UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA

Computational modeling of TGF-β signaling

Márcia Raquel Gomes Carvalho de Abreu

Mestrado Integrado em Engenharia Biomédica e Biofísica
Perfil em Engenharia Clínica e Instrumentação Médica

Dissertação orientada por:
Raquel Conceição
Aurélie Carlier

2020

ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr. Aurélie Carlier for allowing me to be part of the cBITE group in MERLN and tackle this challenge, for constantly pushing me out of my comfort zone in order to become a greater self, for all the knowledge shared with me and deep care to make sure that my wellbeing was never compromised in any situation. I would also like to thank Dr. Raquel Conceição for handling me with such gentleness, patience and flexibility, for her keen eye for detail and for all the guidance provided to me throughout this process.

I would like to express my appreciation to my colleagues from the cBITE group, in MERLN. I could not have asked for a better blend of positive personalities to be surrounded by. In particular, I would like to extend my gratitude to my mentor Kerbaï Saïd for allowing me to take part of his PhD project, for all the assistance provided to me during this project and for his patience to lecture me numerous thorough explanations. Along with Saïd, a special thank you to Zeynep Karagöz and Jasia King, not only for the help provided concerning this work, but also for the amazing, fun times that we shared together and for all the laughter that they granted me with. I would like to expand my appreciation vows to Prof. Dr. Martjin van Griensven for attending my concerns with kindness and consideration and to Dr. Dennie Hebels, who tirelessly helped me to handle less agreeable circumstances which arose along the path.

I would also like to commend the remaining elements of the MERLN institute. I could not have imagined that I would experience such a pleasant work environment, filled with friendly and supportive people. A particular thank you to Dr. Steven Vermeulen for the extremely helpful suggestions, which were pivotal to develop this project, and for exchanging his remarkable knowledge on Biology with me. Further words of acknowledgement to Filipa Lourenço and Daniel Carvalho for their countless generous gestures from the moment I arrived in Maastricht until I came back to Portugal, for all the happy times we spent together, insightful conversations, motivational words and for infecting me with their beautiful, upbeat spirit.

I am forever grateful to my friends Nicole, Sara, Kamalpreet, David, Patrícia, Reiny and Jaime for being by my side anytime needed, even if it was not always possible to do it physically. Some special words are directed to Nicole for the emotional support she offered me during this process and for always providing me outstanding advice and warm words to cheer me up during the worst times.

Lastly, I take this opportunity to thank my parents Lígia and Osvaldo and my siblings Fernando, Virgínia and Sandro for always believing in myself and reminding me that nothing is out of reach.

RESUMO

O fator de transformação do crescimento beta ou TGF-β é uma citocina que consta nos 33 membros da superfamília do TGF-β, envolvida em processos fundamentais da célula, nomeadamente na proliferação, migração, diferenciação das células e apoptose celular, assim como na homeostasia dos tecidos. Consequentemente, o mau funcionamento da via de sinalização do TGF-β está associado a diferentes patologias. Tendo em conta este facto e a grande diversidade de respostas biológicas por si originadas, esta cascata de sinalização é um grande alvo de pesquisas, da parte de diversos autores, através da implementação de diversos modelos computacionais.

A morfologia da célula não é tida em conta na implementação da maioria dos modelos correntes da via de sinalização canónica do TGF-β. Porém, esta característica celular é normalmente vista como uma medida de grande relevância para ditar o modo como a célula responde a estímulos mecânicos externos. A resposta das células ao ambiente extracelular, caracterizada pela conversão de estímulos mecânicos em sinas bioquímicos, tem a designação de mecanotrasdução. As integrinas, que são uma família de receptores de sinais e de adesão células, convertem estímulos mecânicos em sinais bioquímicos, constituindo um meio de comunição entre as células e o ambiente extracelular. Deste modo, as interacções entre a via de sinalização do TGF-β e as integrinas constituem uma forma de mecanotransdução. Um entendimento mais profundo da dinâmica da via de sinalização do TGF-β através de um modo dependente da morfologia da célula e um entendimento do papel que o *crosstalk* entre as integrinas e a via de sinalização canónica do TGF-β tem sobre a expressão genética permitem, consequentemente, perceber melhor a conexão entre esta via e o comportamento celular perante estímulos mecânicos. Tal conhecimento detém o potencial para prever interações entre as células e biomateriais e poderá, possivelmente, ser aplicado a diferentes terapias, tais como terapias associadas ao controlo do desenvolvimento de tumores.

Devido às razões previamente referidas, neste trabalho é apresentado um novo modelo computacional para a via de sinalização canónica do TGF-β, contendo um módulo de *crosstalk* com as integrinas. O modelo construído é espacial, sendo implementado através de equações diferenciais parciais. As equações utilizadas são do tipo reação-difusão.

Os principais objectivos do presente trabalho são avaliar o impacto que a morfologia da célula detém sobre a sinalização *downstream* e perceber qual é a influência da *crosstalk* entre a via de sinalização canónica do TGF-β e as integrinas sobre a expressão genética. Deste modo, o modelo implementado neste trabalho divide-se em dois módulos – Módulo I e Módulo II. O modelo *standard*, que consiste no modelo da via de sinalização canónica do TGF-β, corresponde ao Módulo I. O Módulo I do modelo foi construído com base em dois artigos: *Claus et al*., 2013 e *Nicklas & Saiz*, 2013. As equações que envolvem as diferentes espécies da via de sinalização canónica do TGF-β, assim como as condições de fronteira aplicáveis às membranas citoplasmática e nuclear, foram adaptadas de *Claus et al.*, 2013, e a secção do modelo respeitante ao *trafficking* de receptores foi adaptado de *Nicklas e & Saiz*, 2013. As equações de reação-difusão implementadas no Módulo I traduzem os passos gerais da cascata de sinalização do TGF-β referidos por diversos autores: os recetores do TGF-β são ativados pelo ligando TGF-β, recrutando e fosforilando o Smad2/3 junto à membrana citoplasmática. O Smad2/3 fosforilado, pSmad2/3, pode associar-se ao Smad mediador, o Smad4, formando um complexo, o pSmad2/3 + Smad4. Este complexo migra para o núcleo, onde se associa ao DNA, induzindo expressão genética. Quando o pSmad2/3 não se associa ao Smad4, pode migrar diretamente para o núcleo, tal como moléculas de Smad2/3 e moléculas de Smad4 podem deslocar-se até ao núcleo sem estarem associadas a outras espécies. O processo de formação do complexo pSmad2/3 + Smad4 também pode ocorrer no núcleo da

célula. O Módulo II do modelo diz respeito à implementação de diferentes interações de *crosstalk* entre as integrinas e a via de sinalização do TGF-β. As interações de *crosstalk* incluídas neste módulo são: 1) A *upregulation* dos recetores do TGF-β despoletada pelas integrinas; 2) O aumento da constante catalítica da reacção de fosforilação do Smad2/3 mediada pelo complexo formado pelo TGF-β e os seus recetores, complexo *C*, sendo este aumento despoletado pelas integrinas; 3) A estabilização dos recetores do TGF-β despoletada pelas integrinas, resultando na diminuição da taxa de degradação do complexo *C*; 4) A formação de um complexo composto pelas integrinas e pelo complexo *C*, complexo *IC*, que se liga ao Smad2/3, provocando um aumento da constante catalítica da reação de fosforilação do Smad2/3.

Foram realizadas diferentes experiências *in silico* para atingir os objetivos previamente referidos. Relativamente ao Módulo I, de modo a avaliar o impacto da morfologia celular sobre a expressão genética induzida pelo TGF-β, foram feitas simulações para diferentes geometrias da célula, nomeadamente para elipses e rectângulos com diferentes rácios de largura e altura. Foram ainda realizadas simulações para diferentes áreas da célula, num intervalo de valores que se estende desde a área estabelecida no modelo *standard* até 10 vezes a área do modelo *standard*. A avaliação do impacto da área da célula a nível *downstream* é realizada para as geometrias acima referidas. Adicionalmente, os parâmetros do modelo foram submetidos a uma análise de sensibilidade. No que diz respeito ao Módulo II, realizaram-se diferentes simulações para diferentes tipos de *crosstalk* que se estabelecem entre as integrinas e a via de sinalização do TGF-β para perceber a influência que um aumento dos parâmetros associados a cada tipo de *crosstalk* possui sobre a concentração do complexo pSmad2/3 + Smad4 no núcleo. Esta incrementação dos valores atribuídos a cada parâmetro é feita de modo independente ou para dois ou três parâmetros em simultâneo.

A análise das diferentes simulações realizadas, com o intuito de avaliar apenas o impacto da geometria da célula, indica que a espécie participante desta via de sinalização que mais sofre discrepâncias na sua concentração nuclear por rácio largura/altura, devido a mudanças na geometria da célula, é o Smad2/3. As restantes espécies, nomeadamente o pSmad2/3, o complexo pSmad2/3 + Smad4 e o Smad4, sofrem discrepâncias idênticas por unidade de rácio largura/altura. Constatou-se que existe uma tendência de decréscimo na concentração nuclear do Smad2, pSmad2/3 e complexo pSmad2/3 + Smad4 com o aumento do rácio largura/altura da célula. Por outro lado, constatou-se um aumento da concentração nuclear do Smad4 por unidade de rácio largura/altura. As simulações efectuadas para avaliar o efeito que as dimensões da célula possuem sobre a cascata do TGF-β, a nível *downstream*, indicam que com o aumento do tamanho da célula, existe diminuição da concentração do complexo pSmad2/3 + Smad4 no núcleo da célula, havendo um declínio na expressão genética. Esta conclusão é aplicável a todas as geometrias da célula consideradas. Concretamente, considerando a passagem da área *standard* da célula para uma área 10 vezes superior, constata-se uma diminuição de cerca de 50% na concentração nuclear do complexo pSmad2/3 + Smad4 para todos os formatos da célula. Ainda relativamente a resultados do Módulo I, os diversos parâmetros do modelo foram submetidos a uma análise de sensibilidade que revelou que a concentração do complexo pSmad2/3 + Smad4 no núcleo é mais sensível às reacções do modelo do que a processos de difusão. No entanto, um decréscimo de 100% do valor dos parâmetros de difusão revela uma sensibilidade considerável a nível *downstream* da via de sinalização, o que sugere que ainda assim os parâmetros de difusão devem ser incluídos em modelos que visem estudar a dinâmica da cascata do TGF-β.

As simulações realizadas para estudar o efeito da *crosstalk* entre as integrinas e a via de sinalização do TGF-β sugerem que de todos os tipos de interação implementados no modelo construído, o tipo de interação que causa maior *upregulation* da concentração nuclear do complexo pSmad2/3 + Smad4 é o decréscimo da taxa de degradação dos receptores devido à estabilização dos receptores do TGF-β por parte das integrinas. Por outro lado, o tipo de interacção de *crosstalk* que menos contribui para a expressão genética é o aumento da constante catalítica de fosforilação do Smad2/3 mediada pelo complexo *IC*. Em adição, a expressão genética sofreu mais *upregulation*, em geral, quando houve uma intensificação

simultânea dos efeitos de dois tipos de crosstalk com as integrinas, comparativamente com uma forma de *crosstalk* ou três tipos de interação em conjunto.

Os próximos passos associados ao trabalho apresentado neste documento são o estabelecimento de um conjunto de experiência *in vitro* para analisar o efeito da morfologia da célula e da mecanotransdução, com o intuito de validar os resultados e conclusões obtidos através do conjunto de experiências *in silico*. Adicionalmente, outras formas de *crosstalk* entre as integrinas e a cascata do TGF-β deverão ser incluídas no modelo, tais como a interação da via de sinalização com proteínas reguladas pela actina-G e actina-F do citoesqueleto da célula. Estas proteínas são libertadas quando existe remodelação do citoesqueleto e um desequilíbrio nos seus níveis resulta na expressão de receptores de ligandos da família TGF-β.

