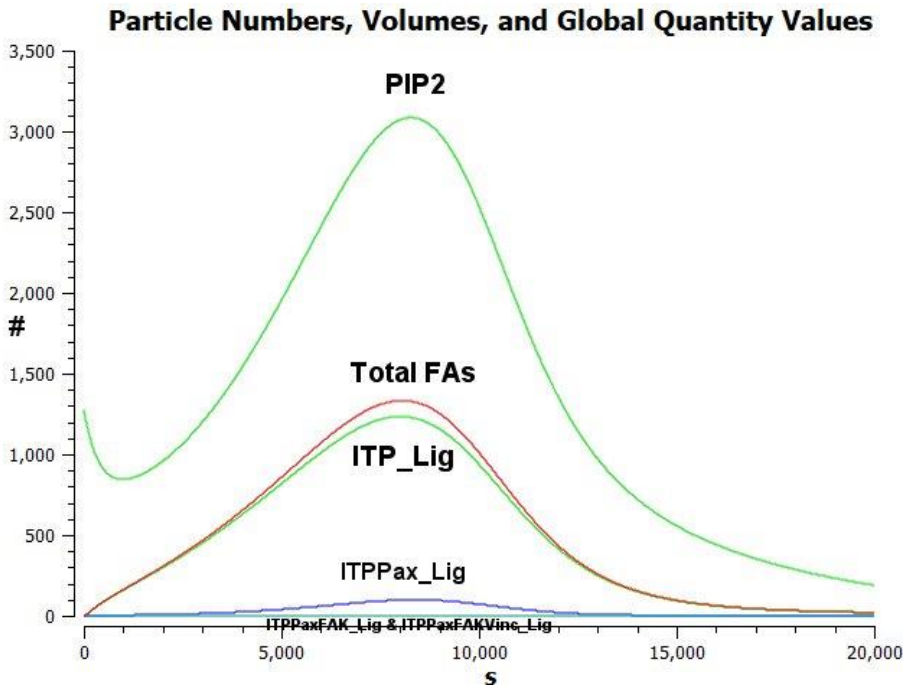

Figure 4.2.Particle numbers for PIP2 and Total FAs, ITP_Lig, ITPPax_Lig, ITPPaxFAK_Lig, ITPPaxFAKVinc_Lig.

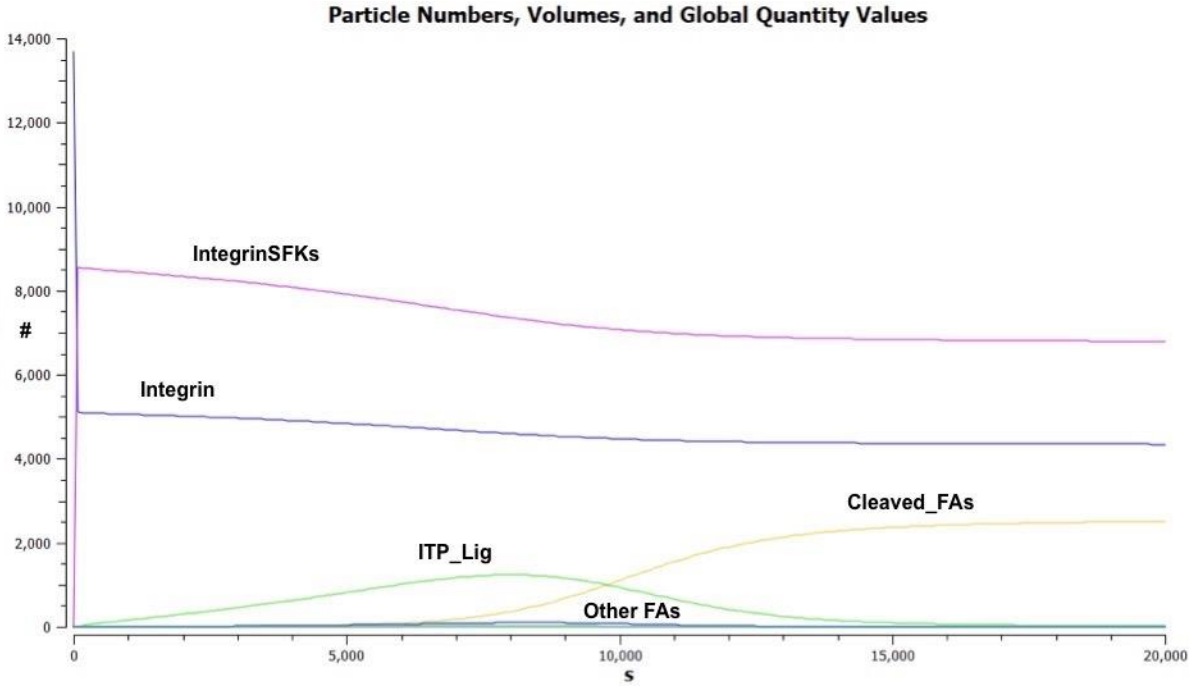

Figure 4.3. Particle numbers for all integrin species.

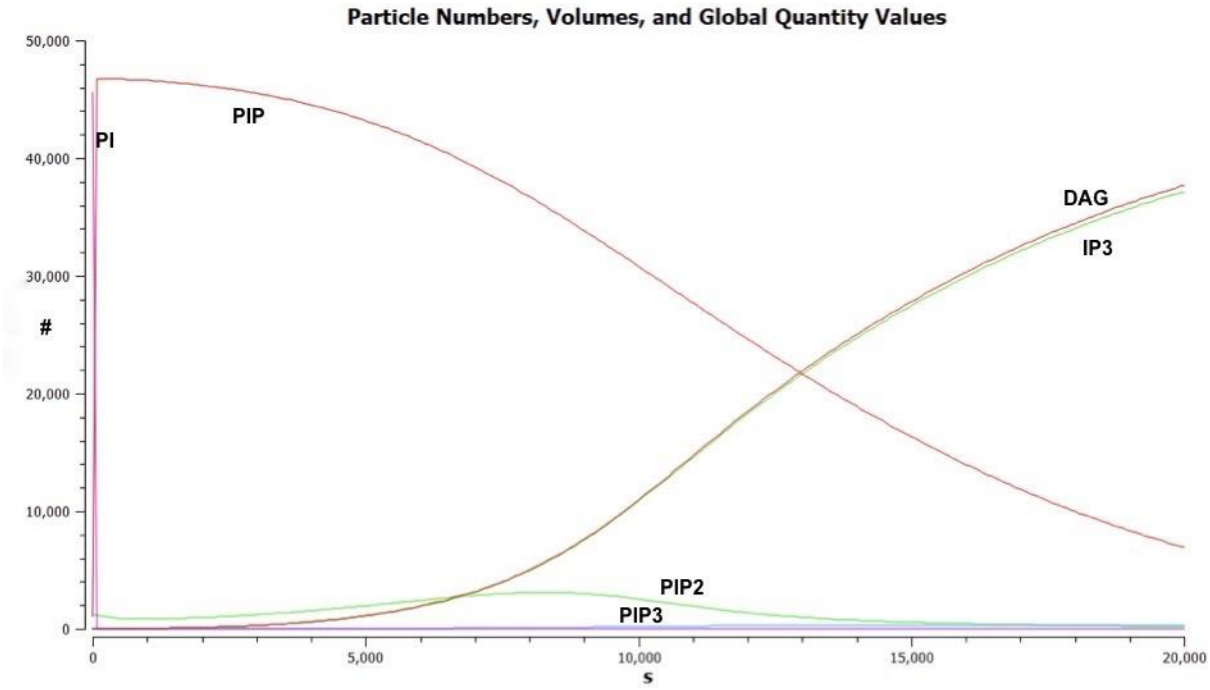

Figure 4.4. Particle numbers for PI, PIP, PIP2, PIP3, IP3 and DAG

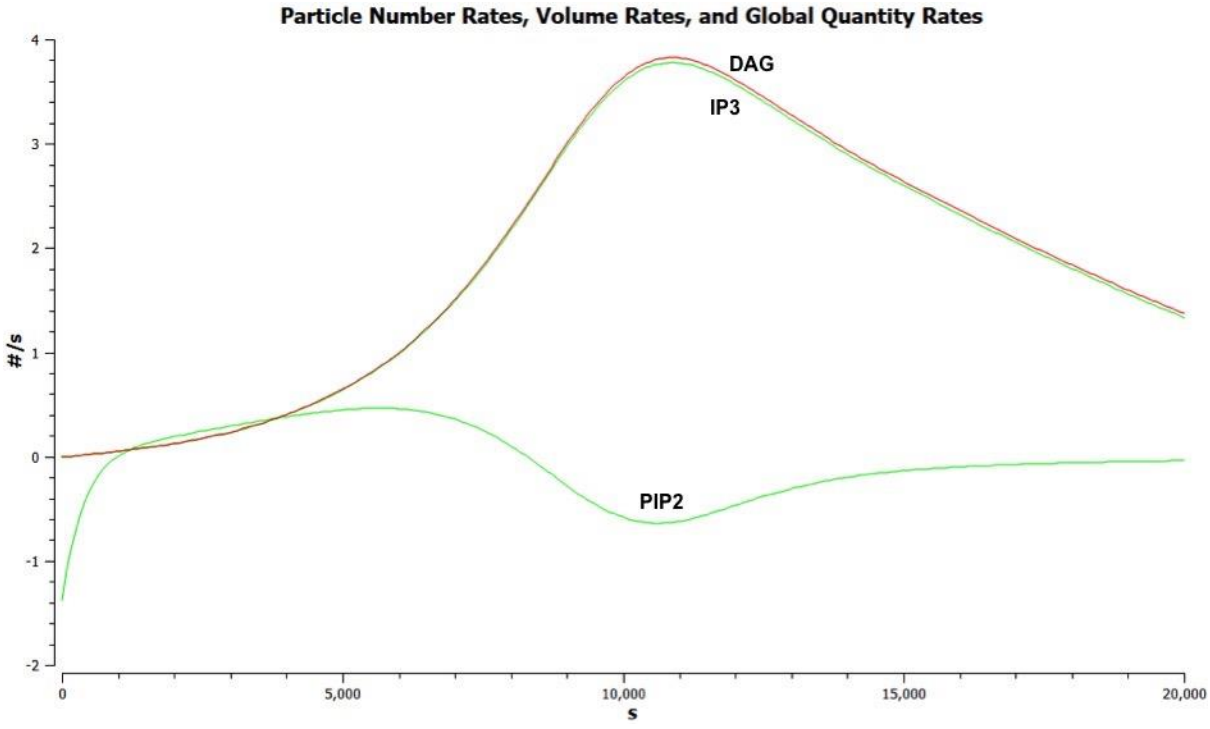

Figure 4.5. Particle formation rates for PIP2, IP3 and DAG

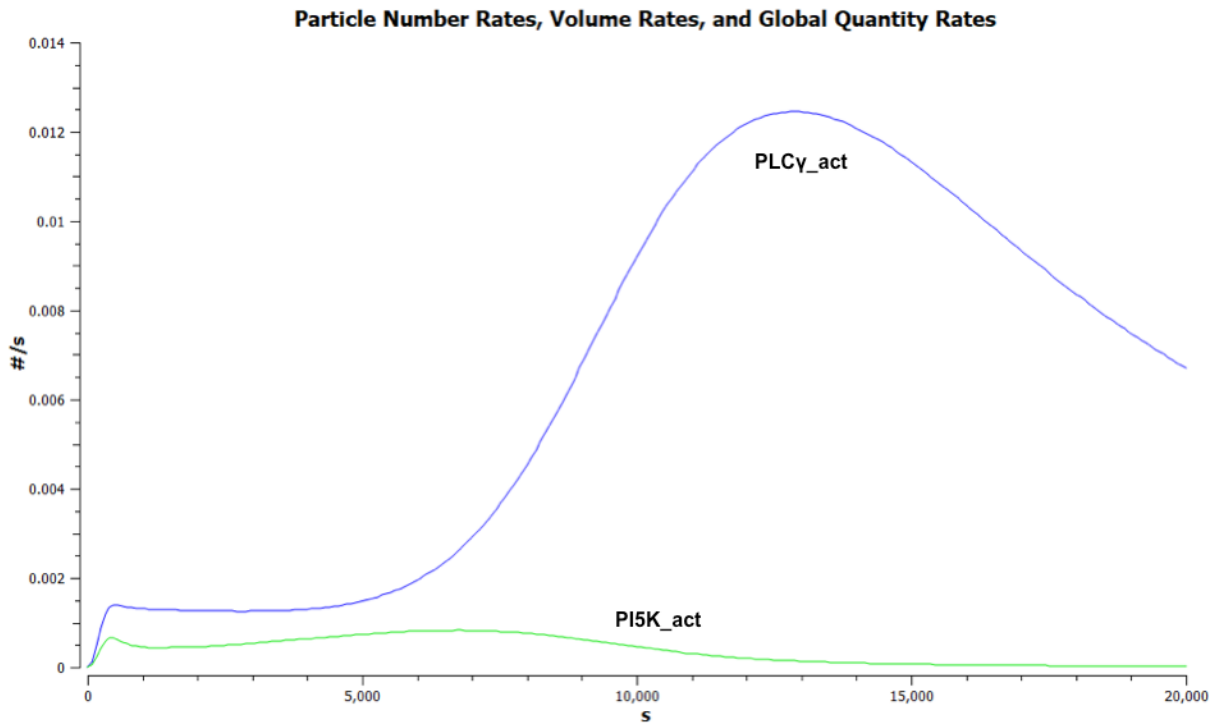

Figure 4.6. Particle formation rates for PI5K_act and PLCγ_act

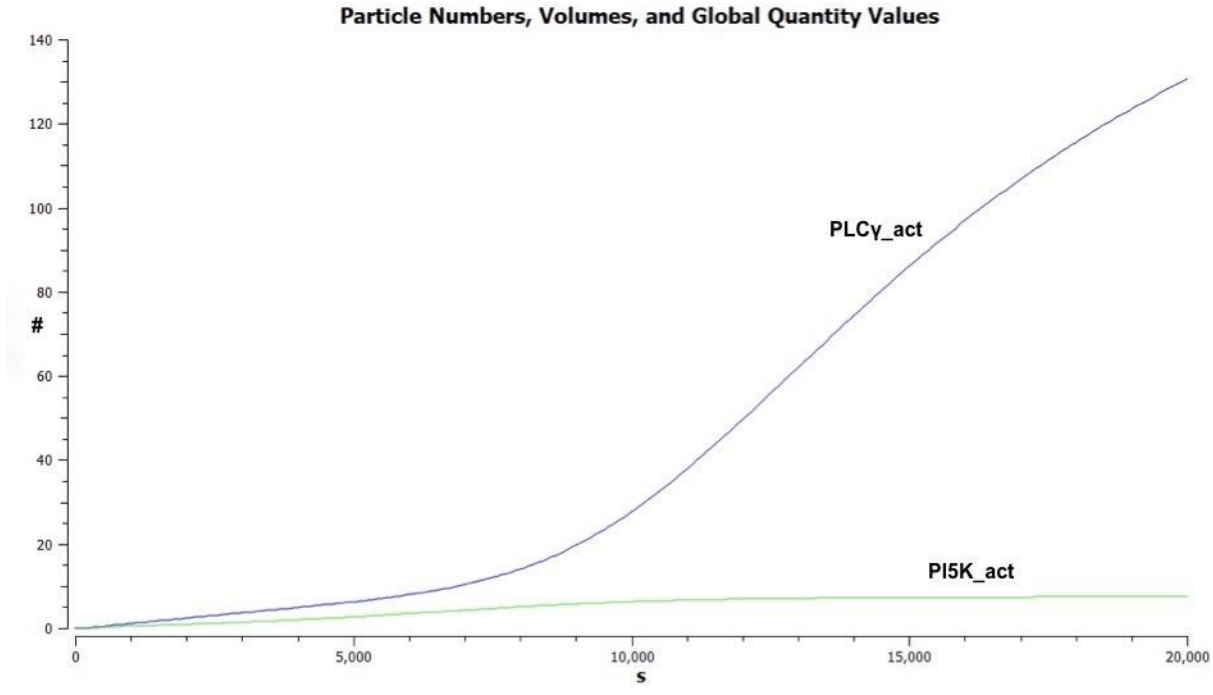

Figure 4.7. Particle numbers for PI5K_act and PLCγ_act

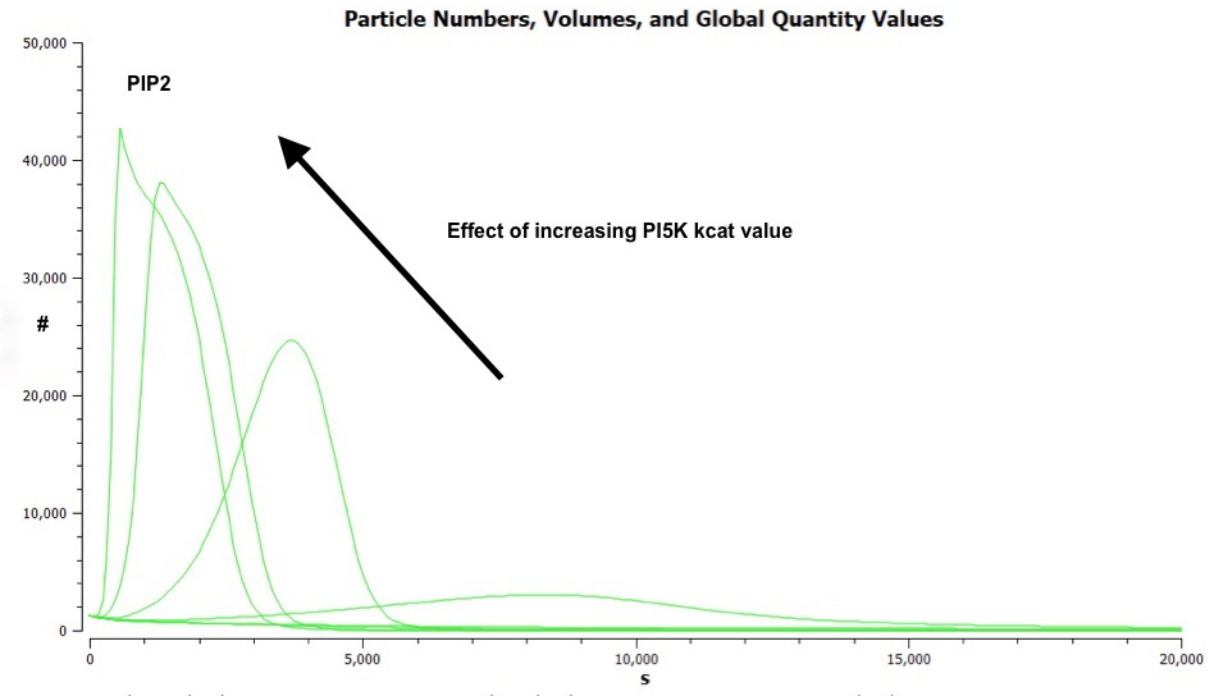

Figure 4.8. Logarithmic parameter scan, showing effects of increasing PI5K kcat value on levels of PIP2. (PI5K kcat values of 0.01, 0.1, 1, 10, 100 and 1000 s^{-1}.)

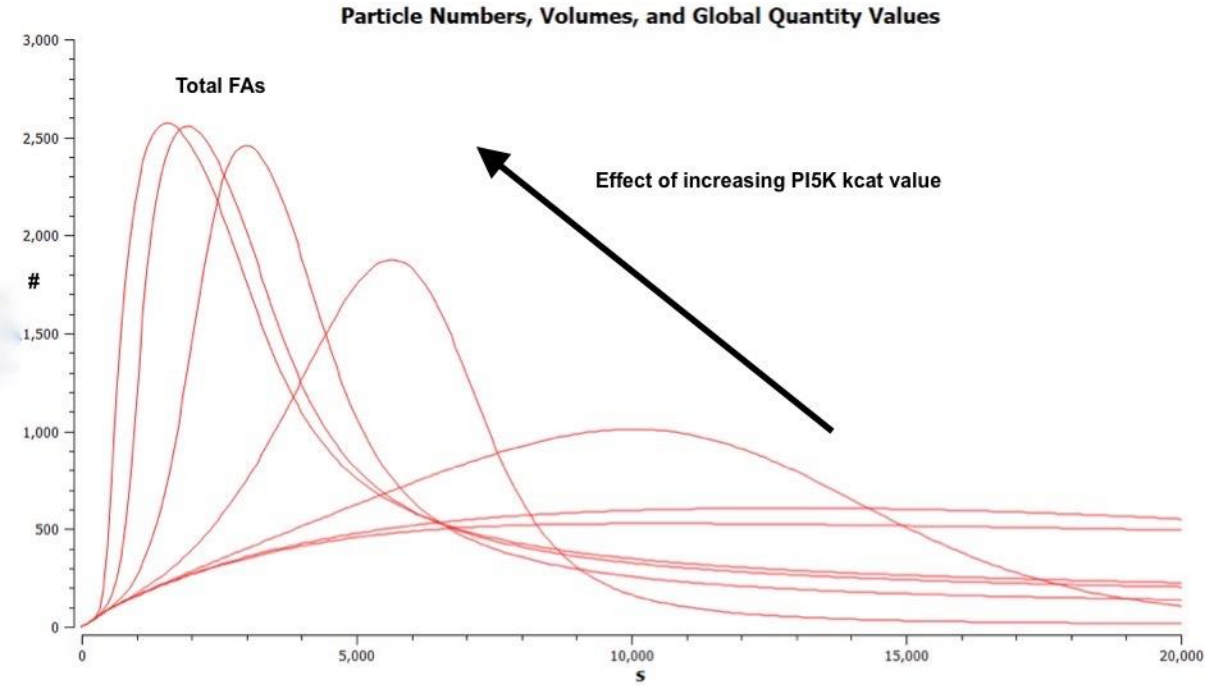

Figure 4.9. Logarithmic parameter scan, showing effects of increasing PI5K kcat value on levels of total FAs. (PI5K kcat values of 0.01, 0.1, 1, 10, 100 and 1000 s^{-1}.)

Since, in this model, integrin must first bind talin (to which it remains bound thenceforward), this means that there is also an upper limit of 3700 FAs possible. Taking into account cleavage by

calpain (represented by the species Cleaved_FAs in this model), amounting to over 2500 FAs after 20,000 seconds have lapsed, this accounts for why a peak level of only just over 1300 FA molecules is achieved at around 8000 seconds elapsed. A further logarithmic parameter scan on the forward rate for fibronectin-Int_Tal_PIP2 binding shows that this peak level for FAs can only be increased from its previous maximum of around 1300 to around 1400 molecules (Figure 4.10) and, even then, only at unrealistic forward rates (well above above what is understood to be thermodynamically possible).

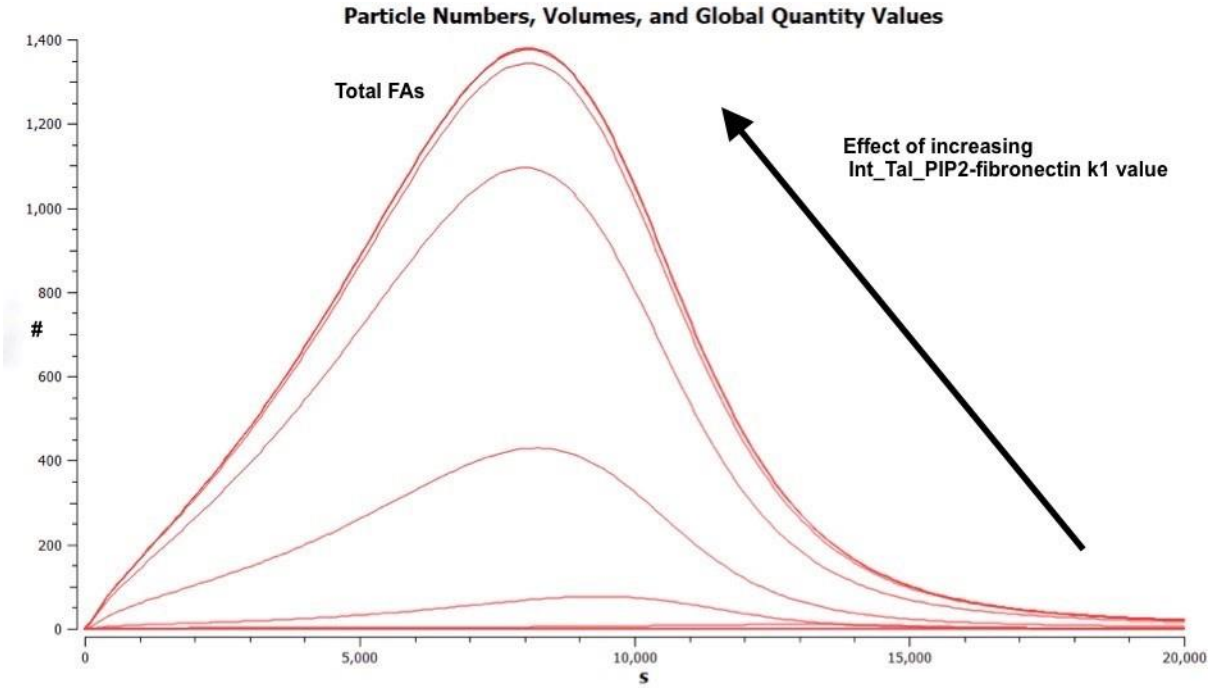

Figure 4.10. Logarithmic parameter scan, showing effects of increasing Int_Tal_PIP2-fibronectin binding forward rate constant (k1) value on levels of total FAs. (k1values in range of1-1e+12$M^{-1}s^{-1}$.)

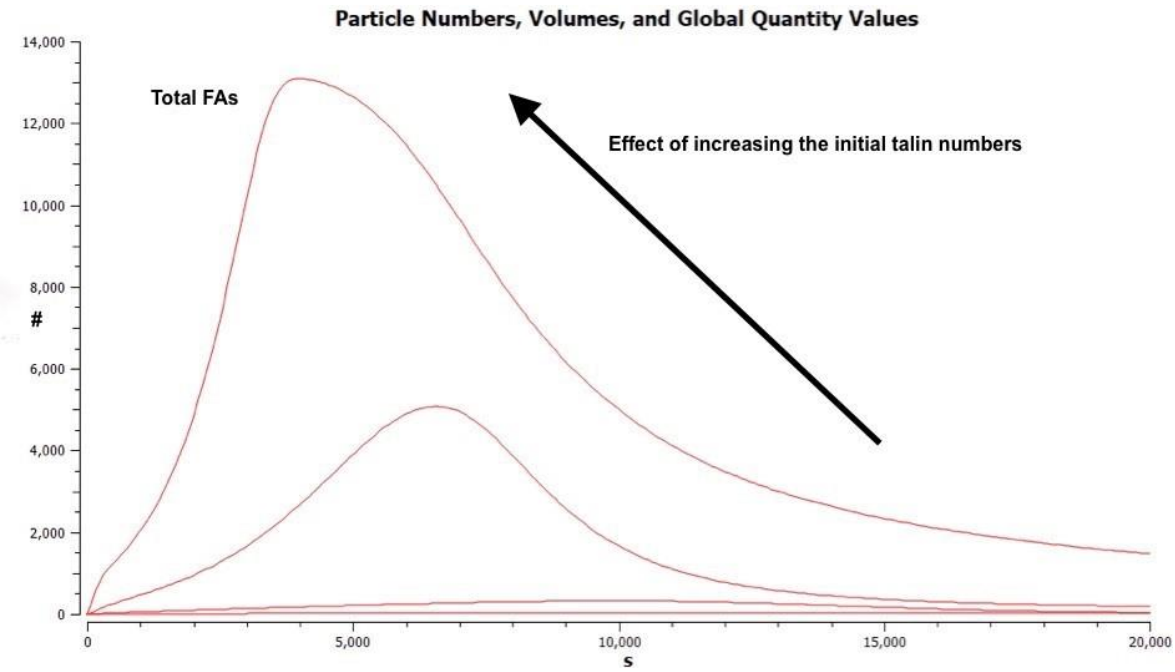

Figure 4.11. Logarithmic parameter scan, showing effects of increasing initial talin particle numbers on levels of total FAs. (Initial particle numbers of 100, 1000, 10,000, 100,000.)

Finally, a logarithmic parameter scan on initial particle numbers of talin confirms that increasing these starting levels greatly influences peak levels of FAs (Figure 4.11). Increasing this number to 100,000 (i.e. several times in excess of initial levels of integrins at around 13,700 molecules) increases peak levels of FAs to around 13,000, i.e. close to the maximum possible. Parameter scans of other model parameters (not shown here, for the sake of brevity) confirms that initial talin numbers are, by far, the most important determinant of peak FA levels.

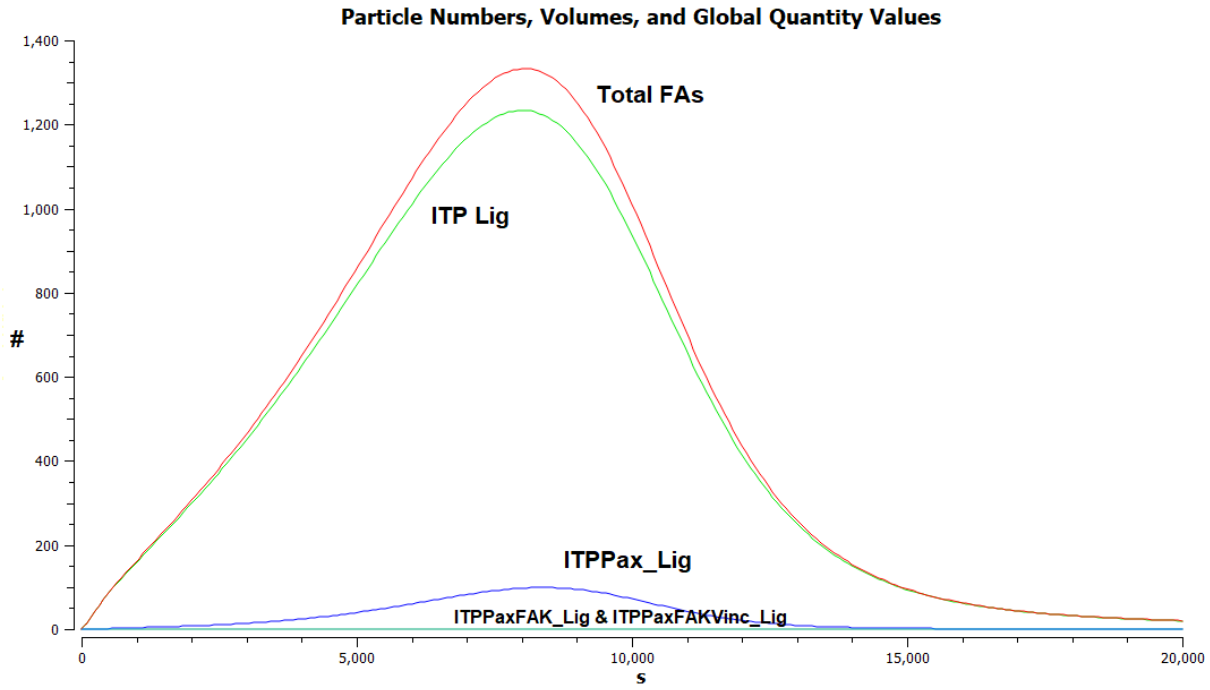

Figure 4.12. Particle numbers for Total FAs, ITP Lig, ITPPax_Lig, ITPPaxFAK_Lig, ITPPaxFAKVinc_Lig

Given this upper limit for FAs, the next question is why ITP_Lig accounts for the vast majority of these in our model, with only a very small representation by other FA species (ITPPax_Lig, ITPPaxFAK_Lig, ITPPaxFAKVinc_Lig, Figure 4.12). Again, an initial sensitivity analysis proves unhelpful, giving contradictory results, many of which are not confirmed by subsequent parameter scans. However, in the case of ITPPax_Lig, its prediction that Paxillin-PIP2 binding rates will influence peak levels of this molecule is confirmed by Figure 4.13. This shows that reducing the Paxillin-PIP2 binding off-rate substantially increases peak levels to around 350 molecules, as compared to the peak level of 100 molecules in this model, without much affecting the peak level of FAs. In other words, at peak FA levels, peak ITPPax_Lig levels go from representing less than 10% of total FAs, to accounting for nearly one quarter of them.

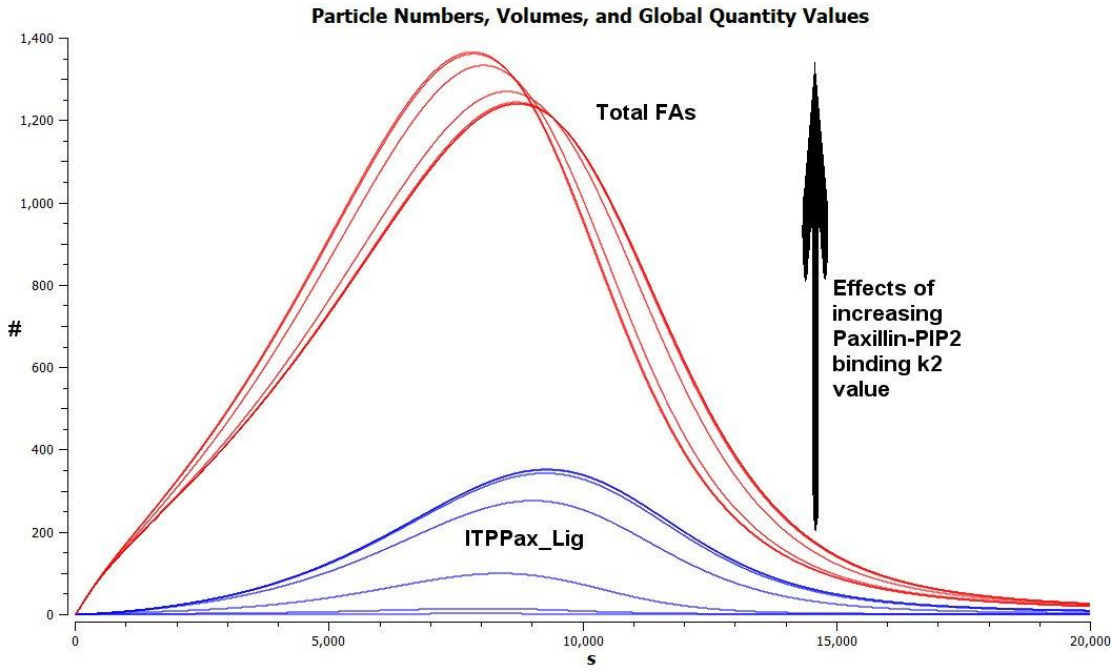

Figure 4.13. Logarithmic parameter scan, showing effects of decreasing Paxillin-PIP2 binding reverse reaction rate constant (k2) on levels of total FAs and ITPPax_Lig. (k2 value varied between 1e-06 and 1 s^{-1}.)

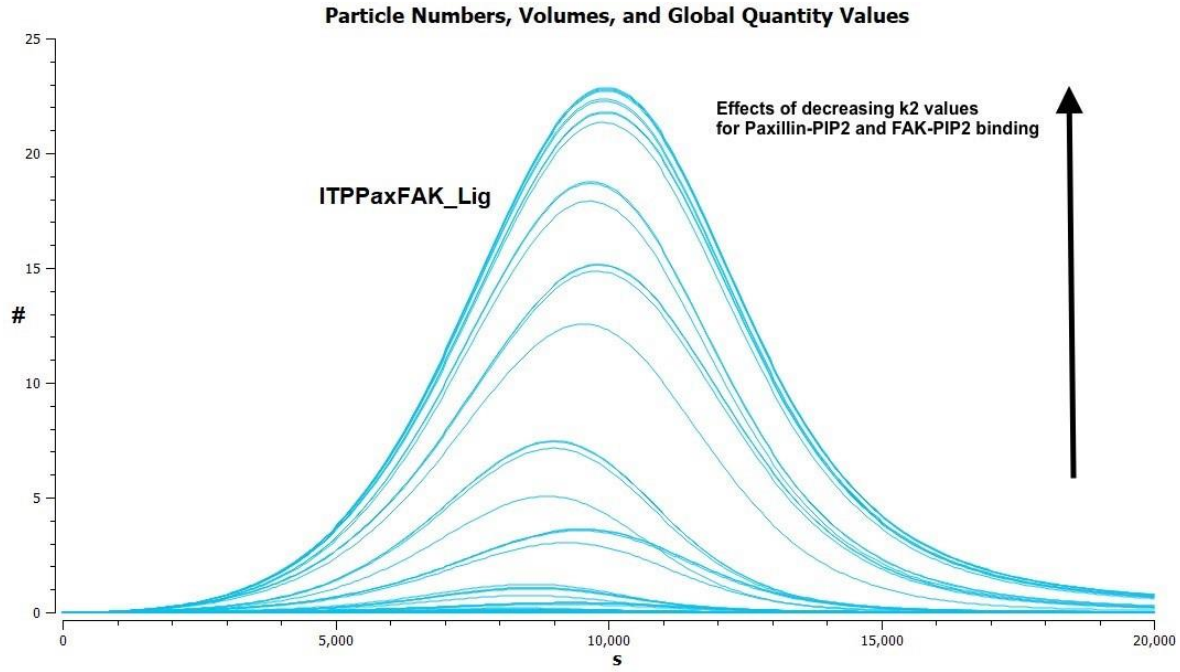

Figure 4.14. Nested logarithmic parameter scan, showing effects of decreasing Paxillin-PIP2 and FAK-PIP2 binding reverse reaction rate constants (k2) on levels of ITPPaxFAK_Lig. (k2 values for both reactions varied between 1e-06 and 1 s^{-1}.) Time course = 20,000 seconds.

However, increasing the FAK-PIP2 binding rate does not raise peak levels of ITPPaxFAK_Lig to the same extent, even after increasing levels of ITPPax_Lig. As shown in Figure 4.14(representing a

nested logarithmic parameter scan of Paxillin-PIP2 and FAK-PIP2 binding reverse rates), maximum peak levels of ITPPaxFAK_Lig are less than 24 molecules, i.e. around only 2% of peak levels of total FAs.

To achieve the highest peak levels of ITPPaxFAK_Lig(of over 500 molecules) requires increasing initial availability of FAK itself, as well as increasing the ITPPax_Lig-FAK_PM binding rate. This is shown in Figure 4.15.

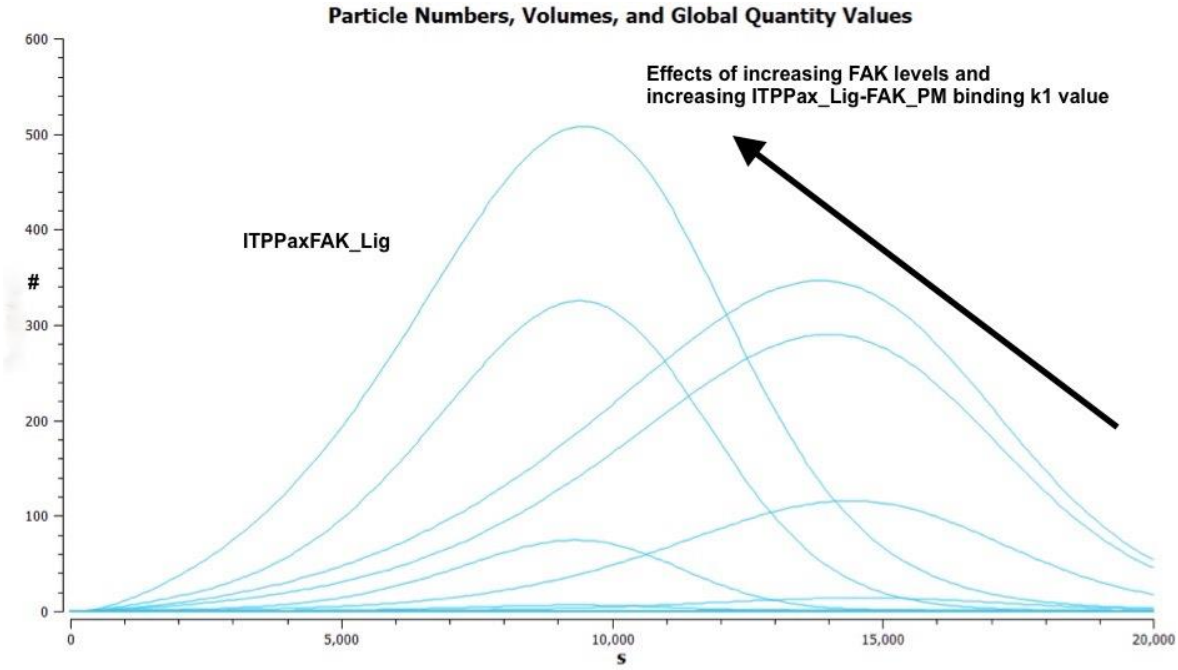

Figure 4.15. Nested logarithmic parameter scan, showing effects of increasing initial levels of FAK and increasing ITPPax_Lig-FAK_PM binding forward rate constant (k1) on levels of ITPPaxFAK_Lig. (FAK initial particle numbers varied between 10,000 and 100,000 molecules,k1 value for ITPPax_Lig-FAK_PM binding reaction varied between 1 and 1e+07 M^{-1} s^{-1}.)

It is only after combining these two factors that levels of ITPPaxFAK_Lig rise (from 10 to 30%) to the point where they represent a substantial percentage of total FAs. For ITPPaxFAKVinc_Lig the picture is more complicated still, requiring increased rates of ITPPaxFAK_Lig-Vinc_PM binding, as well as increased rates of ITPPax_Lig-FAK_PM binding and levels of FAK_PM (both increasing

ITPPax_Lig levels) in order for levels of this FA protein to exceed more than 10% of total FAs (Figure 4.16).

These conclusions suggest that binding rates for these FA proteins may be typically higher than the rather conservative default values assumed here, and that levels of FAK are considerably higher than suggested by the published quantitative proteome used here. Another consideration is that the assumption of sequential binding used here may be faulty. For instance, there is no evidence that vinculin binding requires prior FAK binding to the FA in question, although, as stated earlier, vinculin typically binds to talin by preference over other FA proteins. Similarly, there is some doubt as to which FA proteins act as binding partners to FAK [54].

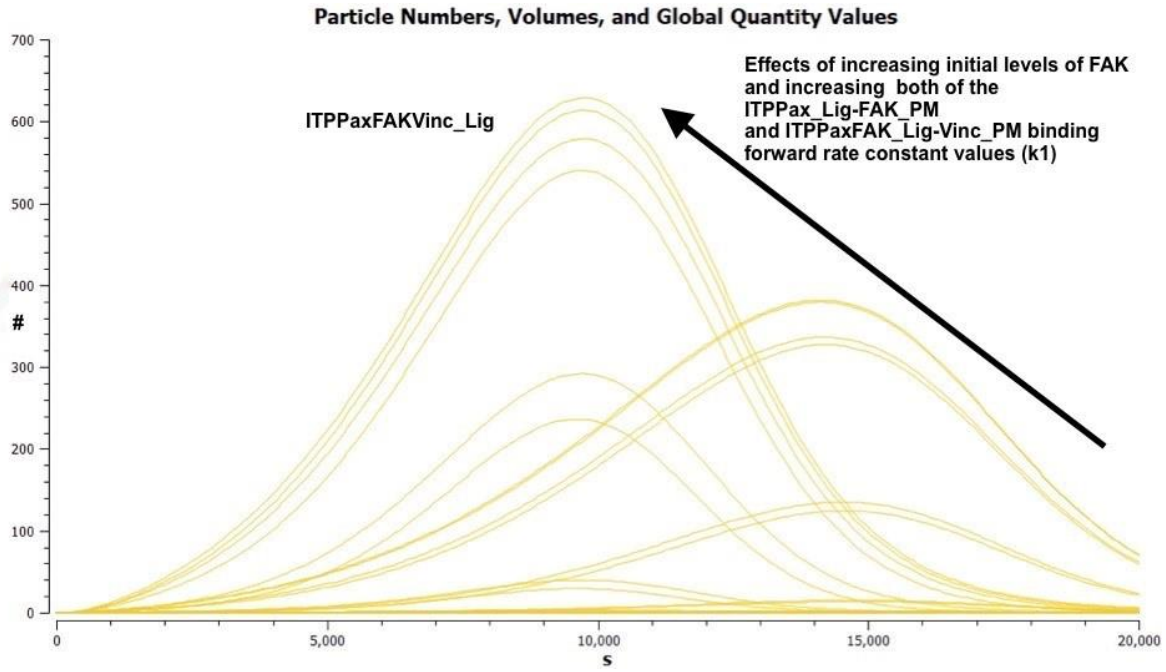

Figure 4.16. Nested logarithmic parameter scan, showing effects of increasing initial levels of FAK and increasing both of the ITPPax_Lig-FAK_PM and ITPPaxFAK_Lig-Vinc_PM binding forward rate constant values (k1) on levels of ITPPaxFAKVinc_Lig. (FAK initial particle numbers varied between 10,000 and 100,000 molecules, k1 value for ITPPax_Lig-FAK_PM binding reaction varied between 1e+06 and 1e+07 M^{-1} s^{-1}, k1 value for ITPPaxFAK_Lig-Vinc_PM binding reaction varied between 1 and 1e+07 M^{-1} s^{-1}.)

Nevertheless, there is a general consensus that talin is an early recruit to FAs, required for integrin activation[217]. So the fact that it is severely rate-limiting in this model, in terms of overall FA levels,

suggests that either initial talin levels in this model are too low, or else talin molecules are in some way shared between integrins. In fact, it has been shown that talin normally exists as an antiparallel dimer, allowing the integrin-binding heads at each end of the dimer to cross-link two adjacent integrins[219]. But this does not numerically improve the situation. It has also been reported that isolated talin head (cleaved from the whole protein) promotes integrin clustering, with the implication that one talin molecule could in some sense "activate" several integrins[303].