Palavras-chave: Sinalização do TGF-β; Morfologia da Célula; Integrinas; Mecanotransdução; *Crosstalk*

ABSTRACT

The TGF-β is a cytokine involved in fundamental cell processes, such as cell migration, proliferation and apoptosis. A deeper insight into the dynamics of the TGF-β pathway through a cell morphology-dependent manner and into the role that the crosstalk with the integrins has upon downstream signaling of this pathway allows a better understanding of the link between the TGF-β pathway and mechanotransduction cues, therefore holding the potential to uncover interactions between biomaterials and cells and apply this knowledge to different therapies.

In this document, a spatial computational model for the canonical TGF-β pathway is presented, including a module for the crosstalk with the integrins. The equations used to model this system are partial differential equations of the reaction-diffusion type. The model was created to evaluate the impact of cell shape and size, as well as crosstalk with the integrins, on genetic expression.

The results of the experiments suggest that a bigger width/height ratio of the cell induces less concentration of the nuclear Smad2/3, nuclear phosphorylated Smad2/3 and nuclear phosphorylated Smad2/3 + Smad4 complex, for elliptical and rectangular cell shapes. Conversely, the concentration of the nuclear Smad4 increases as the width/height ratio of the cell increases, for elliptical and rectangular cell shapes. The nuclear species of the TGF-β pathway that suffers more variation in its nuclear concentration across cell shapes is the Smad2/3. Considering a range of values that go up to 10 times the original area of the cell for different cell shapes, as the area of the cell increases, the concentration of the nuclear phosphorylated Smad2/3 + Smad4 complex decreases by approximately 50%. Sensitivity analysis of the model parameters indicate that genetic expression elicited by the TGF-β when considering diffusive processes differs considerably from results without considering this phenomenon. It is suggested that the way of crosstalk which more heavily increases genetic expression is the decrease of the degradation rate of the receptors due to stabilization from the integrins. Experimental procedures also suggest that the concentration of the nuclear phosphorylated Smad2/3 + Smad4 complex undergoes more upregulation, in general, by the joint behavior of two ways of crosstalk interaction than by the effects of one way of crosstalk independently or the combination of three ways of crosstalk out of the considered interactions.

Keywords: TGF-β Pathway; Cell morphology; Integrins; Mechanotransduction; Crosstalk

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS AND ABBREVIATIONS

1D	One-dimensional
2D	Two-dimensional
4D	Four-dimensional
Akt	Protein Kinase B
ALK	Anaplastic Lymphoma Kinase
BMP	Bone Morphogenetic Protein
C	TGF-β Receptors + TGF-β Complex
$\mathbf{Ca^{2+}}$	Calcium Ion
Co-Smad	Common Mediator Smad
c-Src	Proto-oncogene Tyrosine-protein Kinase
CYT	Cytoplasmic Subdomain
DNA	Deoxyribonucleic Acid
EC	Extracellular Matrix Subdomain
EGF	Epidermal Growth Factor
EGFR	Epidermal Growth Factor Receptor
F-actin	Filamentous Actin
G-actin	Globular Actin
G-protein	Heterotrimeric Guanine Nucleotide-binding Protein
GTP	Guanosine Triphosphate
HaCaT	Immortalized Human Keratinocytes
I	Integrins
IC	Integrins + TGF-β Receptors + TGF-β Complex

Imax	Limit Value of the Concentration of I
Kcat	Catalytic Constant of the Phosphorylation of the SMAD2 Mediated by the C Complex
Kcn_pSMAD2	Nuclear Import Rate of the pSMAD2
Kcn_SMAD2	Nuclear Import Rate of the SMAD2
Kcn_SMAD4	Nuclear Import Rate of the SMAD4
Kcn_TRIMER	Nuclear Import Rate of the TRIMER
Kdeg	Degradation Rate of the C Complex
Kdeg_intg	Degradation Rate of I
Kdephos	Dephosphorylation Rate of the pSMAD2n
Kdiss	Dissociation Rate of the TRIMER and TRIMER_n
Kdiss_IC	Dissociation Rate of the IC Complex
K_extra1	Step the Function for the Increase of the Synthesis Rate of the C Complex
K_extra2	Step of the Function for the Increase of the Catalytic Constant of the Phosphorylation of the SMAD2 Mediated by the C Complex
K_extra3	Step of the Function for the Decrease of the Degradation Rate of the C complex
K_extra4	Step of the Function for the Increase of the Catalytic Constant of the Phosphorylation of the SMAD2 Mediated by the IC complex
Kform_IC	Formation Rate of the IC Complex
Knc_SMAD2	Nuclear export rate of the SMAD2n
Knc_SMAD4	Nuclear export rate of the SMAD4n
Kr	Dephosphorylation Rate of the pSMAD2
Ksyn	Synthesis Rate of the C complex
Ksyn_intg	Synthesis Rate of I
Ktrim	Formation Rate of the TRIMER and TRIMER_n
MAD	Mothers Against Decapentaplegic

MAPK	Mitogenactivated Protein Kinase
NUC	Nuclear Subdomain
ODE	Ordinary Differential Equation
PDE	Partial Differential Equation
PI3K	Phosphatidylinositol-3-kinase
PPM	Metal-dependent Protein Phosphatase
pSMAD2	Cytoplasmic Phosphorylated Smad2
pSmad2/3	Phosphorylated Smad2/3
pSMAD2_diffusionRate	Diffusion Rate of the pSMAD2
pSMAD2n	Nuclear Phosphorylated Smad2
pSMAD2n_diffusionRate	Diffusion Rate of the pSMAD2n
$\mathbf{R^2}$	Coefficient of Determination
$\mathbf{R^2_{adj}}$	Adjusted Coefficient of Determination
SC	(Species Steady-state Concentration in a Given Shape)/(Species Steady-state Concentration in the Circle)
SMAD2	Cytoplasmic Smad2
SMAD2_diffusionRate	Diffusion Rate of the SMAD2
SMAD2n	Nuclear Smad2
SMAD2n_diffusionRate	Diffusion Rate of the SMAD2n
SMAD4	Cytoplasmic Smad4
SMAD4_diffusionRate	Diffusion Rate of the SMAD4
SMAD4n	Nuclear Smad4
SMAD4n_diffusionRate	Diffusion Rate of the SMAD4n
SSR	Sum of Squares
SST	Total Sum of Squares

TGF-β	Transforming Growth Factor Beta
TM	Transmembrane
TRIMER	Cytoplasmic Phosphorylated Smad2 + Smad4 Complex
TRIMER_diffusionRate	Diffusion Rate of the TRIMER
TRIMER_n	Nuclear Phosphorylated Smad2 + Smad4 Complex
TRIMERn_diffusionRate	Diffusion Rate of the TRIMER_n
VCell	Virtual Cell
VEGFR	Vascular Endothelial Growth Factor Receptor

CHAPTER 1 – Introduction

1.1 – Contextualization and Motivation

The transforming growth factor beta (TGF-β) is a cytokine involved in fundamental cell processes, such as cell migration, proliferation, differentiation and tissue homeostasis, development and apoptosis [1], [2]. Consequently, the ill-functioning of this cytokine is associated with different pathologies [1]–[8]. This fact, along with the wide diversity of biological responses which comes from the TGF-β pathway, makes this pathway a target of research through different computational models, presented by different authors [2], [9]–[21]. Nevertheless, the majority of current models of the canonical TGF-β pathway do not take into account cell morphology, which is typically considered to be a very relevant measure of how cells respond to external mechanical cues [22]. This response to the extracellular environment, followed by the conversion of the mechanical cues into biochemical signals is called mechanotransduction [23]. The integrins, which is a family of signaling and cell adhesion receptors, transduce mechanical cues into biochemical signals, constituting a means of communication between the cells and the extracellular environment [24], [25]. Therefore, the interactions between the TGF-β pathway and the integrins constitute a form of mechanotransduction.

A deeper insight into the dynamics of the TGF-β pathway through a morphology-dependent manner and into the role that the crosstalk with the integrins has upon downstream signaling of this pathway allows a better understanding of the link between the TGF-β pathway and mechanotransduction cues, therefore holding the potential to uncover the interactions between biomaterials and cells and apply this knowledge to different therapies.

For these reasons, a new spatial computational model for the canonical TGF-β pathway is presented in this document. The model is implemented in the Virtual Cell modeling software and is modeled through partial differential equations, namely reaction-diffusion equations. The final model contains a module which translates the crosstalk between the TGF-β signaling pathway and the integrins. It is worth noting that the present work was developed under the scope of a PhD project called "Image-based computational modelling of cell-topography induced cell behaviour", developed in the Institute for Technology-Inspired Regenerative Medicine, MERLN, by the PhD student Kerbaï Saïd Eroumé. One aspect of this PhD project concerns the implementation of the YAP/TAZ or Hippo signaling pathway model, which is a signaling pathway related to cell density sensing and arrangement of tissues. The relevance of linking the TGF-β signaling pathway with the Hippo signaling pathway lies in the fact that there is evidence of formation of YAP/TAZ–Smad2/3 complexes in immortalized human keratinocytes [26], known as HaCaT cell line, which suggests that there is crosstalk between both signaling cascades [27].

1.2 – Objectives

The main objectives of this project were: I) to build a spatial, computational model for the canonical TGF-β pathway, built on previous models, in order to assess the effect of cell morphology, particularly of the shape and size of the cell, on downstream signaling; II) to evaluate the influence of the crosstalk between the integrins and the TGF-β canonical pathway on genetic expression. Due to these two main goals, the process of building a new model comprised the implementation of two modules: Module I- The canonical TGF-β signaling pathway computational model; Module II- Crosstalk of the canonical TGF-β signaling pathway with the integrins.

The achievement of these goals required following the plan below:

a) Become acquainted with the Biology concerning the canonical TGF-β pathway;
b) Understand the different topics or sections included in the preexisting computational models of the canonical TGF-β pathway and realize in what manner they can be improved - contribution of this work;
c) Implement the model of the TGF-β canonical pathway – implementation of Module I;
d) Perform experiments to evaluate the effect of the shape and size of the cell on downstream signaling, followed by the analysis of the results;
e) Add, to the previously implemented model, a module representative of the crosstalk between the integrins and the TGF-β pathway – implementation of Module II;
f) Perform experiments to assess the influence of the crosstalk between the integrins and the TGF-β canonical pathway on genetic expression, followed by analysis of results.

1.3 – Dissertation Outline

This section describes the scope of each chapter of this dissertation.

The present chapter provides an overview of the work performed in order to understand the motivation, problems and proposed resolution associated with it. In addition, it discloses the organization of this document.

CHAPTER 2 consists in a literature review of different aspects pertinent to the developed work, including background on: the concept of Tissue Engineering along with applications of this area, the components of the TGF-β family, Biology aspects behind the canonical TGF-β pathway and the existing computational models for the canonical TGF-β pathway, and also on the integrins and their crosstalk with the TGF-β signaling cascade.

CHAPTER 3 displays the materials and methods used in this project, referencing the importance of computational models, identifying the software in which both modules of the model (see Section **1.2 – Objectives**) were implemented, the diagrams which visually display the reactions associated with both modules of the model, the components and parameters which take part of both modules, the geometry of the cell, domain and subdomains of the TGF-β pathway model.

CHAPTER 4 describes parameter fitting procedures of the model, “sanity checks” performed for validation purposes and exhibits the results and discussion concerning the different *in silico* experiments performed to Module I and Module II of the TGF-β model.