However, our current understanding of how talin activates integrins (which requires the talin head to engage with the cytoplasmic domain of the integrin dimer [304]) tends to rule this possibility out. Otherwise, it would suggest that isolated talin head is disengaging from one integrin and successively engaging with its neighbours, without any obvious mechanism for directing this movement. Moreover, in the absence of talin, vinculin has no means of binding directly to integrins (lacking an integrin-binding site[92]). So, unless clustered integrins are able to bind directly to actin, α-actinin or similar actin-associated proteins, without first being activated by talin, the rate-limiting effects of low talin numbers on total FA levels, reported above, remain problematic for this model.

Similar concerns arise with FAK. At 467 molecules, initial FAK particle numbers are considerably lower than those of integrin (13,669), talin and paxillin (both 3266 molecules). Again, given the severe rate-limiting effects described above, either these initial quantities are too low, or one must assume that FAK is not required to be present in every mature FA all of the time. If we assume the former, however, then, as with talin, there is no obvious reason why these proteins should be uniquely underestimated, compared to the other FA proteins. Certainly, it is known that membrane (particularly transmembrane) proteins can be undercounted by the mass spectrometry-based methods currently employed in quantitative proteomics[305306]but then talin and FAK are primarily cytosolic proteins, whereas integrin is a transmembrane protein, so, on that basis, one would expect integrin to be undercounted, rather than talin or FAK.

Yet, the fact that both talin and FAK are both primarily cytosolic proteins might provide a plausible systematic explanation, based on the premise that all cytosolic proteins are underrepresented in this model. This follows from the fact that initial numbers of model species, for this COPASI model, have been estimated by taking their numbers per HeLa cell (as reported in the quantitative proteome) and then dividing this number by the calculated volume for their resident compartment (i.e. in most cases the volume of the cytosol or PM) to arrive at their concentration. These species concentrations are then been multiplied by the volumes calculated for these same compartments in our model (based on a hypothetical lamellipodia), to arrive at the initial particle number for these model species. It could be that the HeLa cell cytosolic volume has been overestimated in the first instance (resulting in too low a concentration for cytosolic molecules), or underestimated in the second instance (resulting in initial particle numbers being lower than they should be, for these same molecules). Either way, assuming that the same mistake hasn't been made for PM-resident species, the result would be that initial particle numbers for cytosolic species will be too low relative to PM species, such as integrins.

In support of this, total copy numbers of talin, paxillin and vinculin are considerably in excess of those of integrin, taking HeLa cells as a whole, but well below those of integrin in our lamellipodia model, as shown in Table 4.5. Similarly, as this table shows, whereas FAK levels are nearly one half those of integrins in HeLa cells (14,000 as against 30,000), their initial numbers are less than 3.5% those of integrins in this model (467 as against 13,689).This is because, whereas the cytosolic volume in the lamellipodia model is one-thirtieth the volume of the HeLa cell cytosol (50 fl as against 1500 fl), the lamellipodial PM volume is only slightly lower than that for the whole HeLa cell (5 fl as against 6.4 fl), as explained later.

Species	HeLa cell copy number	COPASI Lamellipodia model copy number
Integrin	30,000	13,689
Talin	98,000	3267
Paxillin	98,000	3267
Vinculin	310,000	10,333
FAK	14,000	467

Table 4.5. Comparison of species copy numbers, as reported for HeLa cells, and the corresponding initial particle numbers estimated for the COPASI lamellipodia model.

Even allowing for the fact that the tongue-like shape of the lamellipodia will tend to have a much higher surface-to-volume ratio than the more rounded cell from which it is derived, these relative differences seem improbable. However, any reasonable adjustment would likely require the PM volume to be reduced by more than the cytosolic volume is increased, if it were to more accurately mirror reality. Rather than substantially increasing initial levels of talin, FAK and other cytosolic proteins, such an adjustment would result in a reduction of the initial particle numbers for integrins, as well as other important PM species such as phosphoinositides. Nevertheless, by increasing the ratio of cytosolic to PM species numbers, it should ensure that cytosolic species such as talin and FAK are not rate-limiting to the extent they currently are. This in turn should mean that FAs, as a whole, represent a higher proportion of available integrins and, quite likely, more of these FAs will be in the form of ITPPax_Lig, ITPPaxFAK_Lig and ITPPaxFAKVinc_Lig, rather than almost exclusively in the form of ITP_Lig.

Unfortunately, as shown by Figures 4.17 and 4.18, this rebalancing does not occur. Decreasing the PM volume (by a quarter, from 5.475 to 1.37 fl) does indeed increase levels of cytosolic species such as talin, paxillin and FAK, relative to integrins, to the point where they are no longer rate-limiting (Figure 4.17). But as well as decreasing peak levels of total FAs (from around 1400 to 600 molecules), this total is still almost completely dominated by ITP_Lig (Figure 4.18) To change this requires also substantially increasing the on-rates (or decreasing the off-rates) for ITP_Lig-Paxillin_PM, ITPPax_Lig-FAK_PM and ITPPaxFAK_Lig-Vinc_PM binding. It is only after both these

alterations are made that ITPPax_Lig, ITPPaxFAK_Lig and ITPPaxFAKVinc_Lig are represented within total FAs at levels that seem more compatible with what is seen physiologically (Figure 4.19). Similar results are seen if, instead of decreasing PM volume by a quarter, cytosolic volume is increased fourfold (Figure 4. 20). Total FA peak levels are now roughly doubled compared to the standard model but, again, there is a redistribution of model FA species in a manner that seems more physiologically realistic.

But it is hard to justify altering either compartment volume by these amounts. What can more easily be justified is a decrease in the PM volume by one half. This is because the volume of the PM in this model was calculated by multiplying its surface area (730 μm^2) by 7.5 nm, representing the average thickness of a cell membrane. However, all the reactions involving PM-resident species in this model tend to occur at the interface between the cytosol and PM, rather than within the membrane itself. This implies that the effective depth is much less than this, with a corresponding reduction in the effective reaction volume for these PM-associated reactions. In any case, other studies of phospholipid bilayer membranes have reported thicknesses of only 3 nm [307].

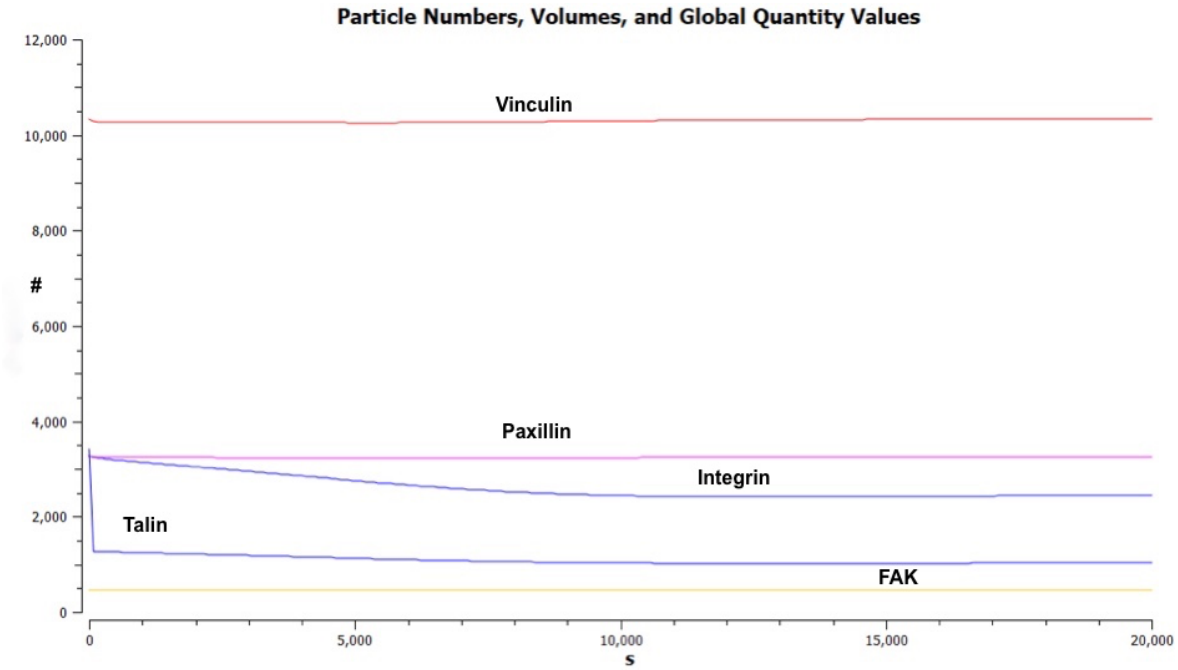

Figure 4.17. Particle numbers for integrin, talin, paxillin, FAK and vinculin after PM volume has been decreased from 5.475 to 1.37 fl.

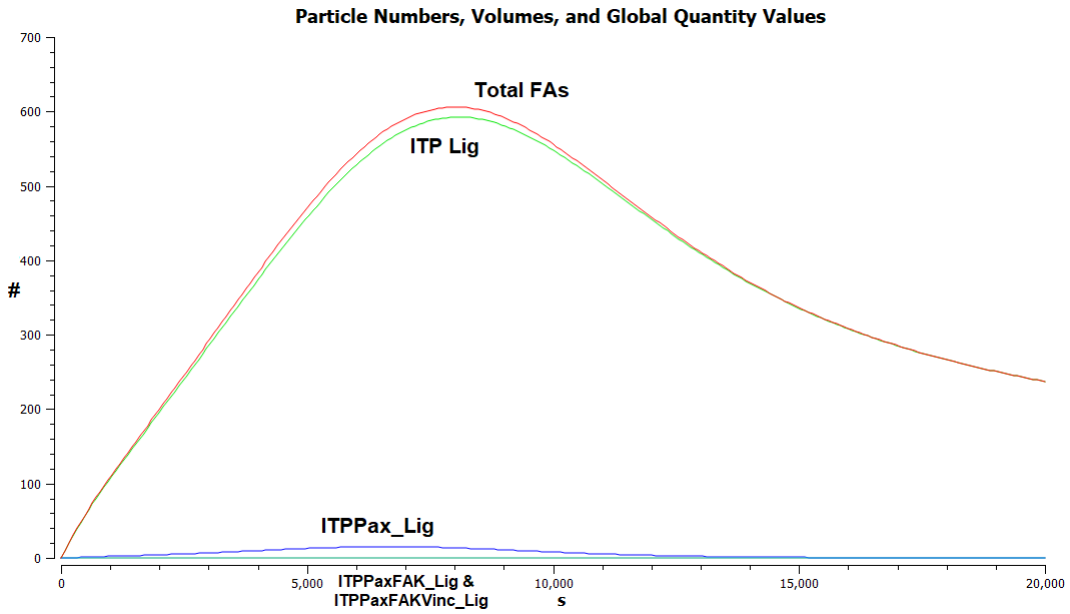

Figure 4.18. Particle numbers for total FAs, ITP_Lig, ITPPax_Lig, ITPPaxFAK_Lig and ITPPaxFAKVinc_Lig after PM volume has been decreased from 5.475 to 1.37 fl.

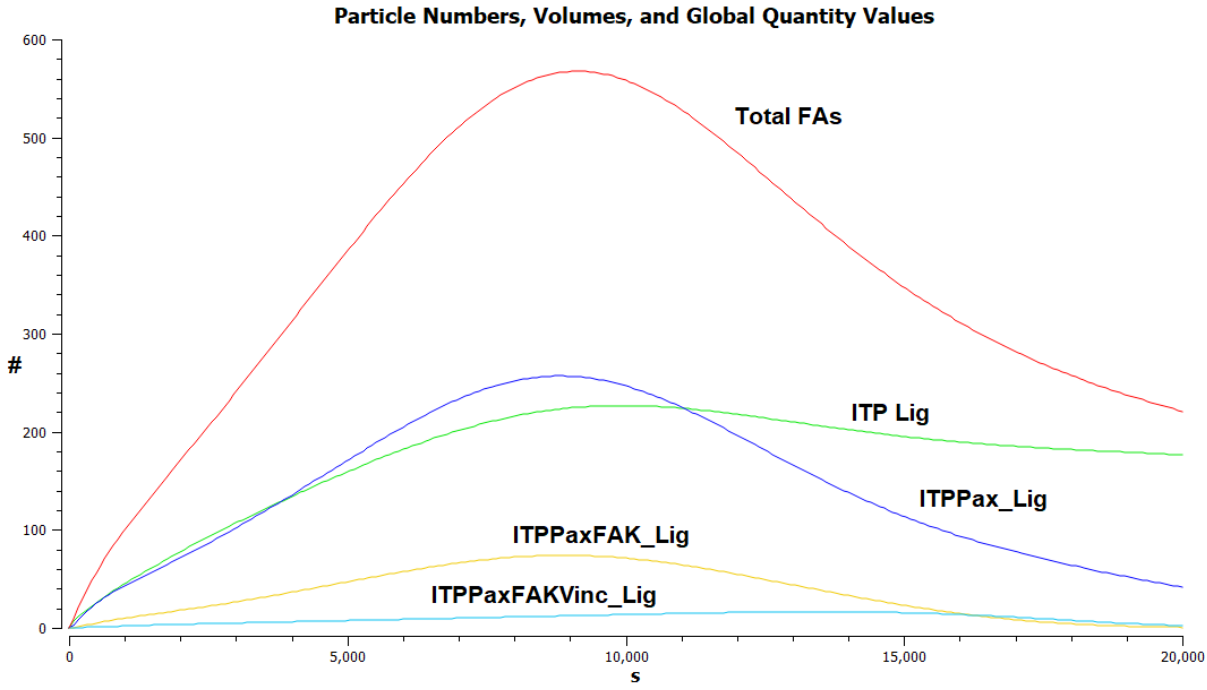

Figure 4.19. Particle numbers for total FAs, ITP_Lig, ITPPax_Lig, ITPPaxFAK_Lig and ITPPaxFAKVinc_Lig, after PM volume has been decreased from 5.475 to 1.37 fl, and on-rates for ITP Lig-Pax_PM, ITPPax_Lig-FAK_PM and ITPPaxFAK_Lig-Vinc_PM binding have been increased from 10,000 to 1,000,000 M^{-1} s^{-1}.

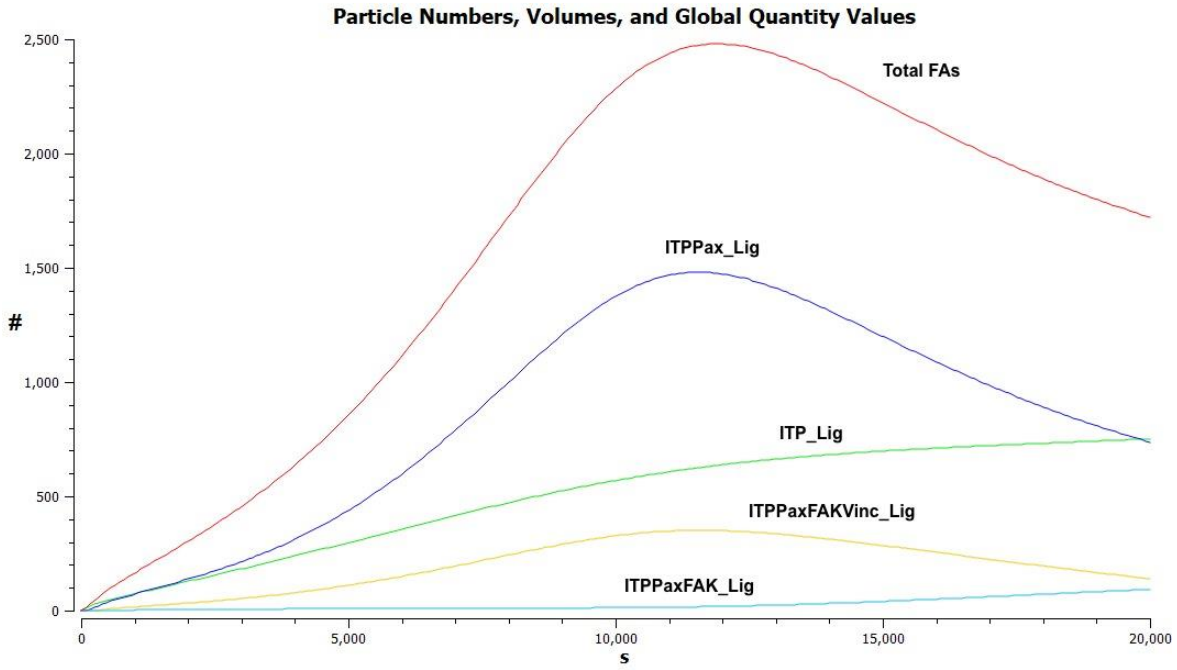

Figure 4.20. Particle numbers for total FAs, ITP_Lig, ITPPax_Lig, ITPPaxFAK_Lig and ITPPaxFAKVinc_Lig after cytosolic volume has been increased fourfold from 50 to 200 fl.

As a compromise, we will reduce the PM volume by a half (rather than a quarter) for the second version of the model. On-rates for ITP_Lig-Pax_PM, ITPPax_Lig-FAK_PM and ITPPaxFAK_Lig-Vinc_PM binding will be increased by an order of magnitude, with off-rates decreased by the same amount.

4.4.1.2 Version 2 (of COPASI model)

Thus, the second version of the model will see the changes seen in Tables 4.6 and 4.7.

Compartment	Version 1 volume	Version 2 volume
PM	5.475 fl	2.74 fl

Table 4.6. Compartment volume changes for the second version of the model.

Reactions	Version 1 rate constant values	Version 2 rate constant values
ITP_Lig + Pax_PM = ITPPax_Lig	k1 = 10,000 $M^{-1} s^{-1}$, k2 = 0.01 s^{-1}	k1 = 100,000 $M^{-1} s^{-1}$, k2 = 0.001 s^{-1}
ITPPax_Lig + FAK_PM = ITPPaxFAK_Lig	k1 = 10,000 $M^{-1} s^{-1}$, k2 = 0.01 s^{-1}	k1 = 100,000 $M^{-1} s^{-1}$, k2 = 0.001 s^{-1}
ITPPaxFAK_Lig + Vinc_PM = ITPPaxFAKVinc_Lig	k1 = 10,000 $M^{-1} s^{-1}$, k2 = 0.01 s^{-1}	k1 = 100,000 $M^{-1} s^{-1}$, k2 = 0.001 s^{-1}

Table 4.7. Reaction rate constant changes for the second version of the model.

This results in a plot of FAs shown in Figure 4.21.

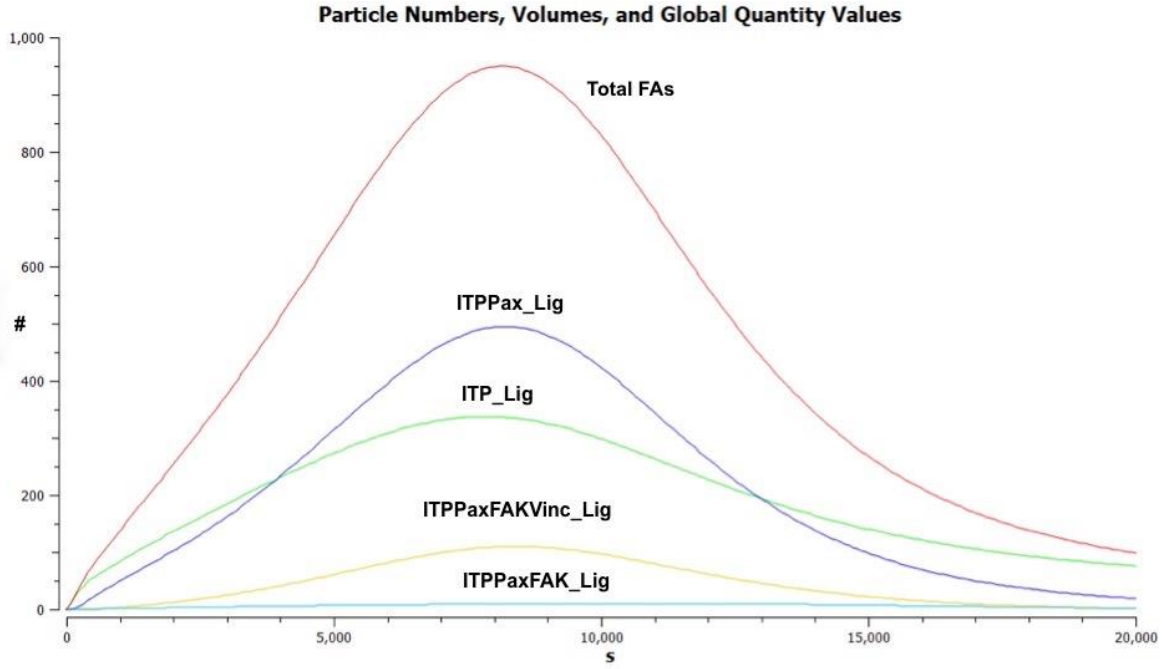

Figure 4.21. Particle numbers for total FAs, ITP_Lig, ITPPax_Lig, ITPPaxFAK_Lig and ITPPaxFAKVinc_Lig, for second version of model. After PM volume has been decreased from 5.475 to 2.74 fl, with on-rates for ITP Lig-Pax_PM, ITPPax_Lig-FAK_PM and ITPPaxFAK_Lig-Vinc_PM binding increased from 10,000 to 100,000 $M^{-1} s^{-1}$ and off-rates decreased from 0.01 to 0.001 s^{-1}.

Bearing in mind these changes were made to default binding rates, why not extend the same changes to all reactions involving these rates? Unfortunately, as Figure 4.22 shows, whilst considerably speeding up overall FA dynamics, these changes also result in peak levels of FAs being substantially reduced, to levels that seem physiologically unlikely if lamellipodia of the size modelled here are to gain sufficient traction for cell locomotion.

Since, as mentioned earlier, FA dynamics in this model appear to be occurring much slower than they would in physiological reality, it may well be that many of these reactions do indeed run much faster than the default rates given to them in this model. Indeed, it may well be that these binding rates lie, on average, somewhere between the values used in this model and the changes suggested here, resulting in faster FA dynamics, but with still sufficient peak FA levels for lamellipodial traction. But, if so, this may only apply to some reactions and, in the circumstances, attempting to speculate what these values might be, in the absence of adequate data as to the levels of individual FA species in lamellipodia, would appear to have little practical benefit.

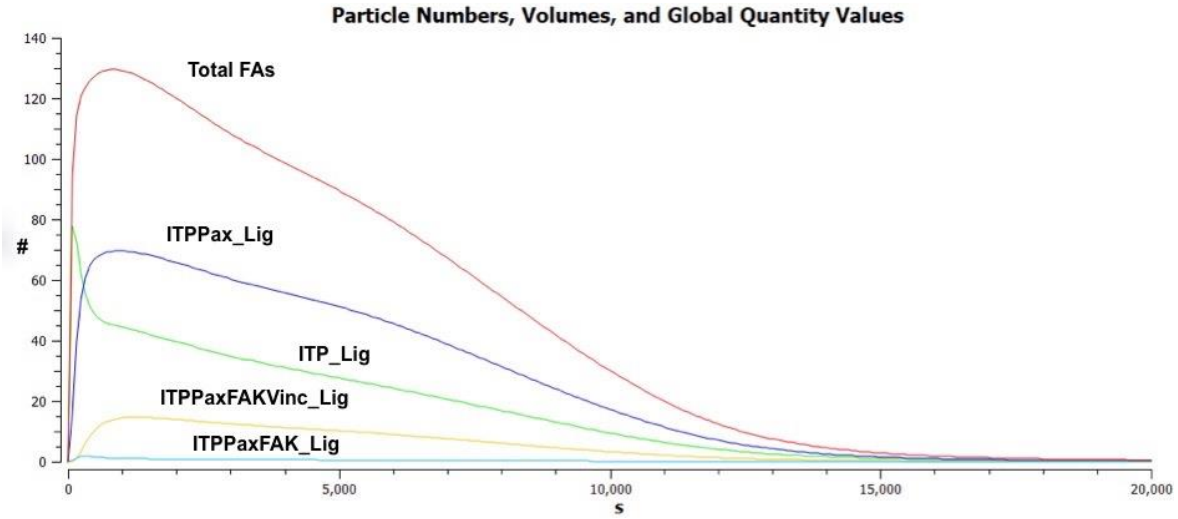

Figure 4.22. Particle numbers for total FAs, ITP_Lig, ITPPax_Lig, ITPPaxFAK_Lig and ITPPaxFAKVinc_Lig, for second version of model. After PM volume has been decreased from 5.475 to 2.74 fl, with on-rates for all reactions using default binding kinetics increased from 10,000 to 100,000 M^{-1} s^{-1} and off-rates decreased from 0.01 to 0.001 s^{-1}.

Overall, despite the relative slowness of the FA dynamics, this would appear to be a useful model, in the sense of having predictive potential and being amenable to interrogation. There is no claim here that it is a particularly faithful reflection of the physiological reality of FA dynamics. Given the complexity of the system being modelled and our incomplete understanding of it, a completely faithful model would probably be impractical, even if it were possible.

Rather, like all useful models, its usefulness lies in the questions that it raises about that physiological reality. Here these questions have primarily revolved around levels of FA complexes, both in aggregate, and relative to each other. This in turn has led us to look into the contributions of PIP2 and of initial levels of talin, FAK and paxillin, of PM and cytosolic volumes and of protein-protein binding rates, in determining these aggregate and relative FA levels. Questions have also arisen about prescriptive, strictly sequential, binding in FA assembly, and whether talin or FAK are obligate FA constituents.

There are, of course, many other questions one could ask of this model; in particular, the relative contributions of calpain cleavage, and of PIP2 hydrolysis and phosphorylation (by, respectively, PLCγ and PI3K), to overall FA levels. In a similar vein, it would be interesting to investigate how changes in the timing and extent of IP3-mediated calcium release affect FA levels (both aggregate and relative).

However, such questions can be included in the Virtual Cell model, to be described in the next section. For comparison purposes (with Version 1), the following plots (Figures 4.23-4.29) of various species' behaviours for the new version of the model, are provided.

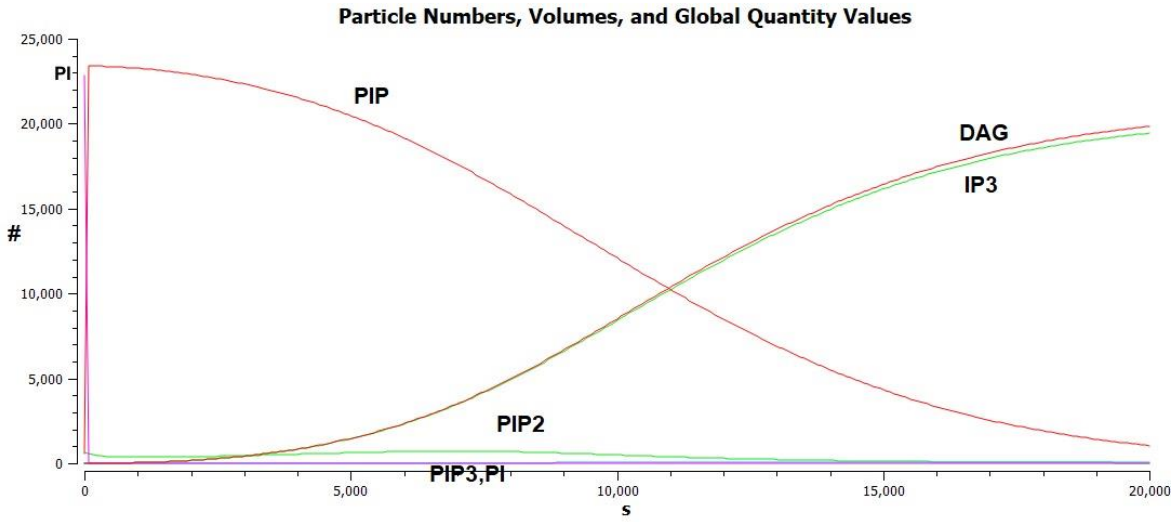

Figure 4.23. Particle numbers for PI, PIP, PIP2, PIP3, IP3 and DAG for second version of COPASI model.

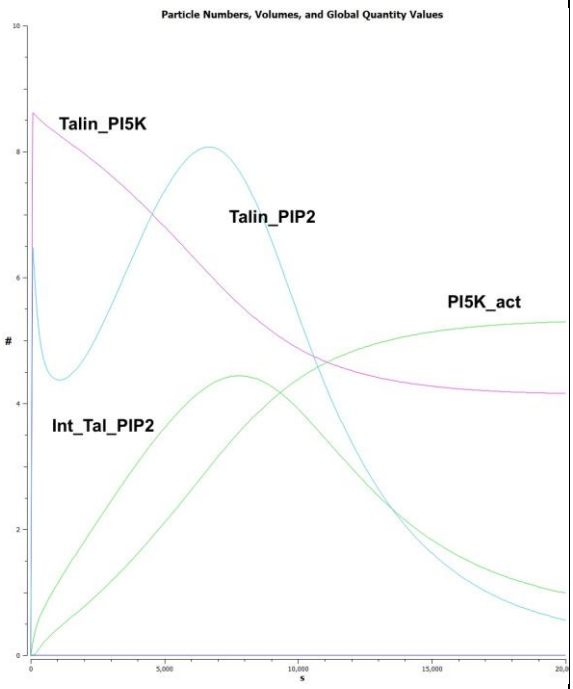

Figure 4.24. Particle numbers for Talin_PIP2, Talin_PI5K, PI5K_act andInt_Tal_PIP2 for second version of COPASI model

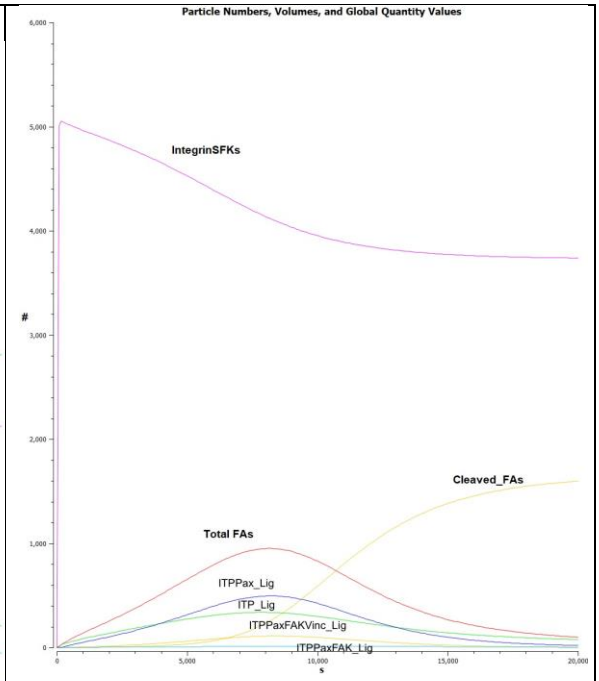

Figure 4.25. Particle numbers for Total FAs, individual FAs, IntegrinSFKs and Cleaved_FAs for second version of COPASI model

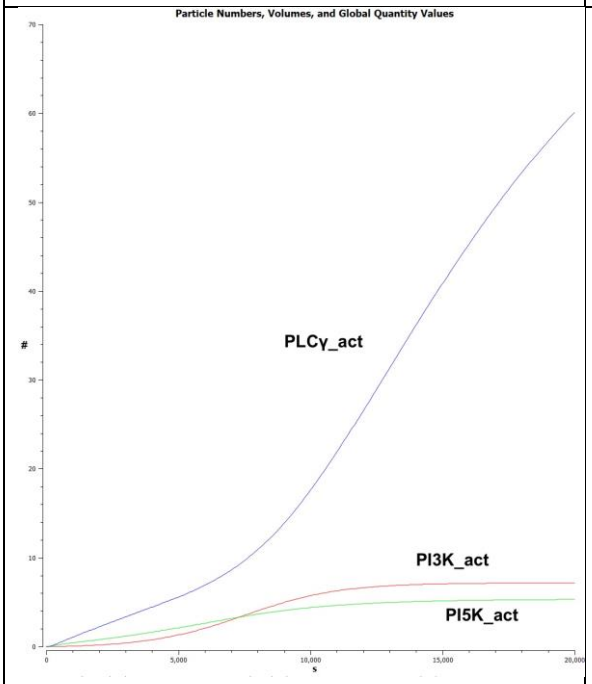

Figure 4.26. Particle numbers for PI5K_act, PLCγ_act and PI3K_act for second version of COPASI model

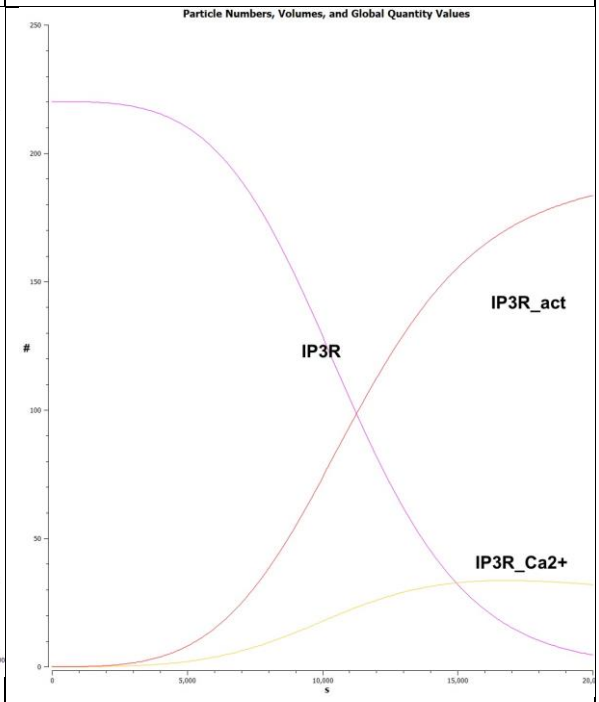

Figure 4.27. Particle numbers for IP3R species for second version of COPASI model

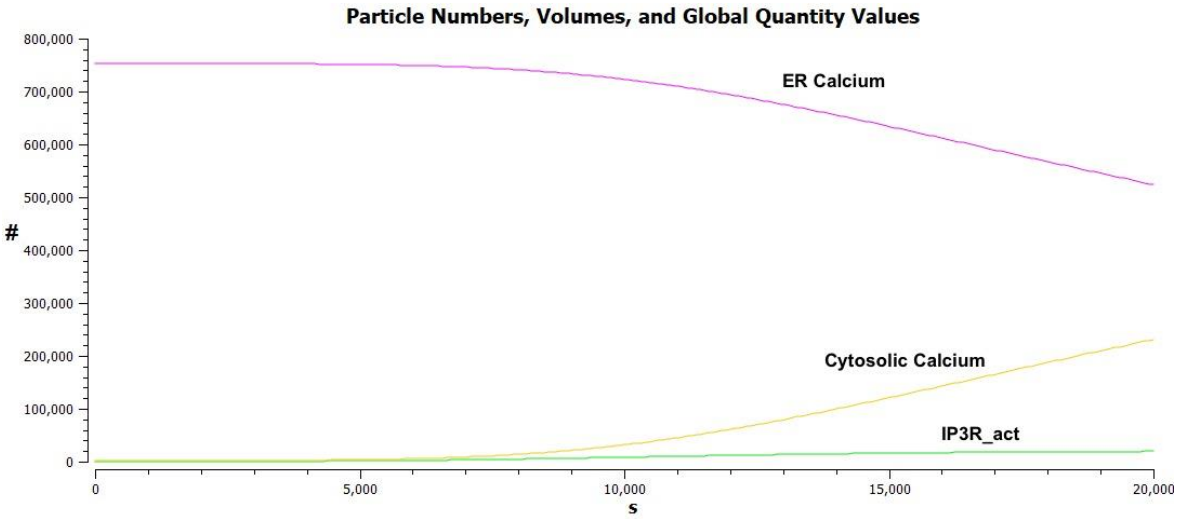

Figure 4.28. Particle numbers for IP3R_act, Ca2+_ER and Ca2+_cyto for second version of COPASI model.

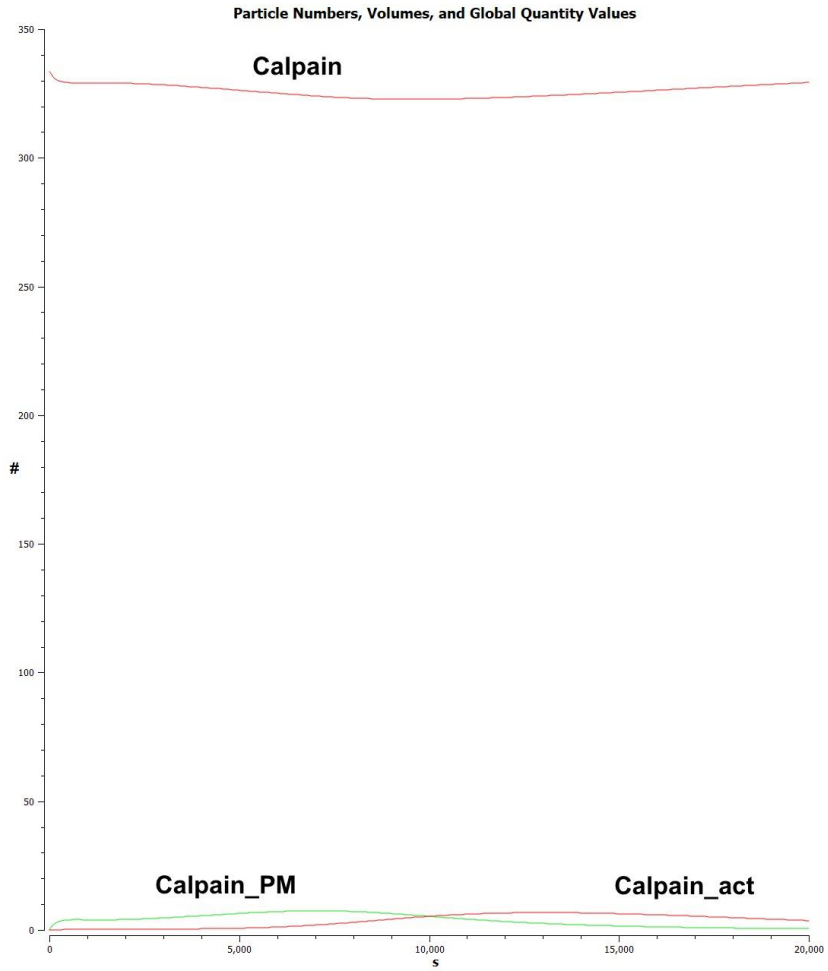

Figure 4.29. Particle numbers for all calpain species for second version of COPASI model.

4.4.2. Virtual Cell model

4.4.2.1 ODE/compartmental application

4.4.2.1.1 Version 1

This Virtual Cell model has implemented the second version of the COPASI model, including the changes summarised in Tables 4.6 and 4.7. However, because Virtual Cell models membranes as two-dimensional surfaces, whereas COPASI models as three-dimensional volumes, a decision has to be made as how best to implement the changes of the COPASI second version model. These halved the PM volume whilst maintaining the concentrations of PM species, thereby halving their copy number. To achieve the same end with the Virtual Cell model, one can either halve the surface area (730 μm^2) originally specified for the PM (as shown in Table 4.1 in the Methods section), whilst retaining the existing densities/concentrations for PM species, or one can maintain the surface area, whilst halving the densities. Since, in Virtual Cell, species quantities have to be specified in terms of densities/concentrations, it is more convenient to halve the surface area, rather than having to alter all the PM species densities.

So, the Virtual Cell model now has the changes to the original specification (as shown in the Methods section) shown in Tables 4.8 and 4.9.

Compartment	Original surface area	New surface area
PM	730 μm^2	365 μm^2

Table 4.8. Compartment volume changes for the Virtual Cell model.

Reactions	Original rate constant values	New rate constant values
ITP_Lig + Pax_PM = ITPPax_Lig	k1 = 0.00221 $\mu m^2.molecules^{-1}.s^{-1}$, k2 = 0.01 s^{-1}	k1 = 0. 0221 $\mu m^2.molecules^{-1}.s^{-1}$, k2 = 0.001 s^{-1}
ITPPax_Lig + FAK_PM = ITPPaxFAK_Lig	k1 = 0.00221 $\mu m^2.molecules^{-1}.s^{-1}$, k2 = 0.01 s^{-1}	k1 = 0. 0221 $\mu m^2.molecules^{-1}.s^{-1}$, k2 = 0.001 s^{-1}
ITPPaxFAK_Lig + Vinc_PM = ITPPaxFAKVinc_Lig	k1 = 0.00221 $\mu m^2.molecules^{-1}.s^{-1}$, k2 = 0.01 s^{-1}	k1 = 0. 0221 $\mu m^2.molecules^{-1}.s^{-1}$, k2 = 0.001 s^{-1}

Table 4.9. Reaction rate constant changes for the Visual Cell model.

Running this model produces the results shown in Figure 4.30. This shows that peak levels of total and individual FAs occur much earlier in the Virtual Cell model (at around 3400 seconds) compared to the COPASI model (at around 8000 seconds, Figure 4.21). However, these peaks levels are also substantially lower, at approximately around are 423, 238, 162, 7 and 20 molecules for, respectively, total FAs, ITP_Lig, ITPPax_Lig, ITPPaxFAK_Lig and ITPPaxFAKVinc_Lig in the Virtual Cell model, as against936, 327, 490, 108 and 8 molecules in the COPASI model. As these values also demonstrate, the relative levels for individual FAs species are also different between the two versions of the model, with peak levels of ITPPax_Lig exceeding ITP_Lig in the COPASI model, but showing lower levels than the latter in the Virtual Cell model. Similarly, peak ITPPaxFAK_Lig levels exceed those of ITPPaxFAKVinc_Lig in the Virtual Cell model, but are lower in the COPASI model.

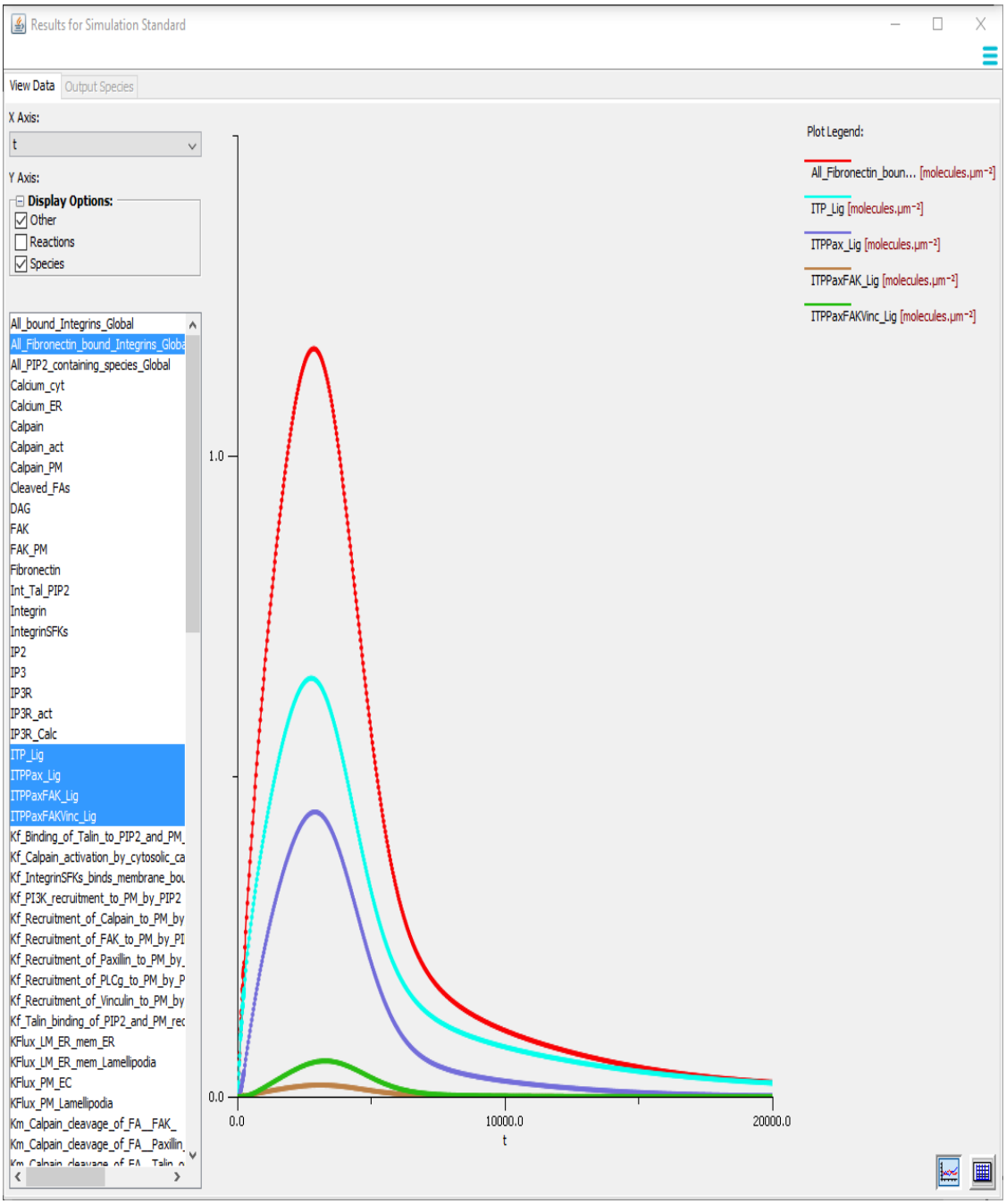

Figure 4.30. Densities (molecules.µm^2) of total FAs, ITP Lig, ITPPax_Lig, ITPPaxFAK_Lig and ITPPaxFAKVinc_Lig in Visual Cell version of the model. Corresponding peak particle levels (in rounded numbers of molecules) are 423, 238, 162, 7 and 20, respectively.