CHAPTER 5 features the conclusions drawn from the different *in silico* experiments, suggestions regarding potential improvements of the model here presented and pertinent steps to take following this work.

CHAPTER 2 – Literature Review

2.1 – Tissue Engineering

Tissue engineering results from the association between engineering principles and life science principles [28]. In particular, fields such as material science, nanotechnology, developmental biology, cell biology and rapid prototyping have been pivotal contributors to the advance of tissue engineering [29], which consists of controlling cell behavior with the intent of improving, regenerating or preserving tissues and organs to treat certain clinical conditions. Simultaneously, it promises technological advances that will potentially lead to treatment and cure of illnesses simply by recurring to drugs, hence eradicating the need for organ transplantation [30]. In addition, it focuses on manipulating the environment surrounding the healing region [28]. There are three main divisions in tissue engineering: cells, materials and architecture, as seen in Figure 2.1. These different components can overlap, performing complementary work with one another, or be engineered individually [29].

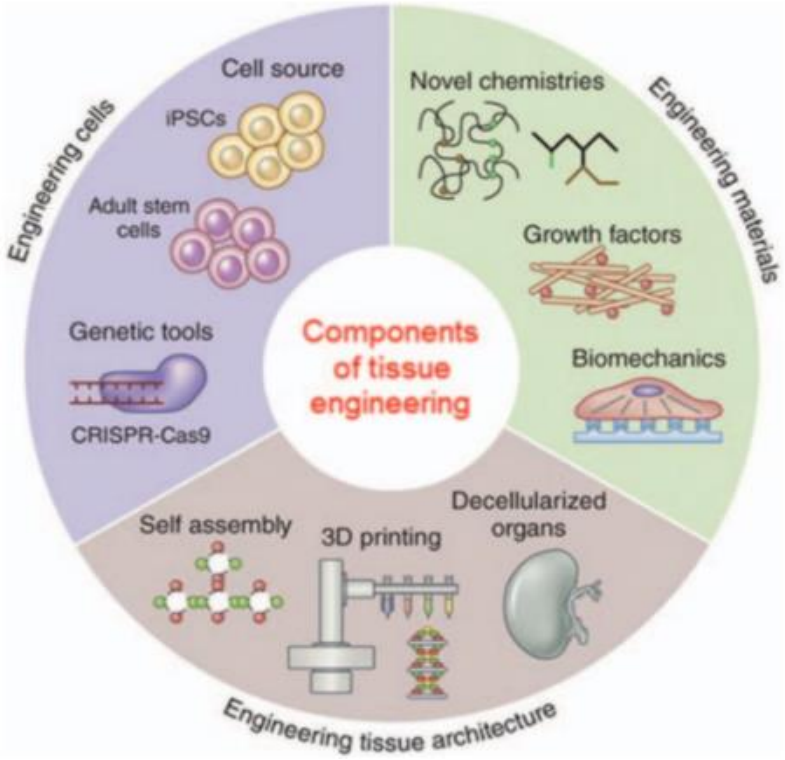

Figure 2.1 – Main sectors in tissue engineering – cells, materials and architecture. Adapted from *Lanza et al.*, 2020 *[29]*.

The employment of tissue engineering techniques usually comprises the placement of stem cells within a porous scaffold (a carrier structure) [31], represented in Figure 2.2, which does or does not release growth factors. Both the scaffold and the growth factors provide physical and chemical signals. These cues prompt the differentiation of stem cells and the secretion of tissue-specific extracellular matrix [30], [32]. However, tissue engineering procedures do not always include scaffolds.

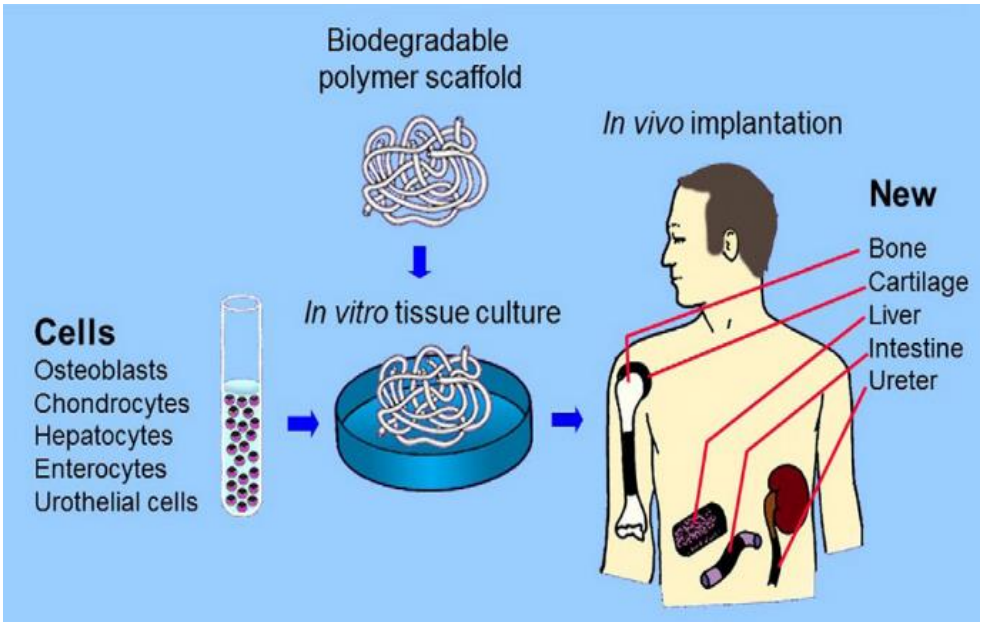

Figure 2.2 – Common tissue engineering procedure. Stem cells are placed on a scaffold made of a biodegradable polymer, which is implemented in the target tissue of the patient. Adapted from *Langer & Vacanti*, 2015 *[31]*.

Currently, three main strategies are explored to treat patients with damaged tissues:
– Implantation of recently isolated or cultured cells, which comprises the injection of cells into the injured tissues, which may have been previously placed in a degradable scaffold *in vitro*;
– Implantation of tissues created *in vitro*, which implies total development of 3D tissues in *in vitro* environment prior to implantation;
– *In situ* tissue regeneration, which consists of promoting tissue restoration by placing a scaffold inside the damaged tissue [30].

There is evidence of different positive outcomes from tissue engineering techniques. 3D tissue has been pivotal to create models for illnesses and to develop drugs [29]. People who suffer from diabetes, skin ulcers and liver conditions have been successfully treated with allogeneic cells, which are cells that originally belong to a genetically similar – but different – individual [33]. There currently is an autologous cell product approved by the FDA used to restore articular cartilage, which consists of a fragment of cartilage taken from a healthy region of a patient's injured knee. Chondrocytes from this section of cartilage are placed in the injured region after being subjected to isolation and expansion in culture. Moreover, adult bone marrow stem cells have the ability to treat various blood disorders. It is known that a bone marrow transplant has cured at least one model of animal liver illness [30].

Even though tissue engineering benefits from major technological advances made recently, namely gene editing, advanced cell screening and cell reprogramming, there are still numerous questions to be

answered and gaps in knowledge to fill. A substantial part of current tissue engineering protocols are equilibrium systems deprived from anisotropy, structural and directional arrangement and heterogeneity, which do not match the reality of natural systems [34]. Furthermore, a source of healthy cells is mandatory for the successful employment of engineered tissues, as well as the development of bioreactors and enhancement of scaffolds [30]. Moreover, there is a lack of fundamental understanding between the crosstalk of the physical and chemical cues.

2.2 – Transforming Growth Factor-β Signaling Pathway

The TGF-β pathway is a signaling process known to have a fundamental role upon various cellular processes, namely proliferation, differentiation, maintenance of tissue homeostasis, development and apoptosis [2]. The ill-functioning of this signaling pathway and modification of any of its components have been linked to different medical conditions, namely cardiovascular diseases, cancer, developmental disorders and cancer [35].

The superfamily of TGF-β is composed of 33 members, including the bone morphogenic protein (BMP), the TGF-β, the nodal growth differentiation factor and the activin [2].

There has been evidence supporting the importance of the TGF-β pathway in tissue engineering and clinical applications in general [36]–[41]. The BMPs have extensively been mentioned as key biological species to promote bone repair and morphogenesis [35]. BMP-7 treatment, along with type I collagen as a carrier, has induced genesis of new bone, via stem cell differentiation, and, consequently, prompted the healing of non-unions in patients [42].

There are two types of TGF-β pathways: the canonical pathway and the non-canonical pathway. The canonical pathway concerns the traditional pathway, hence the pathway in which the effector cytosolic proteins are called Smads [43]. The former designation comes from the combination between the gene Mothers Against Decapentaplegic (MAD) discovered in *Drosophila* and the Sma genes found in *Caenorhabditis elegans*, which were later shown to participate in the TGF-β signaling pathway as fundamental mediators [44], [45]. The non-canonical pathways do not have Smads involved in the signaling cascade. Examples of non-canonical pathways are the mitogenactivated protein kinase pathway (MAPK) [46], the Wnt pathway , the Notch pathway and the phosphoinositide 3-kinase-Akt pathway (PI3K) [43], [46].

The following section will discuss the canonical pathway in more detail.

2.2.1 – Dynamics of the TGF-β Signaling Pathway

The signaling cascade of the TGF-β is triggered when a ligand of the TGF-β superfamily binds to two types of transmembrane serine-threonine kinase receptors lying on the plasma membrane, type I and type II [2], [11], [12]. Following ligand binding, a complex is formed with these two types of receptors and the type I receptor is submitted to transphosphorylation by the type II receptor, which is constitutively active. This sequence of events prompts activation of the kinase of the type I receptor [21] and internalization of this complex [2]. The activated type I receptor phosphorylates and regulates effector

cytosolic proteins Smads or R-Smads, whose subspecification depends on the subfamily of ligands considered – while the type I receptors of anaplastic lymphoma kinase (ALK) 4/5/7 [47] phosphorylate the Smad2/3 and are related to TGF-β, activin and nodal ligands, the type I receptors ALK1/2/3/6 phosphorylate the Smad1, Smad5 and Smad8 and are connected to BMP ligands [2], [21]. Activated R-Smads bind amongst each other and to the Smad4, designated as common mediator Smad or Co-Smad, constituting homomeric complexes and hetero-oligomers, respectively [10], [21]. Consequently, accumulation of Smads in the nucleus is promoted, as opposed to their tendency to accumulate in the cytoplasm in the absence of a ligand [13]. In this process of translocation into the nucleus, these complexes undergo constant nucleocytoplasmic shuttling [2], [18]. The nuclear translocation of the Smad complexes transmits information regarding receptor activity [2]. Having reached the nucleus, these molecules behave as regulators for gene expression, binding to specific promoters and to the deoxyribonucleic acid (DNA) [48], and interacting with transcriptional coactivators, corepressors and transcription factors [14], [15], [18]. This sequence of events can lead to more than 500 different gene responses [18]. Receptors are sensitive to the integral levels of ligands, ratios of several ligands and evolution of ligand concentration across time. Ligands are disassociated from receptors and submitted to degradation in the lysosomes [9]. Dissociation and dephosphorylation of Smad complexes occur in the nucleus by a nuclear phosphatase, such as metal-dependent protein phosphatase 1A (PPM1A) [49], constituting a method for negative regulation of TGF-β signaling in the nucleus [13], [50]. A depiction of the TGF-β signaling pathway can be observed in Figure 2.3. It is worth noting that in this figure, the phosphorylated Smad2/3 molecules are identified by pSmad2/3 and the complexes constituted by the phosphorylated Smad2/3 molecules and the Smad4 is denoted by pSmad2/3 + Smad4. In succeeding sections of this work, particularly from Section **3.2.1 – Diagrams and Components** onwards, the pSmad2/3 + Smad4 complex is referred to as the TRIMER.

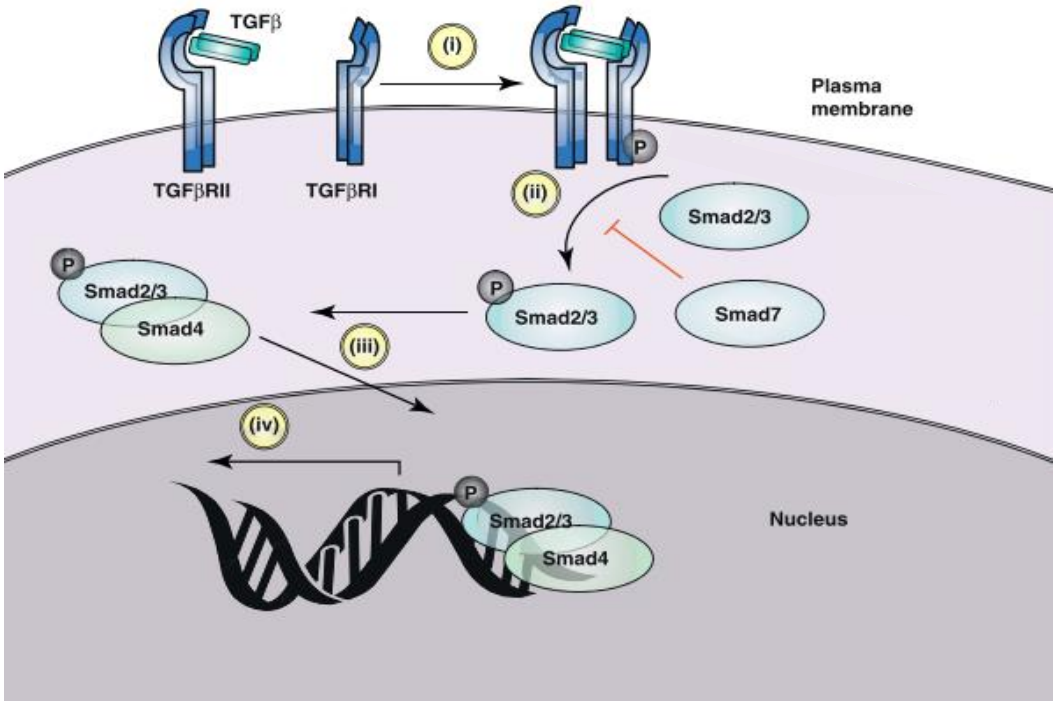

Figure 2.3 – Sequence of events of the TGF-β signaling pathway. Step (i) displays the formation of a complex composed by a ligand bound to type I TGF- β and type II TGF- β receptors, eliciting transphosphorylation of the type I TGF- β receptor and activation of its kinase. Step (ii) concerns the recruitment of Smad2/3 molecules by the type I TGF- β receptor, which phosphorylates them, followed by the formation of a pSmad2/3 + Smad4 complex. Step (iii) displays the translocation of the pSmad2/3 + Smad4 complex towards the nucleus. Step (iv) demonstrates the association of the pSmad2/3 + Smad4 complex with chromatin, where it regulates transcription of various genes. Adapted from *Worthington et al.*, 2011 *[43]*.

The activity of receptors is also regulated by different internalization routes – namely clathrin-dependent endocytosis and caveolar lipid-raft mediated endocytosis- and by receptor turnover [2], [15], [21]. If the receptors actively took part in the signaling process, they have different possible fates after internalization: they face degradation, are used for ligand-binding or used for further internalization processes. Receptors which did not form a ligand-receptor complex go back to the plasma membrane. Regarding internalization routes, the clathrin pathway is considered to be the usual pathway. Internalization through caveolar lipid-raft mediated endocytosis – which implies dissociation within caveolar lipid-raft – occurs when Smad7 and Smurf2 are recruited by ligand-receptor complexes, targeting the latter for degradation purposes [9].

2.2.2 – Computational Models of the TGF-β Pathway

In order to better understand which models have been used to represent the TGF-β signaling pathway, as well as their outcomes and modules used in their implementation, Table A (divided into Table A. 1, Table A. 2, Table A. 3 and Table A. 4), depicted in **APPENDIX A**, was created to compile the most relevant information and compare the models among each other. It is worth mentioning that these tables solely include information regarding models of the canonical TGF-β signaling pathway. Following the

identification of the authors and the year of publication of each article, some criteria to classify the models are mentioned, including: model type, aims and conclusion, topics which are approached in the article, the dimension of the model and length scale in which were implemented.

The category model type concerns the kind of equations which describe the model. The equations are classified as either partial differential equations (PDEs) or ordinary differential equations (ODEs). ODEs have popularly been used to represent reactions of biological systems [51], [52]. By looking at **APPENDIX A**, we can conclude that all the referenced studies created for the TGF-β pathway use ODEs to depict the reactions included in the signaling process. Models built with a high precision degree typically consist in a sequence of elementary reactions represented by ODEs, based on mass action kinetics [53]. ODEs are used to implement systems in time, independently of space. Molecules are considered to be homogeneously mixed in cellular compartments. These equations ease the simulation process and have small computational costs. Conversely, representation of chemical processes with PDEs requires a more rigorous previous mathematical knowledge [52]. Evidence of implementation of PDEs to describe dynamics of the TGF-β pathway is only found in [11] and [20]. In [11], outcomes stemming from two variants of a proposed model – with and deprived of spatial aspects – are compared to conclude if cell geometry has a relevant effect upon signaling from the TGF-β. It is inferred in this study that there is a substantial difference in signaling outcomes between the model with spatial aspects – when PDEs are used to represent cell reactions – and the model described without spatial aspects – reactions are represented by ODEs – considering dendritic cells or cells with several extensions. In the referred work, for the model made up of PDEs, the effect of cell shape was shown to be relevant to the TGF-β signaling pathway as cells with extensions displayed different gene expression patterns than the regularly shaped cells. This finding draws attention for the need to study more extensively the effect of cell morphology upon TGF-β signaling. In [20], PDEs were used to add a trafficking coordinate to the TGF-β signaling pathway. This section of the model aimed to describe the movement of the molecular species across compartments in order to keep track of their location. This study suggests that the insertion of a trafficking coordinate into the TGF-β pathway model allows to accurately replicate the physical properties of trafficking for different cells, while simultaneously describing chemical reactions.

The aims and conclusion section presents the scope of the work described in the article and its most important outcomes. A deeper understanding of tumor growth and proliferation, Smad nucleocytoplasmic shuttling, receptor trafficking and switch-like signaling are recognized as the most common aims of the research performed and presented in **APPENDIX A**. The assessment of the effects of cell shape has solely been done in [11]. Out of the models referenced in **APPENDIX A**, crosstalk between different pathways, which is regarded as a possible explanation for the large diversity of biological responses of the TGF-β signaling pathway, is only studied in [2]. With this study, it can be inferred that there is a large gap in knowledge concerning both the influence of spatial aspects upon the TGF-β signaling pathway and the crosstalk with other signaling pathways.

The topics section displays the different modules of the signaling pathway considered in the implementation of the model. The main topics found across the models are: formation of ligand-receptor complexes; formation, phosphorylation, dephosphorylation and dissociation of Smad complexes and Smad nucleocytoplasmic shuttling. In addition, a receptor trafficking module is used in a big part of the computational models proposed for the TGF-β signaling pathway, particularly in [2], [9], [15], [18], [20], [50]. Once again, the absence of a module of crosstalk between pathways is assessed in almost all the models. The exception consists of [2], where crosstalk between different pathways has been shown to be an important factor for the regulation of signal outcome. Taking into account that the results in [2] suggest that the crosstalk between Smad 1/5/8 and Smad 2/3 channels prompts specific signaling patterns, efforts should be made to include this type of dynamic in TGF-β models. Additionally, and as mentioned before, spatial description should be incorporated into the model.

There is evidence of one-dimensional (1D), two-dimensional (2D) and four-dimensional (4D) models for the TGF-β pathway in **APPENDIX A**. It is pertinent to mention that 1D models depict systems whose response varies solely as a function of time (discarding spatial coordinates), while 2D and 4D models include one or three spatial coordinates, respectively, in addition to the time factor. The only spatial models for the TGF-β pathway displayed in **APPENDIX A** are found in [11] and [20]. Considering the important outcomes of [11], which were mentioned in the previous paragraph, concerning topics of the models, more focus should be directed towards building a spatial model for this pathway.

Most of the work mentioned in **APPENDIX A** is implemented at cell level, meaning that the different locations involved in the signaling process are: the plasma membrane, the cytoplasm and the nucleus. However, one should bear in mind that cells are neighbours to other cells and the cell-neighbours dynamics might potentially affect outcomes from the TGF-β signaling pathway.

It is worth mentioning that **APPENDIX A** only includes information regarding articles which specifically focus on gathering deeper knowledge on the mechanisms behind the TGF-β signaling pathway. However, the TGF-β has been incorporated into mathematical models as a key player for different biological systems. In numerous articles [3]–[8], the important role of the TGF-β upon various physiological processes is highlighted, especially as a suppressive cytokine in tumor growth and proliferation, along with the potential that a deeper understanding of its signaling pathway or way of functioning carry to improve therapies for different pathologies.

After analysis of **APPENDIX A**, we can conclude the vast majority of the work performed to further understand the TGF-β pathway does not take into account spatial aspects of the cell and crosstalk between different pathways (e.g. mechanotransduction pathways), which is thought to contribute heavily for the wide diversity of outcomes of the signaling process [2], [54].

2.3 – Integrins

The integrins consist of a family of heterodimeric transmembrane cell surface adhesion receptors [24], [25], [55] which mediate the attachment of cells to the extracellular matrix, thus guaranteeing a mechanical connection between both and, sometimes, the adhesion between adjacent cells [24], [25], [55]–[58]. Considering that the integrins link the extracellular medium to the intracellular cytoskeleton and to different signaling pathways, they are regarded as a bidirectional route to transduce biochemical and mechanical signals [24], [25].

The importance of the integrins resides in their participation in numerous cell processes, such as migration, motility, proliferation, survival, and numerous cell fate transitions [55], [56]. In particular, integrin activation controls cell adhesion in mechanisms such as leukocyte migration when an immune response is elicited and adhesion among cells and between cells and the extracellular matrix in wound healing [25]. Other examples of cell behaviors which largely depend on integrin signaling are the rise in the intracellular pH, concentration of calcium (Ca^{2+}) [59] and the transduction of immediate-early genes [57]. The integrins constitute an important point of response to inputs from proteins such as growth factors or heterotrimeric guanine nucleotide-binding protein-coupled receptors, also known as G-protein-coupled receptors (G-proteins act as signal transducers for hormones, neurotransmitters and others) [56], [60].

In what concerns their structure, the integrins are composed by α and β subunits [24], [25], [55]–[58]. The β subunit of the integrins comprises between 40 to 70 amino acids, except for the case of the β4

integrin, which possesses 1000 amino acids. The α subunits are generally shorter than the β subunits and tend to have more variety in their isoforms [25]. Currently, there have been found 18 α subunits, 8 β subunits and a total of different 24 $\alpha\beta$ heterodimer combinations have been identified. The intracellular connection of these receptors to the actin cytoskeleton is done via adaptor proteins, such as talin and vinculin [24]. The possible combinations for the heterodimers makes them display overlapping functions, although without redundancy [25].