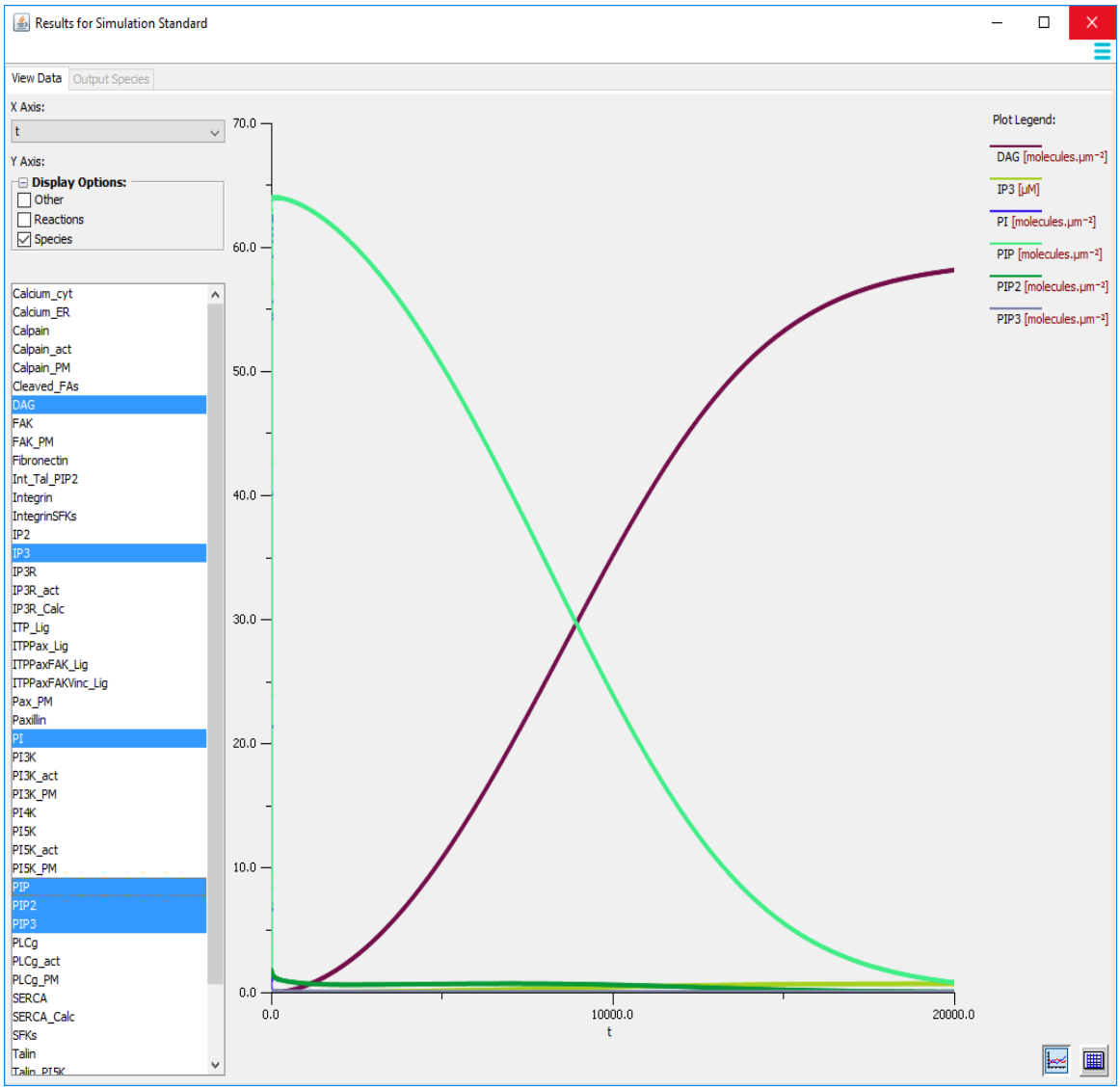

Figure 4.31. Densities (molecules.µm^2) of PI, PIP, PIP2, PIP3, IP3 and DAG for the Virtual Cell model. Peak particle levels: 22,800, 22,850, 639, 14, 20,800 and 21,200 molecules, respectively.

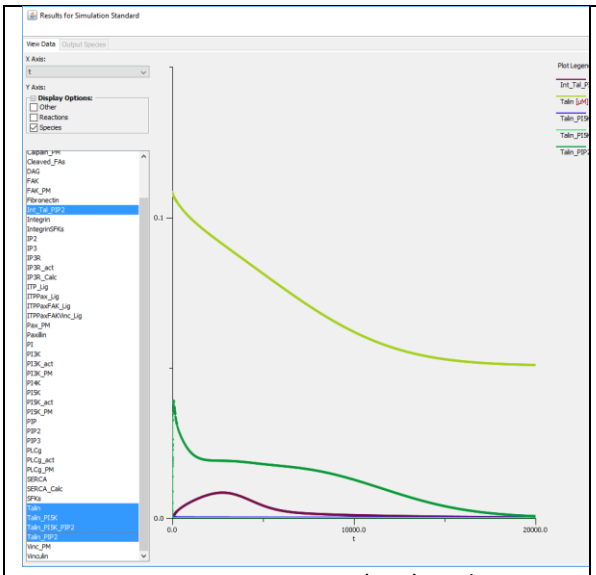

Figure 4.32. Concentrations (µM) or densities (molecules.µm^2) of Talin, Talin_PIP2, Talin_PI5K, Talin_PI5K_PIP2 and Int_Tal_PIP2 for the Virtual Cell model. Peak particle levels: 3312, 28, 1, 8, 558, 6 molecules, respectively

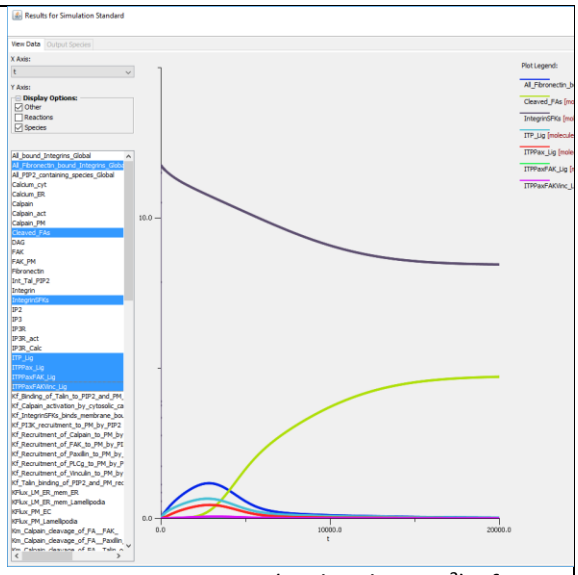

Figure 4.33. Densities (molecules.µm^2) of total FAs, ITP_Lig, ITPPax_Lig, ITPPaxFAK_Lig, ITPPaxFAKVinc_Lig, IntegrinSFKs and Cleaved_FAs for the Virtual Cell model. Peak particle levels: 423, 238, 162, 7, 20, 8500, 3438 molecules, respectively

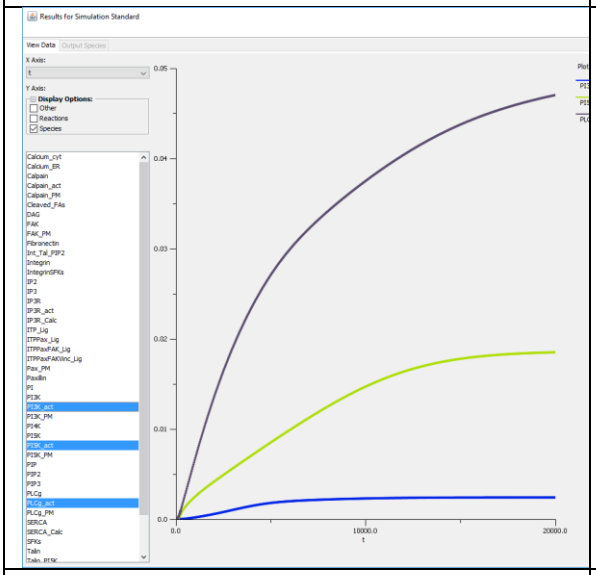

Figure 4.34. Densities (molecules.µm^2) of PI5K_act, PLCγ_act and PI3K_act for the Virtual Cell model. Peak particle levels: 7, 17 and 1 molecules, respectively.

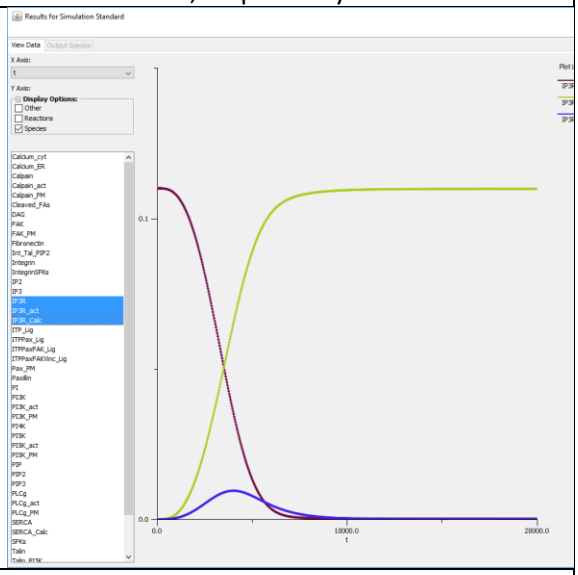

Figure 4.35. Densities (molecules.µm^2) of IP3R species for the Virtual Cell model. Peak particle levels:IP3R = 220, IP3R_Calc = 19, IP3R_act =220 molecules, respectively.

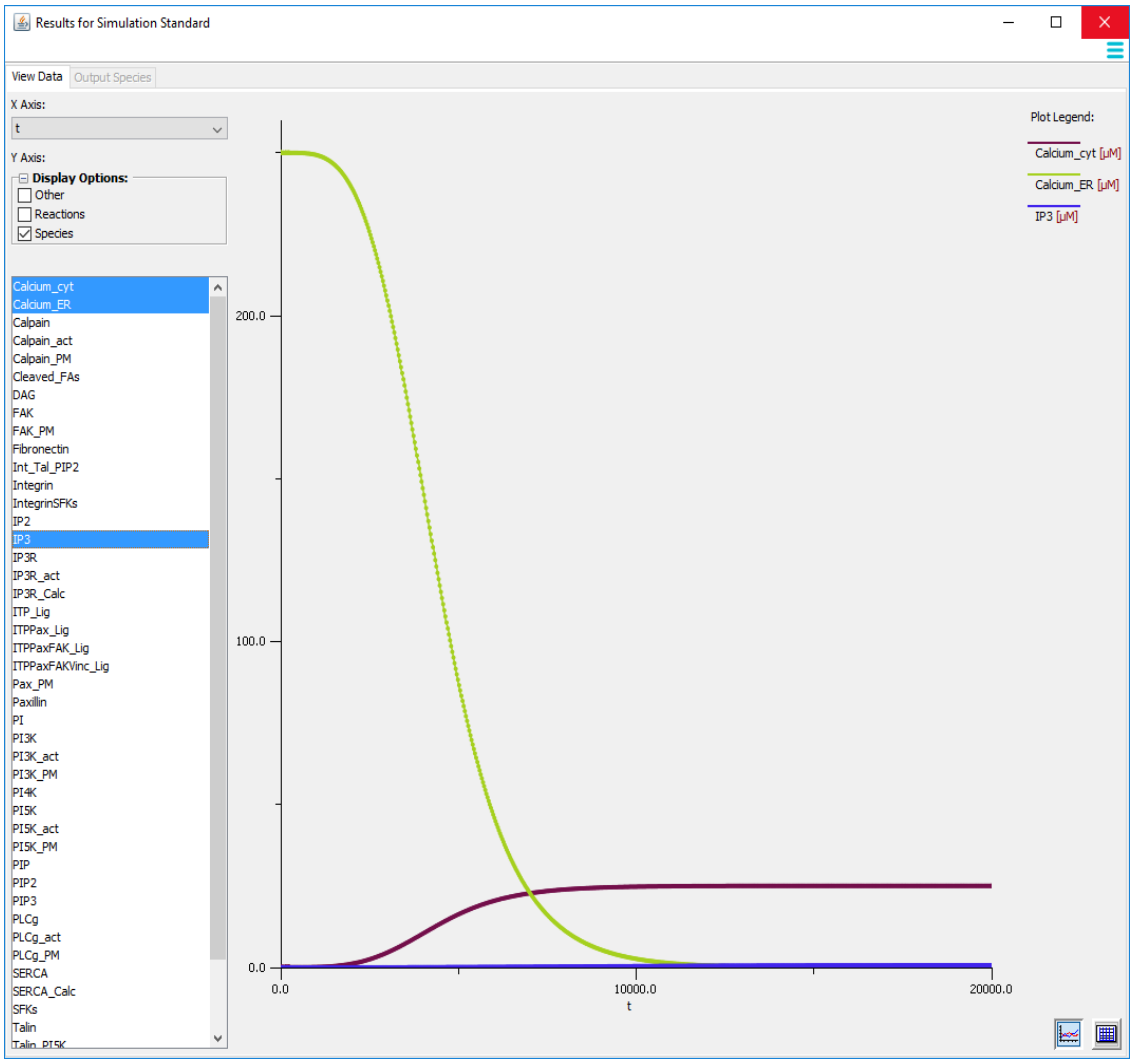

Figure 4.36. Concentrations (µM) for IP3, Ca2+_ER and Ca2+_cyto for the Virtual Cell model. Peak particle levels: 220, 752,750, 752,750 molecules, respectively.

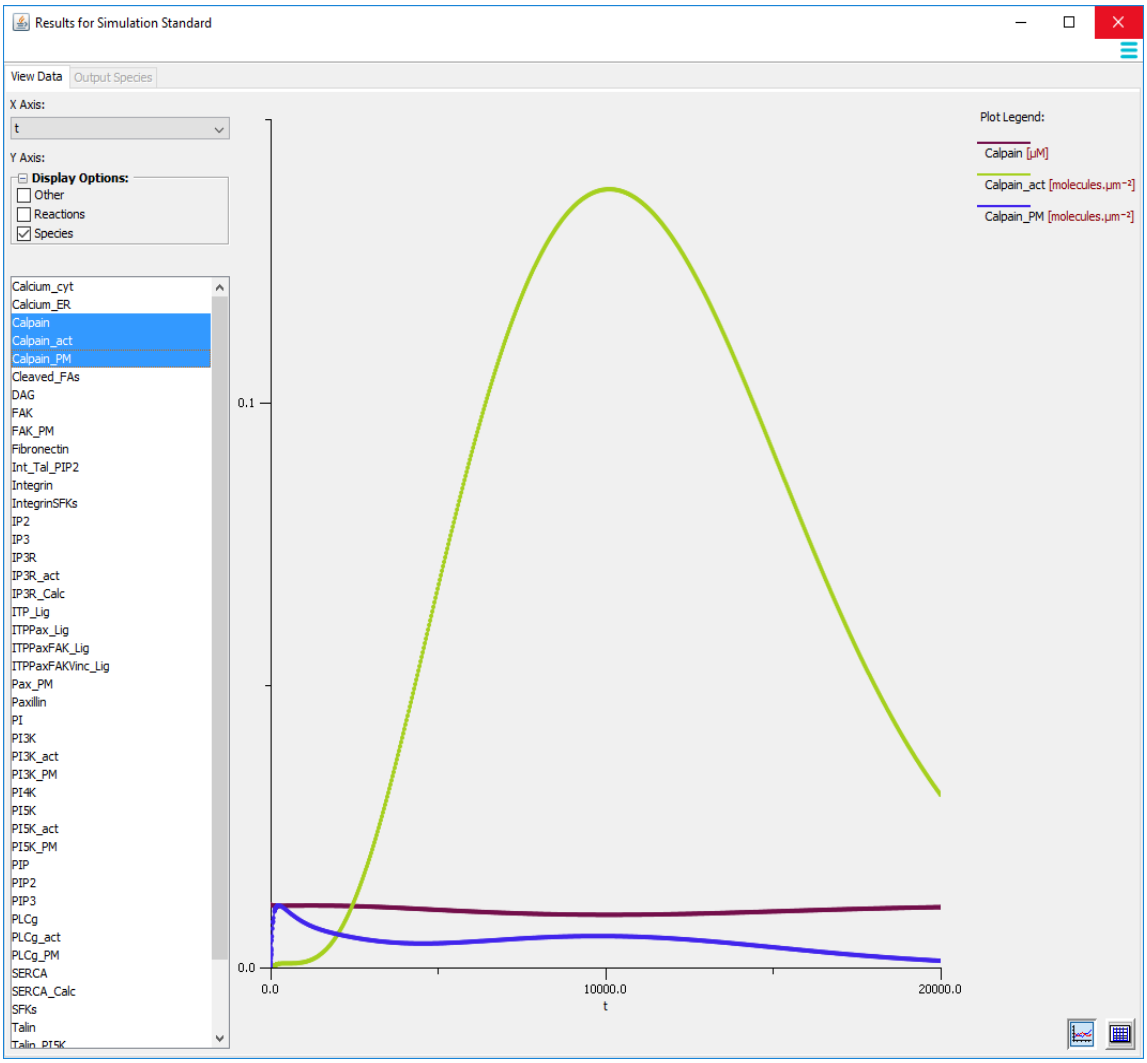

Figure 4.37. Concentrations (µM) or densities (molecules.µm^2) of all calpain species for the Virtual Cell model. Peak particle levels: Calpain = 3312, Calpain_PM = 4, Calpain_act = 50 molecules, respectively.

Understanding the reasons for the substantially lower peak levels of all of these FAs, as well as the differences in distributions, is challenging. As Table 4.10below shows, there are a number of differences, but also of similarities between the two models. For instance, peak levels of PI5K species, and of talin species, are much the same, and differences between peak levels of activated IP3R do not seem enough to explain the FA discrepancies just described. Similarly, the IP3 peak levels are nearly identical, as are the particle number curves (Figures 4.38 and factor 4.39). Despite these similarities, and the relatively small differences in peak levels of activated IP3R just mentioned (185 as against 220 molecules for the COPASI and Virtual Cell models, respectively),

peak levels of cytosolic calcium in the COPASI model (at 240,000 molecules, Figure 4.40) are less than one-third of peak levels in the Virtual Cell model (750,000 molecules, Figure 4.41) after 20,000 seconds elapsed.

Species	COPASI peak level (no. molecules)	COPASI peak time (s)	VCell peak level (no. molecules)	VCell peak time (s)
All FAs	937	8100	423	3000
ITP_Lig	324	7800	238	2800
ITPPax_Lig	492	8160	162	2870
ITPPaxFAK_Lig	10	10,500	7	3100
ITPPaxFAKVinc_Lig	110	8300	20	3300
Int_Tal_PIP2	4	7800	3	2700
IntegrinSFKs	4250	100	4271	100
Talin	3312	0	3312	0
Tal_PIP2	10	6500	14	100
Tal_PI5K	9	100	9	100
Tal_PI5K_PIP2	~0	100 & 6000	~0	100
PI5K	500	0	500	0
PI5K_PM	~0	100	1	20,000
PI5K_act	5	20,000	7	20,000
PLCγ_act	61	20,000	17	20,000
PI3K_act	7	20,000	1	20,000
PI	23,200	0	22,800	0
PIP	23,200	100	22,850	100
PIP2	727	7000	639	0
PIP3	50	13,500	14	0
DAG	20,000	20,000	21,200	20,000
IP3	19,500	20,000	20,800	20,000
IP3R	220	0	220	0
IP3R_Calc	34	17,000	19	4000
IP3R_act	185	20,000	220	10,000
Ca2+_ER	750,000	0	750,000	0
Ca2+_cyto	240,000	20,000	750,000	20,000
Calpain	333	0	333	0
Calpain_PM	8	7500	4	100
Calpain_act	7	13,000	50	10,000
Cleaved_FAs	1574	20,000	1718	20,000

Table 4.10. Comparison of peak levels and the timing of these peak levels of various model species, between the second version of the COPASI model and the Virtual Cell model.

This implies that the ER is releasing more calcium into the cytoplasm, in response to a given level of IP3 stimulation, in the Virtual Cell model than it is releasing in the COPASI model. This is confirmed by a comparison of reaction flux rates for IP3R-IP3 binding on the one hand (Figures

4.42 and 4.43), and for IP3R-mediated ER calcium loading and subsequent unloading into the cytosol (Figures 4.44 and 4.45).

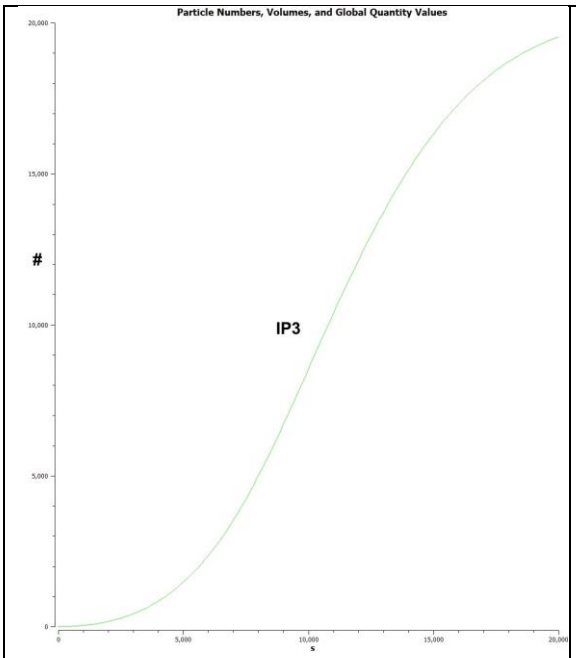

Figure 4.38. Particle numbers for IP3 in COPASI model. Peak particle levels: 19,500 molecules.

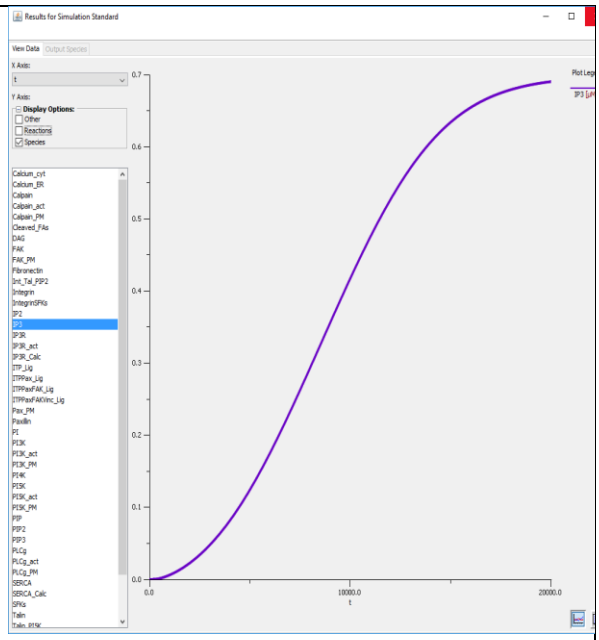

Figure 4.39. Concentrations (µM) of IP3 for Virtual Cell model. Peak particle levels: 20,800 molecules.

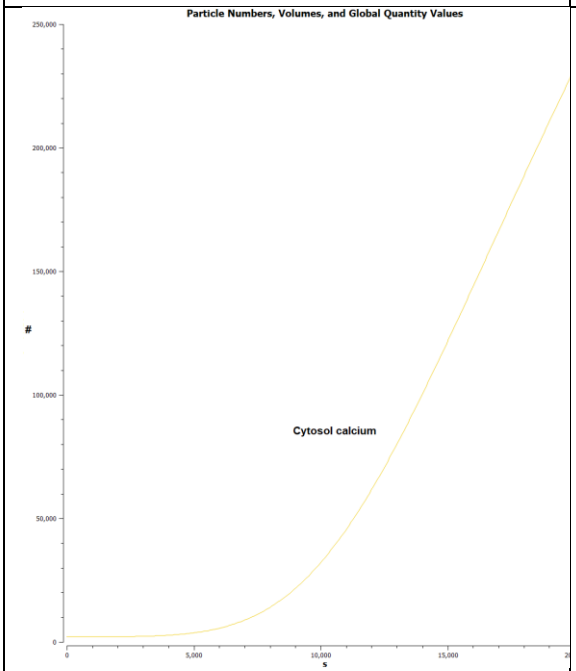

Figure 4.40. Particle numbers for cytosolic calcium in COPASI model. Peak particle levels: 240,000 molecules.

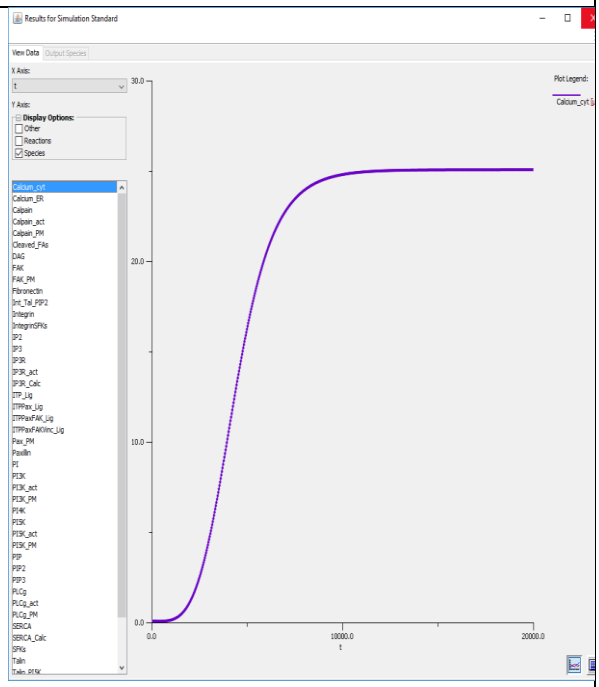

Figure 4.41. Concentrations (µM) of cytosolic calcium in Virtual Cell model. Peak particle levels: 750,000molecules.

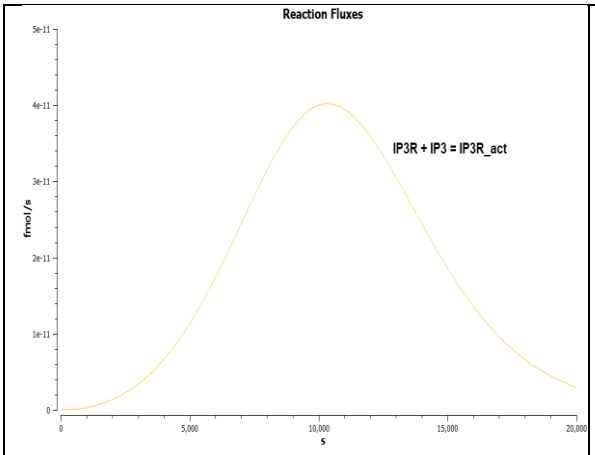

Figure 4.42. Reaction flux (fmol/s) for IP3R activation by IP3 in COPASI model. (Peak flux rate equivalent to 1.20e-05 molecules.µm^{-2}.s^{-1}.)

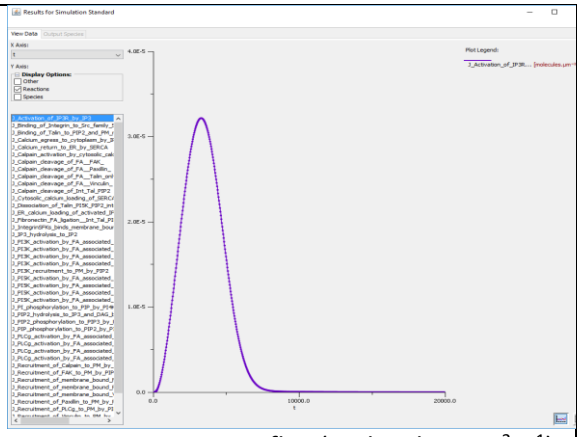

Figure 4.43. Reaction flux (molecules.µm^{-2}.s^{-1}) for IP3R activation by IP3 in Virtual Cell model. (Peak flux rate = 3.21e-05 molecules.µm^{-2}.s^{-1}.)

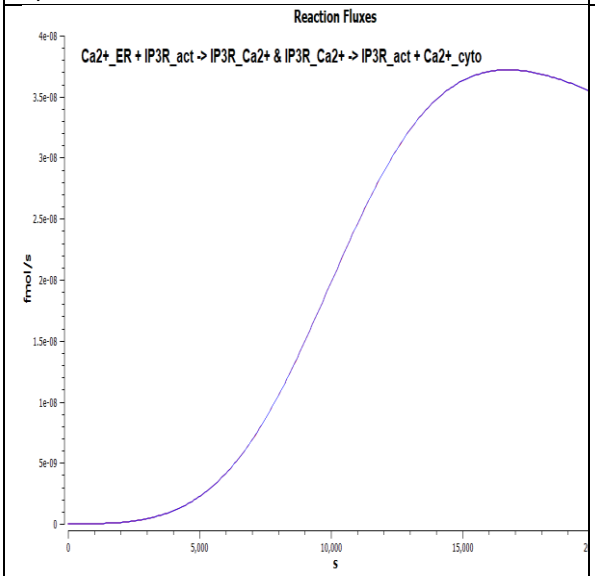

Figure 4.44. Reaction fluxes (fmol/s)for IP3R_act loading by ER calcium and unloading into cytosol in COPASI model. (Identical curves, in both cases, hence superimposed. Peak flux rates equivalent to 1.12e-02 molecules.µm^{-2}.s^{-1}, in both cases)

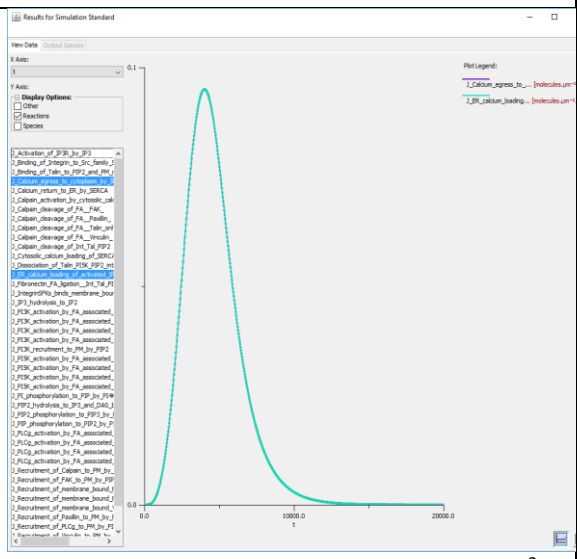

Figure 4.45. Reaction fluxes (molecules.µm^{-2}.s^{-1}) for IP3R_act loading by ER calcium and unloading into cytosol in Virtual Cell model. (Identical curves in both reactions, hence superimposed.) Peak flux rate = 9.5e-02 molecules.µm^{-2}.s^{-1}, in both cases.)

As Figures 4.42-4.45 show, IP3 binding to IP3R occurs at nearly 3 times the rate (3.21e-05 as against = 1.20 e-05 molecules.µm^{-2}.s^{-1}) in the Virtual Cell model as compared with the COPASI model, and IP3R-mediated calcium egress from the ER into the cytosol proceeds at more than eight times the rate (9.5e-05 as against 1.12e-02 molecules.µm^{-2}.s^{-1}). This obviously explains why

calcium egress into the cytosol is much more rapid in the Virtual Cell model, as demonstrated in Figures 4.40 and 4.41.

Given that calpain activation is dependent on calcium binding, this in turn explains why peak levels of active calpain (Calpain_act) are so much higher in the Virtual Cell model (50 molecules, Figure 4.37) than in the COPASI model (7 molecules, Figure 4.46) and these peak are achieved earlier in the Virtual Cell model (10,000 seconds), as compared with the COPASI model (13,000 seconds).

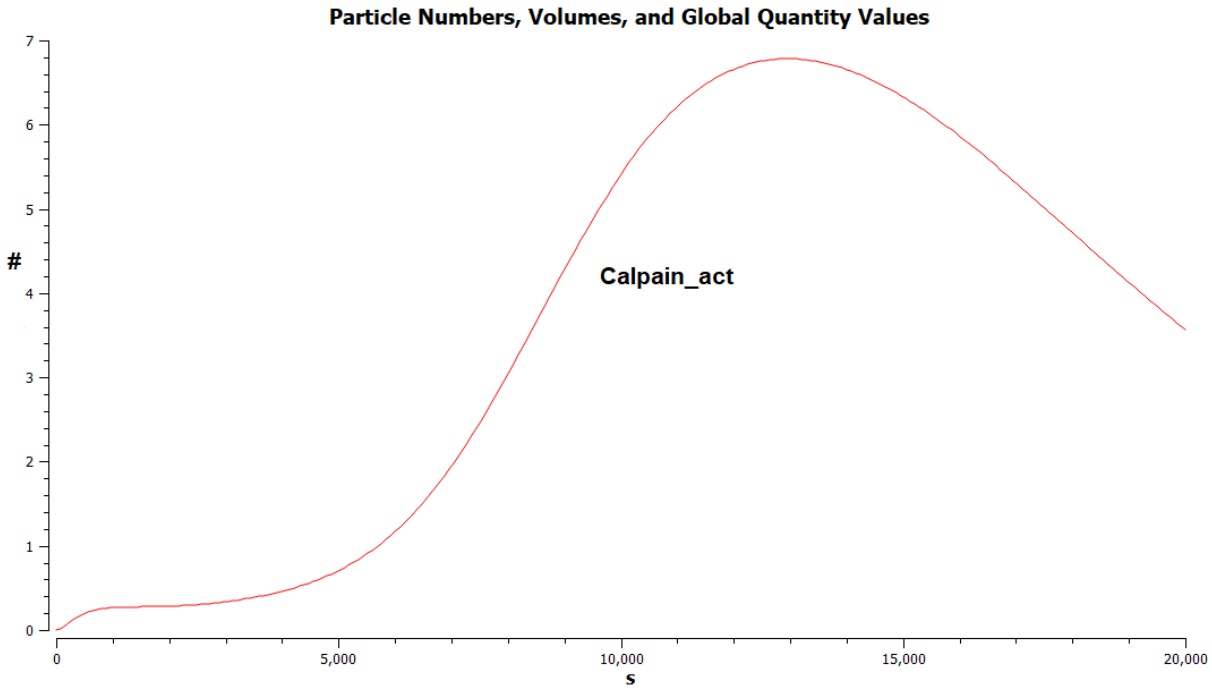

Figure 4.46. Particle numbers for Calpain_act in second version of COPASI model.

This would go a long way to explaining why, overall, FAs achieve higher peak levels in the COPASI model than in its Virtual Cell equivalent. This raises the question, why are these IP3R-mediated reaction rates so different between the two models?

In the case of two of the three reactions involved, the two second-order reactions (the forward reactions for IP3R-IP3 and IP3R_act-Calcium_ER binding) are not intra-membrane reactions, given

that one reactant (IP3, in the first case and ER calcium, in the second, are resident in liquid compartments). Moreover, both the remaining reactions (the reverse rates for IP3R-IP3 binding, and the forward rate for the irreversible calcium-unloading reaction - IP3R_Calc ->Calcium_cyt + IP3R_act) are all first-order (with rate constant values of 1.0E-5 and 10 per second, respectively).

As a consequence, no conversion of any of these four reaction rate values has been necessary in order to implement the original COPASI model reactions in the Virtual Cell model.(Specifically, it has not been necessary to convert volumetric rate constant values in M^{-1} s^{-1} into areametric$\mu m^2.s^{-1}.molecules^{-1}$). So it is not immediately obvious why these reactions are proceeding at a much higher rate in the latter model, compared to the former.

Nevertheless, a logarithmic parameter scan on the forward rate constant (kf) for IP3R-IP3 binding confirms that such alterations have substantial effects both on cytosolic calcium release and on absolute and relative levels of FAs (Figures 4.42 and 4.43). Therefore we choose a kf value that results in a cytosolic calcium particle number curve most closely matching that of the COPASI model (kf = 1.5e-05 μM^{-1} s^{-1}). This gives us a second version of the Virtual Cell compartmental/ODE.

4.4.2.1.2 Version 2 (of Virtual Cell 2D Compartmental/ODE model)

Running a simulation with the change to the IP3R-IP3 binding forward rate (kf) just mentioned, results in FA levels shown in Figure 4.44.

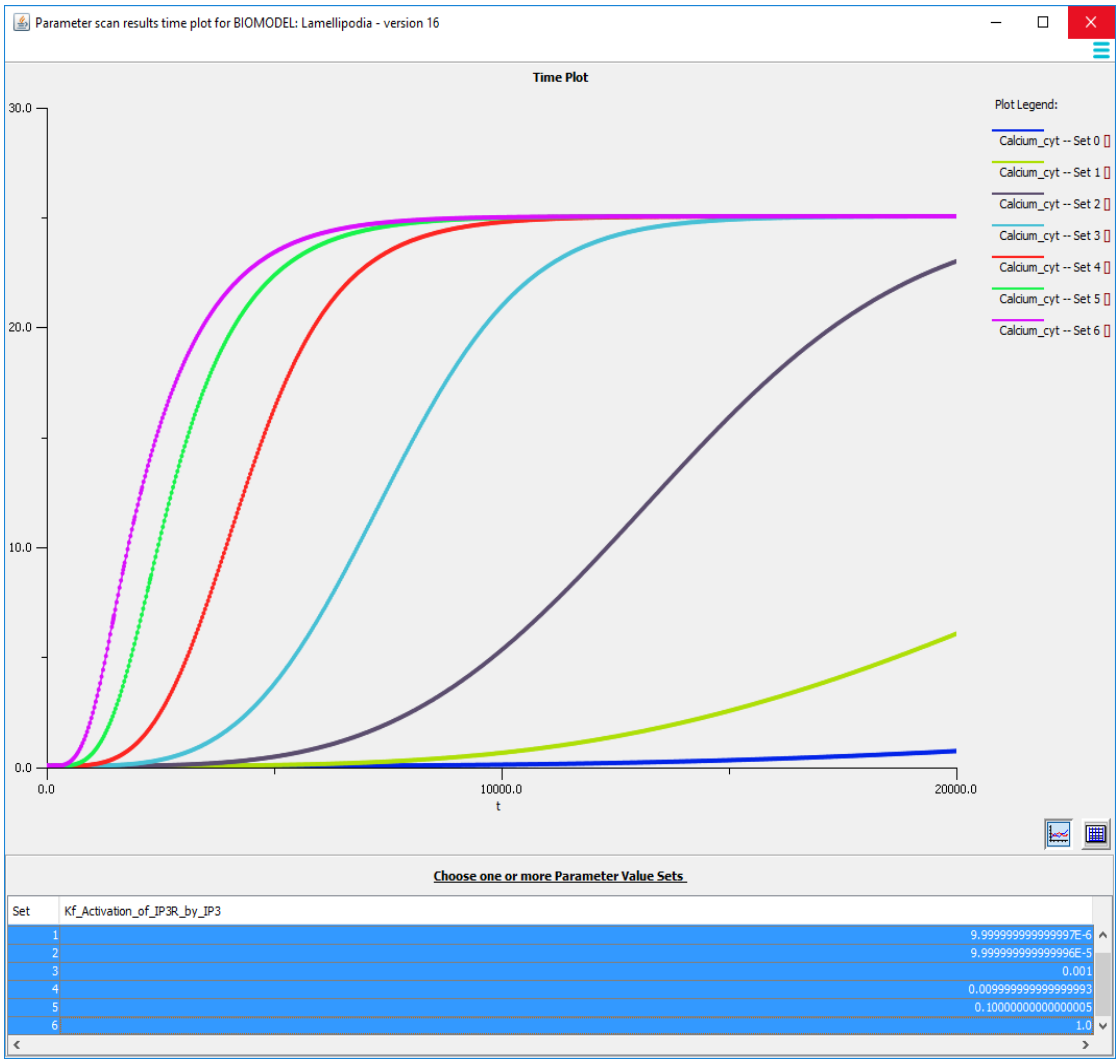

Figure 4.47. Logarithmic parameter scan on forward rate constant (kf) for IP3R-IP3 binding, showing effects on cytosolic calcium (µM) in the Virtual Cell model. 7 kf values: 1e-6, 1e-5, 1e-4, 1e-3, 1e-2, 1e-1, 1.0 $\mu M^{-1}\ s^{-1}$.

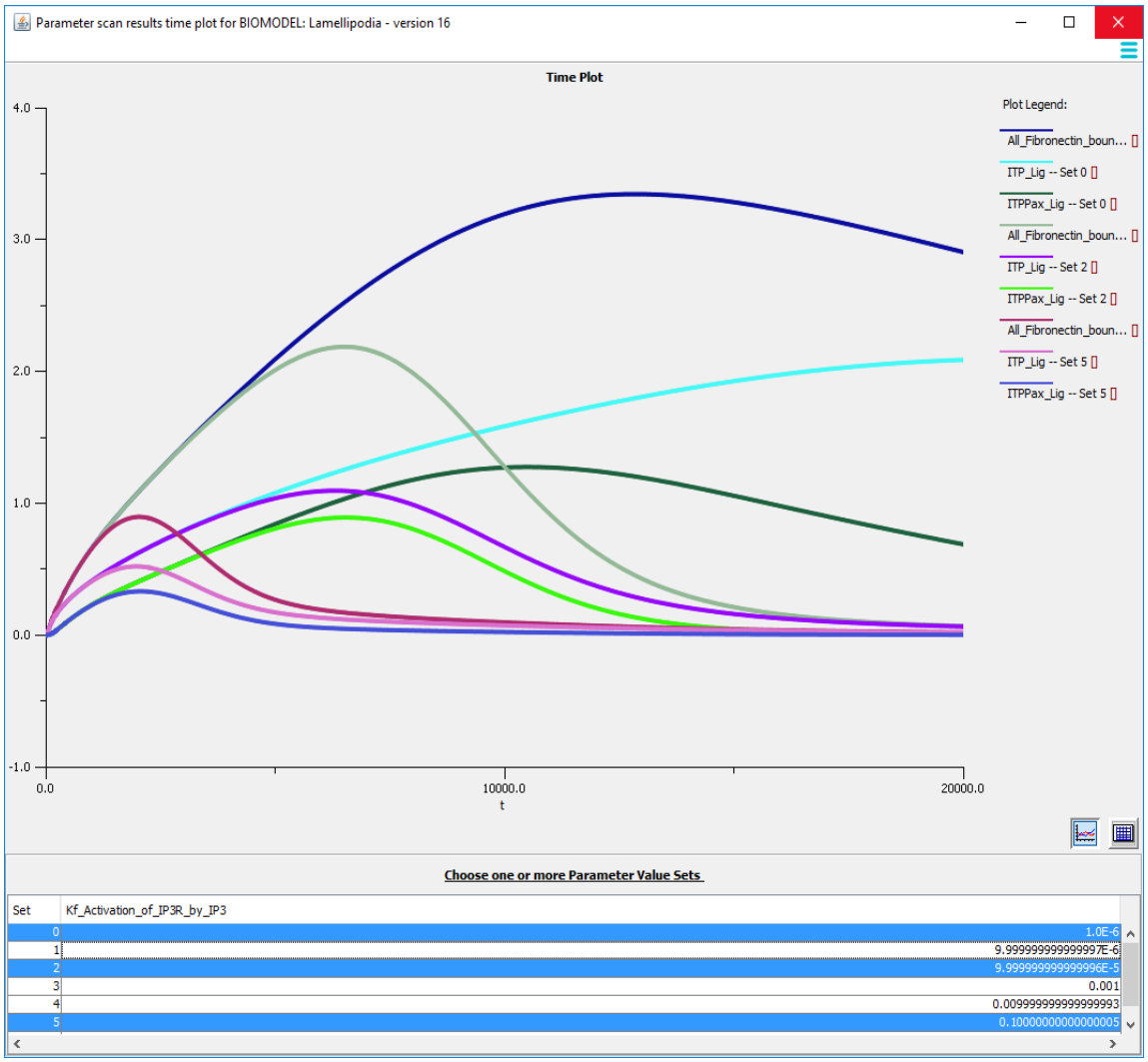

Figure 4.48. Sample results from logarithmic parameter scan on forward rate constant (kf) for IP3R-IP3 binding, showing effects on levels of total FAs, ITP_Lig and ITPPax_Lig (molecules. μM^{-2}) in the Virtual Cell model. 3kf values: 1e-6, 1e-4, 0.1 μM^{-1} s^{-1}.