The transduction of signals along the plasma membrane performed via integrins implies interactions with the extracellular ligands through the extracellular domains of the integrins and, intracellularly, with signaling and cytoskeletal proteins through the cytoplasmic tails of the integrins. The extracellular domains of the integrins are regarded as large, while the intracellular tails are considered to be short [24]. In more detail, the intracellular domain of the integrins, also known as cytoplasmic tails, is where interactions between the integrins and intracellular proteins take place. Both subunits, α and β, are capable of linking to different adaptors. However, there currently is more insight into the species which engage with the β subunit. The extracellular domains of the integrins, also known as ectodomains, are distinct for their capacity to assume different configurations, translating different levels of affinity for ligands. The three principal conformations that the integrins can display are: the bent-closed configuration, which corresponds to a low-affinity arrangement, the extended-closed configuration, for intermediate affinity cases, and extended-open configuration, which is displayed when there is high affinity for ligands. The transition from low affinity configuration to high affinity configuration corresponds to a process of integrin activation. In addition, the integrins have a transmembrane (TM) domain. The TM domains of the integrins are linked to the plasma membrane through non-covalent bonds and serve as a bridge of communication between the ectodomains and the intracellular cytoplasmic tails of the integrins, playing a pivotal role in the transference of configurational modifications which occur in the process of integrin activation [25]. The mechanisms behind integrin activation and signaling comprise different steps, presented in Figure 2.4. Firstly, the cytoplasmic tails of the integrins engage with intracellular adaptors such as talin, in a mechanism called inside-out signaling, which elicits modifications in the configuration of the integrins in order to regulate affinity to extracellular ligands [24], [25], as stated previously in this section. Consequently, the heterodimer interacts with extracellular ligands and links to the force-transducing actin cytoskeleton, leading to complete activation of the integrins and clustering of the integrins – outside-in signaling [24]. These arrangements of the integrins in clusters are denoted as focal adhesions [25]. As a result of the focal adhesions disposition and of the tension caused by internal and external forces, other intracellular binding species are captured, leading to downstream signaling. The inactivation of the integrins happens when endocytic promoters and inhibitory species are engaged [25]. In conclusion, the integrins undergo allosteric conformational modifications in conformity with their functional status [24]. The specific binding proteins with whom the integrins interact with, along with the way that the interactions are controlled mechanical wise, set how integrin signaling is processed and the link to the actin cytoskeleton [25].

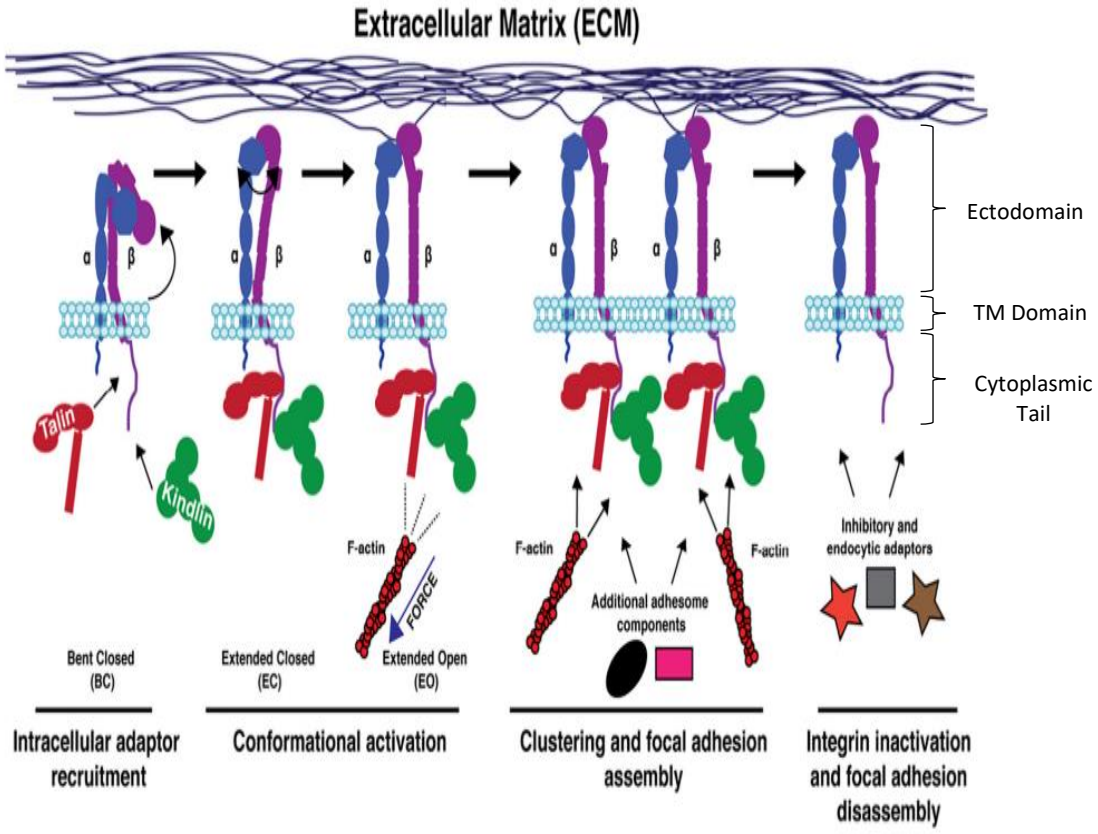

Figure 2.4 – Integrin activation and inactivation process. In the beginning of the activation process, the cytoplasmic tails of the integrins engage with intracellular adaptors, prompting modifications in the configuration of the integrins in order to regulate affinity to extracellular ligand. There is an alteration from bent closed configuration into extended closed configuration. Next, the heterodimer interacts with extracellular ligands and links to the force-transducing actin cytoskeleton, conducting to complete activation of the integrins and to an arrangement of integrin clusters – focal adhesions. The complete activation of the integrins is displayed by a modification in their configuration, particularly from extended closed into extended open configuration. The focal adhesion arrangements are disassociated during the inactivation process of the integrins, which is elicited by the engagement of endocytic promoters and inhibitory species. Adapted from *Kadry & Calderwood*, 2020 *[25]*.

2.3.1 – Crosstalk between the Integrins and the TGF-β Pathway

Mechanotransduction is the concept concerning how cells process mechanical information (from the moment they suffer modifications in the cellular cytoskeleton) due to input sensing, into a specific response [23]. When cells undergo mechanotransductive processes, these changes will most likely reflect in the nucleus, consequently altering how gene expression occurs [61].

There currently are different possibilities that describe the mechanotransductive signaling events which translate the biochemical interaction between the growth factors and the cytoskeleton. Taking into account that numerous signaling molecules take part in both integrin and growth factors signaling [62], the first possibility of interaction consists in the concomitant activation of the same signaling molecules by independent signals, elicited by growth factors and by the integrins. In [46], it is stated that some pathways, such as the MAPK and the PI3K pathway, and the control of the Rho family of guanosine triphosphate-binding (GTP-binding) proteins, the GTPases (which intervene in the

transduction of extracellular signals to the actin cytoskeleton, influencing cell motility [63]–[65]), can be performed through this mechanism. Another possibility of interaction between growth factors and the integrins is the synergistic activation of growth factor-signaling receptors at the focal adhesions. In [46], it is reported that the integrins might gather signaling molecules which build a favorable environment to aid some growth factor receptors, such as the epidermal growth factor receptor (EGFR), the insulin receptor or the vascular endothelial growth factor receptor (VEGFR), in their engagement with downstream signaling species. Another example of a possible synergistic interaction of the integrins with growth factors is the potential physical engagement of the heterodimers with the TGF-β receptors. It is reported that the TGF-β ligand elicits the association of the αvβ3 integrin with the type II TGF-β receptor in pathologies such as lung fibroblasts and breast cancer, which initiates a collaborative type of signaling with the proto-oncogene tyrosine-protein kinase Src (c-Src) [66], and MAPKs [67], [68]. The ability of growth factor stimulation to activate synergistically growth factor-signaling receptors might be a consequence of an arrangement of the receptors as coclusters, at the focal adhesions sites, or it can derive from their association with the actin cytoskeleton [57]. Another form of interaction between the integrins and the growth factor receptors is the activation of the receptors independently of the growth factor ligand. The Epidermal Growth Factor Receptors (EGFR) phosphorylation is a concrete example of this type of interaction since it can be elicited by the integrins without the presence of the Epidermal Growth Factor (EFG) ligand. The influence on the activation of the receptors triggered by the growth factor differs from the impact due to the integrins [46]. Another example of a similar type of interaction is found in [67], where it is stated that the integrins are able to regulate TGF-β signaling indirectly and that the αvβ3 integrin, particularly, modulates the expression of TGF-β receptors, type I TGF-β receptors and type II TGF-β receptors, without dependency of the TGF-β. A different way of interaction between the cytoskeleton and growth factor signaling is the formation of complexes involving common signaling components. Examples of this type of interactions are found in [69], where it is reported that the calponin 1, which is a thin actin-binding protein [70], is able to associate with the Smad or the phosphorylated Smad. Similarly, in [71] it is reported that the cytoskeletal actin-binding protein filamin has the capability of associating with the Smads, thus modulating TGF-β signaling. Other example of this same type of engagement is found in [67], where it is reported that the αvβ3 integrin physically associates with the type II TGF-β receptor and, therefore, controls TGF-β signaling. Also in [67], it is stated that the αvβ5 integrin amplifies TGF-β signaling by also engaging with the type II TGF-β receptor. An additional possibility of interaction between the integrins and growth factors consists in the downstream transcription of the TGF-β and the BMP receptors following the release of some cytoskeletal binding proteins, namely of the globular actin-binding proteins (G-actin-binding proteins) and filamentous actin-binding proteins (F-actin-binding proteins) [72], due to a disturbance in the balance between these proteins associated to cell motility [73].

The importance of having further insight into the crosstalk between the integrins and ligands such as growth factors lies in the potential to lead the development of different therapies or treatment protocols to treat numerous pathologies [67]. In the case of the TGF-β ligand, understanding these interactions promises a new approach to control the growth of tumors and angiogenesis, which concerns the development of new blood vessels from preexisting ones [57], [74]. Previous work performed to assess the crosstalk between pivotal participants of the angiogenesis, including the integrins, encompasses the implementation of one model to map environmental cues to cell phenotypes in [75]. However, in the proposed model from [75], the translation of the different reactions between the participant molecules is not done through ODEs or PDEs but rather by a Boolean signaling network.

CHAPTER 3 – Materials and Methods

3.1 – Computational Modeling

The crosstalk between the physical and chemical cues, delivered by the tissue engineering construct, is difficult to unravel with *in vitro* tools alone, as both pathways always affect each other. Computational models provide an interesting tool to help unravel this crosstalk and improve the design of tissue engineering constructs. A computational model is created to mimic the functioning of a system using previous knowledge regarding the most important components of that system. The process of modeling consists of describing the interaction between state variables through computational algorithms and mathematical relationships [76].

Throughout time, computational models have been proven to be a valuable tool for research and evaluation of molecular mechanisms of biological systems as a whole, as opposed to individual entities. The most specific ones usually consist of in-depth descriptions of reaction mechanisms, through mass action kinetics, and reaction stages are typically represented by ODEs [53]. Values available for state variables constitute input values for the model.

The relevance of computational models lies in their ability to provide outcomes for different values of features or parameters, thus allowing the assessment of input/output relationship to perform virtual experiments and to check hypotheses. In addition, numerous experiments can be carried with the same initial setting, *in silico* observations can be sampled at any time point and any location in respect to *in vitro* experiments, an almost immediate visual depiction and analysis of observations is allowed, and physiological and non-physiological modifications of parameter values within desired extent – from smallest to largest changes – are permitted [76].