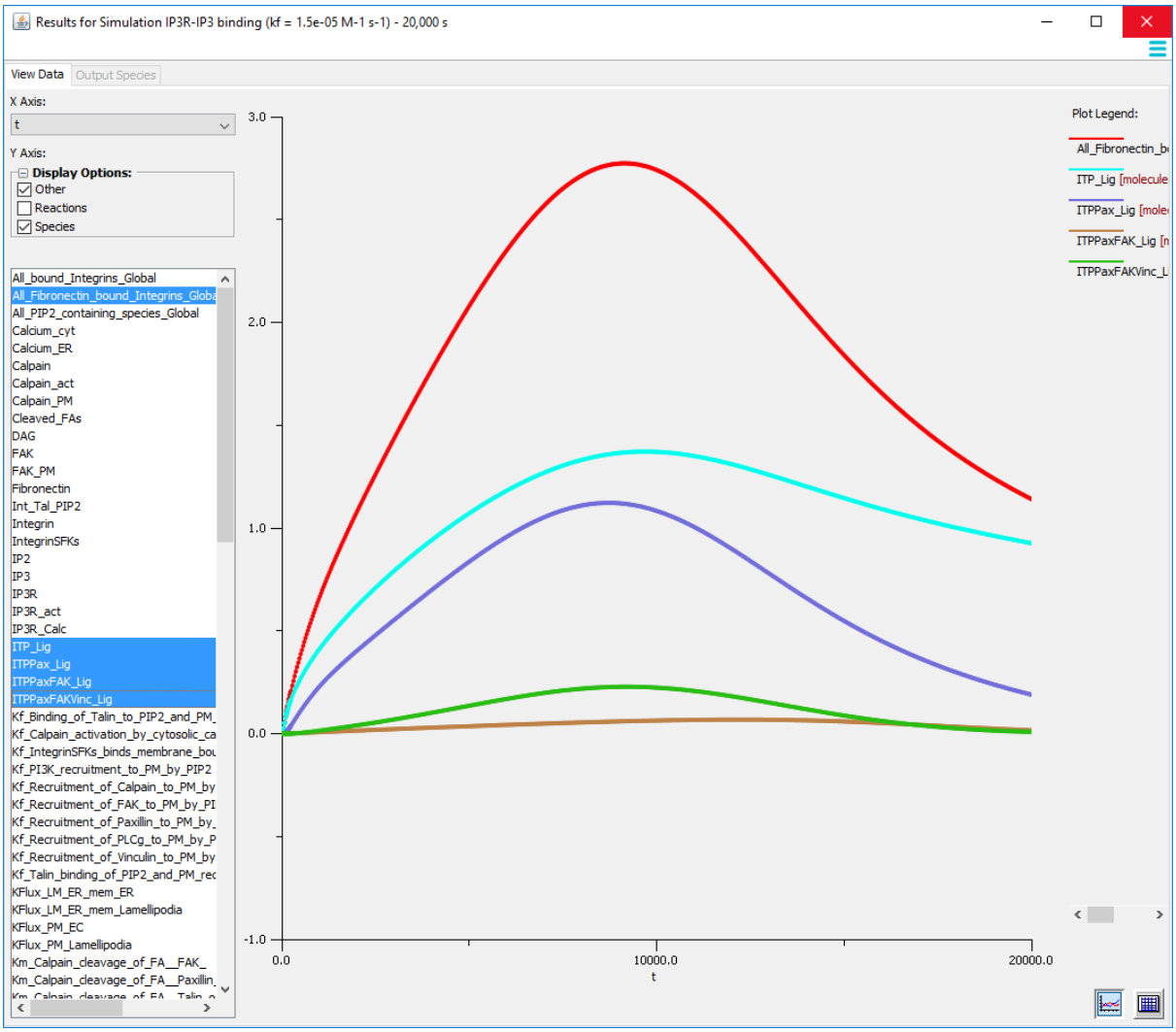

Figure 4.49. Results of reducing forward rate constant value (kf = 1.5e-5 μM^{-1} s^{-1}) for IP3R-IP3 binding on levels of total FAs and individual FAs species (molecules. μM^{-2}) in the Virtual Cell model. Peak particle levels (number of molecules): Total FAs = 1011, ITP_Lig = 500, ITPPax_Lig = 409, ITPPaxFAK_Lig = 26, ITPPaxFAKVinc_Lig = 83.

Comparing Figures 4.21and 4.48, and the results in Table 4.10 above, it can be seen that the new results are now much more similar to the COPASI model. The main difference, of course, is that levels of ITP_Lig are higher than those of ITPPax_Lig in this adjusted version of the Virtual Cell model (500 as against 409 molecules), whereas they are lower in the COPASI model (324 as against 492 molecules).

As one would expect, the main reason for this greater similarity in FA levels is because the reduction in cytosolic calcium leads to reduced calpain activation (Figures 4.50 and 4.51) and

subsequent cleavage of FAs (Figures 4.52 and 4.53), compared to the non-adjusted version of the models and much more in line with the level seen in the COPASI model (Table 4.10).

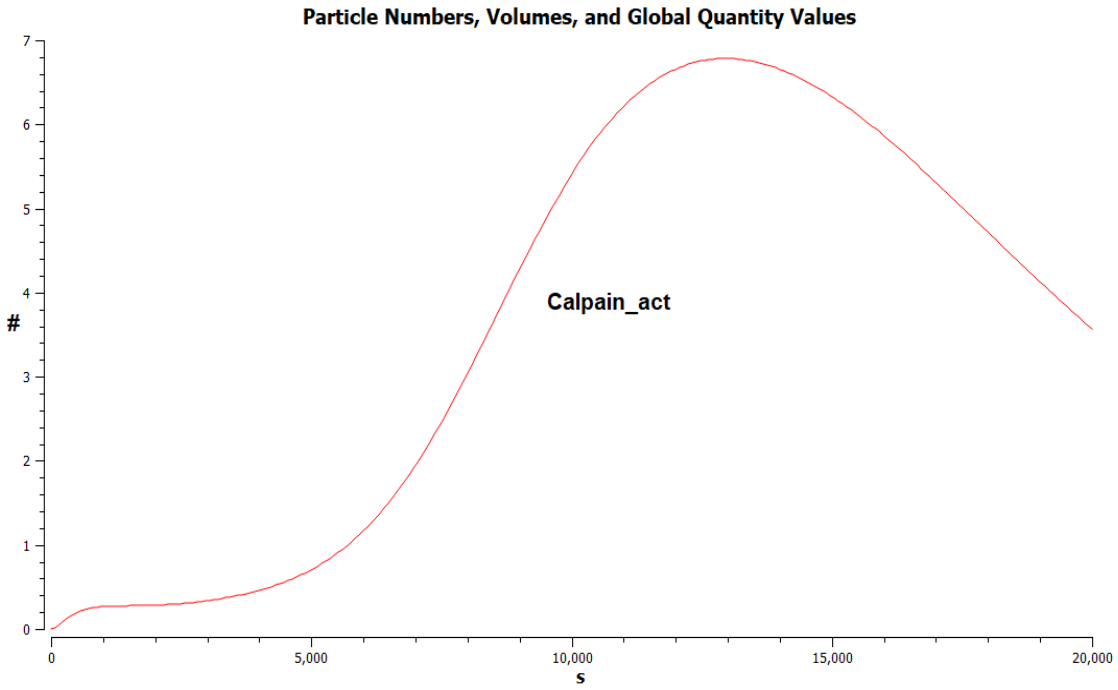

Figure 4.50. Particle levels for Calpain_act in COPASI version of the model.

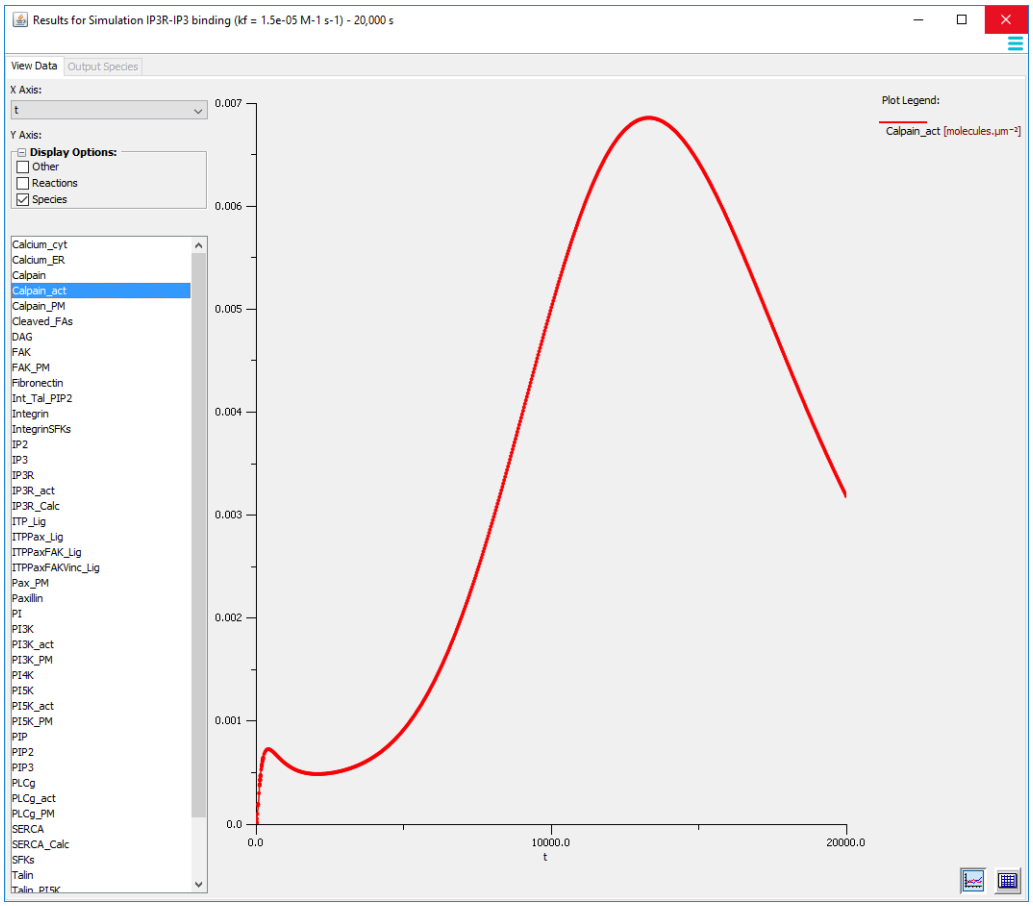

Figure 4.51. Densities (in molecules. μM^{-2}) of Calpain_act in Virtual Cell version of model, adjusted for higher cytosolic calcium release. (Peak particle levels = 3 molecules.)

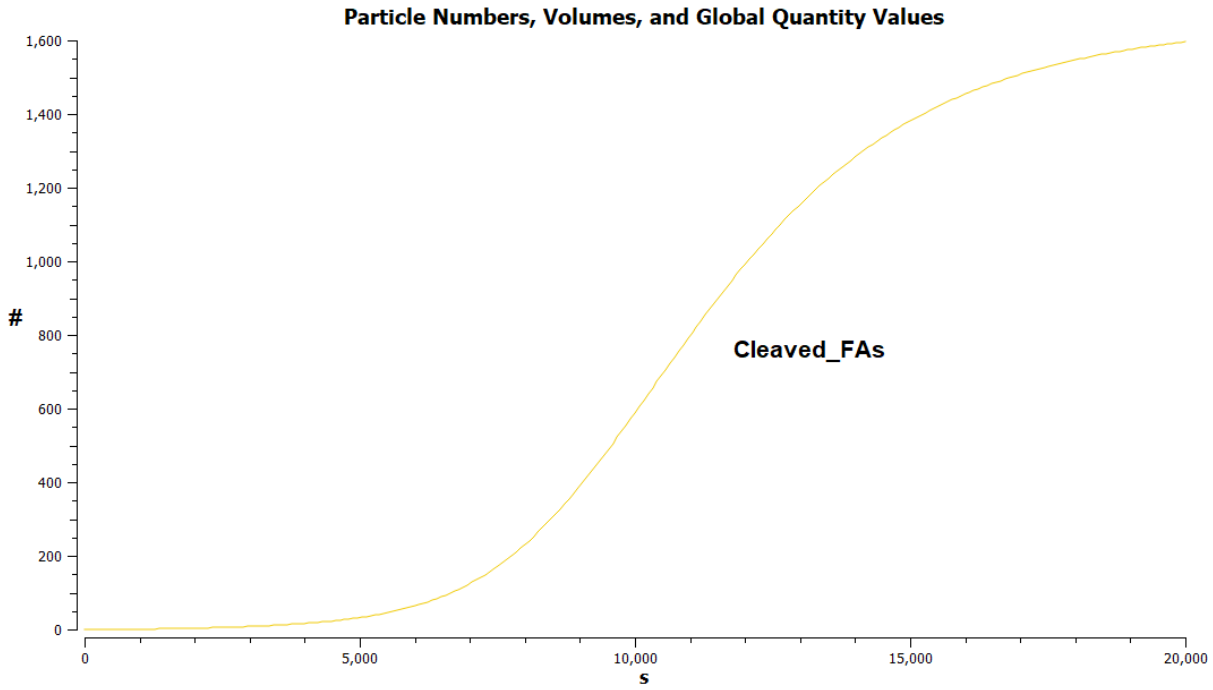

Figure 4.52. Particle levels for Cleaved_FAs in COPASI version of the model.

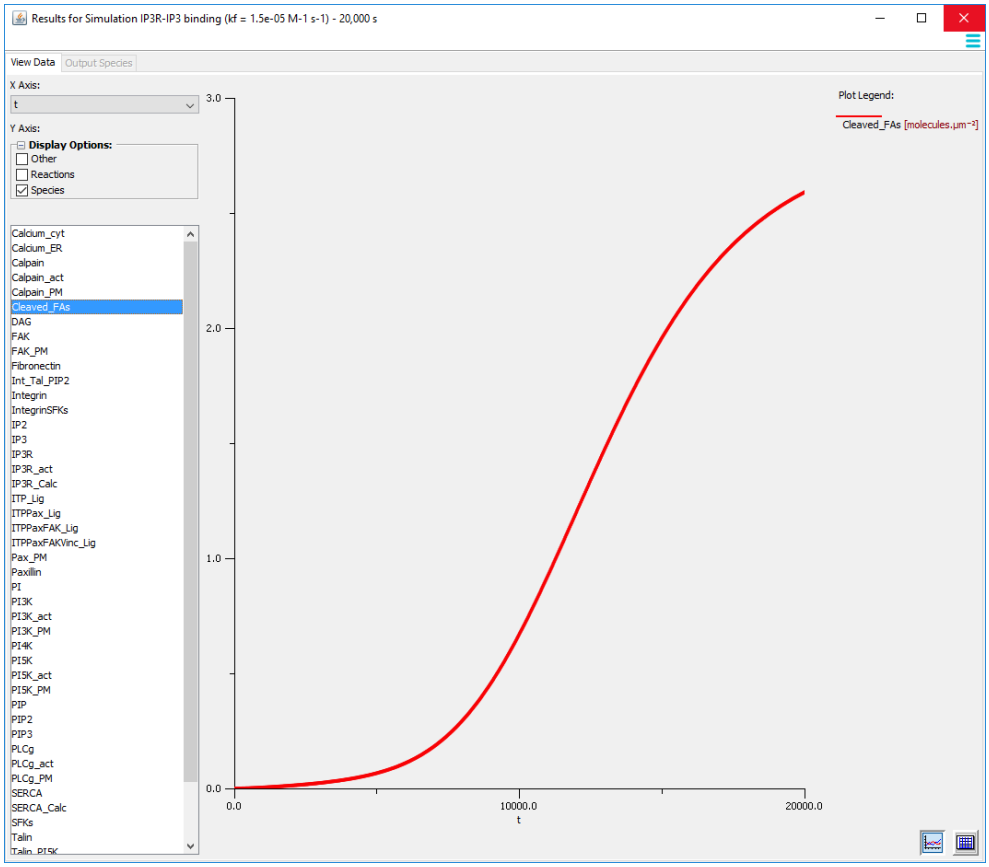

Figure 4.53.Densities (in molecules. μM^{-2}) of Cleaved_FAs in the Virtual Cell version of model. (Peak particle levels = 945 molecules.)

These findings raise interesting questions. Firstly, it seems clear that calpain cleavage is at least as influential on FA levels, in this model, as is PIP2. Looking at Table 4.10,levels of unbound PIP2 are identical in the COPASI model and in the second version of the Virtual Cell model, and a comparison of reaction flux curves for PIP2 creation from PIP for the two models (Figures 4.54 and 4.55) shows that actual levels of PIP2 creation are broadly similar.. Moreover, IP3 creation rates have also been shown to be similar, and peak PIP3 levels would seem to be too low to have much influence on FA levels. Yet these levels are still substantially different between the COPASI and the (unadjusted) Virtual Cell compartmental models.

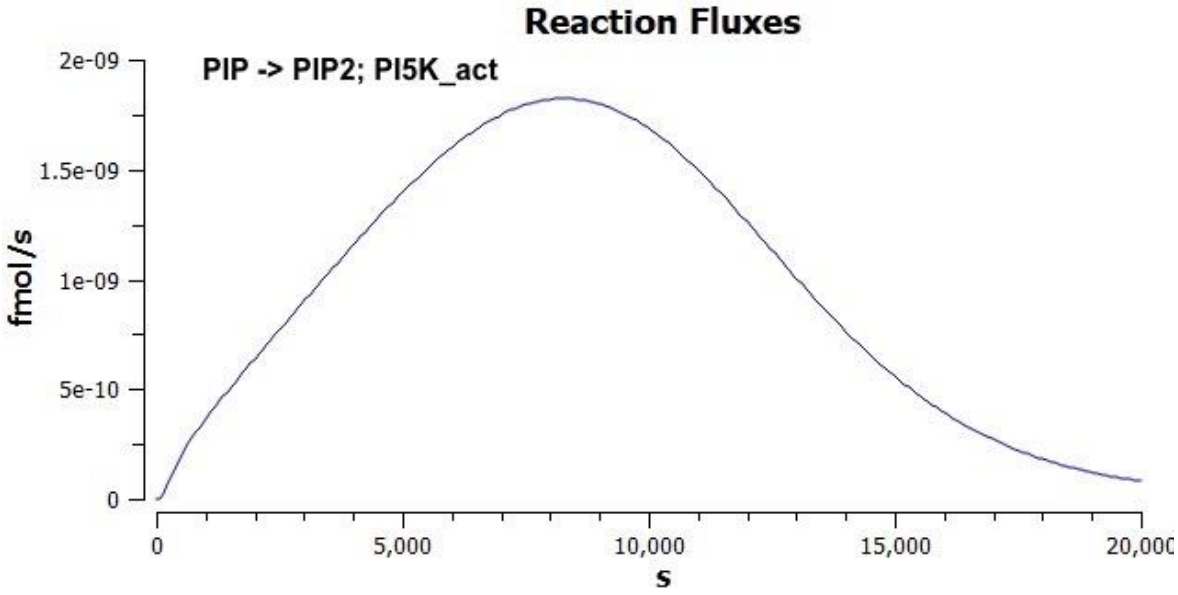

Figure 4.54. Reaction flux curve for PIP2 creation from PIP2 by PI5K in COPASI version of the model. (Peak rate equivalent to 5.21e-03 molecules.$\mu m^{-2}.s^{-1}$ in the Virtual Cell model below.)

Nevertheless, this discrepancy is, to a large extent, restored by equalising calcium release and thereby reducing rates of calpain cleavage in the Virtual Cell model. Despite this equalising effect, differences in relative levels of individual FAs species remain. This raises the possibility that this FA heterogeneity is in some way linked to differences in the way unbound PIP2 is taken up by other species in the two models. However, determining whether this is the case, and the precise mechanism behind it, would require a detailed investigation. Given that this seems likely to be more of a technical matter, probably relating to differences between how COPASI and Virtual Cell implement factors such as membranes, rather than relating to FA biology, such detailed investigation does not seem justified in the circumstances.

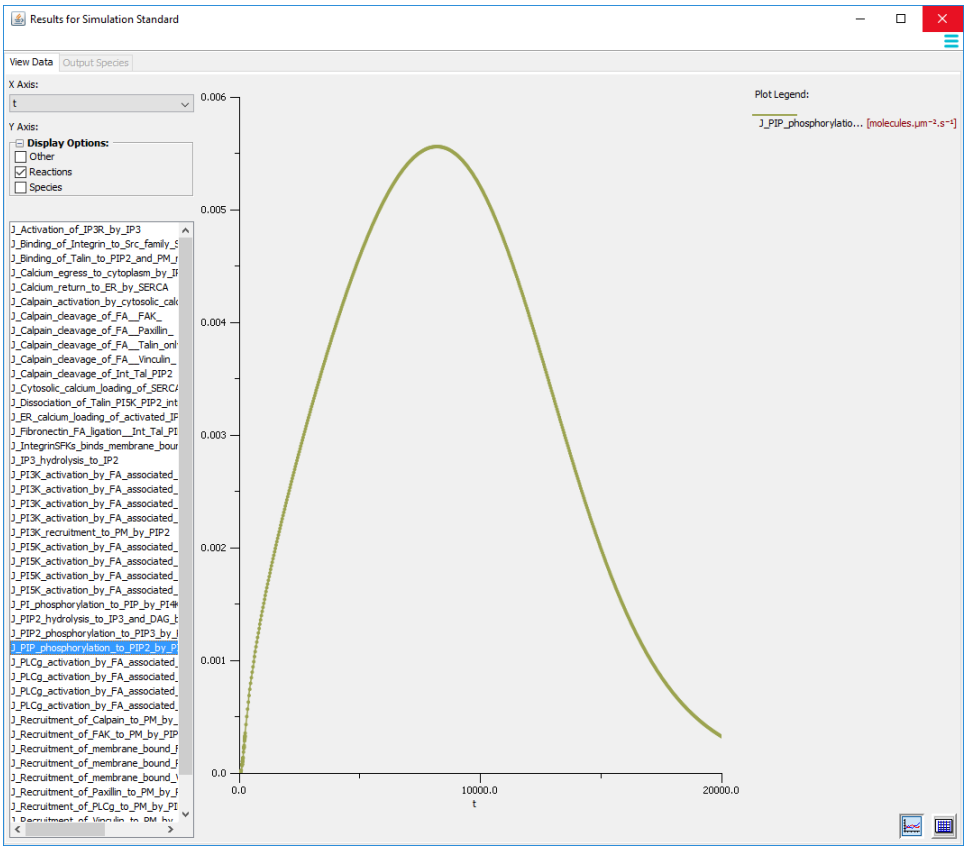

Figure 4.55. Reaction flux curve for PIP2 creation from PIP by PI5K in second Virtual Cell model.(Peak rate: 5.56e-03 molecules.µm^{-2}.s^{-1}.)

Having established the differences between the COPASI and Virtual Cell compartmental models, and seen how such comparisons can give unexpected insights, it would be interesting to see if extending the Virtual Cell model to a 2D model spatial version offers similar insights.

4.4.2.2 Virtual Cell 2D spatial model.

4.4.2.2.1 Version 1

Referring back to Table 4.1 in the Model Setup section, adjusted so that the PM is now half the length originally allotted to it, the 2D spatial version has the desired dimensions shown in Table 4.11.

Compartment	Area/length in Virtual Cell 2D model
Extracellular	~540 μm^2
Plasma membrane	~2433µm *
Cytosolic	325 μm^2
ER membrane	13,333 µm *
Endoplasmic reticulum	~35 μm^2

Table 4.11. Areas and lengths intended for the five model compartments in the Virtual Cell 2D spatial model. Areas are used for Virtual Cell 2D liquid compartments, whilst lengths are used for 2D model membranes. * Owing to the challenges of specifying 2Dmodel geometries of the desired size, using the available geometry specification methods in Virtual Cell, incorrect lengths had to be adopted for membranes and then mapped to the desired values at the simulation specification stage, as explained in the text.

The procedure for creating this 2D spatial model is broadly the same as for the 2D version of the model in Chapter 3, except that there are five compartments here instead of three. Consequently, two analytic equations have to be specified for the corresponding liquid compartments (the cytosol and ER). Also, whereas a circular geometry was sufficient for modelling the whole HeLa cell, we will adopt a tongue-like geometry to represent a lamellipodia, and a simple rectangle to represent the ER, resulting in the following two analytic equations:

Cytosol: ((((x ^ 2.0) / (1.3 ^ 2.0)) + ((y ^ 2.0) / (2.7 ^ 2.0))) <= 65.4)

ER: ((x >= - 1.0) && (x <= 1.0) && (y >= 1.5) && (y <= 18.0))

This results in the model geometry seen in Figure 4.56.

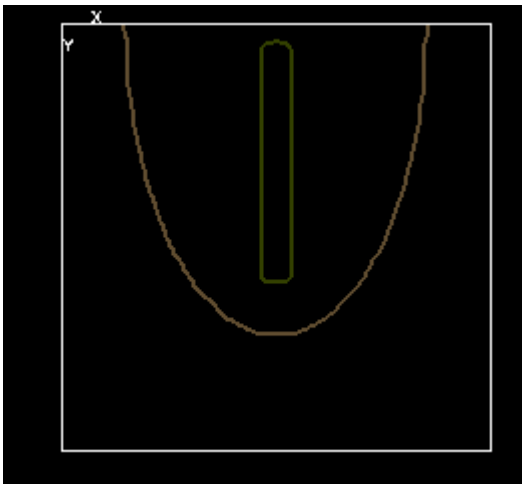

Figure 4.56. Geometry for Virtual Cell 2D spatial model of lamellipodia.

This results in the exercises for the various compartments shown in Table 4.12.

Compartment	Area/length in Virtual Cell 2D model
Extracellular	539.9 μm^2
Plasma membrane	52.6 µm
Cytosolic	324.9 μm^2
ER membrane	36.7 µm
Endoplasmic reticulum	35.3 μm^2

Table 4.12. Actual areas and lengths of the five model compartments in the Virtual Cell 2D spatial model, as specified in geometry definition.

In other words, the three liquid compartments are all close enough to the desired areas, but the two automatically-generated membrane compartments have substantially shorter lengths than desired. As described for the 2D spatial version of the HeLa cell model in Chapter 3 (Section 3.4, Figure 3.62), this must subsequently be remedied by overriding the corresponding AreaPerUnitArea default values in the Parameters section for all simulations of this 2D spatial model, shown in Table 4.13.:

Parameter Name	Default Value	New Value/Expression
AreaPerUnitArea_PM	1.0	46.3
AreaPerUnitArea_ER_membrane	1.0	363.4

Table 4.13. Override values used for converting actual membrane lengths to desired values for Virtual Cell 2D spatial model of lamellipodia FA dynamics.

Having made these changes, the simulation was then run for 20,000 seconds.

Again, as in the HeLa cell model in the previous chapter, running a simulation of this 2D spatial model produces results that are essentially identical to that of the Virtual Cell compartmental model, as just described. This can be seen by comparing Figures 4.57-4.62 with Figures 4.30, 4.31, 4.34, 4.35, 4.39 and 4.41above, respectively. Similarly, after modifying the IP3R-IP3 binding reaction forward rate constant as described above, the results are again identical, as can be seen by comparing Figure 4.63 with 4.49 above, Figures 4.64 with 4.65, and 4.66 with 4.67.

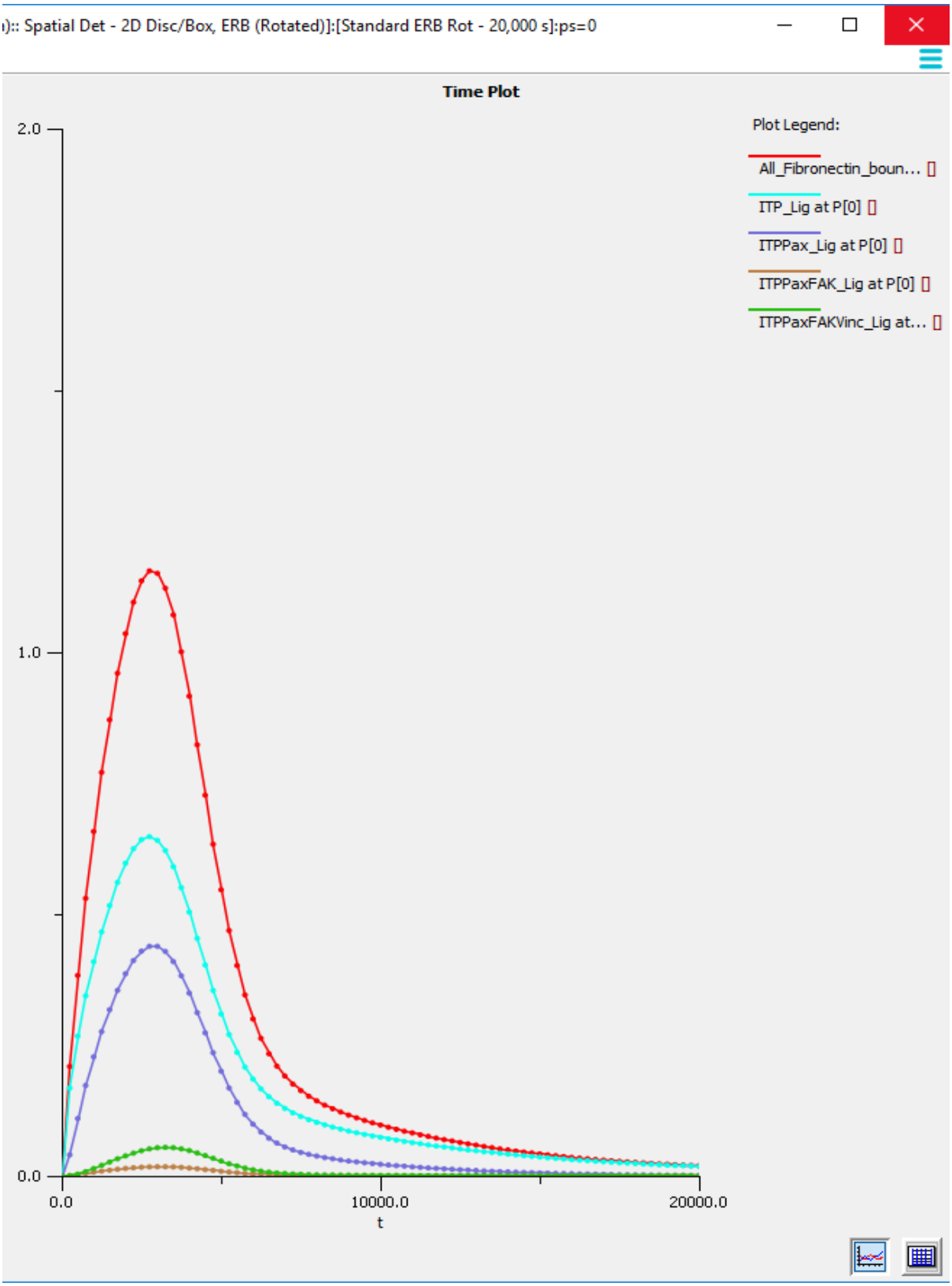

Figure 4.57.Densities (molecules.µm^2) of total FAs, ITP Lig, ITPPax_Lig, ITPPaxFAK_Lig and ITPPaxFAKVinc_Lig in Visual Cell version of the model. Corresponding peak particle levels (in rounded numbers of molecules) are 422, 236, 160, 6 and 21, respectively.

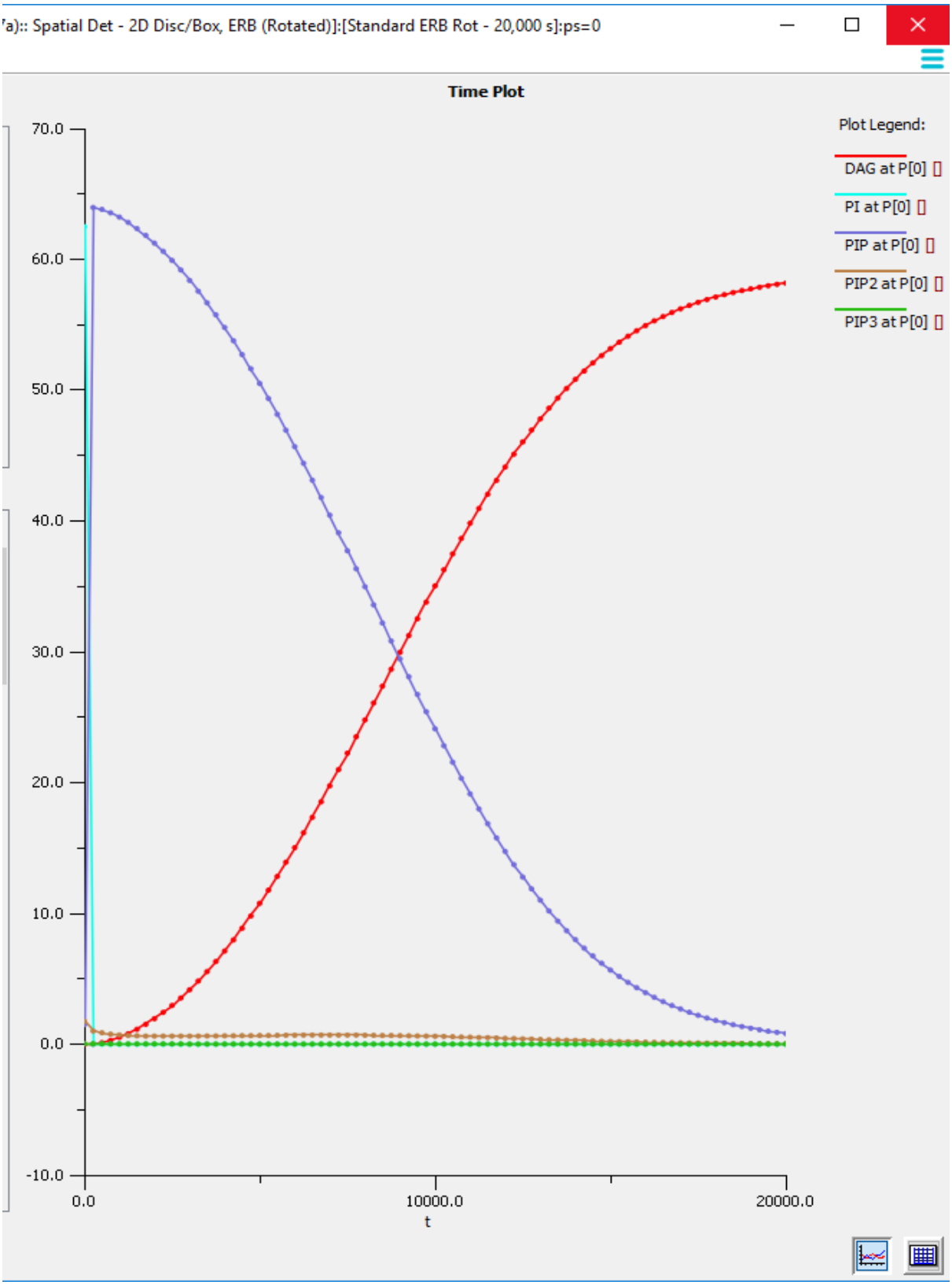

Figure 4.58. Densities (molecules.µm^2) of PI, PIP, PIP2, PIP3, IP3 and DAG for the Virtual Cell 2D spatial model. Peak particle levels: 22,800, 23,350, 639, 13,21,152 and 21,236 molecules, respectively.

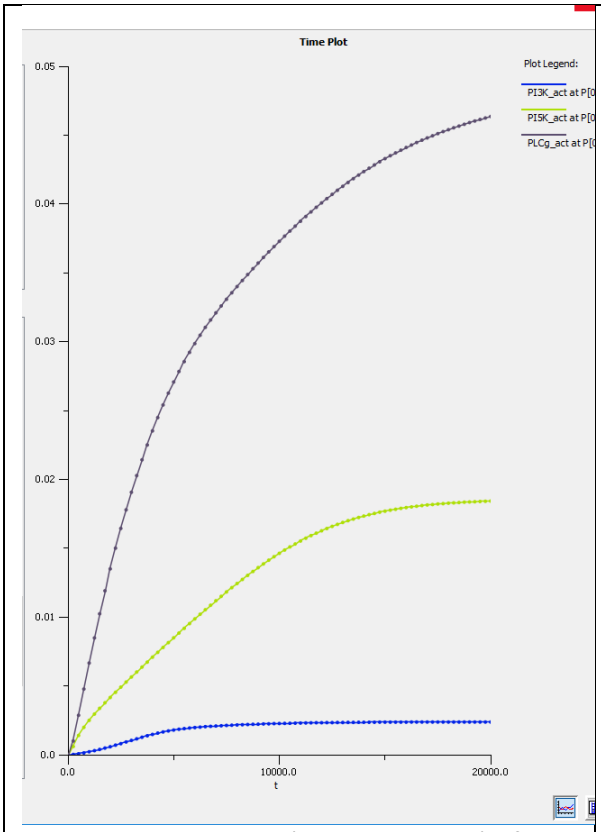

Figure 4.59. Densities (molecules.µm2) of PI5K_act, PLCγ_act and PI3K_act for the Virtual Cell 2D spatial model. Peak particle levels: 7, 17 and 1 molecules, respectively.

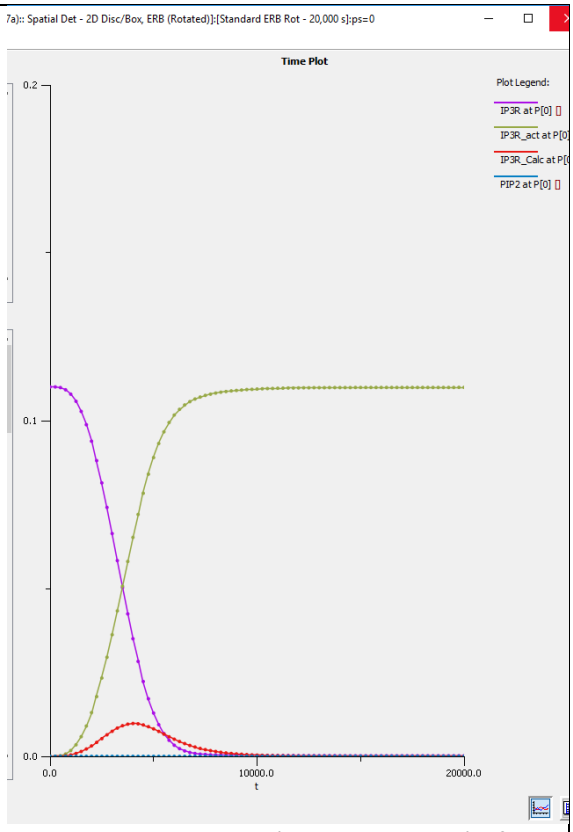

Figure 4.60. Densities (molecules.µm2) of IP3R species for the Virtual Cell 2D spatial model. Peak particle levels: IP3R = 220, IP3R_Calc = 20, IP3R_act = 220 molecules, respectively.

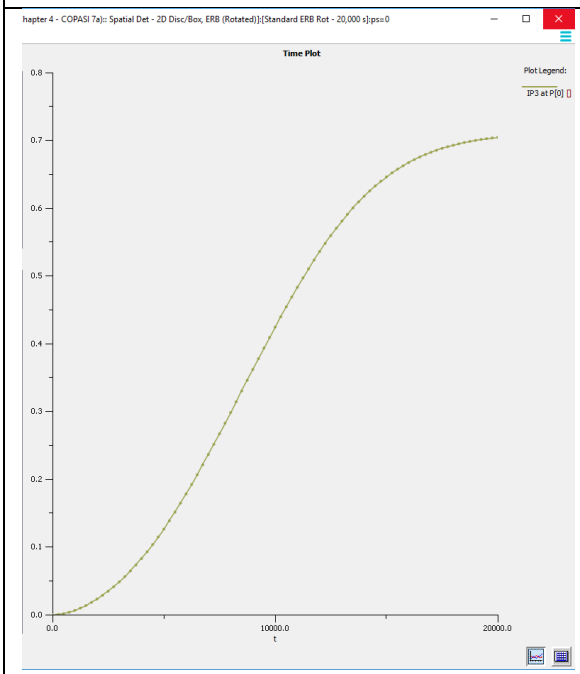

Figure 4.61. Concentrations (µM) of IP3 for Virtual Cell 2D spatial model. Peak particle levels:21,152molecules.

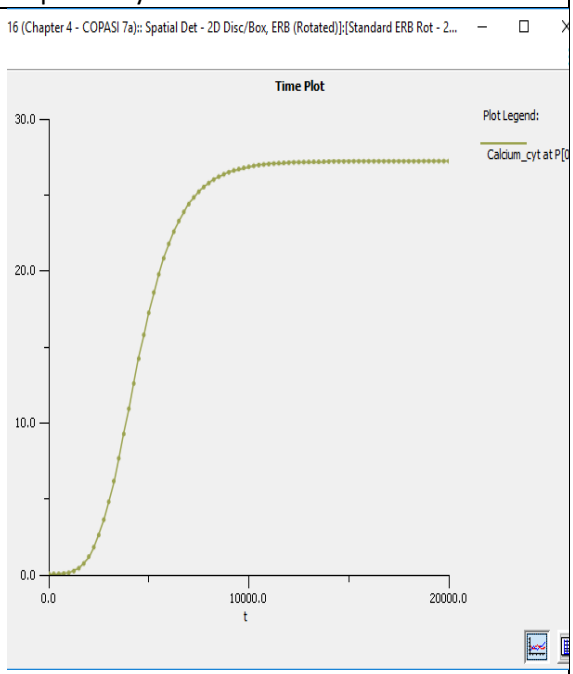

Figure 4.62. Concentrations (µM) of cytosolic calcium for Virtual Cell 2D spatial model. Peak particle levels: 819,600 molecules.

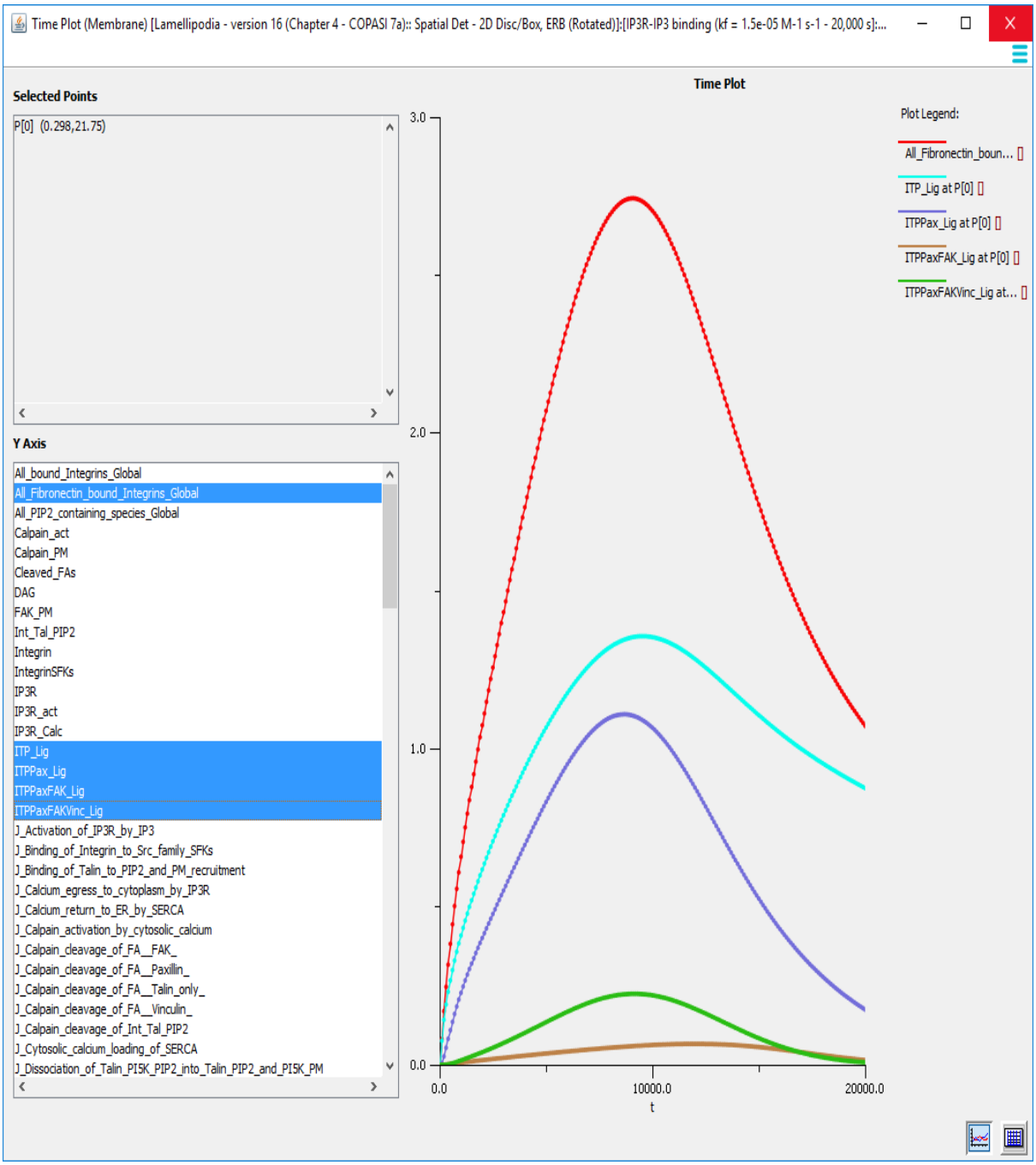

Figure 4.63. Densities (molecules.µm^2) of total FAs, ITP Lig, ITPPax_Lig, ITPPaxFAK_Lig and ITPPaxFAKVinc_Lig in Visual Cell version of the model. Corresponding peak particle levels (in rounded numbers of molecules) are 1001, 495, 405, 28 and 82, respectively.

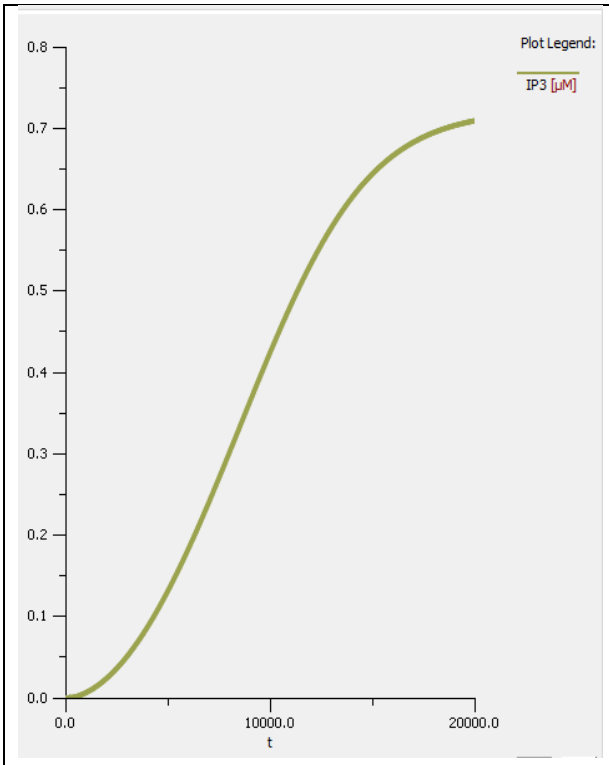

Figure 4.64. Concentrations (µM) of IP3 for Virtual Cell compartmental model, adjusted for higher IP3R-IP3 binding. Peak particle levels:21,350molecules.

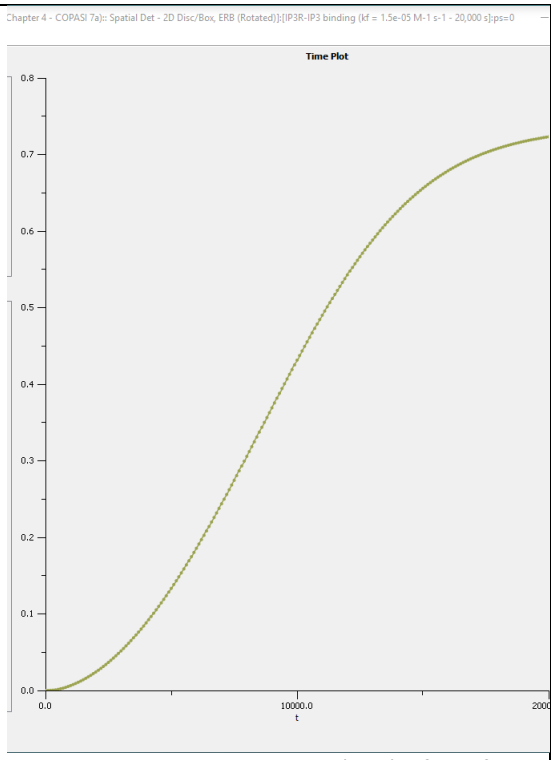

Figure 4.65. Concentrations (µM) of IP3 for Virtual Cell 2D spatial model, adjusted for higher IP3R-IP3 binding. Peak particle levels: 21,770 molecules.

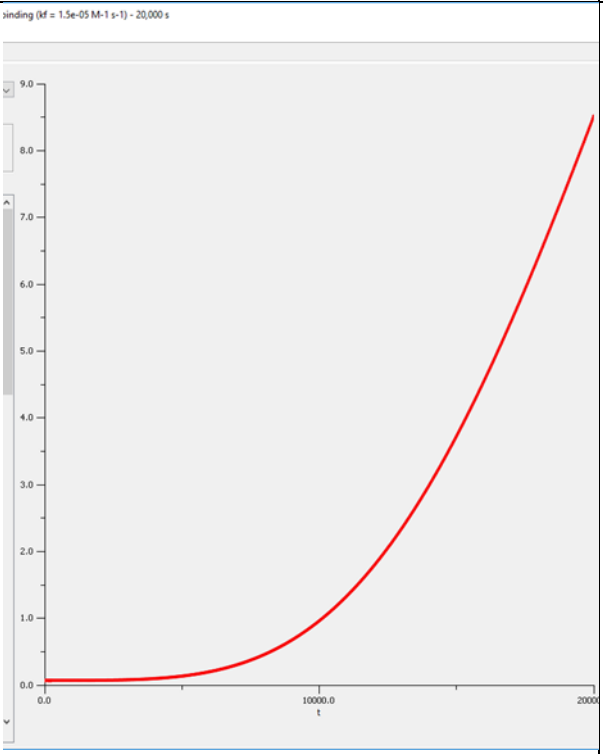

Figure 4.66. Concentrations (µM) of cytosolic calcium for Virtual Cell compartmental model, adjusted for higher IP3R-IP3 binding. Peak particle levels: 256,500 molecules.

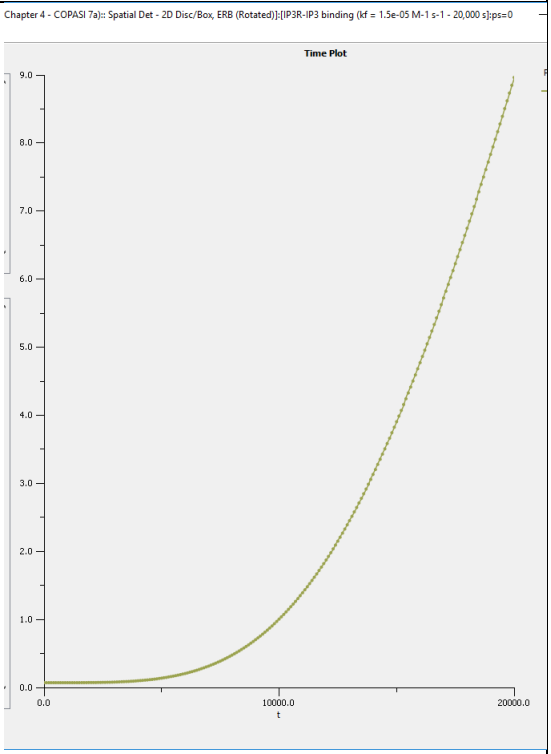

Figure 4.67. Concentrations (µM) of cytosolic calcium for Virtual Cell 2D spatial model, adjusted for higher IP3R-IP3 binding. Peak particle levels:270,000molecules.

Although we have established major similarities between the compartmental and 2D spatial models, there are still obvious advantages of using spatial models. For instance, unlike compartmental models, in which species are assumed to be well-mixed, and with diffusion essentially instantaneous, in spatial models it is possible to vary diffusion rates, in order to determine what effects they have on system behaviours.

So far, default diffusion rate constants have been accepted for all species, which is to say, 0.1 $\mu m^2.s^{-1}$ for membrane-resident species, and 10 $\mu m^2.s^{-1}$ for species resident in liquid compartments. However, it is known that many molecules in aqueous solution have much higher diffusion rates than this, including IP3 and calcium ions, where, amongst other values, rate constants of 280 and 530 $\mu m^2.s^{-1}$ have been reported, respectively[301,300]. As we have just seen that IP3 and cytosolic calcium have such an important role in our Virtual Model in determining levels of FAs, one might reasonably expect these much-increased diffusion rates also to have noticeable effects on these levels. If, for the sake of consistency, we also increase the diffusion constant for ER calcium by the same amount as for cytosolic calcium, one might also reasonably expect this to increase the rate of IP3R-mediated calcium egress and hence calpain cleavage.

4.4.2.2.2 Version 2 (of Virtual Cell 2D Spatial/PDE model)

Having created a new simulation (for this 2D spatial application) we replace the default diffusion constants for these three species with their new values (leaving the IP3R-IP3 binding rate at its original lower value, and retaining the AreaPerUnitArea changes) and then run the simulation for 20,000 seconds. Sample results are shown in Figures 4.68-4.70. They show that, despite increasing the IP3 diffusion rate 23-fold and the calcium rates 58-fold, there is little, if any, discernible difference in FA dynamics, or in IP3 or calcium responses between this model and the original 2D spatial model (with unadjusted IP3R-IP3 binding rates). This is summarised in 4.14.

Species	Peak particle levels for standard model		Peak particle levels for enhanced IP3/calcium diffusion model	
	Peak levels (# molecules)	Peak timings (s)	Peak levels (# molecules)	Peak timings (s)
Total FAs	422	2750	422	2800
ITP_Lig	236	2750	236	3200
ITPPax_Lig	160	2750	160	3000
ITPPaxFAK_Lig	6	3000	5	3000
ITPPaxFAKVinc_Lig	20	3250	20	3400
IP3	21,212	20,000	21,212	20,000
Cytosolic calcium	819,600	20,000	819,600	20,000

Table 4.14. Summary of differences in FA, IP3 and cytosolic calcium peak levels between simulations, using either default diffusion constant values ("standard model") or increased diffusion constant values for IP3, cytosolic and ER-resident calcium ("enhanced diffusion model"), within Virtual Cell 2D Spatial application. (Default diffusion constant values were all 10 $\mu m^2.s^{-1}$ for IP3, Calcium_cyt and Calcium_ER, while the enhanced model used values of 280, 530 and 530 $\mu m^2.s^{-1}$, respectively.)

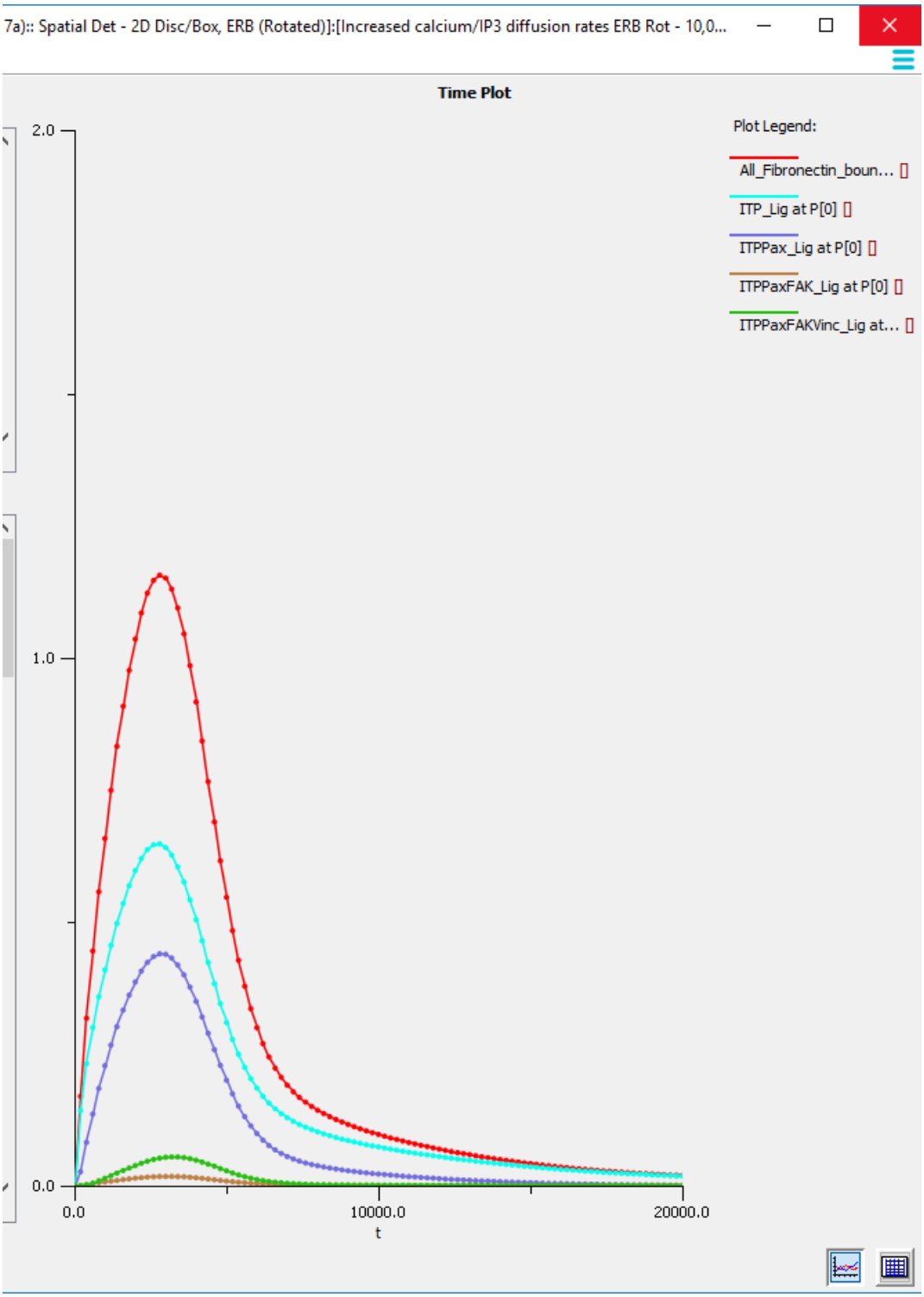

Figure 4.68. Results of increasing diffusion constants for IP3 and calcium ions (cytosolic and ER-resident) on levels of total FAs and individual FAs species (molecules. μM^{-2}) in the Virtual Cell model. Peak particle levels (numbers of molecules): Total FAs = 422, ITP_Lig = 236, ITPPax_Lig = 160, ITPPaxFAK_Lig = 5, ITPPaxFAKVinc_Lig = 20.

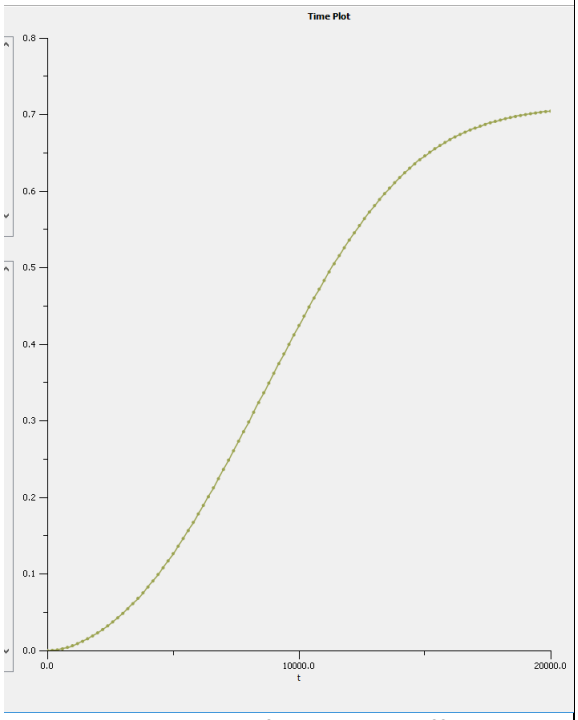

Figure 4.69. Results of increasing diffusion constants for IP3 and calcium ions (cytosolic and ER-resident) on levels of IP3 (µM) in the Virtual Cell 2D Spatial model. Peak particle levels: 21,212 molecules.

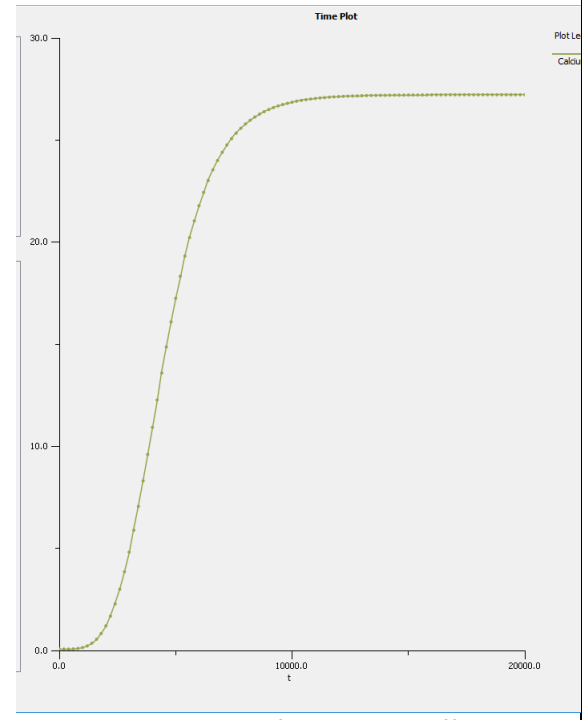

Figure 4.70. Results of increasing diffusion constants for IP3 and calcium ions (cytosolic and ER-resident) on levels of cytosolic calcium (µM) in the Virtual Cell 2D Spatial model. Peak particle levels: 819,600 molecules.

In fact, these results are not quite as surprising as they might first appear, as the peak levels of IP3 and cytosolic calcium are determined largely by the availability of PIP2 (from which IP3 is hydrolysed) and by the initial levels of calcium in the ER, respectively. The simulation runtime of 20,000 seconds is long enough for both of these two to approach their maximal levels at their slower diffusion rates, so the increased diffusion rates were always likely to have marginal influence on the peak levels.

The only real surprise is that the timings of peak FA levels are so little altered. Indeed, some of these differences may be less than they appear, as the time interval between output values was 250 seconds in the case of the original model and 200 seconds in the case of the new model. The most likely reason for this similarity is that the calpain activation curve is also nearly identical in both models, with a peak rate of 2.6e-05 molecules.$\mu m^{-2}.s^{-1}$ at around 4600 seconds, with peak levels of activated calpain only reaching around 50 molecules at this point.

But this is only a very limited exercise, and it may be that larger differences might be seen if diffusion rates were adjusted for all models species, based on published data. However, the similarity of these results confirms the benefits of developing spatial models as compartmental applications in the first instance. And, in fact, if one is only interested in changes in overall levels of models species there is often no benefit in developing a spatial model, unless the model geometry is sufficiently complex, with convolutions and constrictions that make diffusion rates critical in determining outcomes. Certainly this does not seem to be the case in the geometrical representation of a typical lamellipodia used here. Of course, if changes in local concentrations of species are of interest then spatial modelling is the only option.

4.5 Discussion

4.5.1 Introduction

The purpose of this chapter was twofold. Primarily, to use differential equation modelling to investigate the dynamics of FAs, in response to changing levels of PIP2, and to calpain cleavage. Secondly, by the use of two different differential equation modelling tools, and two different approaches (ODE/compartmental as against PDE/spatial methods), to further illustrate the challenges that face the modeller in implementing such models. To this end, and for the sake of continuity, this has involved developing the very simple FA model of the previous chapter into something more complex, using the same modelling software (COPASI and Virtual Cell) and broadly the same methods that were introduced in the previous two chapters.

A number of very interesting insights have arisen from the overall model, as it has been developed.

4.5.2 Insights gained from the lamellipodia FA models

4.5.2.1 Importance of talin availability

Firstly, availability of talin is critical to overall levels of FAs (Figure 4.11). This importance derives from its central role in activating integrins, but also the fact that our current understanding of how it does this and how it fits structurally into FAs, tends to preclude one talin activating more than one integrin. This conclusion is based on the assumption that talin is the primary binding partner for vinculin and must be present for FAs to mature beyond the nascent stage. As this model shows, implicitly assuming this premise means that levels of talin availability must be critical to overall levels of FA formation. Whilst this is not a particularly unexpected finding, it is easy to overlook.

4.5.2.2 Greater importance of concentrations of species rather than particle numbers

Following on from this, it would appear from the results that it is not necessarily the absolute numbers of talin or other FA-associated proteins that is most rate-limiting, but rather their respective concentrations. This does not necessarily mean that local concentrations (i.e. the concentrations of these non-integrin FA proteins in the vicinity of integrins, or of the PM in general) have to be high. As shown by these models, the absolute and relative sizes of the cytosol and PM within lamellipodia also appear to be critical, something that is even more easily overlooked.

4.5.2.3 Importance of PIP2 availability

A further, related, finding of the models is that FA levels appear to be influenced by PIP2 availability, as one would expect. However, this influence is heavily constrained by talin availability and other factors that determine rates of FA assembly (Figures 4.8-9, 4.13-15). Put simply, as PIP2 levels increase, other factors become increasingly rate-limiting on FA formation, setting a clear upper limit to the influence of PIP2 on overall FA levels. Moreover, these models appear to suggest that this limit may be much lower than might be expected, if one simply focuses on a few obvious factors, such as numbers of available integrins.

4.5.2.4 Relative unimportance of integrin-ECM binding rates

In common with PIP2, other single factors, such as binding rates of integrins to their extracellular ligands are shown to only influence FA levels up to a point (Figure 4.10). This is in contrast to the results seen for Version 4 of the early FA model in Chapter 3 (Figures 3.41-3.46), where increasing such binding rates had a very substantial affect both in terms of increasing overall FA levels and in decreasing the overall runtime for FA-associated events. One possible interpretation of these results is that PIP2 and integrin-ECM binding rates may be critical to levels of very early FAs

(however one defines these) but much less so on more mature FA species. Certainly, this would seem to be a prediction of the various models worth testing experimentally.

However, overall, this result adds to the general impression gained from the previously mentioned results, that no single factor is dominant in determining total FA levels. Rather, these models suggest that such levels, as well as overall rates of FA formation, are determined by a relatively large number of factors, some of them likely to be far less obvious.

4.5.2.5 Insights concerning FAK

One such factor, FAK, raised a number of interesting questions. If one assumes, as this model does, a one-to-one relationship between FAK and integrins, and a low abundance relative to the other FA species (as predicted by the HeLa cell quantitative proteome used for this model [182]), this inevitably implies that only a relatively low proportion of FAs can contain FAK. (This is shown in a number of model plots, including Figure 4.19, where FAK-containing FA isoforms represent only around 15% of total FAs at peak levels, even after changes to the model to enhance their levels.) Of course it could be that FAK copy numbers have been underestimated in the quantitative proteome mentioned. However, an even lower copy number value has been reported in a different quantitative proteome, this time for U2OS cells, quoting a figure of 3.29e+03 [305], as against the value of 1.1e+04 quoted for HeLa cells. An alternative possibility is that something specific to FAK causes its numbers to be undercounted relative to other critical FA proteins. But it is hard to think of any property, not shared by any of the other FA proteins used in the model, which would convincingly explain this.

This would appear to leave only two likely explanations. One is that FAK levels are abnormally low in certain cell-lines such as HeLa and U2OS cells. As it happens, both these cell lines are comparatively slow-moving [192,308], so this may prove to be the case, given FAK's known role in promoting FA turnover and migration [309,310]. Further support for this is found in the levels of FAK reported in a human platelet quantitative proteome, where, despite their very small size and their

lack of a nucleus, copy numbers of 3000 are reported [311]. However, this may merely reflect the fact that platelets are unable to generate new copies of FAK, as well as their specialist role, which requires the formation of numerous lamellipodia. Certainly, until a larger number of quantitative proteomes have been published, it cannot be ascertained what typical levels of FAK are in nucleated eukaryotic cells.

Leaving this point aside, phenomena such as cell spreading (where large numbers of simultaneous lamellipodia are observed active around the entire cell periphery) imply that one-to-one FAK-integrin binding would severely limit such spreading in cell lines such as HeLa and U2OS. Even allowing for rapid recycling of FAs and additional FAK expression, there does not appear to be enough FAK for this. One reason for thinking this is that, at any given time, much of this FAK will be resident throughout the cytosol and elsewhere in the cell.

For instance, in the COPASI model, levels of cytosolic FAK never fall below 50% (reaching a minimum of 256 out of the original 457 molecules) at any time. This figure seems likely to be reflected in reality, given the relative sizes of the cytosol and PM. However, it is possible that local levels of FAK may be much higher within lamellipodia than within the cytosol as a whole, as a result of various mechanisms. Depending on how enriched lamellipodia were in this regard, this would mean there was always a ready supply of FAK at hand for FA formation.

Overall, though, the results from this model suggest that FAK involvement in FA dynamics requires that FAK proteins are only temporarily attached to FAs, probably with low residence times, or that other FA-associated proteins can perform the same functions in FAK's absence. There is an alternative to FAK, Protein-tyrosine kinase 2-beta (Uniprot Q14289) with broadly similar properties. However, its numbers as listed in the HeLa quantitative proteome, 3100 [182], add a little less than 30% to the FAK total (~11,000 molecules), so would seem unlikely to affect the issue much. However, at least one other FA-associated protein, ILK (Integrin-linked protein kinase, Uniprot ID Q13418) can perform some of the functions of FAK [56], in terms of being both a

scaffold protein and a protein kinase [312], and are available in much greater numbers (around 210,000 copies) in the HeLa cell [182]. However, ILK is a serine/threonine kinase [313,312], whereas FAK is a tyrosine kinase [56], and does not associate with the same proteins in their scaffold roles [312,56,.313]. Nevertheless, in the form of the ILK-Pinch-Parvin complex [313], it provides an obvious alternative mechanism for linking integrins to the actin cytoskeleton, both in outside-in and inside-out signalling.

4.5.2.6 Criticality of general FA binding requirements and rates

Another general conclusion to be drawn from this model is that, if we assume that FA assembly requires sequential binding, with not much flexibility in the order of this recruitment, binding rates will need to be relatively high. Otherwise, the more downstream proteins tend to be very poorly represented, as a proportion of overall FAs, relative to their abundance. This was seen especially in the case of vinculin in the COPASI model. Despite being highly abundant (over three times the levels of talin and paxillin, and second only to integrins in abundance), levels of vinculin-containing FAs only exceeded 10% of total FAs, after rates for ITPPaxFAK_Lig-Vinc_PM and ITPPax_Lig-FAK_PM binding were both substantially increased (along with levels of FAK_PM, Figure 4.16).

But this effect was not only confined to the most downstream FA component protein in the model. Increasing the ITP_Lig-Paxillin_PM binding rate substantially raised ITPPax_Lig numbers as a portion of overall FAs (Figure 4.13), as did increasing the ITPPax_Lig-Fax_PM binding rate on ITPPaxFAK_Lig representation, as long as the availability of FAK itself was also substantially increased (Figure 4.15).

Finally this was confirmed in the second version of the COPASI model, where all the relevant binding rates were increased (by increasing the forward rate by one order of magnitude and decreasing the reverse rate by the same amount), together with a halving of the PM compartment. (These changes were all carried forward into the subsequent Virtual Cell models.)

In all cases this led to a less uneven representation of FA isoforms compared to what is seen prior to these changes, in which over 90% of FAs take the form of ITP_Lig. This can be seen by comparing Figure 4.12 with Figures 4.21 and 4.30. Given that paxillin-association is commonly used as a FA marker, this more even representation seems the more likely representation of the physiological reality.

4.5.2.7 Differences in behaviour of COPASI and Virtual Cell models

Moving to the Virtual Cell compartmental model, differences between its outputs and those of the COPASI model were greater than one might expect, given that essentially the same parameters and reaction equations were used in both. To a very large extent the problem almost certainly comes down to the way in which COPASI and Virtual Cell represent membranes, as explained earlier.

This means, in turn, that some second-order mass action reaction rate constants and some enzymatic KM values have to be converted so that product formation rates are expressed in areametric rather than volumetric terms (i.e. in the form molecules.$\mu m^2.s^{-1}$ rather than as $\mu M.s^{-1}$). In theory, this should merely require dividing the volumetric rate constant value by the thickness of the membrane in question, this being the ratio of the membrane volume to its surface area. In this case the value of 7.5 nm was chosen, this being the original value assumed for the PM in calculating its volume for the COPASI model.

This value was then used to convert the relevant rate constant and KM values, with results as shown in Figures 4.30-45 and summarised in Table 4.10. Many differences can be seen between the two models, most obviously in the case of FA (both total and individual) and cytosolic calcium levels.

In fact, the rate constant value for the intra-membrane reactions just mentioned, should have been divided by 3.75 nm (half the 7.5 nm actually used), this lower value being consistent with

the halved volume used in the second version of the COPASI model. (It is also closer to many published values for lipid membrane thicknesses of around 4 nm [189,314].) However, when tried this made matters worse, reducing FA levels even further. In any case, as subsequent analysis revealed, the problem primarily appears to lie not with the PM but the ER membrane, as this is where the IP3R species reside. It is their more vigorous response to IP3, pumping out more than three times the amount of calcium ions into the cytosol for the same amount of IP3, that largely accounts for the much higher levels of cytosolic calcium compared to the COPASI model. This leads, in turn, to much higher levels of activated calpain, and much lower levels of FAs as a consequence.

As was seen, suitably lowering just one rate constant value (the forward rate for IP3R-IP3 binding) in the Virtual Cell compartmental model brings cytosolic calcium levels in line with the COPASI model. As a consequence, FA levels are brought much closer in line with those of the equivalent COPASI model, as can be seen by comparing Figures 4.49 and 4.20, and by comparing the FA particle numbers quoted in the caption below Figure 4.49 with the FA levels quoted in Table 4.10. In short, FA levels in this adjusted version of the Virtual Cell compartmental model are more than double the levels they were prior to the adjustment to IP3R-IP3 binding.

However, as commented on at the time, there is no immediately obvious reason why the IP3R response to nearly identical IP3 levels is so different between the COPASI and (unadjusted) Virtual Cell compartmental models. No conversions were required for any of the IP3R-associated second-order reactions.

Without conversion as a possible source of error it is hard to pinpoint why there is such a substantial and important divergence between these two supposedly equivalent models, in terms of their outputs. This illustrates one of the great challenges of modelling, whether involving differential equations or other methods. Developers of different modelling tools have to make difficult decisions as to how to implement cellular structures and various other aspects related to

cell biology, such as intra-membrane reactions and reactions where individual reactants (or reactants and products) are resident in different compartments with different volumes. Inevitably different developers make different decisions.

4.5.2.8 Importance of calpain cleavage in determining FA levels as against PIP2 influence

Leaving aside the specific issues, and taking into account all the results from the three versions of the FA model described so far, the clear prediction of this model is that calpain cleavage is at least as influential in determining FA levels as PIP2. Put simply, on the one hand, the influence of PIP2 in increasing FA assembly is constrained by multiple factors, which quickly take over as rate-limiting factors with increasing FA numbers. By contrast, the influence of calpain cleavage in lowering FA numbers seems less constrained, appearing to be mainly confined to the amount of IP3 produced and the IP3R-mediated calcium response to it. (This is, of course, clearly demonstrated by the initial IP3R/IP3-associated differences seen between the Virtual Cell and COPASI models, as just described.) However, this model has a highly simplified calcium-release mechanism and would clearly benefit from having a much more realistic implementation in future versions of the model. This should then help engender more confidence in the validity of this and the other previously-mentioned predictions of the model.

4.5.2.9 Excessive runtimes of all models

A final issue to be addressed is the very long runtime of the model, mentioned earlier. In the modeller 's experience this is a problem commonly encountered with many models, especially those as large and complex as this one. It was, of course, also a problem in the COPASI SDF1 chemotaxis model in Chapter 2, whose simulations also required a 20,000 second runtime, even though this was, in many ways, a simpler model.

Often this slowness comes down to just one or two factors acting as bottlenecks. For instance, as was shown in Figure 4.9, increasing the rate of PIP2 creation substantially increased not only the overall levels of FAs but also their dynamics, in terms of assembly and disassembly rates. However, increasing fibronectin-integrin binding rates had little effect on these dynamics, with peak FA levels achieved at around 10,000 seconds, even at the highest values (Figure 4.10). Increasing initial talin levels is also shown to increase the rate at which FA assembly peaks and then declines, but only at the highest levels (substantially above the levels predicted by the quantitative proteome) and only fairly moderately (Figure 4.11).

Often, reducing the volume of compartments that contain many model species (typically the cytosol or PM) will speed up the rate of various model behaviours, because (as explained in Chapter 2) the rate of reactions is largely determined by the concentrations of reactants, at least in the vast majority of cases. Keeping the species number constant whilst reducing the reaction volume obviously increases the concentration of the species concerned.

Unfortunately, reducing the PM volume by one half in the COPASI model is seen to have little discernible effect on FA dynamics, at least in terms of peak timings, as can be seen by comparing Figures 4.12 and 4.18. In both cases peak levels are seen to occur at around 7500 seconds. Indeed, given that these peak levels in the reduced-volume version are around a half those for the original PM volume, the overall FA assembly rate has clearly halved - the opposite of what one would have expected. Why this is the case is something of a puzzle. After all, if this same finding was to be extended to proteins grouped together within lipid rafts, it would inevitably raise questions as to what advantage was gained by associating in that way.

Because the focus, when modifying the cytosolic volume, was on its possible effects in increasing relative levels of FA isoforms other than ITP_Lig, and not on speeding up the model, the cytosolic volume was increased, not decreased. What happens if we decrease it instead? Figure 4.71 shows that, just as predicted, if the cytosolic volume is reduced by 75% (from 50 to 12.5 fl) in the

COPASI Model, the time required to reach peak FA levels is reduced by a similar amount (from around 10,000 to around 2500 seconds). Furthermore, this does not come at the expense of FA levels, with total FA numbers actually being a little higher (at nearly 1600 molecules) than in the original COPASI model (where levels were a little below 1400 molecules), as shown in Figure 4.12. ITP_Lig is even a little less dominant than in the original model.

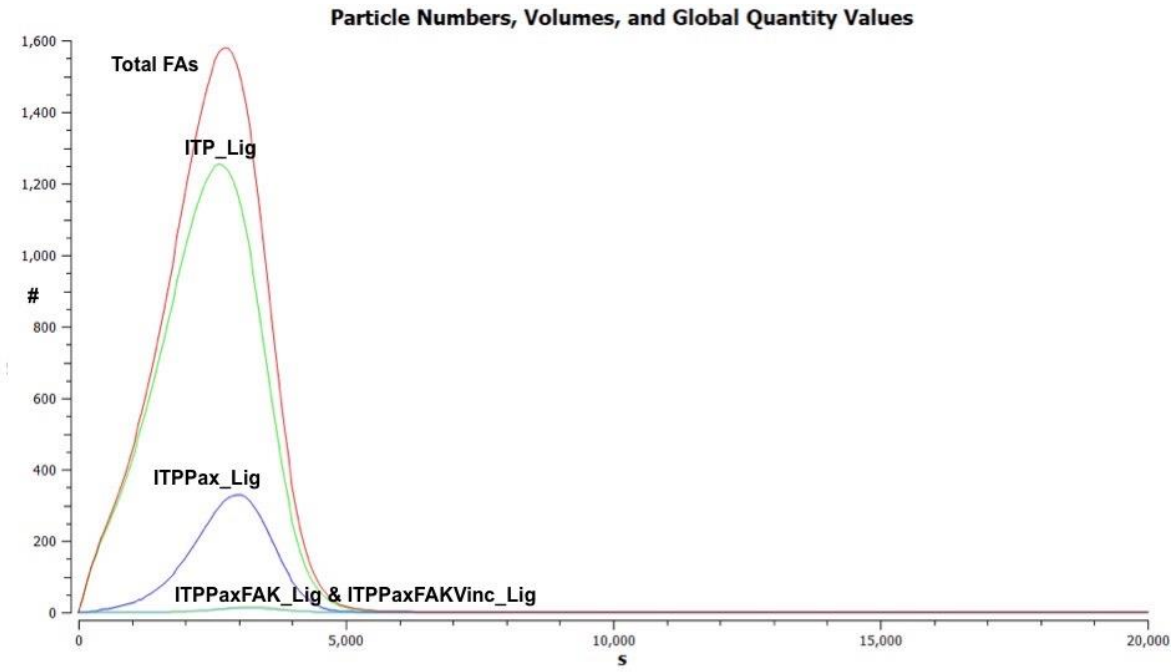

Figure 4.71. Particle numbers for Total FAs, ITP Lig, ITPPax_Lig, ITPPaxFAK_Lig and ITPPaxFAKVinc_Lig in COPASI lamellipodia model, after cytosolic volume has been reduced from 50 to 12.5 fl.

Finally, increasing the rates for ITP_Lig-Pax_PM, ITPPax_Lig-FAK_PM and ITPPaxFAK_Lig-Vinc_PM binding, as described previously, results in a redistribution of FAs, as shown in Figure 4.72.

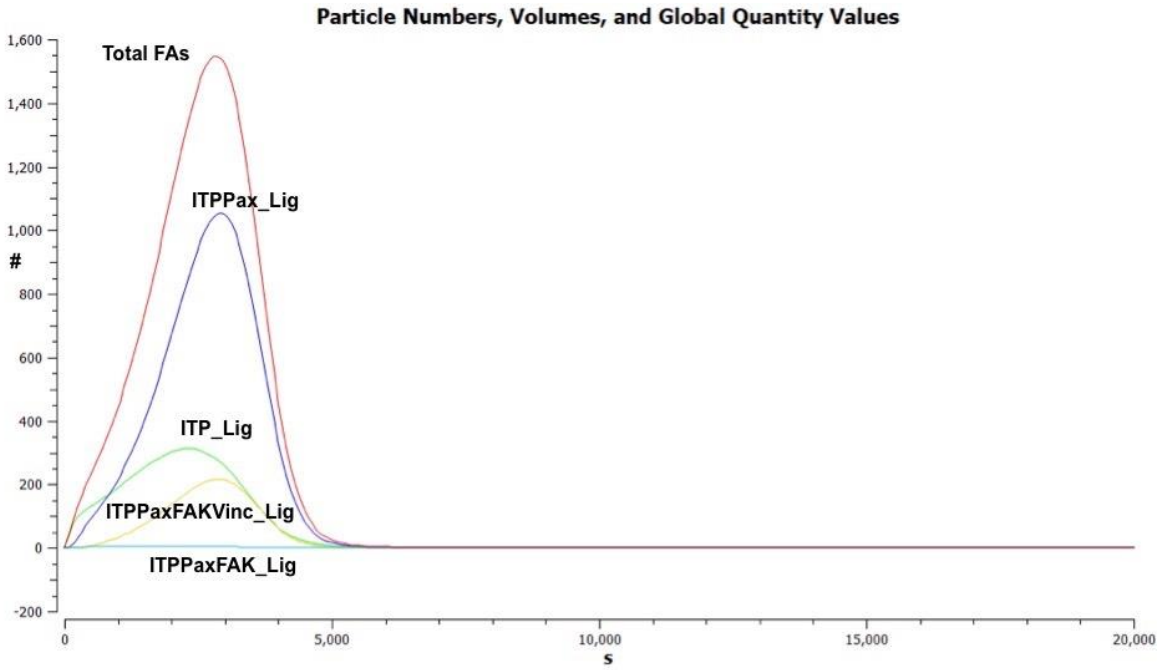

Figure 4.72. Particle numbers for Total FAs, ITP Lig, ITPPax_Lig, ITPPaxFAK_Lig and ITPPaxFAKVinc_Lig in COPASI lamellipodia model, after cytosolic volume has been reduced from 50 to 12.5 fl and binding rates ITP_Lig-Pax_PM, ITPPax_Lig-FAK_PM and ITPPaxFAK_Lig-Vinc_PM binding have been increased from 10,000 to 100,000 M^{-1} s^{-1} for the forward rate, with the reverse rate decreased from 0.01 to 0.001 s^{-1}.

It is clear from the above, then, that the changes in the FA isoform representation described earlier (and shown in Figure 4.21), which resulted from a halving of the PM volume, together with the same changes in the forward and reverse binding rates just used here, are likely to also have been mostly due to the changes in the binding rates, rather than in the PM volume.

Taken together, these results suggest that the reason FA dynamics are so much quicker in reality than in this model is likely to be due to a number of factors. These may include the effective lamellipodial cytosolic volume being much less than the volume originally used here, for reasons described earlier. Given the importance of PIP2 in FA assembly, previously outlined, it seems not unreasonable to assume that rates of PIP2 formation may be substantially higher than the rate assumed here, possibly helped by recycling of breakdown products such as IP3 and DAG, something not included in this model because it was felt the timescale for recycling would be too long to make a noticeable difference, given the FA and lamellipodia timescales described previously. However, PIP2 can also be recycled whole, together with FA-associated (and other

membrane) proteins, by the endocytic recycling pathway [315]. In this case, the timescale may well be relevant, with this option excluded only to keep model complexity within manageable limits. However, it should be considered a priority for any future version possibly in the form of a separate submodel .

Many of these suggestions would also merit further experimental investigation in the lab. This might include seeing how inhibiting or enhancing rates of PIP2 creation affects the rate of FA formation, and whether it affects the relative numbers of different FA isoforms. Put another way, it would be interesting to see if experiments confirm the conclusion drawn here, that PIP2 is more influential in determining the rate of FA assembly rather than in overall numbers and, if so, how much other factors determine overall FA numbers and what other factors determine assembly rates.

In a similar vein, it might be possible to manipulate the cytosolic and PM volumes of lamellipodia to see how such changes affect FA assembly and disassembly rates. However, this could be extremely hard, if not impossible, to achieve, whilst being sure that other aspects of FA formation were not affected in ways that would invalidate the results.

Manipulating levels of FA-associated proteins such as talin, FAK and vinculin would, in principle, also help to inform this model. However, any results from such manipulations would need to ensure that they were not due to protein-specific behaviours not accounted for in this model, such as FAK's tyrosine kinase activity, or talin- or vinculin-mediated mechanotransduction.

As this model hopefully shows, there is still a great deal about FAs that we do not know and therefore much that still remains to be discovered. More information about typical copy numbers of various FA-associated proteins in different cell lines is required, a task that has only just started with the recent development of suitable techniques. As we have seen, many of the binding rates between key FA proteins have still to be measured, and these measurements need to stipulate values for the forward and reverse rates, not just be limited to Kd values. It is only after a number

of such measurements have been made that any kind of consensus can be reached about their actual values.

A much better understanding of the mechanics of FAs is also needed. In particular, there are still a lot of uncertainties remaining about how the key FA-associated proteins bind to existing FAs. More specifically, we need a better understanding of what other FA protein (or proteins), and what binding sites on those proteins, are required to be already present within the FA for such recruitment to take place. Further to this, there are still many uncertainties about the order in which these key FA proteins are recruited, and how much flexibility there is in that order.

It is also far from clear which of the key FA proteins, once bound, must remain permanently in place during the lifetime of a FA (or must be immediately replaced by the same protein type if they leave), and which can leave without unduly affecting FA disassembly. For those FA proteins required to have only temporary association with FAs, typical residency times must be established. As was seen earlier, different assumptions about these factors can result in substantial changes in important model results, so further experiments to answer such questions would greatly benefit this model.

The influence of various post-translational modifications (in this context, phosphorylation, SUMOylation, acetylation, various forms of lipidation, and similar) has not been addressed in much detail (at least, not explicitly) in this model. Again, this has been in order to try and reduce its complexity to manageable levels. Nevertheless, a better understanding of the effects of these, and of the underlying mechanisms behind them, can only improve our overall understanding of how FA assembly and disassembly is regulated. This should help improve the design of models such as this.

In summary, any experiments, or other forms of observation, that furthered our understanding of FA-related matters would also help improve this model in terms of being a better reflection of the

underlying reality being modelled and by increasing ones confidence in any conclusions derived from it. But specifically, experiments and observations that better established typical copy numbers of the various model species within lamellipodia, and the rules determining if and how such species bind to each other (and the rates for such binding, when it happens) would immediately benefit it.

Finally, a great surprise from this chapter is that the differences between the COPASI model and both the Virtual Cell models, based on her were much greater than differences between the two compartmental and the spatial model. Given that the two compartmental models supposedly used very similar parameters and reaction equations, and given the differences between compartmental and spatial modelling illustrated in the previous chapters, one might reasonably expect much greater differences between the spatial model and the two compartmental models, with very little difference between the latter. Instead, the opposite was seen.

4.5.3 Conclusion

The results from the COPASI model and the Virtual Cell compartmental and 2D spatial models are clearly not intended as faithful reflections of the underlying physiology of FA dynamics. Whilst many of the key proteins (integrin, talin, paxillin, SFKs, FAK and vinculin) are included in these models, many of these (especially paxillin, FAK and vinculin) perform a much more limited role (in terms of FA-related behaviour) in the model, compared to in reality.

However, these simplifications and absences are justified for a number of reasons. Firstly there is the need to prevent the model becoming so complex that it is almost impossible to analyse usefully. This is particularly true of models, such as the one here, where there are a number of complex interactions between system components, particularly those involving negative or positive feedback. In the writer's experience this can result in lengthy investigations of certain

model behaviours, where it ultimately proves impossible to determine which of two or more factors are the primary cause of that behaviour, because all act directly on to each other.

Secondly, the larger and more complex model, the greater the likelihood that accumulated errors in its implementation make it such an unfaithful reflection of the underlying reality that it has no usefulness as an investigative tool. Smaller models can still be very useful when their outputs do not mirror those of the underlying system at all well, as long as their design is felt to reflect our current understanding of that system closely enough. In that case, their usefulness can largely lie in determining precisely why the model's behaviour deviates so far from the expected behaviour. For instance, such analysis may reveal that certain interactions must be occurring at a very different rate to that reported in the literature, or it may be that the model behaviour brings into question our current understanding of the system.

However, once a model becomes too large and complex, it becomes much harder to determine where the problem lies, or how well the model implementation actually reflects the real-world mechanism. It can be readily appreciated why this would severely impair its usefulness to anyone trying to understand the underlying system.

Thirdly, the large majority of FAs do not go on to form fully mature adhesions, such as stress fibres. This is thought to be largely because such adhesions fail to gain sufficient traction to progress to these later stages [316], meaning that actin fibres, myosin and other actin-associated proteins (such as α-actinin and zyxin) are not recruited to them. For all these reasons it was felt best to only attempt to model the earlier stages of FA dynamics.

Accepting these restrictions, this project set out to investigate FA dynamics in much of their complexity. This has resulted in a fairly complex model of early FA adhesion, as described in this chapter, leading to a number of conclusions being reached. Amongst these is the inference that, if FA assembly is a fairly prescriptive sequential process (as our current understanding of FA maturation appears to require, and super-resolution imaging appears to confirm [8]), then later-

recruited proteins such as vinculin will tend to be underrepresented, unless average binding rates for FA component proteins are generally relatively high.

Another, unexpected, prediction is that, whilst decreasing the effective cytosolic volume will tend to increase the rate of FA assembly, reducing the effective PM volume appears to have the opposite effect, and certainly lowers overall levels of FAs.

This chapter specifically set out to determine relative influences of PIP2 and of calpain cleavage in influencing FA assembly and disassembly, as well as subsequent FA levels. In this respect, this model appears to be suggesting that the rate of new FA formation is strongly influenced by the rate of PIP2 formation (and possibly by integrin-ECM binding, if the previous model is anything to go by).

However, actual numbers of FAs, both overall and in their relative isoform proportions, depend much more on other factors, especially levels of calpain activity, and overall binding rates. Most of these insights would appear to merit further investigation.