3.1.1 – Virtual Cell Software

The Virtual Cell or VCell [77] is a free software for modeling and analysis of cell biological systems. It is available throughout the internet, aimed at scientists from different fields, such as Physics and Biology.

Models with different levels of complexity, ranging from plain models to test one hypothesis to multi-layered models, can be implemented in VCell.

There are two main types of models which can be implemented in VCell: biological and mathematical models. Biological models are called BioModels. When a BioModel is created, the user implements a model using a graphical, web-based Java interface in which the physiology (kinetics, molecular and structural characteristics) and an application (morphology, initial conditions) or more than one application are attributed. An application can be compartmental or spatial (using a 1D, 2D, 3D or analytical defined geometry or a digital image as a reference for the geometry), and it specifies the BioModel as a rule-based, continuous or stochastic model. Once a BioModel is created, depicted in Figure 3.1, its corresponding mathematical description is automatically generated, depicting partial differential equations or ordinary differential equations. In this modality, the user cannot edit the model code.

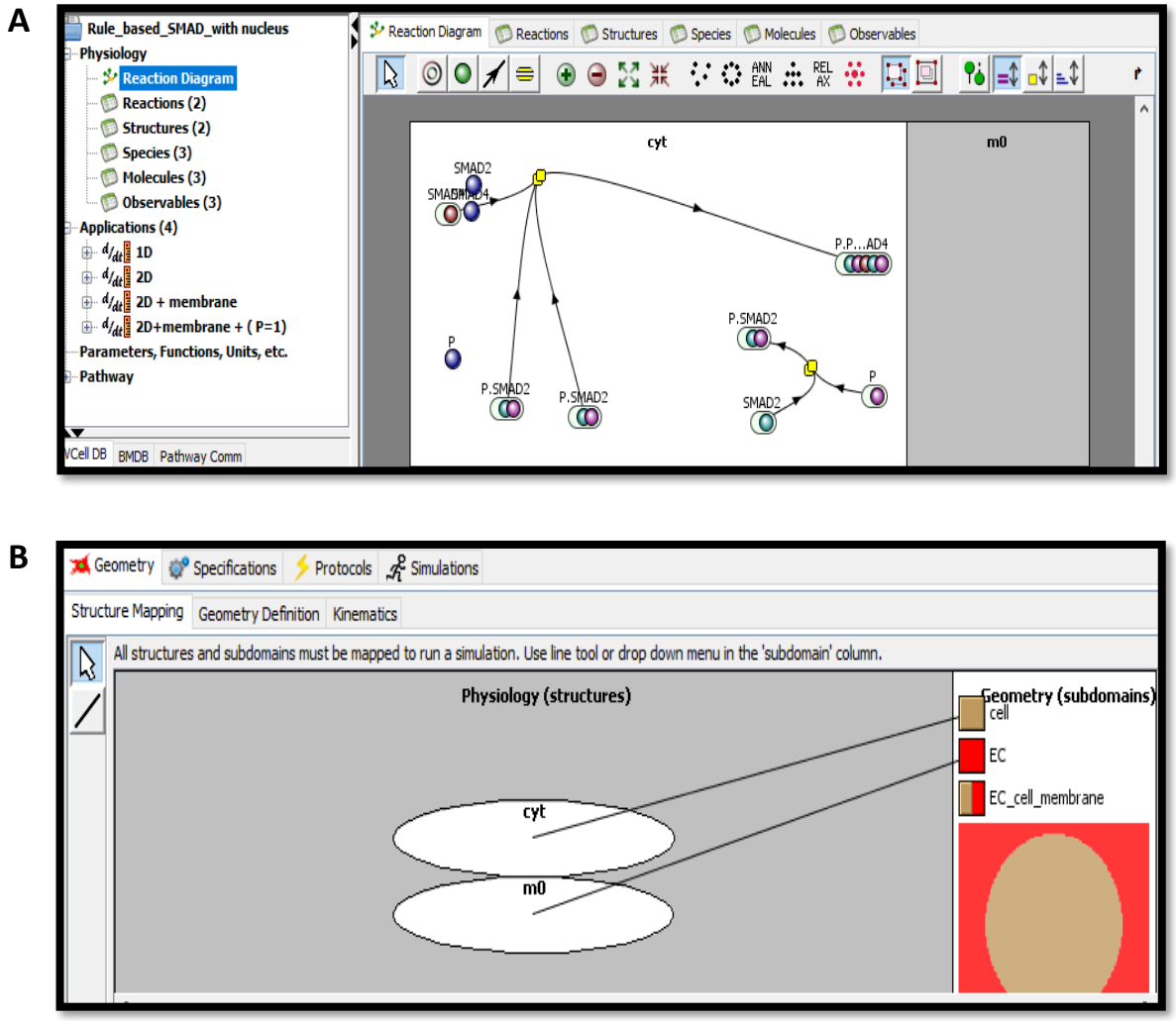

Figure 3.1 – BioModel created on the Virtual Cell software. A – The diagram which translates the system is established by the user through different arrows, geometric shape, symbols and compartments; this environment requires the selection of a physiology and one or more applications. B – A subdomain has to be attributed to each compartment or membrane in the geometry section for each application.

On the other hand, a MathModel, shown in Figure 3.2, allows the user to perform any necessary modifications to the mathematical description of the BioModel or build a model from scratch solely using mathematical knowledge.

The advantages of VCell in comparison to other software for implementation of biological systems are mainly the fact that it automatically generates mathematical encoding for the biological model created by the user via visual interface, as well as for the desired simulations, and the structured presentation and clear separation between the different layers of the model, such as kinetics, geometry, etc. This organization facilitates the process of implementation and promotes the construction of highly rigorous models [78].

The work explained in this document was performed recurring to different MathModels in VCell, version 7.2.0, since numerous simulations were run along the whole process, with few discrepancies from one another. The MathModel modality allows a faster and more direct editing of any part of the system, as opposed to a BioModel, which requires the edition of different characteristics of the system to be completed in different sections of this modality's environment. However, it is worth noting that in the beginning of the implementation of the main model, in order to have a foundational script and build

a more intricate model from there, the chosen modality was the BioModel, since the advantage of automatically generating a mathematical description from a visual representation of the system created by the user is very helpful for individuals who do not yet possess a good level of acquaintance with the software

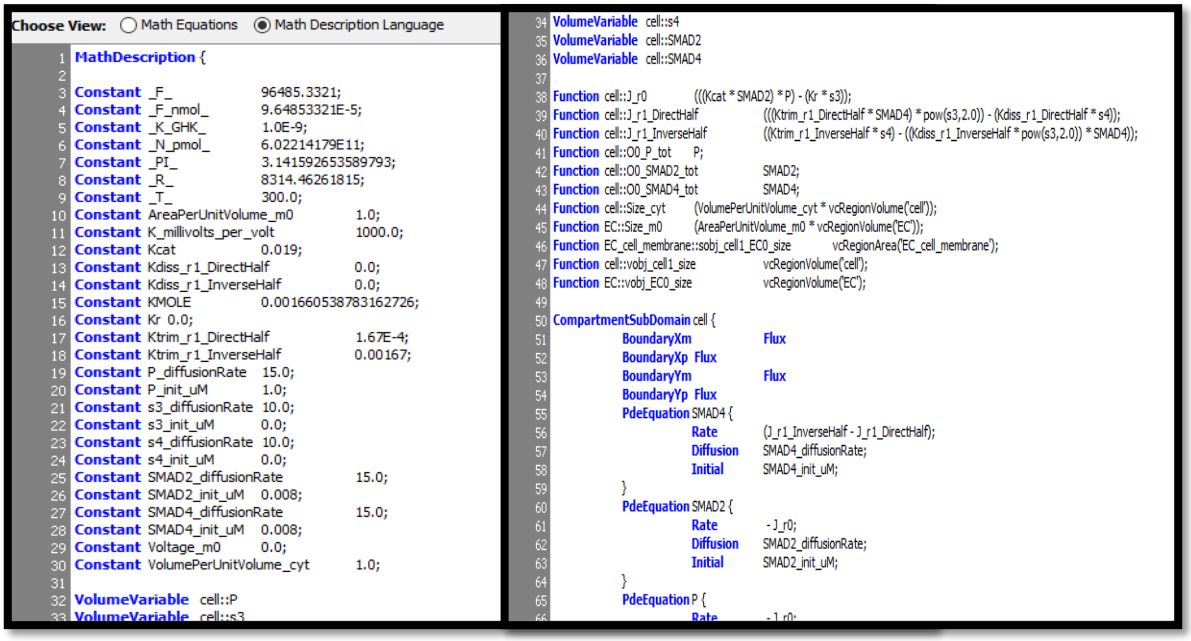

Figure 3.2 – MathModel created on the Virtual Cell software. This modality allows to edit the mathematical description of the system directly in the script, which encompasses constants, variables, functions, boundary conditions and the definition of all of these within the considered compartments.

3.2 – Model Description

3.2.1 – Diagrams and Components

The computational model implemented in this work, consists out of two modules: Module I) The TGF-β signaling pathway computational model and Module II) crosstalk of the TGF-β signaling pathway with the integrins.

Figure 3.3 depicts the reactions and translocations occurring in the implemented computational model of the canonical TGF-β signaling pathway, that is, concerning Module I. In this module, the ligand is assumed to already be bound to its receptor, resulting in a complex – the *C* complex. The *C* complex binds to the SMAD2 in the cytoplasm, in a region near the plasma membrane (it is worth noting that the SMAD2 represents both the Smad2 and Smad3 molecules mentioned in Section **2.2.1 – Dynamics of the TGF-β Signaling Pathway**; these molecules have the same role in the canonical TGF-β signaling pathway). Two molecules of the phosphorylated SMAD2 – pSMAD2- then bind to the co-mediator

SMAD – SMAD4- constituting a complex – the TRIMER. The TRIMER can disassociate into its components – pSMAD2 and SMAD4 – in the cytoplasm, or it can undergo translocation into the nucleus, where it can also be disassociated or it can bind to DNA and prompt genetic expression of certain genes. Besides the TRIMER, every species, except for the *C* complex, are imported into the nucleus at a certain rate. Only SMAD2 and SMAD4 are able to be exported out of the nucleus. The receptors are continuously produced and degraded. In the nucleus, pSMAD2 can also get dephosphorylated. The decision concerning which species should be established in this model as able to move from the cytoplasm into the nucleus and vice-versa was made based on data from literature, compiled in Table B, divided into Table B. 1 and Table B. 2, in **APPENDIX B**. These tables identify which species migrate from one compartment to another according to preexisting models by different authors.

In the created model for Module I, the reactions involving the SMAD2, the pSMAD2, the TRIMER and the SMAD4 are inspired by *Claus et al.*, 2013 [11], while the receptor trafficking section of the model is adapted from *Nicklas & Saiz*, 2013 [2].