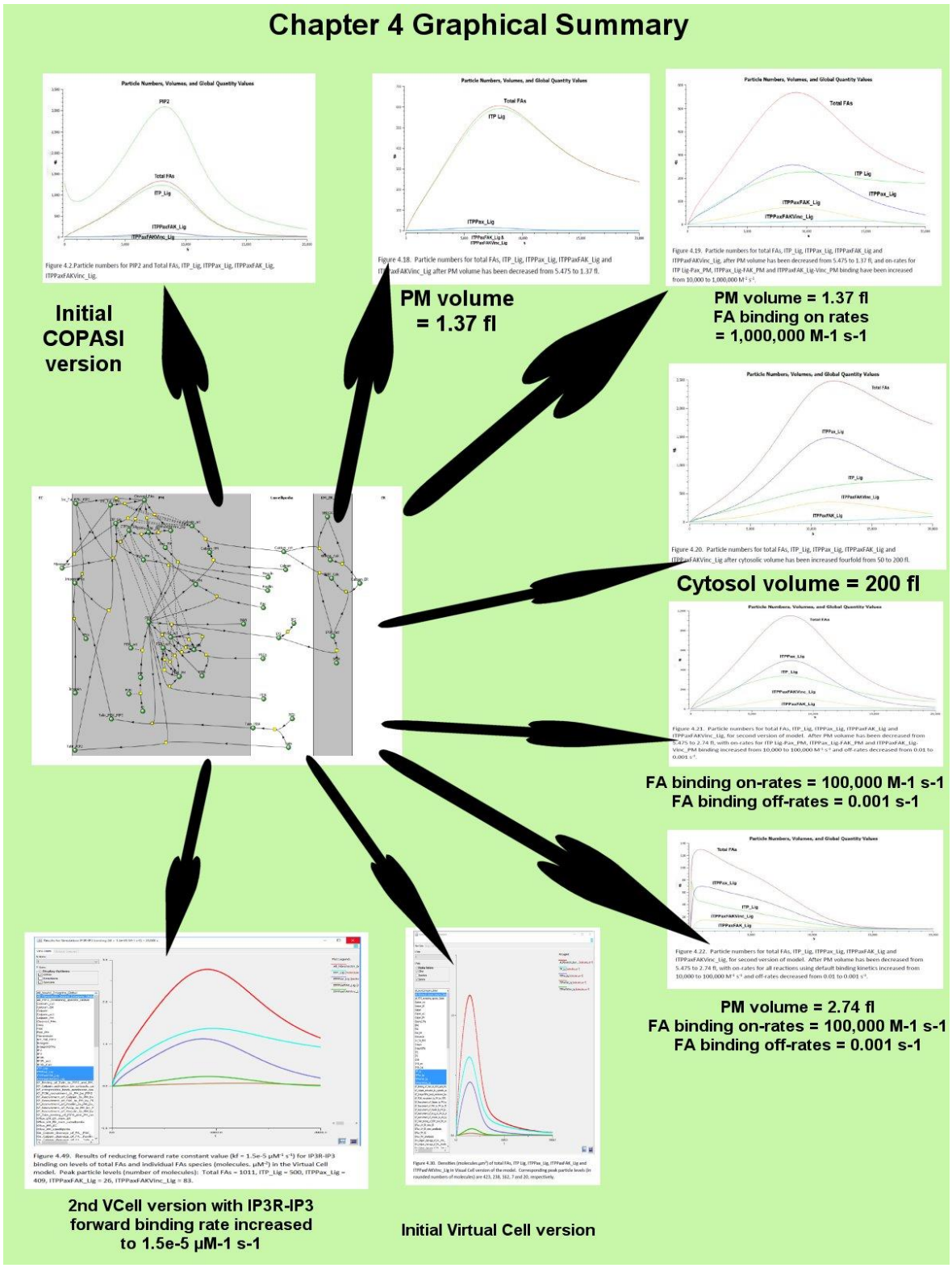

Figure 4.73. Graphical summary of Chapter 4 results

Chapter 5 - General Discussion of the project

5.1 Discussion

5.1.1 Introduction

It has been the intention throughout this project to investigate and illustrate various aspects of early FA assembly and disassembly, including their regulation, by the application of mathematical modelling. The focus here has been entirely on differential equation modelling, but many of the considerations and challenges encountered (most of them quantitative in nature) are also pertinent to other forms of modelling. This discussion will look first at the specific findings of the three different FA-related models, as outlined in Chapters 2, 3 and 4, in the light of the experimental literature around the subject. It will then move on to discuss wider aspects of modelling in cell biology.

5.1.2 The three FA-associated models

In pursuit of this purpose, the last three chapters (i.e. Chapters 2-4) have introduced, respectively, models of SDF1-mediated chemotaxis (focusing on events leading to lamellipodia formation), of integrin activation (representing the earliest stages of FA formation) in whole cells, and of a more complex model of FA dynamics in lamellipodia. All of these have been modelled in HeLa cells, using data from a quantitative proteome [182] to provide all protein copy numbers, and from the wider literature to obtain HeLa-specific compartment volumes [245,188,185]. By basing all three models on one cell line, in this way, it was intended to increase the relevance of the models to each other, and therefore, of any comparisons of results between them. It is hoped that this, in turn, will make any insights drawn from such comparisons more widely applicable to cell biology in general.

Whilst many of the insights and challenges that arose were specific to the model in question, some themes have also arisen that were common to all, many of them applicable to differential equation modelling in general and, indeed, to various other approaches to modelling cell biological behaviour.

5.1.2.1 The COPASI SDF1-CXCR4 lamellipodia model

5.1.2.1.1 Importance of which compartment a model species is resident in

For instance, in the SDF1-chemotaxis model, as implemented in COPASI, it quickly became clear that it is often critical which compartment model species are deemed to be resident in. As we have seen, COPASI does not distinguish between membrane and liquid compartments nor allow explicit specification of any spatial relationship between them, such as Virtual Cell does in both cases. So the importance of which compartment a species is resident in must lie purely in the volume of that compartment, affecting, as it does, the rate of any reactions in which that species features.

Examples of this were seen in the case of the Gi G protein (both in associated form or as disassociated subunits), where, changing it from being cytosolic in version 1 of the model to being membrane-resident in version 2, resulted in substantially more activated PI3K (as shown by comparing Figure 2.11 with Figure 2.3), leading in turn to much more PIP3, Rac1GEF*.PIP3 and Rac1*(shown by comparing Figures 2.12-2.14 to Figures 2.8-2.10). If this were a spatial model then, given that Gβγ is required to activate PI3K (as assumed in this model), the fact that more PI3K is being activated more quickly could be put down to greater proximity of membrane-resident Gβγ to membrane-bound PI3K, as compared to cytosolic Gβγ in the first version. However, being a compartmental model, the reason can only be put down to a concentration effect. Despite there being considerably less Gβγ in the second version of the model (1700

molecules, as seen in Figure 2.11, as against 27,000 in the first version as seen in Figure 2.2 - i.e. a nearly 16-fold difference) there is a greater than 200-fold difference between the volumes of the two compartments, meaning that, overall, concentrations of PM-resident Gβγ are around 15 times higher in the second version than in the first. Consequently, PI5K activation by Gβγ proceeds around 15 times faster in the second version also.

Similarly, changing inactive Rac1GEF from being cytosolic in versions 1 and 2 to being membrane-resident in version 3 led to substantial increases in levels of Rac1GEF*.PIP3 and Rac1*. (Shown by comparing Figures 2.16-17 with Figures 2.13-14). However, in this case numbers of inactive Rac1GEFs were the same in both versions, meaning that the concentrations of Rac1GEFs were around 234 times higher in the second version than in the first, this being the actual difference in volume between the PM (6.4 fl) and the cytosol (1500 fl).Since activated Rac1 is known to be a principal driver of lamellipodia formation [317,114] (as described in Chapter 1), this means that, if this model had been extended to include such formation (as had been the original intention), assumptions about which compartment certain species are resident in would be likely to have important consequences for the rate and extent at which it was seen to occur. This would obviously have had major consequences for the intended FA submodel, into which input such as phosphoinositides and actin-associated data was to be fed from the lamellipodia submodel.

5.1.2.1.2 Importance of PIP2 production in determining downstream species levels

Much as expected, PIP2 levels were shown to have a major influence on downstream species. Changes in version 4 of the model, so that PIP2 levels were no longer fixed (at 2.2e-08 M) but, instead, were the product of PI5K phosphorylation from PIP (itself the product of PI4K phosphorylation of PI) resulted in a more than fivefold increase in peak PIP2 levels (at around 105,000 molecules - as shown in Figure 2.22 - as against the fixed levels of 20,000 molecules in previous versions of the model). Although these were not immediately available, as when fixed, the fact that these peaks were achieved within 30 seconds and largely sustained at these levels

thereafter (by virtue of being greatly in excess of any other reactants in downstream reactions it was involved in), meant that there was a nearly 20-fold increase in peak PIP3 levels, by virtue of a roughly fourfold increase in rates of PIP3 production from PIP2 by PI3K. (This can be seen by comparing Figures 2.24 with 2.25.)

Again, this is largely down to the concentration differences in PIP2 between version 4 and previous versions.

5.1.2.1.3 Issues relating to G protein dynamics

The use of events to introduce SDF1 after a delay of 1500 seconds, coupled with its removal 7000 seconds later (rather than being made available immediately and sustained at a fixed level thereafter), uncovered a fault in the model, relating to G protein dynamics, not evident in previous versions (Figures 2.28-31). Even more surprising, perhaps, was the discovery that much of the reason for the faulty behaviour stemmed from uncritical use of a (G protein) submodel adapted from another published model, in which this submodel appeared to be working acceptably [181]. Although amendments to the more unconventional reactions in the schema resulted in a model with generally more acceptable behaviour (in the sense of producing more Gβγ, following SDF1-mediated activation of CXCR4, as one would expect from a GPCR), it also resulted in some further unexpected behaviour. Most prominently, this included a "double-bounce" effect in which removal of SDF1 was seen to result in increased Gβγ levels, further to the enhanced levels already seen when SDF1 was first introduced (Figure 2.32. Unsurprisingly, this double increase is also seen in downstream products, such as activated Rac1GEFs and Rac1 (Figure 2.33).

Whilst this can be welcomed as a novel testable prediction of the model, the circumstances in which it arose means that it is one that any modeller would want to investigate much further, using different G protein submodel variants, before drawing attention to it more publicly. Clearly

it would have been better to have tested out the submodel in more detail before using it in the model. However, it serves a purpose here as another instructive example of the types of (usually unexpected) pitfalls that can arise when implementing a model.

5.1.2.1.4 Relative values of GTP and GDP

Another somewhat surprising discovery was that the ratio of GTP to GDP had a large effect on subsequent model behaviour, even when numbers of both were hugely in excess of those of other model species. Changing concentrations of GTP and GDP from one millimolar (or around 900,000,000 molecules) each to 500 µM and 50 µM (450,000,000 and 45,000,000 molecules) respectively, resulted in a fourfold increase in peak Gβγ levels (as seen by comparing Figure 2.36 with 2.32). This, in turn, led to increases of around 100% and 60% in peak levels of activated Rac1GEF and Rac1, respectively. (As can be seen by comparing Figure 2.37 with Figure 2.33.) Parameter scans confirmed that this particular case was in conformity with a general trend in which increases in GTP and decreases in GDP both lead to increased peak Gβγ levels (Figures 2.38-9). Moreover this is a trend that is perfectly explicable, given the need for GTP replacement of GDP on the G protein to allow subsequent Gαi and Gβγ subunit dissociation [166,191]. Nevertheless, it is still somewhat surprising to discover that this holds true to the extent that it does, given the numbers just mentioned.

5.1.2.2 The Virtual Cell integrin-ECM ligation/ early FA model

In the next chapter (Chapter 3) a model of Integrin activation and very early FA formation in HeLa cells was implemented, this time using the modelling tool Virtual Cell. At first this was implemented as a compartmental model, which was then further developed into a 2D spatial version.

5.1.2.2.1 Model implementation differences between Virtual Cell and COPASI

In describing how this model was implemented, it was shown how Virtual Cell differs from COPASI, particularly in its representation of membranes as surfaces rather than volumes. This in turn means that reactions in which all reactants are membrane-resident require reaction rate constants for second-and higher-order reactions to be translated from volumetric rates (in the case of second-order, meaning units of M^{-1} s^{-1}) to areametric rates (molecules^{-1}.µm^{-2}.s^{-1}), achieved by dividing the volumetric rate by an appropriate value for relevant membrane thickness. Strictly speaking this requires knowledge of the lipid membrane used in the experimental setting in which the rate in question was measured, but such details are seldom available in sufficient detail, meaning that an estimate based on typical PM values is arguably the next best option.

5.1.2.2.2 Phosphoinositides as principal drivers of key model events

Having developed the model to a workable first version, the results showed it to be working acceptably. That is to say, following a lag phase of around two minutes, integrin-ECM ligation is sufficient for the model to enter a second phase characterised by exponential increase in levels of species such as PI5K_act, PIP2 and Int_Tal_PIP2 (Figures 3.12, 3.13 and 3.7, respectively). Various species were seen to have reached near-minimal or -maximal levels within 300-500 seconds, e.g. ITP_Lig Lig (Figure 3.8), PI5K_act (Figure 3.12), PIP (Figure 3.14) and PIP2 (Figure 3.13). This timescale (i.e. around 5-8 minutes) appears to be consistent with various reports for average lifetimes for lamellipodia-associated FAs (ranging from seconds to minutes, [207,208,209]), particularly one very recent (unpublished) study of PIP2 co-localisation with individual FAs. This reports typical FA assembly-disassembly timescales of between 100 and 400 seconds as shown in Figure 5.1.

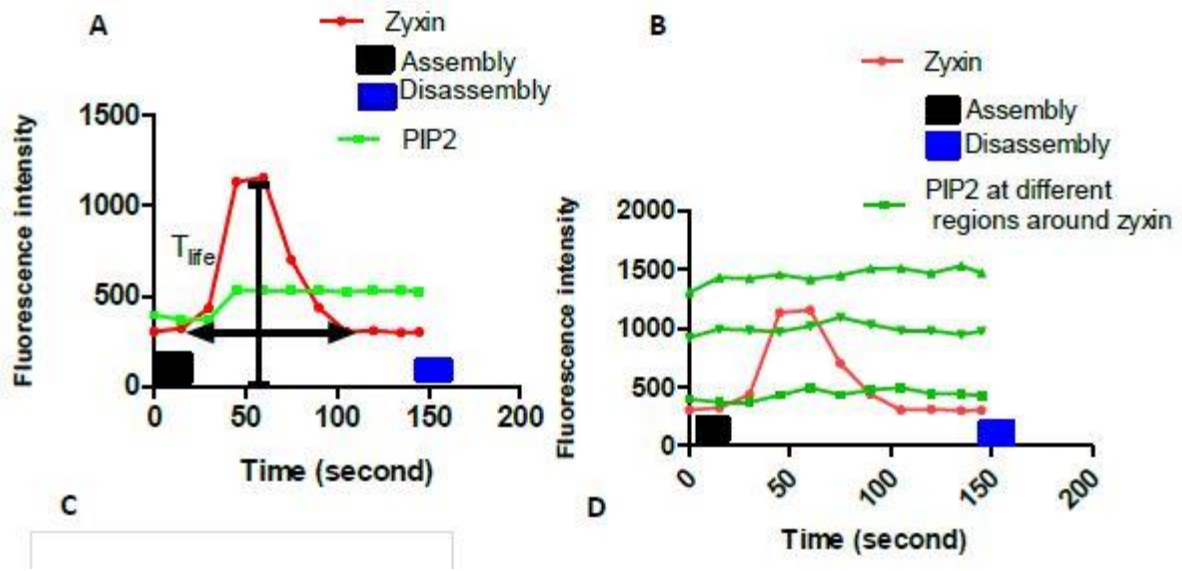

Figure 5.1. Dynamic range of local levels of PIP2 during zyxin turnover. A &B) Quantification of PIP2 within and around zyxin-associated FAs during assembly and disassembly. Key: Green curve = PIP2-associated fluorescent intensity; Red curve = zyxin-associated fluorescent intensity. Black squares = Assembly phase; Blue squares = Disassembly phase. Local levels of PIP2 constant during assembly and disassembly process. Fluorescent intensity recorded in 10 migrating MDA-MB-231 cells, n = 8. [Source: Dhurgham Al-Fahad (2017) Regulation of focal adhesions by PI(4,5)P2 and PI(3,4,5)P3 in cancer cell migration. PhD thesis, University of Reading].

The Virtual Cell integrin-binding model timescale also seems reasonably consistent with lifetimes of individual lamellipodia themselves (in the order of 3-10 minutes [203,204,205]), and with many reports for the timescale of cell spreading, which are of the same order [206,175]. Given that our current understanding of lamellipodia formation is that the majority fail to gain traction and are quickly retracted [292], this implies that most observable lamellipodia will be characterised by nascent adhesions [318]. The fact that this model demonstrates how outside-in signalling can generate very rapid integrin activation and ECM binding in response to PIP2 signalling, suggests that its basic underlying mechanism of operation is consistent with what is seen physiologically.

A number of findings point to exhaustion of PIP, and therefore of PIP2 creation, as the key factor in ending the served exponential phase of ECM-integrin ligation, and to PIP2 as the principal driver of key model events, such as talin and integrin activation. This is also consistent with a key result of the fourth version of the COPASI SDF1 inside-out signalling model, which was able to produce 105,000 molecules of PIP2 within 30 seconds of initial SDF1 signalling, suggesting that

that model would have been able to output PIP2 to the intended FA submodel at a sufficient rate to drive early FA assembly at a similar rate to what is seen in this outside-in model. Both models also raise important questions in terms of PIP availability for conversion into PIP2. If we assume that PI is being generated simultaneously to the events described in both models, with PIP being rapidly generated from PI by activated PI4K (as it clearly is in the first COPASI model), rather than assuming that PI levels (or PIP levels, as in this Virtual Cell model) are fixed, then it may well be the case that PIP2 formation can be extended to much higher levels than seen in both models. Obviously, this should allow even more rapid and extended integrin activation and ECM-ligation.

However, this would require the inclusion of a complete phosphoinositide cycle, similar to that used in the Purvis platelet signalling model [181] from which the G protein reaction schema was taken. The added complexity that this would involve can be seen from the schematic diagram (taken from this quoted model) in Figure 5.2, which only includes one of the many enzymes (essentially one per molecular transition) that are also involved.

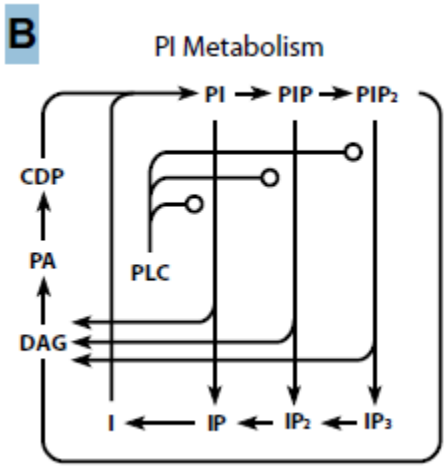

Figure 5.2. A schematic diagram of the Phosphoinositide module from a platelet signalling model, showing the various molecular stages in the phosphoinositide cycle. (Apart from PLC, none of the enzymes involved are included in this diagram.) Source: [181].

It is not clear that this Virtual Cell model justifies such an additional layer of complexity, given that, in some respects, its relative simplicity can be seen as a virtue (certainly in terms of being

readily amenable to analysis). However, its underlying assumption that PIP (and, by implication, PI) exhaustion is the dominant factor in integrin activation and ECM ligation, does not appear to have any real support in the literature. Therefore, any future version of this model needs to address this aspect, at the very least by having PIP availability more closely matching the physiological reality, as reflected in the literature.

5.1.2.2.3 Problems with modelling reactions

As it was, attempting to make PIP2 availability more physiologically realistic (in Version 2) led to an unanticipated problem. Specifically, instead of being membrane-resident, PI5K was divided between cytosolic and PM populations, with twice as many PI5K_cyto as PI5K_PM species, and with talin recruiting the former to the PM. (In fact, based on the relative sizes of the two compartments, there probably should have been a much higher proportion of cytosolic PI5K.) In the absence of any reaction data for this talin-mediated PI5K transport, an existing reaction (for talin-PIP2 binding, for which published rate constants were available), was adapted for the purpose. Unfortunately this resulted in the model running considerably slower, with key events, such as PIP exhaustion, timings for peak PIP2, Int_Tal_PIP2, ITP Lig and other species seen to be occurring after several thousand seconds had elapsed (Figures 3.15-3.26), rather than 400-600 seconds, as in the first version.

Subsequent investigation showed this slowing down was primarily because of the way talin-mediated PI5K recruitment had been modelled as a single reaction. Breaking this down into two simpler reactions (Tal_PI5K + PIP2 = Tal_PI5K_PIP2 and Tal_PI5K_PIP2 -> Tal_PIP2 + PI5K_PM) in place of the original reaction (Tal_PI5K + PIP2 = Tal_PI5K + PI5K_PM), resulted in simulations in which the same events, described above, occurred on a much faster timescale, similar to the original version. In the writer's experience it is generally best to avoid conflating multistep

reactions into one simplified version, as above, as a result is often a substantial slowing down, as seen here.

In the COPASI SDF1-CXCR4 lamellipodia model described previously, reliance on a published reaction schema proved to be misplaced. Here, the same appears to be true in the case of published rate constant values for fibronectin binding to α5β1 integrins, which were used in the generic ECM-integrin binding reaction for this model. Whilst the model appeared to be running acceptably in terms of the timescale of reported events, a critical defect remained, namely that far too low a proportion of activated integrins (in the form of Int_Tal_PIP2) were being ligated by ECM species.

Comparison with other published experimental investigations of such binding subsequently led to the conclusion that the reverse rate constant used in the model was likely to be too high by several orders of magnitude. Reducing this value from 49 s^{-1} to 0.1 s^{-1} led to a considerable increase in integrin ligation by ECM species, with peak levels of ITP_Lig increasing from 185 molecules to 22,480 molecules, out of a possible total of 23,600.

As well as increasing integrin-ECM binding levels, this change to the reverse rate in the sixth version of the model also led to a further gain in overall model speed. Now, for instance PIP exhaustion occurred at 200 seconds after model commencement, rather than 400 seconds, as seen in Version 3. This illustrates that, even in this rather simplified model of very early FA dynamics, the positive feedback nature of the relationship between integrin-ECM ligation, PIP2 generation and subsequent talin-mediated integrin activation, means that every factor affecting one part of this overall mechanism is likely to have effects on the whole. Thus if PIP2 has emerged from this model as the principal driver of events in very early FA dynamics, it is far from being the only one.

It also illustrates the point that even simplified models of physiologically complex mechanisms in cell biology can give subtle insights into their behaviour. In Chapter 4 this model has been

extended to incorporate many other factors in early FA dynamics. Whilst still a fairly simplified model of the underlying physiology, it is, nevertheless, a much more complex model. Whether this is beneficial, in terms of understanding this underlying physiology, will be addressed in the relevant discussion.

5.1.2.2.4 A comparison between the compartmental and 2D spatial versions of this integrin-activation/early FA model

Finally, concerning this model, once it was established that the sixth version was generally behaving acceptably (in terms of still meeting the same criteria mentioned earlier), 2D spatial version of it was implemented (Version 7). After adjusting PM scaling rates to ensure that all compartments sizes were a reasonable match for the corresponding sizes in the compartmental model, results from simulation runs were also seen to be broadly very similar to results from the compartmental version. Whilst reassuring in one respect, it also raises the question whether such spatial modelling adds sufficient value to justify the obvious overheads, both in terms of additional computational requirements and in the additional time required for adequately specifying such models.

Obviously, the answer to this depends entirely on the questions being asked. Compartmental models only show total levels of a species (as concentrations or particle numbers within their resident compartment) at a given time, whereas spatial models can show these, but also local concentrations at specific points in the model geometry. Consequently, any meaningful comparison of results between compartmental and spatial models can only involve compartmental totals (or concentrations), as in this case. Clearly, such a comparison, on its own, will lead to unfair conclusions.

For this reason, amongst others, a further brief comparison of results from two Virtual Cell compartmental and 2D spatial versions of the same model is included in the next chapter. In this

case, though, a rather more complex spatial geometry was used, and changes in the diffusion rates of two critical model species are tested. This latter addition was included in order to help determine if assumptions about diffusion rates are critical to overall model behaviour. It was also hoped that it will help determine at what stage (or in what cases), in future model development, if at all, it may make sense to resort to spatial rather than compartmental modelling.

5.1.2.3 More complex model of early FA assembly and disassembly in lamellipodia

5.1.2.3.1 Recapitulation of the FA assembly/disassembly model

In Chapter 4 the integrin activation model just described was developed into a much more complex model of early FA assembly and disassembly, incorporating many of the species seen in the previous two models. (I.e. the SDF1-CXCR4 chemotaxis lamellipodia model and the integrin activation model.) These included species such as ECM (a generic integrin ligand), integrin itself, talin, PI4K, PI5K, and PI3K, as well as the phosphoinositides PI, PIP, PIP2 and PIP3. However, this new model extends the phosphoinositide mechanism of the earlier versions, with the addition of the PIP2-hydrolysing protein PLCγ, together with its breakdown products IP3 and DAG [319].

This was included not so much to make the phosphoinositide dynamics more realistic, but primarily because IP3 was needed to trigger the rather simplified mechanism of calcium release [319]. This required the addition of a further internal calcium-storing compartment (the endoplasmic reticulum or ER [296]) and two pump proteins, IP3R (which releases calcium into the cytosol, in response to IP3 binding) and SERCA (which constitutively pumps cytosolic calcium back into the ER) [320].

This release of cytosolic calcium was required to activate calpain [321], with this activated calpain cleaving the various FA isoforms [322]. (Such calpain cleavage having been widely reported to be very important in terms of FA dynamics, certainly at an early stage [322,269,271].) By having rates of

calcium release determined by IP3 levels, cleavage rates were indirectly tied to FA levels, by virtue of the fact that PLCγ had to be activated by these SFK-containing FAs [279] before it could hydrolyse PIP2. It is known that such SFK-mediated activation of PLCγ, requires that SFKs have themselves been activated, in response to integrin ligation by fibronectin [279] and other ECM ligands. Therefore IP3-driven calcium release (and, in turn, calpain cleavage of FAs) is indirectly tied to the level of integrin-ECM ligation, and so also to early FA formation. Finally, before PLCγ can be activated by SFKs, it must first be recruited to the PM by PIP3 [276,275], meaning that levels of activated PLCγ are indirectly also tied to levels of PIP2, this being phosphorylated by PI3K to PIP3. So, altogether a very complex network of interdependencies was evident in this final model, yet one that was well-supported by the literature.

Finally, in addition to being bound by talin (required for integrin activation, as in the Chapter 3 model), integrin has first to be bound by SFKs, then (according to this model) in strict sequence, talin, paxillin, FAK and vinculin. This assumption of very prescriptive recruitment was included mainly in order to test how such assumptions dictate overall model behaviour.

5.1.2.3.2 Role of PIP2 in early FA dynamics

Having seen how PIP2 independently affects integrin dynamics in the Chapter 3 model, this new model was implemented to understand how PIP2 and calpain cleavage together might regulate FA assembly and disassembly. As can be appreciated from the description above (and explained in more detail in the introduction to Chapter 4), this is a challenging question, involving, as it does, both aspects of positive and negative feedback in both cases. By incorporating both these mechanisms of FA regulation it was felt that such a model might yield interesting insights and predictions, some of them unexpected.

One expected result of the Chapter 4 FA assembly/disassembly model was that PIP2 was seen to primarily influence the FA assembly rate, in that the rate of PIP2 formation appeared to directly influence the rate of FA formation. In this respect this finding concurred with the findings of the

Virtual Cell integrin activation model. However, once PIP2 levels exceed a certain threshold level (which appear to be quite low), other factors are seen to become more important in determining overall FAs levels.

5.1.2.3.3 Role of calpain cleavage in early FA dynamics

Whilst such factors include availability of FA building blocks such as integrin, talin, FAK, as well as overall FA-related binding rates, the main determinant was shown to be calpain cleavage. Amongst the strongest evidence in support of this conclusion was seen when the compartmental version of the model was ported from COPASI into Virtual Cell, resulting in a substantial reduction in overall FA numbers, from 937 molecules to 423, a fall of 55%.

This was subsequently determined to be the result of greater sensitivity to IP3 in the latter model, leading to faster calcium release and therefore of calcium-induced calpain activation, leading in turn to greater rates of calpain cleavage of FAs. Once this IP3 sensitivity issue was addressed FA levels in the Virtual Cell version increased to match those seen in the COPASI version.

5.1.2.3.4 Role of average FA binding rates in early FA dynamics

Average FA binding rates, by contrast, were seen to be most important in determining the individual representation of FA isoforms within the overall total. A lower binding rate (with the forward rate constant set at 10,000 M^{-1} s^{-1} and the reverse rate at 0.01 s^{-1}) resulted in ITP_Lig (i.e. FAs without paxillin, FAK or vinculin) representing over 90% of total FAs (Figure 4.11). By contrast, a higher binding rate (with the forward rate constant increased to 1,000,000 M^{-1} s^{-1}, and the reverse rate unchanged) was associated with a more even distribution (Figure 4.18). However, given the various assumptions of the model, including prescriptive sequential binding, this finding should be treated with particular caution, and would be very hard to test experimentally. In any case, the proposed average forward rate constant value of 1,000,000 M^{-1} s^{-1} is so high that it would require firm support in the literature before taking this finding seriously.

5.1.2.3.5 Role of PM and cytosolic compartment volumes in early FA dynamics

The effects of altering the volume of the PM and cytosol compartments in the Chapter 4 COPASI model provided an interesting comparison, in that reducing the PM volume by half showed little discernible effect on the timings of peak levels of FA isoforms (Figures 4.11 and 4.17). By contrast, reducing the cytosolic volume by 75% resulted in a similar reduction in these timings (Figures 4.11 and 4.70). Exactly why this difference is seen is not clear at this time. Certainly one might reasonably expect that rates would increase in all reactions involving PM species, given that reducing the PM volume results in a corresponding increase in the concentration of all PM-resident species. This certainly seems to be the best explanation for the results from reducing the cytosolic volume.

However, the opposite is seen. Whilst peak timings remain the same, peak levels are halved, implying that the relevant reaction rates have fallen by half on average, also. In other words, halving the PM volume appears to have halved the average rates of these reactions - the opposite of what is seen in the cytosol. This may be the result of the more complex interdependencies between PM-resident species (including positive and negative feedback), as compared to the relative lack of such interdependencies between cytosolic species. But only a further, detailed investigation could establish this.

That aside, these findings do suggest that the relatively rapid rate of lamellipodia formation seen, for instance, during cell spreading, may be possible, in part, because the rapid accretion of actin filaments, microtubules, peripheral ER and other molecular structures characteristic of such formation substantially reduces the effective volume of lamellipodia. This, in turn, should result in the rates of all reactions involving species within that volume (including those related to FA formation) tending to run much faster than they would otherwise do. Of course, this only holds true if the structures mentioned do not substantially impede access of FA-related species to the leading edge of lamellipodia, where FAs are forming.

5.1.2.3.6 Implications of the findings concerning FAK

Another interesting finding concerned FAK. If we assume that the copy numbers for FAK taken from the HeLa cell quantitative proteome are reasonably accurate, then their low apparent availability is problematic if we also assume highly prescriptive FA binding, as indicated by super-resolution microscopy (Figure 1.1) and our current understanding of FA assembly. Certainly it appears that FAK must be very efficiently used in HeLa cell lamellipodia formation, perhaps through directed transport to lamellipodia, rapid re-use or recycling, both highly localised. The only alternatives to this appear to be rapid increases in FAK protein expression (implying some form of signalling from lamellipodia to the nucleus) or high levels of redundancy, with some other protein or proteins performing the same roles as FAK. This has been discussed in more detail in section 4.5, but would appear to have wider implications.

5.1.2.3.7 The wider implications of assuming inflexible FA binding mechanisms and limited redundancy for key FA proteins

As just explained, the problems relating to FAK largely come down to assumptions about its role in FA assembly and disassembly, whether this role can be performed by other proteins, as well as assumptions about the manner in which it binds to FAs. However, this also can also be said to apply to other key FA proteins. Whilst a number of assumptions in the model may justifiably be called into question, the problems that arise from them, particularly the gross underrepresentation of more mature FA forms (containing FAK or vinculin, but even FAs containing only talin and paxillin) in the overall FA total, are not in doubt. The value of such models as this often lie in determining which of the assumptions best explain the aberrant behaviour (assuming one is sure the behaviour is aberrant) and what, if anything, that says about the underlying physiology, or our current understanding of it. Having discounted some

possibilities, including absolute and relative compartment volumes, and PIP2 availability, this has still to be determined, and would seem to merit further investigation.

5.1.2.4 A summary of findings and insights from the three models

Whatever the explanation, the fact that this question has arisen at all gives a further illustration of how the highly quantitative nature of mathematical modelling offers a different way of looking at systems in cell biology, often raising questions that might not result from other investigative approaches.

5.1.2.4.1 The critical importance of PIP2 in FA and lamellipodial dynamics generally

A clear theme from all models is that rapid PIP2 availability is critical to all aspects of both lamellipodia and FA dynamics. In particular, without it being available in large quantities (and, just as importantly, very high local concentrations), and rapidly so in response to appropriate stimuli, neither inside-out nor outside-in signalling is likely to occur on a timescale commensurate with what is seen physiologically, in terms of observed rates of FA assembly and disassembly, lamellipodia formation and cell spreading. In turn, this means that all aspects determining this PIP2 availability, including PI and PIP availability, PI4K and PI5K activation rates (and activation mechanisms), are likely to be critical to these commonly occurring cellular behaviours also.

5.1.2.4.2 Scale, complexity and timescales

There is a clear problem seen here in terms of the overall timescale of FA-critical events in the two more complex, large-scale models (i.e. the SDF1 model described in Chapter 2 and the Chapter 4 FA model). Whereas the smaller-scale integrin-activation model in Chapter 3 was observed to operate over a timescale of a few minutes (with rates of integrin activation, integrin-ECM binding and PIP2 formation apparently commensurate with physiological requirements), the two more complex models, just mentioned, consistently operated over timescales of several

hours or more. Obviously this does not seem to be a very likely reflection of the dynamics that they are supposed to model. The fact that the simpler integrin activation model did not successfully scale up in this regard, clearly raises questions about how complex systems, such as FA dynamics, with their numerous participants and highly complex interactions, operate at the rapid rates that are observed.

5.1.2.4.3 Concentrations and compartment volumes

Part of the answer to this may lie in another common finding of the three models, namely that local (meaning in this context, predominantly compartmental) concentrations of key system species can be critical to overall system performance, as was seen in the case of G proteins and inactive Rac1GEFs in the SDF1 chemotaxis model. Undoubtedly, the fact that so many FA-related proteins and lipids are PM-associated, in some form or other, prior to (or very soon after) FA assembly begins, is a critical prerequisite if such assembly is to occur over the short timescales (seemingly seconds or minutes) observed.

However, this does not seem to be enough on its own. Clearly, the fact that many FA-critical proteins are recruited directly to existing FAs, rather than just to the PM, is a very important factor, meaning that their local concentrations will tend to be much higher than the corresponding PM concentrations. This is an obvious limitation of having the PM as the smallest possible compartment in which FA assembly can occur. The fact that increasing evidence points to FA formation being confined to much smaller lipid rafts and other PM microdomains [323,324,325,89,326], provides an obvious means whereby local concentrations of FA-critical proteins and lipids can be much higher.

To some extent this should have been addressed by the 2D spatial models, both of which assumed that PI5K has to be activated by ECM-ligated FAs, meaning that PIP2 should be higher in the vicinity of FAs. (Obviously something not possible in the compartmental models, despite the same underlying assumption.) However, in both the 2D models, activated PI5K and PIP2 were

able to diffuse freely within the PM, contrary to what one would expect if their activities were confined within microdomains. On this basis, extending the FA lamellipodia model to include such microdomains would appear to be a priority for future model development.

However, it should be remembered that reducing the volume of the PM by half in this model did not have the desired effects, effectively halving the rate of FA formation and thus leaving overall numbers largely unchanged. (Compare Figures 4.18 and 4.12.) Whether this finding has any applicability physiologically, it does suggest that increasing local concentrations of interacting species, by simply reducing the effective volumes in which they interact, may not be sufficient to accelerate reaction rates, where there are complex relationships between these species, particularly where negative feedback is involved.

The fact that such complex interactions are less in evidence among the cytosolic species, by contrast, may explain why reducing the volume of this compartment by 75% had the expected effect of reducing the time for FA species to reach peak levels by a similar amount, without affecting those levels adversely. (Compare Figure 4.71 with 4.12.) However, 2500 seconds for such peaks to occur, still represents around 40 minutes, which seems to be on the high side of estimates for FA assembly, lamellipodia formation and cell spreading, referred to earlier.

More critically, estimates, by the writer, of how much the effective volume of lamellipodia could be reduced by the presence of actin fibres, peripheral ER and microtubules suggests that a reduction of 25% would seem more appropriate. Obviously, however, this did not take into account other structures, such as mitochondria and ribosomes, as well as the very many proteins and other molecules in the generally crowded space of the lamellipodial cytosol. Certainly, the findings of all three models strongly suggest that this is an important subject for further investigation by microscopy and other investigative means. Similarly, determining whether FAs are indeed exclusively confined within microdomains, and what the average sizes of these are, is another area of obvious importance.

5.1.3 General modelling considerations

A number of issues have arisen repeatedly in the course of developing the models, described above, that clearly applied generally. These include the general lack of modelling-relevant data in the literature, at least in a form that can be used in models.

5.1.3.1 Lack of quantitative proteomes

In particular, the question concerning FAK highlighted the current lack of published cellular quantitative proteomes. As a result, it was not possible to determine whether the low levels of FAK reported in the HeLa cell (as well as the U2OS) proteome could be relied on, or whether they might have been substantially underestimated for some reason. Without many more such proteomes it will be difficult to arrive at a consensus view concerning such protein copy numbers, limiting the reliability of models that use such numbers.

The FAK issue also raises the question whether cellular quantitative proteomes are sufficient on their own for modelling purposes. It was suggested that one explanation of how the low cellular FAK numbers might be consistent with the large numbers of FAK-containing FAs seen in lamellipodia might be that FAK numbers are somehow enriched within lamellipodia, as compared with the rest of the cytosol. Indeed, such explanations may well be true for other proteins in other contexts. Yet this still raises the question as to why this should apply to these particular proteins and not to others. For instance, why are Rac1, talin and vinculin numbers so high at a cellular level, when they could presumably also use the same localised enrichment mechanisms hypothesised for FAK?

For these reasons modellers also need quantitative proteomes specific to particular cellular structures and behaviours, such as lamellipodia. Without these, it is not possible to be sure about the actual levels of species within (or in the immediate vicinity of) these structures, or in the immediate locality of these behaviours. In fact, such specialist proteomes are now beginning to

be produced, but unfortunately, at least in the case of FAs, mainly seem to be largely semi quantitative in nature [327]. Whilst informative in themselves, in terms of relative numbers of some species, they don't currently provide the absolute numbers needed by modellers in terms of providing initial quantities for their models. These are urgently needed if we are to have more confidence in the models that use them.

5.1.3.2 Lack of reaction rate data

Another deficiency that emerges from these models is the general lack of published data relating to reactions, particularly rate constant values. As a consequence, many of the reactions used in these models relied on ad hoc rate constant values, usually based on estimates of the likely range they might fall within. As with behaviour-specific quantitative proteomes, there are FA-relevant interactomes [328], but, again, these do not provide the relevant reaction rates as hard numbers.

5.1.3.2.1 Wide variability of published reaction rate data

Moreover, where such data is available, the few reported values may vary so widely that it is hard to determine the most appropriate value to use. This can be readily appreciated by consulting the BRENDA enzyme database at random [329]. While some of this variation is probably unavoidable, given all the variables that can affect reactions, a large part of the problem appears to be that different studies use different experimental methodologies, inevitably reporting widely different results as a consequence [246]. As a minimum, standardised methodologies need to be agreed for experiments of this type, no doubt requiring different standards for different protein classes and reaction types. It should then be a requirement that published results for such reactions should have adopted these methodologies. Of course, this would not just be of benefit to the modelling community. Any kind of consensus around reaction rates is unlikely in the absence of such standards.

5.1.3.2.2 Limited usefulness of Kd values for reaction-modelling purposes

On the same theme, it has been mentioned a number of times that Kd values are of limited benefit to modellers when it comes to modelling binding reactions. In most cases these need to be modelled in the form of reversible mass action reactions, but this requires the relevant rate constants for the component forward and reverse reactions, and these values are not provided by Kds. As explained earlier, because Kd values are obtained by dividing the reverse rate by the forward rate for the relevant binding reaction, this means that the same Kd value can be obtained from an infinite combination of forward and reverse rates. Randomly adopting one particular combination for ones model is generally not advisable as it often turns out (at some point in model development, if not immediately) that different model outputs results from different combinations.

Thus, for modelling purposes, Kd values are not sufficient on their own. One might argue that they are not very helpful for the wider field of cell biology either, since they can hide much important information. Two different proteins with similar Kd values for a particular ligand may have very different relationships with the ligand in question if one protein has very high on- and off-rates, whereas both these rates are very low for the other protein.

Kd values taken in isolation can be very misleading for these reasons. After all, a large number of relatively weak and transitory interactions can sum up to a much stronger overall interaction. A good example of this is seen in the case of highly glycosylated proteins such as the selectins, whose polysaccharide chains "bind weakly but with rapid on-rates to ligands exposed at sites of blood vessel injury" [330]. More relevantly to FAs are syndecans, such as syndecan-4 that perform a similar role in wounded dermis, binding in the same manner as selectins to fibronectin and vitronectin, and triggering Rac1 activation in the process [330]. Such binding might well be missed by relying on individual Kd values alone, showing why they are of such limited value for modelling purposes.

5.1.3.2.3 Challenges of converting published reaction rate data for modelling purposes

Continuing with the challenges of obtaining the data required for the three models from the literature, another challenge faced was converting many second-order volumetric rate constant values (obtained from the literature) into areametric equivalents, suitable for use in the Virtual Cell models. In this case a simple solution was used, which was to divide such values by 7.5 nm, chosen as a suitable value for PM thickness, meaning it was also judged a suitable volume-to-surface ratio for PMs. However, there is no guarantee that such a value is applicable to the lipid bilayers, let alone lipid monolayers, used in any experiments in which the original volumetric rate constant was obtained. Unfortunately such information is not usually provided with the published data for such experiments. Therefore it is not clear how reliable the areametric rates used in these examples were. Again, this is an area that would benefit greatly from the establishment of standards.

5.1.3.3 Cell membrane representations

An obvious question arising from this issue, is why COPASI and Virtual Cell model membrane so differently. Certainly, modelling membranes as volumes, as COPASI does, means that one does not have to convert volumetric reaction rates. Also, the fact that COPASI does not permit spatial modelling means that there is a less obvious need to represent membranes in accordance with their physiological spatial dimensions. In contrast, by representing membranes as surfaces, Virtual Cell makes the progression from a compartmental to a spatial model relatively straightforward, as was demonstrated in the case of the Chapter 3 and 4 models.

But there is arguably a wider merit in the way that Virtual Cell models membranes as surfaces, providing a more realistic context for membrane-based reactions. Clearly these are different from reactions occurring in liquid compartments, in terms of the more limited access reactants have to

each other. This is most obvious in cases where all reactants are membrane-associated, where crowding-out effects can be particularly rate-limiting [331,332,333]. This is why Virtual Cell issues warnings against using mass action kinetics for such reactions, providing two General reaction types (one for lumped kinetics) in which the user is forced to enter a relevant expression, rather than simply entering the usual rate constant values. By this method, the user is essentially guided to consider alternative kinetics that take into account crowding-out and other membrane-specific effects.

Is this necessarily a better approach than that adopted by COPASI? In fact most reactions involving membrane-associated reactants would appear to take place on the surface of the membrane, rather than within it [334], even in cases where all reactants are considered membrane-resident. This is certainly the case with FA-related reactions. However, given their role as a structural link between membrane-embedded ECM-ligated integrins and the cytosolic cytoskeleton, FAs are a rather poor example.

Nevertheless, there are good reasons to believe that most membrane-reactions will occur on the surface rather than within the membrane. Firstly, the faster diffusion rates in the cytosol compared to the PM, together with reduced scope for crowding-out effects, mean that such reactions can proceed much more rapidly on the surface. Secondly, most forms of membrane-recruitment of cytosolic proteins involve attachment to the membrane surface by various forms of lipidation, as well as by hydrophobic and electrostatic attachment to integral membrane lipids and proteins [335], rather than by embedding within it. (Apart from anything else, the usually temporary nature of such recruitments would suggest that attachment is far more efficient than embedding.)

Therefore it seems likely that most membrane-associated reactions will occur within a narrow band close to the surface of membranes, probably not much deeper in most cases than the thickness of the membrane itself. This implies an effective reaction volume similar to that of the

membrane. Thus it may be that in most cases (if possibly only fortuitously) the COPASI volumetric approach to membrane-associated reactions is the more appropriate one. However, given the importance of such reactions to cell functioning, this is something that deserves wider discussion within the cell biology modelling community, in the hope of deriving some sort of consensus. This in turn, might hopefully lead to a standard (or set of standards for different cases) for the handling of such reactions.

5.1.4 Cell biology and mathematical modelling

Moving on from the challenges faced by modellers and developers of modelling software, an obvious question is what benefits modelling can bring to cell biology, and why it is not used more widely. It is hoped that this report has shown some of the benefits, in terms of addressing subjects of interest in a more rigorously quantitative fashion, and in generating original insights and questions that are less likely to arise from other less quantitative approaches. This is not to claim that mathematical modelling is better than other common approaches to cell biology, merely very different from them, and should therefore be generally welcomed for this reason alone. In this respect its relationship to other modes of investigation is rather like the relationship of microscopy, and other forms of imaging, to lab-based experimental approaches; neither being better or worse than the other, in terms of the insights or questions generated, but beneficial to cell biology, and to each other, by virtue of their differences.

5.1.4.1 Hodgkin-Huxley model of the squid giant axon

But, in order to be more widely accepted, mathematical modelling needs to deliver what it promises. Until comparatively recently it has been greatly hampered by a lack of suitable software tools and methodologies, and inadequate computational resources, as well by other factors, some of which will be addressed later. For this reason its impact on cell biology has not

been as extensive as it might have been. However, if a single example may serve to show its potential, it is that of the Hodgkin-Huxley model of the squid giant axon. First published in 1952, this model was the result of extensive lab-based investigation, principally involving voltage clamping experiments, which informed a slowly-evolving mathematical model, generated with the most basic of computational resources [336]. The insights and impact of this Nobel prize-winning model can perhaps best be summed up by the following two quotes:

"The modern history of ion channels began in 1952 when Hodgkin and Huxley published their seminal papers on the theory of the action potential in the squid giant axon…"[336]

"Not only was this the first quantitative description of electrical excitability in nerve cells, but it also incorporated physical correlates that elegantly predicted recent results concerning the gating structures of ion channels. It remains one of the best examples of how phenomenological description with mathematical modelling can reveal mechanisms long before they can be directly observed."[336]

Sadly, it is hard to find other examples within cell biology that come close, in terms of their impact, to this example. Undoubtedly, one reason for this is a general lack of understanding of what mathematical modelling is, or what relevance it has to the wider discipline. This may help explain the comparative lack of modelling-related research being undertaken, and a tendency for such research to be poorly-integrated with other forms of research. Regrettably, regarding the latter, too many published models appear to be more interested in the underlying mathematics than in the underlying biology, and, even when this is not the case, rather too many such papers seem poorly informed by whatever aspect of cell biology they are supposed to be modelling. In general, the field can seem rather semi-detached from the mainstream of the subject. Obviously this does not help in terms of achieving wider acceptance.

5.1.4.2 Comparison with the important role of mathematical modelling in physics

This is in complete contrast to physics, where mathematical modelling is so well accepted and integrated that the term could almost be said to be redundant. Instead, its closest correlate would be “theoretical physics”, a discipline that has indisputably contributed hugely to the

subject, and to our understanding of the world. Newton's laws, Maxwell's equations, Einstein's field equations and much of quantum mechanics are just a few of the more obvious examples of mathematical models that have had a huge impact on physics. Moreover, far from emerging in isolation from what was going on in the rest of physics, these models were addressing issues of great concern at the time, providing answers to pressing questions that had recently emerged from experiments and other forms of observation. Yet, in doing so, these models generated further questions, which required new experiments and observations to answer.

5.1.4.3 Possible reasons for the lack of modelling in cell biology

Why has mathematical modelling yet to have the same impact on cell biology? Part of the reason, undoubtedly, is that cells are so complex, both in terms of their structure and in their chemistry, something that can perhaps be most readily appreciated by consulting various relevant databases such as UniProtKB[337], KEGG Pathway [338], BRENDA [329], and similar. This may help explain why so few of the equations regularly used in cell biology have originated from the discipline itself, Michaelis-Menten enzyme kinetics being one of the few examples. (By contrast, the majority, e.g. Fick's laws of diffusion, the Nernst and Henderson-Hasselbalch equations, have emerged from other fields, mostly chemistry and physics.)

Perhaps because of this complexity, it seems fair to say that there is a general perception that cells and their behaviour are not very amenable to useful mathematical analysis. To some extent this may be true, and there is a clear risk from attempting to over-apply mathematics to cell behaviour. To take a celebrated quote from another science that attempts to understand very complex behaviour, "Mathematics brought rigor to economics. Unfortunately it also brought mortis." [339]. However, there is no reason to think that cell biology is in any imminent danger in this respect, and every reason to think that it could benefit greatly from the additional rigour that mathematical modelling, with its highly quantitative focus, can bring.

Put briefly, cell biology is not as quantitative as, probably, most cell biologist suppose. To apply a fairly crude but hopefully instructive metric, if one takes a random sample of journal articles in the field and then places the various figures and tables into one of three categories, qualitative, semi-quantitative and quantitative, typically one finds that the great majority fall into the first two categories. (Where "qualitative" designates any figure not containing numbers, "quantitative" denotes any that yield absolute quantitative information about the object of interest and "semi-quantitative" denotes any that yield only relative values of one sort or another, including surrogate values, such as fluorescent intensities and densities, used as indicators of relative numbers.)

Of course the primary explanation for this is technical. It is extremely hard to measure things inside cells, especially if obtaining this information requires not disrupting the cell's normal behaviour. Often uniquely identifying the thing that one is trying to measure is hard enough. But, as shown by some of the examples above, concerning reaction rate constants, Kd values and quantitative proteomes, there also seems to be a lack of interest in providing detailed quantitative information in many areas. The hard reality seems to be that much of the information that modellers need is not currently available because it is not much needed by those working in laboratories or looking down microscopes. If physics is anything to go by, cell biology is all the poorer for this state of affairs.

5.1.4.4 The way forward

By its constant demands for reliable quantitative information, modelling can help redress this balance. The best way to achieve this is collaboratively. If lab-based researchers and microscopists are working on the same projects with modellers, it will be much easier to convince them of the reasons for providing the information the modeller needs. As has hopefully been

shown by the example of an incrementally developed model of FA dynamics described in this report, there is every reason to believe that the benefits of such collaboration are likely to be in both directions.

Bibliography

(1) Frantz, C.; Stewart, K. M.; Weaver, V. M. The Extracellular Matrix at a Glance. *J. Cell Sci.***2010**, *123* (24), 4195–4200.

(2) Yanagishita, M. Function of Proteoglycans in the Extracellular Matrix. *Acta Pathol. Jpn.***1993**, *43* (6), 283–293.

(3) Leblond, C. P.; Inoue, S. Structure, Composition, and Assembly of Basement Membrane. *Am. J. Anat.***1989**, *185* (4), 367–390.

(4) Bonnans, C.; Chou, J.; Werb, Z. Remodelling the Extracellular Matrix in Development and Disease. *Nat. Rev. Mol. Cell Biol.***2014**, *15* (12), 786–801.

(5) Lorentzen, A.; Bamber, J.; Sadok, A.; Elson-Schwab, I.; Marshall, C. J. An Ezrin-Rich, Rigid Uropod-like Structure Directs Movement of Amoeboid Blebbing Cells. *J. Cell Sci.***2011**, *124* (8), 1256–1267.

(6) Visse, R.; Nagase, H. Matrix Metalloproteinases and Tissue Inhibitors of Metalloproteinases Structure, Function, and Biochemistry. *Circ. Res.***2003**, *92* (8), 827–839.

(7) Lawson, C.; Schlaepfer, D. D. Integrin Adhesions. *Cell Adhes. Migr.***2012**, *6* (4), 302–306.

(8) Kanchanawong, P.; Shtengel, G.; Pasapera, A. M.; Ramko, E. B.; Davidson, M. W.; Hess, H. F.; Waterman, C. M. Nanoscale Architecture of Integrin-Based Cell Adhesions. *Nature***2010**, *468* (7323), 580–584.

(9) Huttenlocher, A.; Horwitz, A. R. Integrins in Cell Migration. *Cold Spring Harb. Perspect. Biol.***2011**, *3* (9), a005074.

(10) Beningo, K. A.; Dembo, M.; Kaverina, I.; Small, J. V.; Wang, Y. Nascent Focal Adhesions Are Responsible for the Generation of Strong Propulsive Forces in Migrating Fibroblasts. *J. Cell Biol.***2001**, *153* (4), 881–888.

(11) Zaidel-Bar, R.; Itzkovitz, S.; Ma'ayan, A.; Iyengar, R.; Geiger, B. Functional Atlas of the Integrin Adhesome. *Nat. Cell Biol.***2007**, *9* (8), 858–867.

(12) Jaffer, Z. M.; Chernoff, J. P21-Activated Kinases: Three More Join the Pak. *Int. J. Biochem. Cell Biol.***2002**, *34* (7), 713–717.

(13) Amano, M.; Nakayama, M.; Kaibuchi, K. Rho-Kinase/ROCK: A Key Regulator of the Cytoskeleton and Cell Polarity. *Cytoskelet. Hoboken NJ***2010**, *67* (9), 545–554.

(14) Besson, A.; Assoian, R. K.; Roberts, J. M. Regulation of the Cytoskeleton: An Oncogenic Function for Cdk Inhibitors? *Nat. Rev. Cancer***2004**, *4* (12), 948–955.

(15) Nagano, M.; Hoshino, D.; Koshikawa, N.; Akizawa, T.; Seiki, M. Turnover of Focal Adhesions and Cancer Cell Migration. *Int. J. Cell Biol.***2012**, *2012*.

(16) Sharma, S. V. Rapid Recruitment of P120RasGAP and Its Associated Protein, P190RhoGAP, to the Cytoskeleton during Integrin Mediated Cell-Substrate Interaction. *Oncogene***1998**, *17* (3), 271–281.

(17) Deakin, N. O.; Turner, C. E. Paxillin Comes of Age. *J. Cell Sci.***2008**, *121* (15), 2435–2444.

(18) Petit, V.; Thiery, J.-P. Focal Adhesions: Structure and Dynamics. *Biol. Cell***2000**, *92* (7), 477–494.

(19) Hynes, R. O. Integrins: Bidirectional, Allosteric Signaling Machines. *Cell***2002**, *110* (6), 673–687.

(20) Calvete, J. J. Structures of Integrin Domains and Concerted Conformational Changes in the Bidirectional Signaling Mechanism of AlphaIIbbeta3. *Exp. Biol. Med. Maywood NJ***2004**, *229* (8), 732–744.

(21) Takada, Y.; Ye, X.; Simon, S. The Integrins. *Genome Biol.***2007**, *8* (5), 215.

(22) Humphries, J. D.; Byron, A.; Humphries, M. J. Integrin Ligands at a Glance. *J. Cell Sci.***2006**, *119* (19), 3901–3903.

(23) Liljas, L. Integrins http://xray.bmc.uu.se/lars/Practicals/Integrin/integrin.html (accessed Oct 19, 2013).

(24) Xiong, J.-P.; Stehle, T.; Goodman, S. L.; Arnaout, M. A. A Novel Adaptation of the Integrin PSI Domain Revealed from Its Crystal Structure. *J. Biol. Chem.***2004**, *279* (39), 40252–40254.
(25) NIH. DBP6: Integrin http://www.ks.uiuc.edu/Research/dbps/dbp6.html (accessed Oct 22, 2013).
(26) Hytönen, V. P.; Vogel, V. How Force Might Activate Talin's Vinculin Binding Sites: SMD Reveals a Structural Mechanism. *PLoS Comput Biol***2008**, *4* (2), e24.
(27) Critchley, D. R.; Gingras, A. R. Talin at a Glance. *J. Cell Sci.***2008**, *121* (9), 1345–1347.
(28) Ulmer, T. S.; Calderwood, D. A.; Ginsberg, M. H.; Campbell, I. D. Domain-Specific Interactions of Talin with the Membrane-Proximal Region of the Integrin Beta3 Subunit. *Biochemistry (Mosc.)***2003**, *42* (27), 8307–8312.
(29) Shattil, S. J.; Kim, C.; Ginsberg, M. H. The Final Steps of Integrin Activation: The End Game. *Nat. Rev. Mol. Cell Biol.***2010**, *11* (4), 288–300.
(30) Fillingham, I.; Gingras, A. R.; Papagrigoriou, E.; Patel, B.; Emsley, J.; Critchley, D. R.; Roberts, G. C. K.; Barsukov, I. L. A Vinculin Binding Domain from the Talin Rod Unfolds to Form a Complex with the Vinculin Head. *Structure***2005**, *13* (1), 65–74.
(31) Price, G. J.; Jones, P.; Davison, M. D.; Patel, B.; Bendori, R.; Geiger, B.; Critchley, D. R. Primary Sequence and Domain Structure of Chicken Vinculin. *Biochem. J.***1989**, *259* (2), 453–461.
(32) Roberts, G. C. K.; Critchley, D. R. Structural and Biophysical Properties of the Integrin-Associated Cytoskeletal Protein Talin. *Biophys. Rev.***2009**, *1* (2), 61–69.
(33) Song, X.; Yang, J.; Hirbawi, J.; Ye, S.; Perera, H. D.; Goksoy, E.; Dwivedi, P.; Plow, E. F.; Zhang, R.; Qin, J. A Novel Membrane-Dependent on/off Switch Mechanism of Talin FERM Domain at Sites of Cell Adhesion. *Cell Res.***2012**, *22* (11), 1533–1545.
(34) Martel, V.; Racaud-Sultan, C.; Dupe, S.; Marie, C.; Paulhe, F.; Galmiche, A.; Block, M. R.; Albiges-Rizo, C. Conformation, Localization, and Integrin Binding of Talin Depend on Its Interaction with Phosphoinositides. *J. Biol. Chem.***2001**, *276* (24), 21217–21227.
(35) Ling, K.; Doughman, R. L.; Firestone, A. J.; Bunce, M. W.; Anderson, R. A. Type I Gamma Phosphatidylinositol Phosphate Kinase Targets and Regulates Focal Adhesions. *Nature***2002**, *420* (6911), 89–93.
(36) Ling, K.; Doughman, R. L.; Iyer, V. V.; Firestone, A. J.; Bairstow, S. F.; Mosher, D. F.; Schaller, M. D.; Anderson, R. A. Tyrosine Phosphorylation of Type Iγ Phosphatidylinositol Phosphate Kinase by Src Regulates an Integrin-Talin Switch. *J. Cell Biol.***2003**, *163* (6), 1339–1349.
(37) Cance, W. G.; Kurenova, E.; Marlowe, T.; Golubovskaya, V. Disrupting the Scaffold to Improve Focal Adhesion Kinase-Targeted Cancer Therapeutics. *Sci. Signal.***2013**, *6* (268), pe10.
(38) Schwock, J.; Dhani, N. PTK2 (PTK2 Protein Tyrosine Kinase 2). *Atlas Genet Cytogenet Oncol Haematol.*; 2011.
(39) Lietha, D.; Cai, X.; Ceccarelli, D. F. J.; Li, Y.; Schaller, M. D.; Eck, M. J. Structural Basis for the Autoinhibition of Focal Adhesion Kinase. *Cell***2007**, *129* (6), 1177–1187.
(40) Papusheva, E.; Queiroz, F. M. de; Dalous, J.; Han, Y.; Esposito, A.; Jares-Erijmanxa, E. A.; Jovin, T. M.; Bunt, G. Dynamic Conformational Changes in the FERM Domain of FAK Are Involved in Focal-Adhesion Behavior during Cell Spreading and Motility. *J. Cell Sci.***2009**, *122* (5), 656–666.
(41) Chen, T.-H.; Chan, P.-C.; Chen, C.-L.; Chen, H.-C. Phosphorylation of Focal Adhesion Kinase on Tyrosine 194 by Met Leads to Its Activation through Relief of Autoinhibition. *Oncogene***2011**, *30* (2), 153–166.
(42) Kadaré, G.; Toutant, M.; Formstecher, E.; Corvol, J.-C.; Carnaud, M.; Boutterin, M.-C.; Girault, J.-A. PIAS1-Mediated Sumoylation of Focal Adhesion Kinase Activates Its Autophosphorylationn. *J. Biol. Chem.***2003**, *278* (48), 47434–47440.

(43) Bellis, S. L.; Miller, J. T.; Turner, C. E. Characterization of Tyrosine Phosphorylation of Paxillin in Vitro by Focal Adhesion Kinase. *J. Biol. Chem.***1995**, *270* (29), 17437–17441.
(44) Tachibana, K.; Urano, T.; Fujita, H.; Ohashi, Y.; Kamiguchi, K.; Iwata, S.; Hirai, H.; Morimoto, C. Tyrosine Phosphorylation of Crk-Associated Substrates by Focal Adhesion Kinase. A Putative Mechanism for the Integrin-Mediated Tyrosine Phosphorylation of Crk-Associated Substrates. *J. Biol. Chem.***1997**, *272* (46), 29083–29090.
(45) Calalb, M. B.; Polte, T. R.; Hanks, S. K. Tyrosine Phosphorylation of Focal Adhesion Kinase at Sites in the Catalytic Domain Regulates Kinase Activity: A Role for Src Family Kinases. *Mol. Cell. Biol.***1995**, *15* (2), 954–963.
(46) Chen, H.-C.; Appeddu, P. A.; Isoda, H.; Guan, J.-L. Phosphorylation of Tyrosine 397 in Focal Adhesion Kinase Is Required for Binding Phosphatidylinositol 3-Kinase. *J. Biol. Chem.***1996**, *271* (42), 26329–26334.
(47) Calalb, M. B.; Zhang, X.; Polte, T. R.; Hanks, S. K. Focal Adhesion Kinase Tyrosine-861 Is a Major Site of Phosphorylation by Src. *Biochem. Biophys. Res. Commun.***1996**, *228* (3), 662–668.
(48) Schaller, M. D. Cellular Functions of FAK Kinases: Insight into Molecular Mechanisms and Novel Functions. *J. Cell Sci.***2010**, *123* (7), 1007–1013.
(49) Schlaepfer, D. D.; Hanks, S. K.; Hunter, T.; van der Geer, P. Integrin-Mediated Signal Transduction Linked to Ras Pathway by GRB2 Binding to Focal Adhesion Kinase. *Nature***1994**, *372* (6508), 786–791.
(50) Ezratty, E. J.; Partridge, M. A.; Gundersen, G. G. Microtubule-Induced Focal Adhesion Disassembly Is Mediated by Dynamin and Focal Adhesion Kinase. *Nat. Cell Biol.***2005**, *7* (6), 581–590.
(51) Deramaudt, T. B.; Dujardin, D.; Hamadi, A.; Noulet, F.; Kolli, K.; De Mey, J.; Takeda, K.; Ronde, P. FAK Phosphorylation at Tyr-925 Regulates Cross-Talk between Focal Adhesion Turnover and Cell Protrusion. *Mol. Biol. Cell***2011**, *22* (7), 964–975.
(52) Arold, S. T. How Focal Adhesion Kinase Achieves Regulation by Linking Ligand Binding, Localization and Action. *Curr. Opin. Struct. Biol.***2011**, *21* (6), 808–813.
(53) Hayashi, I.; Vuori, K.; Liddington, R. C. The Focal Adhesion Targeting (FAT) Region of Focal Adhesion Kinase Is a Four-Helix Bundle That Binds Paxillin. *Nat. Struct. Biol.***2002**, *9* (2), 101–106.
(54) Lawson, C.; Lim, S.-T.; Uryu, S.; Chen, X. L.; Calderwood, D. A.; Schlaepfer, D. D. FAK Promotes Recruitment of Talin to Nascent Adhesions to Control Cell Motility. *J. Cell Biol.***2012**, *196* (2), 223–232.
(55) Garces, C. A.; Kurenova, E. V.; Golubovskaya, V. M.; Cance, W. G. Vascular Endothelial Growth Factor Receptor-3 and Focal Adhesion Kinase Bind and Suppress Apoptosis in Breast Cancer Cells. *Cancer Res.***2006**, *66* (3), 1446–1454.
(56) Mitra, S. K.; Hanson, D. A.; Schlaepfer, D. D. Focal Adhesion Kinase: In Command and Control of Cell Motility. *Nat. Rev. Mol. Cell Biol.***2005**, *6* (1), 56–68.
(57) Xie, Z.; Sanada, K.; Samuels, B. A.; Shih, H.; Tsai, L. H. Serine 732 Phosphorylation of FAK by Cdk5 Is Important for Microtubule Organization, Nuclear Movement, and Neuronal Migration. *Cell***2003**, *114* (4), 469–482.
(58) Park, A. Y. J.; Shen, T.-L.; Chien, S.; Guan, J.-L. Role of Focal Adhesion Kinase Ser-732 Phosphorylation in Centrosome Function during Mitosis. *J. Biol. Chem.***2009**, *284* (14), 9418–9425.
(59) Golubovskaya, V. M.; Finch, R.; Cance, W. G. Direct Interaction of the N-Terminal Domain of Focal Adhesion Kinase with the N-Terminal Transactivation Domain of P53. *J. Biol. Chem.***2005**, *280* (26), 25008–25021.
(60) Lim, S.-T. S. Nuclear FAK: A New Mode of Gene Regulation from Cellular Adhesions. *Mol. Cells***2013**, *36* (1), 1–6.
(61) Sen, B.; Johnson, F. M. Regulation of Src Family Kinases in Human Cancers. *J. Signal Transduct.***2011**, *2011*.

(62) Guarino, M. Src Signaling in Cancer Invasion. *J. Cell. Physiol.***2010**, *223* (1), 14–26.
(63) Frame, M. C. Src in Cancer: Deregulation and Consequences for Cell Behaviour. *Biochim. Biophys. Acta***2002**, *1602* (2), 114–130.
(64) Kim, L. C.; Song, L.; Haura, E. B. Src Kinases as Therapeutic Targets for Cancer. *Nat. Rev. Clin. Oncol.***2009**, *6* (10), 587–595.
(65) Twamley-Stein, G. M.; Pepperkok, R.; Ansorge, W.; Courtneidge, S. A. The Src Family Tyrosine Kinases Are Required for Platelet-Derived Growth Factor-Mediated Signal Transduction in NIH 3T3 Cells. *Proc. Natl. Acad. Sci.***1993**, *90*, 7696–7700.
(66) Dubrovskyi, O.; Tian, X.; Poroyko, V.; Yakubov, B.; Birukova, A. A.; Birukov, K. G. Identification of Paxilin Domains Interacting with β-Catenin. *FEBS Lett.***2012**, *586* (16), 2294–2299.
(67) Brown, M. C.; Perrotta, J. A.; Turner, C. E. Serine and Threonine Phosphorylation of the Paxillin LIM Domains Regulates Paxillin Focal Adhesion Localization and Cell Adhesion to Fibronectin. *Mol. Biol. Cell***1998**, *9* (7), 1803–1816.
(68) Turner, C. E. Paxillin and Focal Adhesion Signalling. *Nat. Cell Biol.***2000**, *2* (12), E231–E236.
(69) Schaller, M. D. Paxillin: A Focal Adhesion-Associated Adaptor Protein. *Publ. Online 02 Oct. 2001 Doi101038sjonc1204786***2001**, *20* (44).
(70) Herreros, L.; Rodríguez-Fernández, J. L.; Brown, M. C.; Alonso-Lebrero, J. L.; Cabañas, C.; Sánchez-Madrid, F.; Longo, N.; Turner, C. E.; Sánchez-Mateos, P. Paxillin Localizes to the Lymphocyte Microtubule Organizing Center and Associates with the Microtubule Cytoskeleton. *J. Biol. Chem.***2000**, *275* (34), 26436–26440.
(71) Efimov, A.; Schiefermeier, N.; Grigoriev, I.; Brown, M. C.; Turner, C. E.; Small, J. V.; Kaverina, I. Paxillin-Dependent Stimulation of Microtubule Catastrophes at Focal Adhesion Sites. *J. Cell Sci.***2008**, *121* (2), 196–204.
(72) Jamieson, J. S.; Tumbarello, D. A.; Hallé, M.; Brown, M. C.; Tremblay, M. L.; Turner, C. E. Paxillin Is Essential for PTP-PEST-Dependent Regulation of Cell Spreading and Motility: A Role for Paxillin Kinase Linker. *J. Cell Sci.***2005**, *118* (Pt 24), 5835–5847.
(73) Zaidel-Bar, R.; Milo, R.; Kam, Z.; Geiger, B. A Paxillin Tyrosine Phosphorylation Switch Regulates the Assembly and Form of Cell-Matrix Adhesions. *J. Cell Sci.***2007**, *120* (1), 137–148.
(74) Jeon, K. W. *International Review Of Cell and Molecular Biology*; Academic Press, 2011.
(75) Smith, M. A.; Blankman, E.; Deakin, N. O.; Hoffman, L. M.; Jensen, C. C.; Turner, C. E.; Beckerle, M. C. LIM Domains Target Actin Regulators Paxillin and Zyxin to Sites of Stress Fiber Strain. *PLoS ONE***2013**, *8* (8), e69378.
(76) Zhao, Z.; Manser, E.; Loo, T.-H.; Lim, L. Coupling of PAK-Interacting Exchange Factor PIX to GIT1 Promotes Focal Complex Disassembly. *Mol. Cell. Biol.***2000**, *20* (17), 6354–6363.
(77) Turner, C. E.; Miller, J. T. Primary Sequence of Paxillin Contains Putative SH2 and SH3 Domain Binding Motifs and Multiple LIM Domains: Identification of a Vinculin and Pp125Fak-Binding Region. *J. Cell Sci.***1994**, *107* (6), 1583–1591.
(78) Chua, G.-L.; Patra, A. T.; Tan, S.-M.; Bhattacharjya, S. NMR Structure of Integrin A4 Cytosolic Tail and Its Interactions with Paxillin. *PLoS ONE***2013**, *8* (1), e55184.
(79) Liu, S.; Kiosses, W. B.; Rose, D. M.; Slepak, M.; Salgia, R.; Griffin, J. D.; Turner, C. E.; Schwartz, M. A.; Ginsberg, M. H. A Fragment of Paxillin Binds the A4 Integrin Cytoplasmic Domain (Tail) and Selectively Inhibits A4-Mediated Cell Migration. *J. Biol. Chem.***2002**, *277* (23), 20887–20894.
(80) Schaller, M. D.; Parsons, J. T. Pp125FAK-Dependent Tyrosine Phosphorylation of Paxillin Creates a High-Affinity Binding Site for Crk. *Mol. Cell. Biol.***1995**, *15* (5), 2635–2645.
(81) Iwasaki, T.; Nakata, A.; Mukai, M.; Shinkai, K.; Yano, H.; Sabe, H.; Schaefer, E.; Tatsuta, M.; Tsujimura, T.; Terada, N.; et al. Involvement of Phosphorylation of Tyr-31 and Tyr-118 of Paxillin in MM1 Cancer Cell Migration. *Int. J. Cancer J. Int. Cancer***2002**, *97* (3), 330–335.

(82) Petit, V.; Boyer, B.; Lentz, D.; Turner, C. E.; Thiery, J. P.; Vallés, A. M. Phosphorylation of Tyrosine Residues 31 and 118 on Paxillin Regulates Cell Migration through an Association with Crk in Nbt-II Cells. *J. Cell Biol.***2000**, *148* (5), 957–970.
(83) Brugnera, E.; Haney, L.; Grimsley, C.; Lu, M.; Walk, S. F.; Tosello-Trampont, A.-C.; Macara, I. G.; Madhani, H.; Fink, G. R.; Ravichandran, K. S. Unconventional Rac-GEF Activity Is Mediated through the Dock180-ELMO Complex. *Nat. Cell Biol.***2002**, *4* (8), 574–582.
(84) Tsubouchi, A.; Sakakura, J.; Yagi, R.; Mazaki, Y.; Schaefer, E.; Yano, H.; Sabe, H. Localized Suppression of RhoA Activity by Tyr31/118-Phosphorylated Paxillin in Cell Adhesion and Migration. *J. Cell Biol.***2002**, *159* (4), 673–683.
(85) Thomas, J. W.; Cooley, M. A.; Broome, J. M.; Salgia, R.; Griffin, J. D.; Lombardo, C. R.; Schaller, M. D. The Role of Focal Adhesion Kinase Binding in the Regulation of Tyrosine Phosphorylation of Paxillin. *J. Biol. Chem.***1999**, *274* (51), 36684–36692.
(86) Turner, C. E.; Brown, M. C.; Perrotta, J. A.; Riedy, M. C.; Nikolopoulos, S. N.; McDonald, A. R.; Bagrodia, S.; Thomas, S.; Leventhal, P. S. Paxillin LD4 Motif Binds PAK and PIX through a Novel 95-KD Ankyrin Repeat, ARF–GAP Protein: A Role in Cytoskeletal Remodeling. *J. Cell Biol.***1999**, *145* (4), 851–863.
(87) Nayal, A.; Webb, D. J.; Brown, C. M.; Schaefer, E. M.; Vicente-Manzanares, M.; Horwitz, A. R. Paxillin Phosphorylation at Ser273 Localizes a GIT1–PIX–PAK Complex and Regulates Adhesion and Protrusion Dynamics. *J. Cell Biol.***2006**, *173* (4), 587–589.
(88) Scheswohl, D. M.; Harrell, J. R.; Rajfur, Z.; Gao, G.; Campbell, S. L.; Schaller, M. D. Multiple Paxillin Binding Sites Regulate FAK Function. *J. Mol. Signal.***2008**, *3* (1), 1.
(89) Palazzo, A. F.; Eng, C. H.; Schlaepfer, D. D.; Marcantonio, E. E.; Gundersen, G. G. Localized Stabilization of Microtubules by Integrin- and FAK-Facilitated Rho Signaling. *Science***2004**, *303* (5659), 836–839.
(90) What is Vinculin? https://www.mechanobio.info/what-is-mechanosignaling/what-is-the-extracellular-matrix-and-the-basal-lamina/what-are-focal-adhesions/what-is-vinculin/ (accessed Jun 3, 2018).
(91) Goldmann, W. H.; Ingber, D. E. Intact Vinculin Protein Is Required for Control of Cell Shape, Cell Mechanics, and Rac-Dependent Lamellipodia Formation. *Biochem. Biophys. Res. Commun.***2002**, *290* (2), 749–755.
(92) Ziegler, W. H.; Liddington, R. C.; Critchley, D. R. The Structure and Regulation of Vinculin. *Trends Cell Biol.***2006**, *16* (9), 453–460.
(93) Thievessen, I.; Thompson, P. M.; Berlemont, S.; Plevock, K. M.; Plotnikov, S. V.; Zemljic-Harpf, A.; Ross, R. S.; Davidson, M. W.; Danuser, G.; Campbell, S. L.; et al. Vinculin–Actin Interaction Couples Actin Retrograde Flow to Focal Adhesions, but Is Dispensable for Focal Adhesion Growth. *J Cell Biol***2013**, *202* (1), 163–177.
(94) Bear, J. E.; Gertler, F. B. Ena/VASP: Towards Resolving a Pointed Controversy at the Barbed End. *J. Cell Sci.***2009**, *122* (12), 1947–1953.
(95) Beckham, Y.; Vasquez, R. J.; Stricker, J.; Sayegh, K.; Campillo, C.; Gardel, M. L. Arp2/3 Inhibition Induces Amoeboid-like Protrusions in MCF10A Epithelial Cells by Reduced Cytoskeletal-Membrane Coupling and Focal Adhesion Assembly. *PloS One***2014**, *9* (6), e100943.
(96) Ridley, A. J. Rho GTPases and Cell Migration. *J. Cell Sci.***2001**, *114* (15), 2713–2722.
(97) Ridley, A. J. Rho GTPases and Actin Dynamics in Membrane Protrusions and Vesicle Trafficking. *Trends Cell Biol.***2006**, *16* (10), 522–529.
(98) Hall, A. Rho GTPases and the Actin Cytoskeleton. *Science***1998**, *279* (5350), 509–514.
(99) Zhang, B.; Zheng, Y. Regulation of RhoA GTP Hydrolysis by the GTPase-Activating Proteins P190, P50RhoGAP, Bcr, and 3BP-1. *Biochemistry (Mosc.)***1998**, *37* (15), 5249–5257.
(100) Price, L. S.; Leng, J.; Schwartz, M. A.; Bokoch, G. M. Activation of Rac and Cdc42 by Integrins Mediates Cell Spreading. *Mol. Biol. Cell***1998**, *9* (7), 1863–1871.

(101) Sander, E. E.; Klooster, J. P. ten; Delft, S. van; Kammen, R. A. van der; Collard, J. G. Rac Downregulates Rho Activity Reciprocal Balance between Both Gtpases Determines Cellular Morphology and Migratory Behavior. *J. Cell Biol.***1999**, *147* (5), 1009–1022.
(102) Rottner, K.; Hall, A.; Small, J. V. Interplay between Rac and Rho in the Control of Substrate Contact Dynamics. *Curr. Biol. CB***1999**, *9* (12), 640–648.
(103) Worthylake, R. A.; Burridge, K. RhoA and ROCK Promote Migration by Limiting Membrane Protrusions. *J. Biol. Chem.***2003**, *278* (15), 13578–13584.
(104) Alblas, J.; Ulfman, L.; Hordijk, P.; Koenderman, L. Activation of RhoA and ROCK Are Essential for Detachment of Migrating Leukocytes. *Mol. Biol. Cell***2001**, *12* (7), 2137–2145.
(105) Worthylake, R. A.; Lemoine, S.; Watson, J. M.; Burridge, K. RhoA Is Required for Monocyte Tail Retraction during Transendothelial Migration. *J. Cell Biol.***2001**, *154* (1), 147–160.
(106) Schoenwaelder, S. M.; Burridge, K. Bidirectional Signaling between the Cytoskeleton and Integrins. *Curr. Opin. Cell Biol.***1999**, *11* (2), 274–286.
(107) Rodríguez-Fernández, J. L.; Sánchez-Martín, L.; Rey, M.; Vicente-Manzanares, M.; Narumiya, S.; Teixidó, J.; Sánchez-Madrid, F.; Cabañas, C. Rho and Rho-Associated Kinase Modulate the Tyrosine Kinase PYK2 in T-Cells through Regulation of the Activity of the Integrin LFA-1. *J. Biol. Chem.***2001**, *276* (44), 40518–40527.
(108) Williams, L. M.; Ridley, A. J. Lipopolysaccharide Induces Actin Reorganization and Tyrosine Phosphorylation of Pyk2 and Paxillin in Monocytes and Macrophages. *J. Immunol.***2000**, *164* (4), 2028–2036.
(109) Duong, L. T.; Rodan, G. A. PYK2 Is an Adhesion Kinase in Macrophages, Localized in Podosomes and Activated by B2-Integrin Ligation. *Cell Motil. Cytoskeleton***2000**, *47* (3), 174–188.
(110) Maekawa, M.; Ishizaki, T.; Boku, S.; Watanabe, N.; Fujita, A.; Iwamatsu, A.; Obinata, T.; Ohashi, K.; Mizuno, K.; Narumiya, S. Signaling from Rho to the Actin Cytoskeleton through Protein Kinases ROCK and LIM-Kinase. *Science***1999**, *285* (5429), 895–898.
(111) Watanabe, N.; Kato, T.; Fujita, A.; Ishizaki, T.; Narumiya, S. Cooperation between MDia1 and ROCK in Rho-Induced Actin Reorganization. *Nat. Cell Biol.***1999**, *1* (3), 136–143.
(112) Burridge, K.; Wennerberg, K. Rho and Rac Take Center Stage. *Cell***2004**, *116* (2), 167–179.
(113) D'Souza-Schorey, C.; Boshans, R. L.; McDonough, M.; Stahl, P. D.; Van Aelst, L. A Role for POR1, a Rac1-Interacting Protein, in ARF6-Mediated Cytoskeletal Rearrangements. *EMBO J.***1997**, *16* (17), 5445–5454.
(114) Van Aelst, L.; Joneson, T.; Bar-Sagi, D. Identification of a Novel Rac1-Interacting Protein Involved in Membrane Ruffling. *EMBO J.***1996**, *15* (15), 3778–3786.
(115) Watanabe, T.; Wang, S.; Noritake, J.; Sato, K.; Fukata, M.; Takefuji, M.; Nakagawa, M.; Izumi, N.; Akiyama, T.; Kaibuchi, K. Interaction with IQGAP1 Links APC to Rac1, Cdc42, and Actin Filaments during Cell Polarization and Migration. *Dev. Cell***2004**, *7* (6), 871–883.
(116) Westendorf, J. J. The Formin/Diaphanous-Related Protein, FHOS, Interacts with Rac1 and Activates Transcription from the Serum Response Element. *J. Biol. Chem.***2001**, *276* (49), 46453–46459.
(117) Aspenström, P.; Richnau, N.; Johansson, A.-S. The Diaphanous-Related Formin DAAM1 Collaborates with the Rho GTPases RhoA and Cdc42, CIP4 and Src in Regulating Cell Morphogenesis and Actin Dynamics. *Exp. Cell Res.***2006**, *312* (12), 2180–2194.
(118) Szczepanowska, J. Involvement of Rac/Cdc42/PAK Pathway in Cytoskeletal Rearrangements. *Acta Biochim. Pol.***2009**, *56* (2), 225–234.
(119) Vidal, C.; Geny, B.; Melle, J.; Jandrot-Perrus, M.; Fontenay-Roupie, M. Cdc42/Rac1-Dependent Activation of the P21-Activated Kinase (PAK) Regulates Human Platelet Lamellipodia Spreading: Implication of the Cortical-Actin Binding Protein Cortactin. *Blood***2002**, *100* (13), 4462–4469.
(120) Weston, L.; Coutts, A. S.; Thangue, N. B. L. Actin Nucleators in the Nucleus: An Emerging Theme. *J. Cell Sci.***2012**, *125* (15), 3519–3527.

(121) Smith, M. A.; Blankman, E.; Gardel, M. L.; Luettjohann, L.; Waterman, C. M.; Beckerle, M. C. A Zyxin-Mediated Mechanism for Actin Stress Fiber Maintenance and Repair. *Dev. Cell***2010**, *19* (3), 365–376.
(122) Harbeck, B.; Hüttelmaier, S.; Schlüter, K.; Jockusch, B. M.; Illenberger, S. Phosphorylation of the Vasodilator-Stimulated Phosphoprotein Regulates Its Interaction with Actin. *J. Biol. Chem.***2000**, *275* (40), 30817–30825.
(123) Reinhard, M.; Halbrugge, M.; Scheer, U.; Wiegand, C.; Jockusch, B. M.; Walter, U. The 46/50 KDa Phosphoprotein VASP Purified from Human Platelets Is a Novel Protein Associated with Actin Filaments and Focal Contacts. *EMBO J.***1992**, *11* (6), 2063–2070.
(124) Sjöblom, B.; Salmazo, A.; Djinović-Carugo, K. Alpha-Actinin Structure and Regulation. *Cell. Mol. Life Sci. CMLS***2008**, *65* (17), 2688–2701.
(125) Rangamani, P.; Fardin, M.-A.; Xiong, Y.; Lipshtat, A.; Rossier, O.; Sheetz, M. P.; Iyengar, R. Signaling Network Triggers and Membrane Physical Properties Control the Actin Cytoskeleton-Driven Isotropic Phase of Cell Spreading. *Biophys. J.***2011**, *100* (4), 845–857.
(126) Gilmore, A. P.; Burridge, K. Regulation of Vinculin Binding to Talin and Actin by Phosphatidyl-Inositol-4-5-Bisphosphate. *Nature***1996**, *381* (6582), 531–535.
(127) Eykhoff, P. *System Identification: Parameter and State Estimation*; Wiley-Interscience, 1974.
(128) Friedman, N. Inferring Cellular Networks Using Probabilistic Graphical Models. *Science***2004**, *303* (5659), 799–805.
(129) Voit, E. O. *A First Course in Systems Biology*; Garland Science: New York, 2013.
(130) Ko, C.-L.; Voit, E. O.; Wang, F.-S. Estimating Parameters for Generalized Mass Action Models with Connectivity Information. *BMC Bioinformatics***2009**, *10* (1), 140.
(131) Weisstein, E. W. SIR Model -- from Wolfram MathWorld http://mathworld.wolfram.com/SIRModel.html (accessed Oct 9, 2013).
(132) Weisstein, E. W. Lotka-Volterra Equations -- from Wolfram MathWorld http://mathworld.wolfram.com/Lotka-VolterraEquations.html (accessed Oct 8, 2013).
(133) Sobie, E. A. An Introduction to MATLAB. *Sci. Signal.***2011**, *4* (191), tr7.
(134) Wolfram Mathematica: Modern Technical Computing https://www.wolfram.com/mathematica/ (accessed Jun 2, 2018).
(135) Hoops, S.; Sahle, S.; Gauges, R.; Lee, C.; Pahle, J.; Simus, N.; Singhal, M.; Xu, L.; Mendes, P.; Kummer, U. COPASI—a COmplex PAthway SImulator. *Bioinformatics***2006**, *22* (24), 3067–3074.
(136) Moraru, I. I.; Schaff, J. C.; Slepchenko, B. M.; Blinov, M.; Morgan, F.; Lakshminarayana, A.; Gao, F.; Li, Y.; Loew, L. M. The Virtual Cell Modeling and Simulation Software Environment. *IET Syst. Biol.***2008**, *2* (5), 352–362.
(137) Allen, W. J.; Balius, T. E.; Mukherjee, S.; Brozell, S. R.; Moustakas, D. T.; Lang, P. T.; Case, D. A.; Kuntz, I. D.; Rizzo, R. C. DOCK 6: Impact of New Features and Current Docking Performance. *J. Comput. Chem.***2015**, *36* (15), 1132–1156.
(138) Brooks, B. R.; Brooks, C. L.; MacKerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.***2009**, *30* (10), 1545–1614.
(139) TMHMM Server, v. 2.0 http://www.cbs.dtu.dk/services/TMHMM/ (accessed Oct 9, 2013).
(140) Sachs, K.; Perez, O.; Pe'er, D.; Lauffenburger, D. A.; Nolan, G. P. Causal Protein-Signaling Networks Derived from Multiparameter Single-Cell Data. *Science***2005**, *308* (5721), 523–529.
(141) Müssel, C.; Hopfensitz, M.; Kestler, H. A. BoolNet—an R Package for Generation, Reconstruction and Analysis of Boolean Networks. *Bioinformatics***2010**, *26* (10), 1378–1380.
(142) Singh, A.; Nascimento, J. M.; Kowar, S.; Busch, H.; Boerries, M. Boolean Approach to Signalling Pathway Modelling in HGF-Induced Keratinocyte Migration. *Bioinformatics***2012**, *28* (18), i495–i501.

(143) Karlebach, G.; Shamir, R. Modelling and Analysis of Gene Regulatory Networks. *Nat. Rev. Mol. Cell Biol.***2008**, *9* (10), 770–780.
(144) Fauré, A.; Naldi, A.; Chaouiya, C.; Thieffry, D. Dynamical Analysis of a Generic Boolean Model for the Control of the Mammalian Cell Cycle. *Bioinformatics***2006**, *22* (14), e124–e131.
(145) The Process Diagram http://www.celldesigner.org/documents/ProcessDiagram.html (accessed Oct 10, 2013).
(146) Castiglione, F. Agent Based Modeling. *Scholarpedia***2006**, *1* (10), 1562.
(147) Brown, B. N.; Price, I. M.; Toapanta, F. R.; DeAlmeida, D. R.; Wiley, C. A.; Ross, T. M.; Oury, T. D.; Vodovotz, Y. An Agent-Based Model of Inflammation and Fibrosis Following Particulate Exposure in the Lung. *Math. Biosci.***2011**, *231* (2), 186–196.
(148) Mi, Q.; Rivière, B.; Clermont, G.; Steed, D. L.; Vodovotz, Y. Agent-Based Model of Inflammation and Wound Healing: Insights into Diabetic Foot Ulcer Pathology and the Role of Transforming Growth Factor-Beta1. *Wound Repair Regen. Off. Publ. Wound Heal. Soc. Eur. Tissue Repair Soc.***2007**, *15* (5), 671–682.
(149) Dong, X.; Foteinou, P. T.; Calvano, S. E.; Lowry, S. F.; Androulakis, I. P. Agent-Based Modeling of Endotoxin-Induced Acute Inflammatory Response in Human Blood Leukocytes. *PLoS ONE***2010**, *5* (2), e9249.
(150) Mansury, Y.; Kimura, M.; Lobo, J.; Deisboeck, T. S. Emerging Patterns in Tumor Systems: Simulating the Dynamics of Multicellular Clusters with an Agent-Based Spatial Agglomeration Model. *J. Theor. Biol.***2002**, *219* (3), 343–370.
(151) Zhang, L.; Wang, Z.; Sagotsky, J. A.; Deisboeck, T. S. Multiscale Agent-Based Cancer Modeling. *J. Math. Biol.***2009**, *58* (4–5), 545–559.
(152) Mansury, Y.; Diggory, M.; Deisboeck, T. S. Evolutionary Game Theory in an Agent-Based Brain Tumor Model: Exploring the 'Genotype–Phenotype' Link. *J. Theor. Biol.***2006**, *238* (1), 146–156.
(153) Mogilner, A.; Edelstein-Keshet, L. Regulation of Actin Dynamics in Rapidly Moving Cells: A Quantitative Analysis. *Biophys. J.***2002**, *83* (3), 1237–1258.
(154) Ditlev, J. A.; Vacanti, N. M.; Novak, I. L.; Loew, L. M. An Open Model of Actin Dendritic Nucleation. *Biophys. J.***2009**, *96* (9), 3529–3542.
(155) Kapustina, M.; Vitriol, E.; Elston, T. C.; Loew, L. M.; Jacobson, K. Modeling Capping Protein FRAP and CALI Experiments Reveals in Vivo Regulation of Actin Dynamics. *Cytoskelet. Hoboken NJ***2010**, *67* (8), 519–534.
(156) Kraikivski, P.; Slepchenko, B. M. Quantifying a Pathway: Kinetic Analysis of Actin Dendritic Nucleation. *Biophys. J.***2010**, *99* (3), 708–715.
(157) Jamali, Y.; Jamali, T.; Mofrad, M. R. K. An Agent Based Model of Integrin Clustering: Exploring the Role of Ligand Clustering, Integrin Homo-Oligomerization, Integrin–Ligand Affinity, Membrane Crowdedness and Ligand Mobility. *J. Comput. Phys.***2013**, *244*, 264–278.
(158) Helikar, T.; Kochi, N.; Kowal, B.; Dimri, M.; Naramura, M.; Raja, S. M.; Band, V.; Band, H.; Rogers, J. A. A Comprehensive, Multi-Scale Dynamical Model of ErbB Receptor Signal Transduction in Human Mammary Epithelial Cells. *PLoS ONE***2013**, *8* (4), e61757.
(159) Mischnik, M.; Boyanova, D.; Hubertus, K.; Geiger, J.; Philippi, N.; Dittrich, M.; Wangorsch, G.; Timmer, J.; Dandekar, T. A Boolean View Separates Platelet Activatory and Inhibitory Signalling as Verified by Phosphorylation Monitoring Including Threshold Behaviour and Integrin Modulation. *Mol. Biosyst.***2013**, *9* (6), 1326–1339.
(160) Geier, F.; Fengos, G.; Iber, D. A Computational Analysis of the Dynamic Roles of Talin, Dok1, and PIPKI for Integrin Activation. *PLoS ONE***2011**, *6* (11), e24808.
(161) Cirit, M.; Krajcovic, M.; Choi, C. K.; Welf, E. S.; Horwitz, A. F.; Haugh, J. M. Stochastic Model of Integrin-Mediated Signaling and Adhesion Dynamics at the Leading Edges of Migrating Cells. *PLoS Comput Biol***2010**, *6* (2), e1000688.