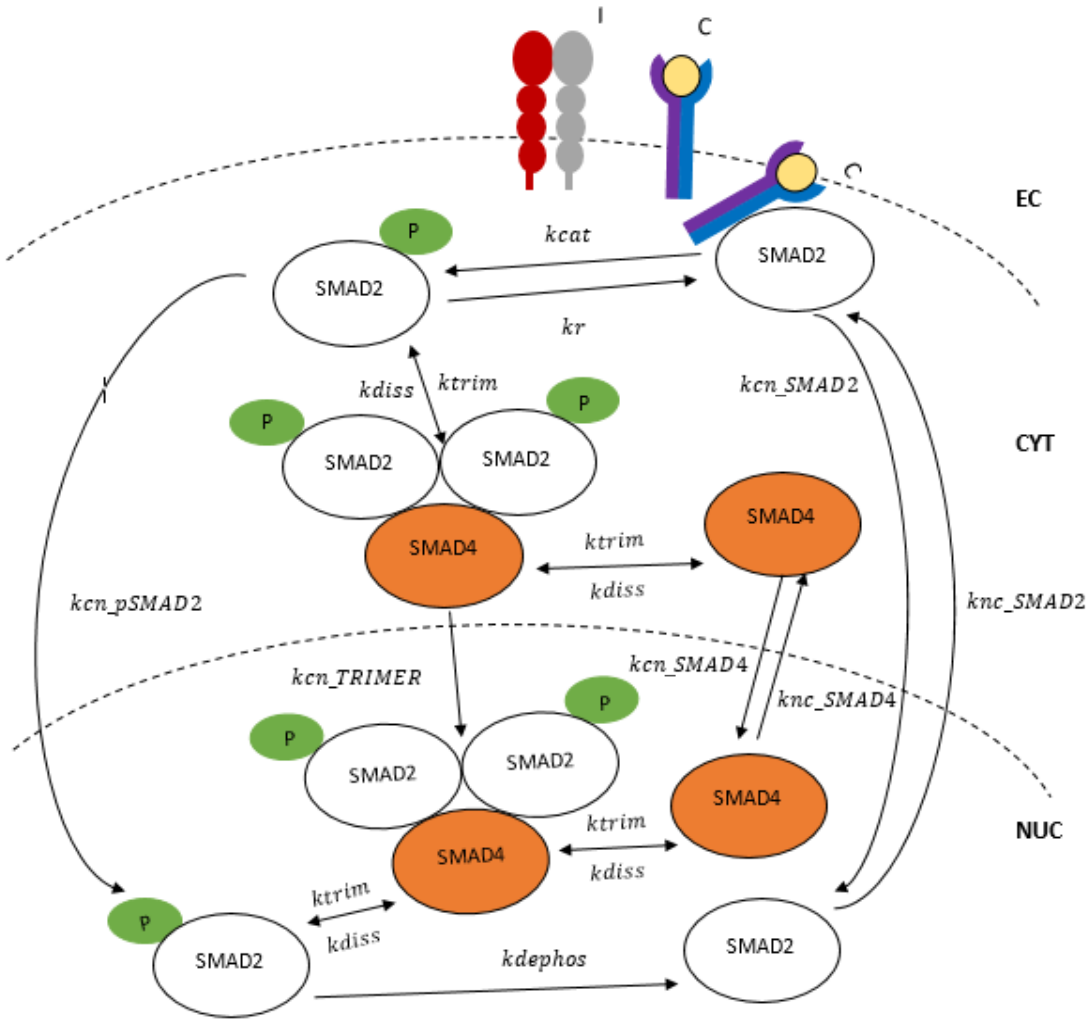

Figure 3.3 – Scheme of the TGF-β signaling pathway. In this model, the cell is surrounded by the extracellular matrix, denoted by *EC*. The integrins, *I*, are able to interact with the extracellular matrix. A complex made up by the TGF-β ligand and the TGF-β receptors, *C* complex, is synthesized near the cytoplasmic membrane and binds to the SMAD2 in the cytoplasm, *CYT*, prompting its phosphorylation, forming the pSMAD2, which can remain in the cytoplasm or migrate into the nucleus, *NUC*. When the pSMAD2 remains in the nucleus, it takes part in the formation of a complex, the TRIMER, made up by two pSMAD2 molecules and one SMAD4 molecule. Similarly, the TRIMER can remain in the nucleus or undergo translocation into the nucleus. When remaining in the nucleus, it can dissociate into its components pSMAD2 and SMAD4. The SMAD4 can also migrate into the nucleus or stay in the cytoplasm. The SMAD2 and the SMAD4 are able to leave the nucleus and undergo translocation into the cytoplasm, while the TRIMER and the pSMAD2 have to dissociate and dephosphorylate, respectively, in order to be possible for these species to leave the nucleus.

Figure 3.4 presents the crosstalk between the integrins and the TGF-β signaling pathway, concerning Module II. This module establishes four different possible ways in which the integrins and the TGF-β signaling pathway are able to interact. The first way, depicted in Figure 3.4.A, consists in the upregulation of the TGF-β receptors elicited by the integrins; another way of interaction, represented in Figure 3.4.B, consists in the increase of the catalytic constant of the phosphorylation of the SMAD2 mediated by the *C* complex, whose synthesis is regulated by the integrins; the third way, depicted in Figure 3.4.C, consists in the stabilization of the TGF-β receptors due to the integrins, causing a decrease in the degradation of the receptors; finally, the fourth interaction, seen in Figure 3.4.D, concerns the formation of a complex made up by the integrins, *I*, and the TGF-β receptors, the *IC* complex, which mediates the phosphorylation of the SMAD2, resulting in an increase of the catalytic constant of the phosphorylation

of the SMAD2 complex. In this model, the integrins are considered to be synthesized and degraded in the cytoplasm. While the production of the *C* complex is dependent on the integrins, the formation and dissociation of the *IC* complex, which occur in the cytoplasm (near the plasma membrane), depend on its components – the integrins and the *C* complex. Since the *IC* complex elicits the phosphorylation of the SMAD2, two species in Module II mediate the phosphorylation of the SMAD2: the *IC* complex and the *C* complex. Just like in the standard model, the *C* complex binds to the SMAD2 in the cytoplasm, near the plasma membrane, eliciting the phosphorylation of the SMAD2. Two molecules of the pSMAD2 bind to the co-mediator SMAD, the SMAD4, constituting a complex – the TRIMER. The TRIMER can disassociate into its components, pSMAD2 and SMAD4, in the cytoplasm, or it can undergo translocation into the nucleus, where it can also be disassociated, or it can bind to DNA and prompt genetic expression of certain genes. Besides the TRIMER, every species, except for the *C* complex, are able to move into the nucleus at a certain rate. Only SMAD2 and SMAD4 are able to be exported out of the nucleus. The receptors are continuously produced and degraded. In the nucleus, pSMAD2 also can be dephosphorylated.

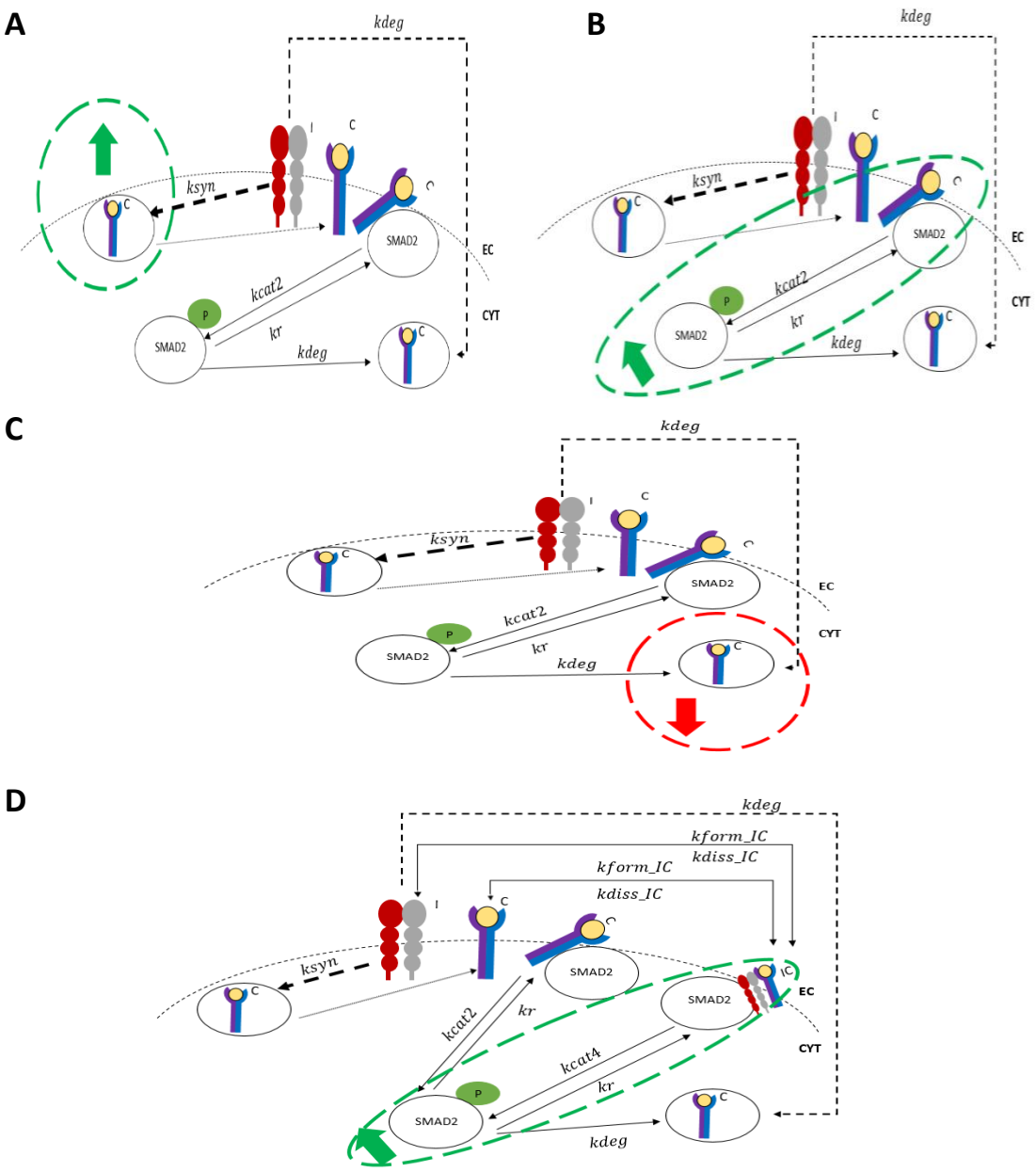

Figure 3.4 – Schematic of the TGF-β signaling pathway crosstalk with the integrins. The thick dashed arrows denote the influence of the integrins, *I*, upon the synthesis of the receptors, the lighter shade of black arrows represent acts of translocation, the thin dashed arrows stand for the influence of the integrins upon the degradation of the *C* complex and the regular arrows translate reactions. The domain and the subdomains are the same from the standard model. There are four different ways of crosstalk with the integrins: (A) the integrins elicit the production of TGF-β receptors – the vesicle which contains the newly synthesized receptors is circled in green, standing next to a green arrow which represents the increase of the synthesis of the receptors; (B) the integrins prompt the synthesis of receptors which take part in the *C* complex and this complex binds to the SMAD2, eliciting more phosphorylation of the SMAD2, denoted by the green ellipse and the green arrow; (C) the integrins elicit the stabilization of the TGF-β receptors, which results in less degradation of the TGF-β receptors – the vesicle which contains the receptors about to be degraded is circled in red, near a red arrow which indicates the decrease of the degradation rate of the receptors; (D) the integrins form a complex with the TGF-β receptors, the *IC* complex, which binds to the SMAD2, eliciting an increase in the phosphorylation of the SMAD2, represented by the green ellipse and the green arrow. All the processes for the TGF-β signaling pathway which take place in the standard model – formation and dissociation of the TRIMER complex in the cytoplasm and in the nucleus, migration of every species into the nucleus, dephosphorylation of the SMAD2 in the nucleus, migration of the SMAD2 and SMAD4 into the cytoplasm- remain valid for this module.