(162) Falkenberg, C. V.; Loew, L. M. Computational Analysis of Rho GTPase Cycling. *PLoS Comput. Biol.***2013**, *9* (1).
(163) Paul, R.; Heil, P.; Spatz, J. P.; Schwarz, U. S. Propagation of Mechanical Stress through the Actin Cytoskeleton toward Focal Adhesions: Model and Experiment. *Biophys. J.***2008**, *94* (4), 1470–1482.
(164) Kim, M.-C.; Neal, D. M.; Kamm, R. D.; Asada, H. H. Dynamic Modeling of Cell Migration and Spreading Behaviors on Fibronectin Coated Planar Substrates and Micropatterned Geometries. *PLoS Comput. Biol.***2013**, *9* (2).
(165) Vila-Coro, A. J.; Rodríguez-Frade, J. M.; Martín De Ana, A.; Moreno-Ortíz, M. C.; Martínez-A, C.; Mellado, M. The Chemokine SDF-1alpha Triggers CXCR4 Receptor Dimerization and Activates the JAK/STAT Pathway. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.***1999**, *13* (13), 1699–1710.
(166) Tuteja, N. Signaling through G Protein Coupled Receptors. *Plant Signal. Behav.***2009**, *4* (10), 942–947.
(167) Rosenbaum, D. M.; Rasmussen, S. G. F.; Kobilka, B. K. The Structure and Function of G-Protein-Coupled Receptors. *Nature***2009**, *459* (7245), 356–363.
(168) Calebiro, D.; Nikolaev, V. O.; Gagliani, M. C.; de Filippis, T.; Dees, C.; Tacchetti, C.; Persani, L.; Lohse, M. J. Persistent CAMP-Signals Triggered by Internalized G-Protein–Coupled Receptors. *PLoS Biol.***2009**, *7* (8).
(169) New, D. C.; Wong, Y. H. Molecular Mechanisms Mediating the G Protein-Coupled Receptor Regulation of Cell Cycle Progression. *J. Mol. Signal.***2007**, *2*, 2.
(170) Ngai, J.; Inngjerdingen, M.; Berge, T.; Taskén, K. Interplay between the Heterotrimeric G-Protein Subunits Galphaq and Galphai2 Sets the Threshold for Chemotaxis and TCR Activation. *BMC Immunol.***2009**, *10*, 27.
(171) Schwindinger, W. F.; Robishaw, J. D. Heterotrimeric G-Protein Betagamma-Dimers in Growth and Differentiation. *Oncogene***2001**, *20* (13), 1653–1660.
(172) Cao, S.; Cao, R.; Liu, X.; Luo, X.; Zhong, W. Design, Synthesis and Biological Evaluation of Novel Benzothiazole Derivatives as Selective PI3Kβ Inhibitors. *Molecules***2016**, *21* (7), 876.
(173) Dbouk, H. A.; Vadas, O.; Shymanets, A.; Burke, J. E.; Salamon, R. S.; Khalil, B. D.; Barrett, M. O.; Waldo, G. L.; Surve, C.; Hsueh, C.; et al. G Protein–Coupled Receptor–Mediated Activation of P110β by Gβγ Is Required for Cellular Transformation and Invasiveness. *Sci. Signal.***2012**, *5* (253), ra89–ra89.
(174) Barber, M. A.; Welch, H. C. E. PI3K and RAC Signalling in Leukocyte and Cancer Cell Migration. *Bull. Cancer (Paris)***2006**, *93* (5), E44-52.
(175) Giannone, G.; Dubin-Thaler, B. J.; Döbereiner, H.-G.; Kieffer, N.; Bresnick, A. R.; Sheetz, M. P. Periodic Lamellipodial Contractions Correlate with Rearward Actin Waves. *Cell***2004**, *116* (3), 431–443.
(176) Berrier, A. L.; Mastrangelo, A. M.; Downward, J.; Ginsberg, M.; LaFlamme, S. E. Activated R-Ras, Rac1, Pi 3-Kinase and Pkcε Can Each Restore Cell Spreading Inhibited by Isolated Integrin B1 Cytoplasmic Domains. *J. Cell Biol.***2000**, *151* (7), 1549–1560.
(177) Chen, B.; Chou, H.-T.; Brautigam, C. A.; Xing, W.; Yang, S.; Henry, L.; Doolittle, L. K.; Walz, T.; Rosen, M. K. Rac1 GTPase Activates the WAVE Regulatory Complex through Two Distinct Binding Sites. *eLife6*.
(178) Sit, S.-T.; Manser, E. Rho GTPases and Their Role in Organizing the Actin Cytoskeleton. *J Cell Sci***2011**, *124* (5), 679–683.
(179) Li, Z.; Dong, X.; Dong, X.; Wang, Z.; Liu, W.; Deng, N.; Ding, Y.; Tang, L.; Hla, T.; Zeng, R.; et al. Regulation of PTEN by Rho Small GTPases. *Nat. Cell Biol.***2005**, *7* (4), 399–404.
(180) Kinzer-Ursem, T. L.; Linderman, J. J. Both Ligand- and Cell-Specific Parameters Control Ligand Agonism in a Kinetic Model of G Protein–Coupled Receptor Signaling. *PLOS Comput. Biol.***2007**, *3* (1), e6.

(181) Purvis, J. E.; Chatterjee, M. S.; Brass, L. F.; Diamond, S. L. A Molecular Signaling Model of Platelet Phosphoinositide and Calcium Regulation during Homeostasis and P2Y1 Activation. *Blood***2008**, *112* (10), 4069–4079.
(182) Nagaraj, N.; Wisniewski, J. R.; Geiger, T.; Cox, J.; Kircher, M.; Kelso, J.; Pääbo, S.; Mann, M. Deep Proteome and Transcriptome Mapping of a Human Cancer Cell Line. *Mol. Syst. Biol.***2011**, *7*, 548.
(183) Aiuti, A.; Webb, I. J.; Bleul, C.; Springer, T.; Gutierrez-Ramos, J. C. The Chemokine SDF-1 Is a Chemoattractant for Human CD34+ Hematopoietic Progenitor Cells and Provides a New Mechanism to Explain the Mobilization of CD34+ Progenitors to Peripheral Blood. *J. Exp. Med.***1997**, *185* (1), 111–120.
(184) Roskams, J.; Rodgers, L. *Lab Ref: A Handbook of Recipes, Reagents, and Other Reference Tools for Use at the Bench*; Cold Spring Harbor Laboratory Press, 2002.
(185) Maul, G. G.; Deaven, L. Quantitative Determination of Nuclear Pore Complexes in Cycling Cells with Differing DNA Content. *J. Cell Biol.***1977**, *73* (3), 748–760.
(186) Lewin, B.; Cassimeris, L.; Plopper, G. *Cells*; Jones & Bartlett Learning, 2007.
(187) » How big is the endoplasmic reticulum of cells? http://book.bionumbers.org/how-big-is-the-endoplasmic-reticulum-of-cells/ (accessed Jul 16, 2018).
(188) Milo, R.; Phillips, R. Spread surface area of HeLa cell - Human Homo sapiens - BNID 103718 http://bionumbers.hms.harvard.edu//bionumber.aspx?id=103718&ver=1 (accessed Jul 16, 2018).
(189) Milo, R.; Phillips, R. *Cell Biology by the Numbers*; Garland Science, 2015.
(190) Brock, C.; Schaefer, M.; Reusch, H. P.; Czupalla, C.; Michalke, M.; Spicher, K.; Schultz, G.; Nürnberg, B. Roles of Gβγ in Membrane Recruitment and Activation of P110γ/P101 Phosphoinositide 3-Kinase γ. *J. Cell Biol.***2003**, *160* (1), 89–99.
(191) Vögler, O.; Barceló, J. M.; Ribas, C.; Escribá, P. V. Membrane Interactions of G Proteins and Other Related Proteins. *Biochim. Biophys. Acta BBA - Biomembr.***2008**, *1778* (7), 1640–1652.
(192) Glenn, H. L.; Messner, J.; Meldrum, D. R. A Simple Non-Perturbing Cell Migration Assay Insensitive to Proliferation Effects. *Sci. Rep.***2016**, *6*, 31694.
(193) Debreceni, B.; Gao, Y.; Guo, F.; Zhu, K.; Jia, B.; Zheng, Y. Mechanisms of Guanine Nucleotide Exchange and Rac-Mediated Signaling Revealed by a Dominant Negative Trio Mutant. *J. Biol. Chem.***2004**, *279* (5), 3777–3786.
(194) Marei, H.; Malliri, A. GEFs: Dual Regulation of Rac1 Signaling. *Small GTPases***2016**, *8* (2), 90–99.
(195) Bos, J. L.; Rehmann, H.; Wittinghofer, A. GEFs and GAPs: Critical Elements in the Control of Small G Proteins. *Cell***2007**, *129* (5), 865–877.
(196) Xu, C.; Watras, J.; Loew, L. M. Kinetic Analysis of Receptor-Activated Phosphoinositide Turnover. *J. Cell Biol.***2003**, *161* (4), 779–791.
(197) Coleman, D. E.; Berghuis, A. M.; Lee, E.; Linder, M. E.; Gilman, A. G.; others. Structures of Active Conformations of Gi Alpha 1 and the Mechanism of GTP Hydrolysis. *Science***1994**, *265* (5177), 1405–1412.
(198) Albe, K. R.; Butler, M. H.; Wright, B. E. Cellular Concentrations of Enzymes and Their Substrates. *J. Theor. Biol.***1990**, *143* (2), 163–195.
(199) Janes, K. A.; Lauffenburger, D. A. Models of Signalling Networks – What Cell Biologists Can Gain from Them and Give to Them. *J. Cell Sci.***2013**, *126* (9), 1913–1921.
(200) Bentele, M.; Lavrik, I.; Ulrich, M.; Stößer, S.; Heermann, D. W.; Kalthoff, H.; Krammer, P. H.; Eils, R. Mathematical Modeling Reveals Threshold Mechanism in CD95-Induced Apoptosis. *J. Cell Biol.***2004**, *166* (6), 839–851.
(201) Chen, W. W.; Schoeberl, B.; Jasper, P. J.; Niepel, M.; Nielsen, U. B.; Lauffenburger, D. A.; Sorger, P. K. Input-Output Behavior of ErbB Signaling Pathways as Revealed by a Mass Action Model Trained against Dynamic Data. *Mol. Syst. Biol.***2009**, *5*, 239.

(202) Nakakuki, T.; Birtwistle, M. R.; Saeki, Y.; Yumoto, N.; Ide, K.; Nagashima, T.; Brusch, L.; Ogunnaike, B. A.; Okada-Hatakeyama, M.; Kholodenko, B. N. Ligand-Specific c-Fos Expression Emerges from the Spatiotemporal Control of ErbB Network Dynamics. *Cell***2010**, *141* (5), 884–896.
(203) Zaidel-Bar, R. Early Molecular Events in the Assembly of Matrix Adhesions at the Leading Edge of Migrating Cells. *J. Cell Sci.***2003**, *116* (22), 4605–4613.
(204) Burnette, D. T.; Manley, S.; Sengupta, P.; Sougrat, R.; Davidson, M. W.; Kachar, B.; Lippincott-Schwartz, J. A Role for Actin Arcs in the Leading-Edge Advance of Migrating Cells. *Nat. Cell Biol.***2011**, *13* (4), 371–382.
(205) Zimmermann, J.; Falcke, M. Formation of Transient Lamellipodia. *PLoS ONE***2014**, *9* (2).
(206) Albert, P. Dissertation submitted to the Combined Faculties for the Natural Sciences and for Mathematics of the Ruperto-Carola University of Heidelberg, Germany for the degree of Doctor of Natural Sciences http://archiv.ub.uni-heidelberg.de/volltextserver/19570/1/main_print.pdf (accessed Jul 17, 2018).
(207) Tan, L.; Meyer, T.; Pfau, B.; Hofmann, T.; Tan, T. W.; Jones, D. Rapid Vinculin Exchange Dynamics at Focal Adhesions in Primary Osteoblasts Following Shear Flow Stimulation. *J. Musculoskelet. Neuronal Interact.***2010**, *10* (1), 92–99.
(208) Hu, Y.-L.; Lu, S.; Szeto, K. W.; Sun, J.; Wang, Y.; Lasheras, J. C.; Chien, S. FAK and Paxillin Dynamics at Focal Adhesions in the Protrusions of Migrating Cells. *Sci. Rep.***2014**, *4*.
(209) Hadzic, E.; Catillon, M.; Halavatyi, A.; Medves, S.; Van Troys, M.; Moes, M.; Baird, M. A.; Davidson, M. W.; Schaffner-Reckinger, E.; Ampe, C.; et al. Delineating the Tes Interaction Site in Zyxin and Studying Cellular Effects of Its Disruption. *PLoS ONE***2015**, *10* (10).
(210) Guanine nucleotide exchange factor VAV3 - Homo sapiens (Human) http://www.uniprot.org/uniprot/Q9UKW4 (accessed Nov 24, 2013).
(211) Cantley, L. C. The Phosphoinositide 3-Kinase Pathway. *Science***2002**, *296* (5573), 1655–1657.
(212) Yin, H. L.; Janmey, P. A. Phosphoinositide Regulation of the Actin Cytoskeleton. *Annu. Rev. Physiol.***2003**, *65* (1), 761–789.
(213) Yin, H. L. Gelsolin: Calcium- and Polyphosphoinositide-Regulated Actin- Modulating Protein. *BioEssays***1987**, *7* (4), 176–179.
(214) Rameh, L. E.; Mackey, A. M. IQGAP1 Makes PI(3)K Signalling as Easy as PIP, PIP2, PIP3. *Nat. Cell Biol.***2016**, *18* (12), 1263–1265.
(215) Das, M.; Subbayya Ithychanda, S.; Qin, J.; Plow, E. F. Mechanisms of Talin-Dependent Integrin Signaling and Crosstalk. *Biochim. Biophys. Acta BBA - Biomembr.***2014**, *1838* (2), 579–588.
(216) Law, D. A.; Nannizzi-Alaimo, L.; Phillips, D. R. Outside-in Integrin Signal Transduction. Alpha IIb Beta 3-(GP IIb IIIa) Tyrosine Phosphorylation Induced by Platelet Aggregation. *J. Biol. Chem.***1996**, *271* (18), 10811–10815.
(217) Calderwood, D. A. Talin Controls Integrin Activation. *Biochem. Soc. Trans.***2004**, *32* (Pt3), 434–437.
(218) Wegener, K. L.; Partridge, A. W.; Han, J.; Pickford, A. R.; Liddington, R. C.; Ginsberg, M. H.; Campbell, I. D. Structural Basis of Integrin Activation by Talin. *Cell***2007**, *128* (1), 171–182.
(219) Golji, J.; Mofrad, M. R. K. The Talin Dimer Structure Orientation Is Mechanically Regulated. *Biophys. J.***2014**, *107* (8), 1802–1809.
(220) Durrant, T. N.; van den Bosch, M. T.; Hers, I. Integrin $\alpha_{IIb}\beta_3$ Outside-in Signaling. *Blood***2017**, blood-2017-03-773614.
(221) Saltel, F.; Mortier, E.; Hytonen, V. P.; Jacquier, M.-C.; Zimmermann, P.; Vogel, V.; Liu, W.; Wehrle-Haller, B. New PI(4,5)P2- and Membrane Proximal Integrin-Binding Motifs in the Talin Head Control ?3-Integrin Clustering. *J. Cell Biol.***2009**, *187* (5), 715–731.
(222) Kalli, A. C.; Campbell, I. D.; Sansom, M. S. P. Conformational Changes in Talin on Binding to Anionic Phospholipid Membranes Facilitate Signaling by Integrin Transmembrane Helices. *PLoS Comput Biol***2013**, *9* (10), e1003316.

(223) Bout, I. van den; Divecha, N. PIP5K-Driven PtdIns(4,5)P2 Synthesis: Regulation and Cellular Functions. *J. Cell Sci.***2009**, *122* (21), 3837–3850.
(224) Klinghoffer, R. A.; Sachsenmaier, C.; Cooper, J. A.; Soriano, P. Src Family Kinases Are Required for Integrin but Not PDGFR Signal Transduction. *EMBO J.***1999**, *18* (9), 2459–2471.
(225) Moore, D. T.; Nygren, P.; Jo, H.; Boesze-Battaglia, K.; Bennett, J. S.; DeGrado, W. F. Affinity of Talin-1 for the B3-Integrin Cytosolic Domain Is Modulated by Its Phospholipid Bilayer Environment. *Proc. Natl. Acad. Sci.***2012**, *109* (3), 793–798.
(226) Elosegui-Artola, A.; Bazellières, E.; Allen, M. D.; Andreu, I.; Oria, R.; Sunyer, R.; Gomm, J. J.; Marshall, J. F.; Jones, J. L.; Trepat, X.; et al. Rigidity Sensing and Adaptation through Regulation of Integrin Types. *Nat. Mater.***2014**, *13* (6), 631–637.
(227) Calderwood, D. A.; Yan, B.; Pereda, J. M. de; Alvarez, B. G.; Fujioka, Y.; Liddington, R. C.; Ginsberg, M. H. The Phosphotyrosine Binding-like Domain of Talin Activates Integrins. *J. Biol. Chem.***2002**, *277* (24), 21749–21758.
(228) Lele, T. P.; Thodeti, C. K.; Pendse, J.; Ingber, D. E. Investigating Complexity of Protein-Protein Interactions in Focal Adhesions. *Biochem. Biophys. Res. Commun.***2008**, *369* (3), 929–934.
(229) Songyang, Z.; Carraway, K. L.; Eck, M. J.; Harrison, S. C.; Feldman, R. A.; Mohammadi, M.; Schlessinger, J.; Hubbard, S. R.; Smith, D. P.; Eng, C. Catalytic Specificity of Protein-Tyrosine Kinases Is Critical for Selective Signalling. *Nature***1995**, *373* (6514), 536–539.
(230) Yan, B.; Calderwood, D. A.; Yaspan, B.; Ginsberg, M. H. Calpain Cleavage Promotes Talin Binding to the B3Integrin Cytoplasmic Domain. *J. Biol. Chem.***2001**, *276* (30), 28164–28170.
(231) Ye, X.; McLean, M. A.; Sligar, S. G. Phosphatidylinositol 4,5-Bisphosphate Modulates the Affinity of Talin-1 for Phospholipid Bilayers and Activates Its Autoinhibited Form. *Biochemistry (Mosc.)***2016**, *55* (36), 5038–5048.
(232) Cuerrier, C. M.; Pelling, A. E. *Cells, Forces, and the Microenvironment*; CRC Press, 2015.
(233) Harms, B. D.; Bassi, G. M.; Horwitz, A. R.; Lauffenburger, D. A. Directional Persistence of EGF-Induced Cell Migration Is Associated with Stabilization of Lamellipodial Protrusions. *Biophys. J.***2005**, *88* (2), 1479–1488.
(234) Bear, J. E.; Svitkina, T. M.; Krause, M.; Schafer, D. A.; Loureiro, J. J.; Strasser, G. A.; Maly, I. V.; Chaga, O. Y.; Cooper, J. A.; Borisy, G. G.; et al. Antagonism between Ena/VASP Proteins and Actin Filament Capping Regulates Fibroblast Motility. *Cell***2002**, *109* (4), 509–521.
(235) Choi, S.; Thapa, N.; Tan, X.; Hedman, A. C.; Anderson, R. A. PIP Kinases Define PI4,5P2 Signaling Specificity by Association with Effectors. *Biochim. Biophys. Acta***2015**, *1851* (6), 711–723.
(236) Pereda, J. M. de; Wegener, K. L.; Santelli, E.; Bate, N.; Ginsberg, M. H.; Critchley, D. R.; Campbell, I. D.; Liddington, R. C. Structural Basis for Phosphatidylinositol Phosphate Kinase Type Iγ Binding to Talin at Focal Adhesions. *J. Biol. Chem.***2005**, *280* (9), 8381–8386.
(237) Di Paolo, G.; Pellegrini, L.; Letinic, K.; Cestra, G.; Zoncu, R.; Voronov, S.; Chang, S.; Guo, J.; Wenk, M. R.; De Camilli, P. Recruitment and Regulation of Phosphatidylinositol Phosphate Kinase Type 1 Gamma by the FERM Domain of Talin. *Nature***2002**, *420* (6911), 85–89.
(238) Ling, K.; Schill, N. J.; Wagoner, M. P.; Sun, Y.; Anderson, R. A. Movin' on up: The Role of PtdIns(4,5)P(2) in Cell Migration. *Trends Cell Biol.***2006**, *16* (6), 276–284.
(239) Barsukov, I. L.; Prescot, A.; Bate, N.; Patel, B.; Floyd, D. N.; Bhanji, N.; Bagshaw, C. R.; Letinic, K.; Di Paolo, G.; De Camilli, P.; et al. Phosphatidylinositol Phosphate Kinase Type 1 and 1-Integrin Cytoplasmic Domain Bind to the Same Region in the Talin FERM Domain. *J. Biol. Chem.***2003**, *278* (33), 31202–31209.
(240) Wolfenson, H.; Iskratsch, T.; Sheetz, M. P. Early Events in Cell Spreading as a Model for Quantitative Analysis of Biomechanical Events. *Biophys. J.***2014**, *107* (11), 2508–2514.

(241) Baldassarre, M.; Razinia, Z.; Burande, C. F.; Lamsoul, I.; Lutz, P. G.; Calderwood, D. A. Filamins Regulate Cell Spreading and Initiation of Cell Migration. *PLOS ONE***2009**, *4* (11), e7830.
(242) Cuvelier, D.; Théry, M.; Chu, Y.-S.; Dufour, S.; Thiéry, J.-P.; Bornens, M.; Nassoy, P.; Mahadevan, L. The Universal Dynamics of Cell Spreading. *Curr. Biol.***2007**, *17* (8), 694–699.
(243) Takagi, J.; Strokovich, K.; Springer, T. A.; Walz, T. Structure of Integrin A5β1 in Complex with Fibronectin. *EMBO J.***2003**, *22* (18), 4607–4615.
(244) Chen, L. L.; Whitty, A.; Lobb, R. R.; Adams, S. P.; Pepinsky, R. B. Multiple Activation States of Integrin ⌴4ᴺ⌞1 Detected through Their Different Affinities for a Small Molecule Ligand. 10.
(245) Philips, R. M. & R. » How Big Is a Human Cell?
(246) Nilsson, A.; Nielsen, J.; Palsson, B. O. Metabolic Models of Protein Allocation Call for the Kinetome. *Cell Syst.***2017**, *5* (6), 538–541.
(247) Wasinger, V. C.; Zeng, M.; Yau, Y. Current Status and Advances in Quantitative Proteomic Mass Spectrometry. *Int. J. Proteomics***2013**, *2013*, e180605.
(248) COPASI: Support/User Manual/Importing and Exporting/Importing and Exporting SBML Files http://copasi.org/Support/User_Manual/Importing_and_Exporting/Importing_and_Exporting_SBML_Files/ (accessed Jul 10, 2018).
(249) Bonabeau, E. Agent-Based Modeling: Methods and Techniques for Simulating Human Systems. *Proc. Natl. Acad. Sci.***2002**, *99* (suppl 3), 7280–7287.
(250) Molecular dynamics - Latest research and news | Nature https://www.nature.com/subjects/molecular-dynamics (accessed Jul 4, 2018).
(251) Brown, S.-A.; Morgan, F.; Watras, J.; Loew, L. M. Analysis of Phosphatidylinositol-4,5-Bisphosphate Signaling in Cerebellar Purkinje Spines. *Biophys. J.***2008**, *95* (4), 1795–1812.
(252) Wehrle-Haller, B. Structure and Function of Focal Adhesions. *Curr. Opin. Cell Biol.***2012**, *24* (1), 116–124.
(253) Liu, J.; Wang, Y.; Goh, W. I.; Goh, H.; Baird, M. A.; Ruehland, S.; Teo, S.; Bate, N.; Critchley, D. R.; Davidson, M. W.; et al. Talin Determines the Nanoscale Architecture of Focal Adhesions. *Proc. Natl. Acad. Sci.***2015**, *112* (35), E4864–E4873.
(254) Reiske, H. R.; Kao, S.-C.; Cary, L. A.; Guan, J.-L.; Lai, J.-F.; Chen, H.-C. Requirement of Phosphatidylinositol 3-Kinase in Focal Adhesion Kinase-Promoted Cell Migration. *J. Biol. Chem.***1999**, *274* (18), 12361–12366.
(255) Schaller, M. D.; Hildebrand, J. D.; Shannon, J. D.; Fox, J. W.; Vines, R. R.; Parsons, J. T. Autophosphorylation of the Focal Adhesion Kinase, Pp125FAK, Directs SH2-Dependent Binding of Pp60src. *Mol. Cell. Biol.***1994**, *14* (3), 1680–1688.
(256) Hamadi, A.; Bouali, M.; Dontenwill, M.; Stoeckel, H.; Takeda, K.; Rondé, P. Regulation of Focal Adhesion Dynamics and Disassembly by Phosphorylation of FAK at Tyrosine 397. *J. Cell Sci.***2005**, *118* (19), 4415–4425.
(257) Mitra, S. K.; Schlaepfer, D. D. Integrin-Regulated FAK–Src Signaling in Normal and Cancer Cells. *Curr. Opin. Cell Biol.***2006**, *18* (5), 516–523.
(258) Bass, M. D.; Morgan, M. R.; Roach, K. A.; Settleman, J.; Goryachev, A. B.; Humphries, M. J. P190RhoGAP Is the Convergence Point of Adhesion Signals from Alpha 5 Beta 1 Integrin and Syndecan-4. *J. Cell Biol.***2008**, *181* (6), 1013–1026.
(259) Huveneers, S.; Danen, E. H. J. Adhesion Signaling – Crosstalk between Integrins, Src and Rho. *J. Cell Sci.***2009**, *122* (8), 1059–1069.
(260) Izard, T.; Brown, D. T. Mechanisms and Functions of Vinculin Interactions with Phospholipids at Cell Adhesion Sites. *J. Biol. Chem.***2016**, *291* (6), 2548–2555.
(261) Wu, Z.; Li, X.; Sunkara, M.; Spearman, H.; Morris, A. J.; Huang, C. PIPKI? Regulates Focal Adhesion Dynamics and Colon Cancer Cell Invasion. *PLoS ONE***2011**, *6* (9).

(262) Chandrasekar, I.; Stradal, T. E. B.; Holt, M. R.; Entschladen, F.; Jockusch, B. M.; Ziegler, W. H. Vinculin Acts as a Sensor in Lipid Regulation of Adhesion-Site Turnover. *J. Cell Sci.***2005**, *118* (7), 1461–1472.
(263) Ezratty, E. J.; Bertaux, C.; Marcantonio, E. E.; Gundersen, G. G. Clathrin Mediates Integrin Endocytosis for Focal Adhesion Disassembly in Migrating Cells. *J. Cell Biol.***2009**, *187* (5), 733–747.
(264) Chao, W.-T.; Kunz, J. Focal Adhesion Disassembly Requires Clathrin-Dependent Endocytosis of Integrins. *FEBS Lett.***2009**, *583* (8), 1337–1343.
(265) Ueno, T.; Falkenburger, B. H.; Pohlmeyer, C.; Inoue, T. Triggering Actin Comets versus Membrane Ruffles: Distinctive Effects of Phosphoinositides on Actin Reorganization. *Sci. Signal.***2011**, *4* (203), ra87.
(266) Schmid, S. L.; McNiven, M. A.; De Camilli, P. Dynamin and Its Partners: A Progress Report. *Curr. Opin. Cell Biol.***1998**, *10* (4), 504–512.
(267) Chao, W.-T.; Ashcroft, F.; Daquinag, A. C.; Vadakkan, T.; Wei, Z.; Zhang, P.; Dickinson, M. E.; Kunz, J. Type I Phosphatidylinositol Phosphate Kinase Beta Regulates Focal Adhesion Disassembly by Promoting Beta1 Integrin Endocytosis. *Mol. Cell. Biol.***2010**, *30* (18), 4463–4479.
(268) Franco, S. J.; Rodgers, M. A.; Perrin, B. J.; Han, J.; Bennin, D. A.; Critchley, D. R.; Huttenlocher, A. Calpain-Mediated Proteolysis of Talin Regulates Adhesion Dynamics. *Nat. Cell Biol.***2004**, *6* (10), 977–983.
(269) Kerstein, P.; M Patel, K.; M Gomez, T. *Calpain-Mediated Proteolysis of Talin and FAK Regulates Adhesion Dynamics Necessary for Axon Guidance*; 2017; Vol. 37.
(270) Chan, K. T.; Bennin, D. A.; Huttenlocher, A. Regulation of Adhesion Dynamics by Calpain-Mediated Proteolysis of Focal Adhesion Kinase (FAK). *J. Biol. Chem.***2010**, *285* (15), 11418–11426.
(271) Chang, S.-J.; Chen, Y.-C.; Yang, C.-H.; Huang, S.-C.; Huang, H.-K.; Li, C.-C.; Harn, H. I.-C.; Chiu, W.-T. Revealing the Three Dimensional Architecture of Focal Adhesion Components to Explain Ca2 +-Mediated Turnover of Focal Adhesions. *Biochim. Biophys. Acta BBA - Gen. Subj.***2017**, *1861* (3), 624–635.
(272) Serrano, K.; Devine, D. V. Vinculin Is Proteolyzed by Calpain during Platelet Aggregation: 95 KDa Cleavage Fragment Associates with the Platelet Cytoskeleton. *Cell Motil. Cytoskeleton***2004**, *58* (4), 242–252.
(273) Sprague, C. R.; Fraley, T. S.; Jang, H. S.; Lal, S.; Greenwood, J. A. Phosphoinositide Binding to the Substrate Regulates Susceptibility to Proteolysis by Calpain. *J. Biol. Chem.***2008**, *283* (14), 9217–9223.
(274) Xi, X.; Flevaris, P.; Stojanovic, A.; Chishti, A.; Phillips, D. R.; Lam, S. C. T.; Du, X. Tyrosine Phosphorylation of the Integrin B3 Subunit Regulates B3 Cleavage by Calpain. *J. Biol. Chem.***2006**, *281* (40), 29426–29430.
(275) Bae, Y. S.; Cantley, L. G.; Chen, C. S.; Kim, S. R.; Kwon, K. S.; Rhee, S. G. Activation of Phospholipase C-Gamma by Phosphatidylinositol 3,4,5-Trisphosphate. *J. Biol. Chem.***1998**, *273* (8), 4465–4469.
(276) Falasca, M.; Logan, S. K.; Lehto, V. P.; Baccante, G.; Lemmon, M. A.; Schlessinger, J. Activation of Phospholipase C Gamma by PI 3-Kinase-Induced PH Domain-Mediated Membrane Targeting. *EMBO J.***1998**, *17* (2), 414–422.
(277) Chen, H. C.; Guan, J. L. Association of Focal Adhesion Kinase with Its Potential Substrate Phosphatidylinositol 3-Kinase. *Proc. Natl. Acad. Sci. U. S. A.***1994**, *91* (21), 10148–10152.
(278) Armulik, A.; Velling, T.; Johansson, S. The Integrin B1 Subunit Transmembrane Domain Regulates Phosphatidylinositol 3-Kinase-Dependent Tyrosine Phosphorylation of Crk-Associated Substrate. *Mol. Biol. Cell***2004**, *15* (6), 2558–2567.
(279) Tvorogov, D.; Wang, X.-J.; Zent, R.; Carpenter, G. Integrin-Dependent PLC-Γ1 Phosphorylation Mediates Fibronectin-Dependent Adhesion. *J. Cell Sci.***2005**, *118* (3), 601–610.

(280) Kanner, S. B.; Grosmaire, L. S.; Ledbetter, J. A.; Damle, N. K. Beta 2-Integrin LFA-1 Signaling through Phospholipase C-Gamma 1 Activation. *Proc. Natl. Acad. Sci. U. S. A.***1993**, *90* (15), 7099–7103.
(281) Clark, E. A.; Brugge, J. S. Integrins and Signal Transduction Pathways: The Road Taken. *Sci.-N. Y. THEN Wash.-***1995**, 233–233.
(282) Raftopoulou, M.; Hall, A. Cell Migration: Rho GTPases Lead the Way. *Dev. Biol.***2004**, *265* (1), 23–32.
(283) Nader, G. P. F.; Ezratty, E. J.; Gundersen, G. G. FAK, Talin and PIPKIγ Regulate Endocytosed Integrin Activation to Polarize Focal Adhesion Assembly. *Nat. Cell Biol.***2016**, *18* (5), 491–503.
(284) Rubashkin, M. G.; Cassereau, L.; Bainer, R.; DuFort, C. C.; Yui, Y.; Ou, G.; Paszek, M. J.; Davidson, M. W.; Chen, Y.-Y.; Weaver, V. M. Force Engages Vinculin and Promotes Tumor Progression by Enhancing PI3K Activation of Phosphatidylinositol (3,4,5)-Triphosphate. *Cancer Res.***2014**, *74* (17), 4597–4611.
(285) Raucher, D.; Sheetz, M. P. Cell Spreading and Lamellipodial Extension Rate Is Regulated by Membrane Tension. *J. Cell Biol.***2000**, *148* (1), 127–136.
(286) Flevaris, P.; Stojanovic, A.; Gong, H.; Chishti, A.; Welch, E.; Du, X. A Molecular Switch That Controls Cell Spreading and Retraction. *J. Cell Biol.***2007**, *179* (3), 553–565.
(287) Schaller, M. D.; Otey, C. A.; Hildebrand, J. D.; Parsons, J. T. Focal Adhesion Kinase and Paxillin Bind to Peptides Mimicking Beta Integrin Cytoplasmic Domains. *J. Cell Biol.***1995**, *130* (5), 1181–1187.
(288) Schaller, M. D. Paxillin: A Focal Adhesion-Associated Adaptor Protein. *Publ. Online 02 Oct. 2001 Doi101038sjonc1204786***2001**, *20* (44).
(289) Welf, E. S.; Naik, U. P.; Ogunnaike, B. A. A Spatial Model for Integrin Clustering as a Result of Feedback between Integrin Activation and Integrin Binding. *Biophys. J.***2012**, *103* (6), 1379–1389.
(290) Vinzenz, M.; Nemethova, M.; Schur, F.; Mueller, J.; Narita, A.; Urban, E.; Winkler, C.; Schmeiser, C.; Koestler, S. A.; Rottner, K.; et al. Actin Branching in the Initiation and Maintenance of Lamellipodia. *J Cell Sci***2012**, *125* (11), 2775–2785.
(291) The Cytoskeleton and Cell Migration - Electron Tomography http://cellix.imba.oeaw.ac.at/lamellipodia/et (accessed Aug 26, 2017).
(292) Tsai, F.-C.; Meyer, T. Ca2+ Pulses Control Local Cycles of Lamellipodia Retraction and Adhesion along the Front of Migrating Cells. *Curr. Biol. CB***2012**, *22* (9), 837–842.
(293) Puhka, M.; Vihinen, H.; Joensuu, M.; Jokitalo, E. Endoplasmic Reticulum Remains Continuous and Undergoes Sheet-to-Tubule Transformation during Cell Division in Mammalian Cells. *J. Cell Biol.***2007**, *179* (5), 895–909.
(294) Philips, R. M. & R. » How big is the endoplasmic reticulum of cells? http://book.bionumbers.org/how-big-is-the-endoplasmic-reticulum-of-cells/ (accessed Sep 7, 2017).
(295) Shemesh, T.; Klemm, R. W.; Romano, F. B.; Wang, S.; Vaughan, J.; Zhuang, X.; Tukachinsky, H.; Kozlov, M. M.; Rapoport, T. A. A Model for the Generation and Interconversion of ER Morphologies. *Proc. Natl. Acad. Sci.***2014**, *111* (49), E5243–E5251.
(296) Schwarz, D. S.; Blower, M. D. The Endoplasmic Reticulum: Structure, Function and Response to Cellular Signaling. *Cell. Mol. Life Sci.***2016**, *73*, 79–94.
(297) Bygrave, F. L.; Benedetti, A. What Is the Concentration of Calcium Ions in the Endoplasmic Reticulum? *Cell Calcium***1996**, *19* (6), 547–551.
(298) Sage, S. O.; Pugh, N.; Mason, M. J.; Harper, A. G. S. Monitoring the Intracellular Store Ca2+ Concentration in Agonist-Stimulated, Intact Human Platelets by Using Fluo-5N. *J. Thromb. Haemost. JTH***2011**, *9* (3), 540–551.
(299) Rink, T. J. Cytosolic Calcium in Platelet Activation. *Experientia***1988**, *44* (2), 97–100.
(300) Donahue, B.; Abercrombie, R. Free Diffusion Coefficient of Ionic Calcium in Cytoplasm. *Cell Calcium***1988**, *8*, 437–448.

(301) Dickinson, G. D.; Ellefsen, K. L.; Dawson, S. P.; Pearson, J. E.; Parker, I. Hindered Cytoplasmic Diffusion of Inositol Trisphosphate Restricts Its Cellular Range of Action. *Sci. Signal.***2016**, *9* (453), ra108.
(302) Sprague, B. L.; Pego, R. L.; Stavreva, D. A.; McNally, J. G. Analysis of Binding Reactions by Fluorescence Recovery after Photobleaching. *Biophys. J.***2004**, *86* (6), 3473–3495.
(303) Cluzel, C.; Saltel, F.; Lussi, J.; Paulhe, F.; Imhof, B. A.; Wehrle-Haller, B. The Mechanisms and Dynamics of Avβ3 Integrin Clustering in Living Cells. *J. Cell Biol.***2005**, *171* (2), 383–392.
(304) Wang, J. Pull and Push: Talin Activation for Integrin Signaling. *Cell Res.***2012**, *22* (11), 1512–1514.
(305) Beck, M.; Schmidt, A.; Malmstroem, J.; Claassen, M.; Ori, A.; Szymborska, A.; Herzog, F.; Rinner, O.; Ellenberg, J.; Aebersold, R. The Quantitative Proteome of a Human Cell Line. *Mol. Syst. Biol.***2011**, *7*, 549.
(306) Schrimpf, S. P.; Weiss, M.; Reiter, L.; Ahrens, C. H.; Jovanovic, M.; Malmström, J.; Brunner, E.; Mohanty, S.; Lercher, M. J.; Hunziker, P. E.; et al. Comparative Functional Analysis of the Caenorhabditis Elegans and Drosophila Melanogaster Proteomes. *PLoS Biol.***2009**, *7* (3), e1000048.
(307) van Meer, G.; Voelker, D. R.; Feigenson, G. W. Membrane Lipids: Where They Are and How They Behave. *Nat. Rev. Mol. Cell Biol.***2008**, *9* (2), 112–124.
(308) Fallica, B.; Maffei, J. S.; Villa, S.; Makin, G.; Zaman, M. Alteration of Cellular Behavior and Response to PI3K Pathway Inhibition by Culture in 3D Collagen Gels. *PLOS ONE***2012**, *7* (10), e48024.
(309) Choi, C. K.; Zareno, J.; Digman, M. A.; Gratton, E.; Horwitz, A. R. Cross-Correlated Fluctuation Analysis Reveals Phosphorylation-Regulated Paxillin-FAK Complexes in Nascent Adhesions. *Biophys. J.***2011**, *100* (3), 583–592.
(310) Parsons, J. T. Focal Adhesion Kinase: The First Ten Years. *J. Cell Sci.***2003**, *116* (8), 1409–1416.
(311) Burkhart, J. M.; Vaudel, M.; Gambaryan, S.; Radau, S.; Walter, U.; Martens, L.; Geiger, J.; Sickmann, A.; Zahedi, R. P. The First Comprehensive and Quantitative Analysis of Human Platelet Protein Composition Allows the Comparative Analysis of Structural and Functional Pathways. *Blood***2012**, *120* (15), e73-82.
(312) Sakai, T.; Li, S.; Docheva, D.; Grashoff, C.; Sakai, K.; Kostka, G.; Braun, A.; Pfeifer, A.; Yurchenco, P. D.; Fässler, R. Integrin-Linked Kinase (ILK) Is Required for Polarizing the Epiblast, Cell Adhesion, and Controlling Actin Accumulation. *Genes Dev.***2003**, *17* (7), 926–940.
(313) Wickström, S. A.; Lange, A.; Montanez, E.; Fässler, R. The ILK/PINCH/Parvin Complex: The Kinase Is Dead, Long Live the Pseudokinase! *EMBO J.***2010**, *29* (2), 281–291.
(314) Philips, R. M. & R. » What Is the Thickness of the Cell Membrane?
(315) Maxfield, F. R.; McGraw, T. E. Endocytic Recycling. *Nat. Rev. Mol. Cell Biol.***2004**, *5* (2), 121–132.
(316) Yeh, Y.-C.; Ling, J.-Y.; Chen, W.-C.; Lin, H.-H.; Tang, M.-J. Mechanotransduction of Matrix Stiffness in Regulation of Focal Adhesion Size and Number: Reciprocal Regulation of Caveolin-1 and B1 Integrin. *Sci. Rep.***2017**, *7* (1), 15008.
(317) Nobes, C. D.; Hall, A. Rho, Rac, and Cdc42 GTPases Regulate the Assembly of Multimolecular Focal Complexes Associated with Actin Stress Fibers, Lamellipodia, and Filopodia. *Cell***1995**, *81* (1), 53–62.
(318) Steffen, A.; Ladwein, M.; Dimchev, G. A.; Hein, A.; Schwenkmezger, L.; Arens, S.; Ladwein, K. I.; Holleboom, J. M.; Schur, F.; Small, J. V.; et al. Rac Function Is Critical for Cell Migration but Not Required for Spreading and Focal Adhesion Formation. *J. Cell Sci.***2013**.
(319) Putney, J. W.; Tomita, T. Phospholipase C Signaling and Calcium Influx. *Adv. Biol. Regul.***2012**, *52* (1), 152–164.

(320) Johny, J. P.; Plank, M. J.; David, T. Importance of Altered Levels of SERCA, IP3R, and RyR in Vascular Smooth Muscle Cell. *Biophys. J.***2017**, *112* (2), 265–287.
(321) Khorchid, A.; Ikura, M. How Calpain Is Activated by Calcium. *Nat. Struct. Mol. Biol.***2002**, *9* (4), 239–241.
(322) Giannone, G.; Rondé, P.; Gaire, M.; Beaudouin, J.; Haiech, J.; Ellenberg, J.; Takeda, K. Calcium Rises Locally Trigger Focal Adhesion Disassembly and Enhance Residency of Focal Adhesion Kinase at Focal Adhesions. *J. Biol. Chem.***2004**, *279* (27), 28715–28723.
(323) Berditchevski, F. Complexes of Tetraspanins with Integrins: More than Meets the Eye. *J. Cell Sci.***2001**, *114* (Pt 23), 4143–4151.
(324) Grande-García, A.; Echarri, A.; Del Pozo, M. A. Integrin Regulation of Membrane Domain Trafficking and Rac Targeting. *Biochem. Soc. Trans.***2005**, *33* (4), 609–613.
(325) Pozo, M. A. del; Alderson, N. B.; Kiosses, W. B.; Chiang, H.-H.; Anderson, R. G. W.; Schwartz, M. A. Integrins Regulate Rac Targeting by Internalization of Membrane Domains. *Science***2004**, *303* (5659), 839–842.
(326) Yin, H. L.; Janmey, P. A. Phosphoinositide Regulation of the Actin Cytoskeleton. *Annu. Rev. Physiol.***2003**, *65* (1), 761–789.
(327) Schiller, H. B.; Friedel, C. C.; Boulegue, C.; Fässler, R. Quantitative Proteomics of the Integrin Adhesome Show a Myosin II-Dependent Recruitment of LIM Domain Proteins. *EMBO Rep.***2011**, *12* (3), 259–266.
(328) Kuo, J.-C.; Han, X.; Hsiao, C.-T.; Yates, J. R.; Waterman, C. M. Analysis of the MyosinII-Responsive Focal Adhesion Proteome Reveals a Role for β-Pix in Negative Regulation of Focal Adhesion Maturation. *Nat. Cell Biol.***2011**, *13* (4), 383–393.
(329) Enzyme Database - BRENDA https://www.brenda-enzymes.org/ (accessed Jul 14, 2018).
(330) Bass, M. D.; Williamson, R. C.; Nunan, R. D.; Humphries, J. D.; Byron, A.; Morgan, M. R.; Martin, P.; Humphries, M. J. A Syndecan-4 Hair Trigger Initiates Wound Healing through Caveolin- and RhoG-Regulated Integrin Endocytosis. *Dev. Cell***2011**, *21* (4), 681–693.
(331) Zhou, H.-X. Crowding Effects of Membrane Proteins. *J. Phys. Chem. B***2009**, *113* (23), 7995–8005.
(332) Kim, J. S.; Yethiraj, A. Crowding Effects on Association Reactions at Membranes. *Biophys. J.***2010**, *98* (6), 951–958.
(333) Guigas, G.; Weiss, M. Effects of Protein Crowding on Membrane Systems. *Biochim. Biophys. Acta BBA - Biomembr.***2016**, *1858* (10), 2441–2450.
(334) Monine, M. I.; Haugh, J. M. Reactions on Cell Membranes: Comparison of Continuum Theory and Brownian Dynamics Simulations. *J. Chem. Phys.***2005**, *123* (7), 074908.
(335) Triola, G. The Protein Lipidation and Its Analysis. *J. Glycomics Lipidomics***2010**, *2* (2), 1–14.
(336) Schwiening, C. J. A Brief Historical Perspective: Hodgkin and Huxley. *J. Physiol.***2012**, *590* (Pt 11), 2571–2575.
(337) UniProt https://www.uniprot.org/ (accessed Jul 14, 2018).
(338) KEGG PATHWAY Database https://www.genome.jp/kegg/pathway.html (accessed Jul 14, 2018).
(339) Economics - An A-Z of business quotations https://www.economist.com/schumpeter/2012/07/20/economics (accessed Jul 14, 2018).