In Table 3.1, all species taken into account during the implementation of the computational model, for Module I and Module II of the work, are presented. Molecules of the SMAD2, pSMAD2, TRIMER complex and SMAD4 in two different compartments, in the cytoplasm and the nucleus, are listed as components in Table 3.1. We assume that, initially, the molecules are uniformly distributed throughout the cytoplasm and the nucleus. Additionally, molecules of the complex made up by the ligand TGF-β and its receptors, the *C* complex, molecules of the integrins, denoted by *I* and, the complex made up by the integrins and the receptors, the *IC* complex, are also presented. The initial concentration of each species, which is also stated in Table 3.1, is mentioned after the designation of the component, followed by its units and literature reference where the value was found.

It is worth noting that even though the initial concentration of the *C* complex in this model is 0.0 μM, there is an activation of the signaling cascade since there is synthesis of this complex right after initial instance set for the model. This same principle is applied in the model for the crosstalk with the integrins, in which the activation of the signaling cascade is only possible because in the first instance there is solely synthesis of *I*, instead of synthesis and degradation.

Table 3.1 – Components which take part in the TGF-β signaling pathway, along with their initial concentration in the cell, respective unit and literature reference, when applicable.

Component	Initial value (μM)	
SMAD2 Cytoplasmic Smad2	0.08	[11]
pSMAD2 Cytoplasmic Phosphorylated Smad2	0.0	[11]
TRIMER Cytoplasmic Phosphorylated Smad2 + Smad4 Complex	0.0	[11]
SMAD4 Cytoplasmic Smad4	0.08	[11]
SMAD2n Nuclear Smad2	0.093	[11]
pSMAD2n Nuclear Phosphorylated Smad2	0.0	[11]
TRIMER_n Nuclear Phosphorylated Smad2 + Smad4 Complex	0.0	[11]
SMAD4n Nuclear Smad4	0.016	[11]
C TGF-β Receptors + TGF-β Complex	0.0	[2]
I Integrins	0.0	
IC Integrins + TGF-β Receptors +TGF-β Complex	0.0	

3.2.2 – Parameters

All the parameters used in the equations which translate the reactions depicted in the diagrams of Module I and Module II (see Sections **3.2.1 – Diagrams and Components** and **3.2.3 – Equations and Boundary Conditions** to observe the diagrams of both stages of the work and the equations, respectively) are presented in Table 3.2 and Table 3.3. Following the name of the parameter, its description,

value and corresponding unit are mentioned, as well as literature source, when applicable. The value of the parameter kr equals 0.0 s^{-1} since in the model, the dephosphorylation of the pSMAD2 is only considered to occur within the nucleus.

The reason behind the selection of the diffusion coefficient of the species *C* and *I* resides in the fact that their value was found to be close to 0.0 $\mu m^2 s^{-1}$ in literature and did not hold an influence upon the results in simulations. A value of 0.049 $\mu m^2 s^{-1}$ for the diffusion coefficient of *C* was used in [79] and for *I*, a diffusion coefficient of 0.01 $\mu m^2 s^{-1}$ was found in [80]. Lastly, the diffusion coefficient of the *IC* complex is set to 0.0 μM. Since there were not found literature sources stating a value for the diffusion coefficient of this complex, by default, it was selected an identical diffusion coefficient to those of *I* and *C*.

Table 3.2 – Parameters used in Modules I and II of the model of the TGF-β signaling pathway – part 1.

Parameter	Description	Value	Unit	Literature
$kcat$	Catalytic rate constant of the phosphorylation of the SMAD2 mediated by the *C* complex	800	s^{-1}	Fitted to experimental data (see Section **4.1 – Parameter Fitting**)
$ktrim$	Formation rate of the TRIMER and TRIMER_n	0.167	$s^{-1}\mu M^{-2}$	[11]
$kdiss$	Dissociation rate of the TRIMER and TRIMER_n	0.00167	s^{-1}	[11]
kcn_SMAD2	Nuclear import rate of the SMAD2	0.7030	$s^{-1}\mu m$	Fitted to experimental data (see Section **4.1 – Parameter Fitting**)
kcn_pSMAD2	Nuclear import rate of the pSMAD2	1.6616	$s^{-1}\mu m$	Fitted to experimental data (see Section **4.1 – Parameter Fitting**)
kcn_TRIMER	Nuclear import rate of the TRIMER	1.1503	$s^{-1}\mu m$	Fitted to experimental data (see Section **4.1 – Parameter Fitting**)
kcn_SMAD4	Nuclear import rate of the SMAD4	0.0447	$s^{-1}\mu m$	Fitted to experimental data (see Section **4.1 – Parameter Fitting**)
knc_SMAD2	Nuclear export rate of the SMAD2n	0.5112	$s^{-1}\mu m$	Fitted to experimental data (see Section **4.1 – Parameter Fitting**)
knc_SMAD4	Nuclear export rate of the SMAD4n	0.0256	$s^{-1}\mu m$	Fitted to experimental data (see Section **4.1 – Parameter Fitting**)
$ksyn$	Synthesis rate of the *C* complex	1.467E-7	$\mu M\ s^{-1}$	[2]
$kdeg$	Degradation rate of the *C* complex	4.63E-4	s^{-1}	[2]
$ksyn\_intg$	Synthesis rate of *I*	1.5E-6	$\mu M\ s^{-1}$	Fitted to experimental data (see Section **4.1 – Parameter Fitting**)
$kdeg\_intg$	Degradation rate of *I*	4.6E-4	s^{-1}	Fitted to experimental data (see Section **4.1 – Parameter Fitting**)
$kform\_IC$	Formation rate of the *IC* complex	0.00146	$s^{-1}\mu M^{-1}$	Fitted to experimental data (see Section **4.1 – Parameter Fitting**)
$kdiss\_IC$	Dissociation rate of the *IC* complex	0.016	s^{-1}	Fitted to experimental data (see Section **4.1 – Parameter Fitting**)
$Imax$	Limit value of the concentration of *I*	0.004	μM	Fitted to experimental data (see Section **4.1 – Parameter Fitting**)
$kdephos$	Dephosphorylation rate of the pSMAD2n	0.08	s^{-1}	Fitted to experimental data (see Section **4.1 – Parameter Fitting**)

Table 3.3 – Parameters used in Modules I and II of the model of the TGF-β signaling pathway – part 2.

Parameter	Description	Value	Unit	Literature
kr	Dephosphorylation rate of the pSMAD2	0.0	s^{-1}	[11]
$SMAD2\_diffusionRate$	Diffusion coefficient of the SMAD2	15.0	$\mu m^2 s^{-1}$	[11]
$pSMAD2\_diffusionRate$	Diffusion coefficient of the pSMAD2	15.0	$\mu m^2 s^{-1}$	[11]
$SMAD4\_diffusionRate$	Diffusion coefficient of the SMAD4	15.0	$\mu m^2 s^{-1}$	[11]
$SMAD2n\_diffusionRate$	Diffusion coefficient of the SMAD2n	15.0	$\mu m^2 s^{-1}$	[11]
$pSMAD2n\_diffusionRate$	Diffusion coefficient of the pSMAD2n	15.0	$\mu m^2 s^{-1}$	[11]
$SMAD4n\_diffusionRate$	Diffusion coefficient of the SMAD4n	15.0	$\mu m^2 s^{-1}$	[11]
$TRIMER\_diffusionRate$	Diffusion coefficient of the TRIMER	1.0	$\mu m^2 s^{-1}$	[11]
$TRIMERn\_diffusionRate$	Diffusion coefficient of the TRIMER_n	1.0	$\mu m^2 s^{-1}$	[11]
$C\_diffusionRate$	Diffusion coefficient of the *C* complex	0.0	$\mu m^2 s^{-1}$	[79]
$I\_diffusionRate$	Diffusion coefficient of *I*	0.0	$\mu m^2 s^{-1}$	[80], [81]
$IC\_diffusionRate$	Diffusion coefficient of the *IC* complex	0.0	$\mu m^2 s^{-1}$	----------
k_{extra_1}	Step of the function for the increase of the synthesis rate of the *C* complex	0.1		Fitted to experimental data (see Section **4.1 – Parameter Fitting**)
k_{extra_2}	Step of the function for the increase of the catalytic constant of the phosphorylation of the SMAD2 mediated by the *C* complex	0.1		Fitted to experimental data (see Section **4.1 – Parameter Fitting**)
k_{extra_3}	Step of the function for the decrease of the degradation rate of the *C* complex	0.1		Fitted to experimental data (see Section **4.1 – Parameter Fitting**)
k_{extra_4}	Step of the function for the increase of the catalytic constant of the phosphorylation of the SMAD2 mediated by the *IC* complex	0.1		Fitted to experimental data (see Section **4.1 – Parameter Fitting**)

3.2.3 – Equations and Boundary Conditions

Equations (3.1) – (3.18) concern the reactions and boundary conditions applicable in the cytoplasm. These equations are reaction-diffusion equations, $x \epsilon\ \Omega_{cyt},\ t \epsilon\ (0, T]$, translating the time and space dependent variation of the concentration of all the species mentioned in Section **3.2.1 – Diagrams and Components**, Table 3.1. Equations (3.1) and (3.2) were inspired by [11] but were adapted to this model by introducing a catalytic reaction term, $kcat2 \cdot SMAD2 \cdot C$, and a dephosphorylation term, $kr \cdot pSMAD2$. Equations (3.5) and (3.6) were also inspired by [11]. Equation (3.8) is based on the diagram for the TGF-β signaling pathway presented in [2] and was introduced to this model in order to include a receptor trafficking section, which is absent in [11]. In Equation (3.1), which describes the variation of the concentration of SMAD2, a term was introduced in order to integrate the phosphorylation of the SMAD2 elicited by the *IC* complex, meaning that the phosphorylation of this species is now under the influence of both the *C* complex, through function $kcat2$, and the *IC* complex, through function $kcat4$. It is worth noting that $kcat2$ is defined from $kcat$ (see Section **3.2.2 – Parameters**, Table 3.2), which stands for the catalytic constant of the phosphorylation of the SMAD2, and is restricted to an area delimited by a circumference of radius 12 μm, in its inner part, and a circumference of radius 15 μm on its outer part; $kcat4$ is also established using $kcat$, being defined within the previously mentioned area. The parameters $kcat2$ and $kcat4$ were both created instead of just using one variable for the equations, reflecting two types of phosphorylation- mediated by the *C* complex and mediated by the *IC* complex – in order to facilitate the comprehension of the script (to different equations are attributed different variables) and allow the eventual interplay, during simulations, with different values for each $kcat$. The variation of the concentration of the pSMAD2, which is represented by Equation (3.2), also reflects the phosphorylation of the SMAD2 by the two previously mentioned complexes. The equation which translates the change in concentration of the *C* complex, Equation (3.8), takes into account the influence of the integrins on the synthesis and degradation of this complex, through functions $ksyn2$ and $kdeg2$, as well as the influence of the formation and dissociation of the *IC* complex. Equations (3.7) and (3.11) translate the variation of the concentration of *I* and of the *IC* complex, respectively. In Equation (3.7), it can be seen that the concentration of the integrins, besides depending on its own synthesis and degradation, is also influenced by the formation and dissociation of the *IC* complex. By its turn, the change in concentration of the *IC*, Equation (3.11), varies according to its formation rate and dissociation rate, taking into account that this complex is made up by the integrins and the *C* complex.

It is worth noting that in the present section, the Laplacian, $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$, is represented by $\triangle$.

$$\partial_t SMAD2 = SMAD2\_diffusionRate \cdot \triangle\, SMAD2 - kcat2 \cdot SMAD2 \cdot C + kr \cdot pSMAD2 - kcat4 \cdot SMAD2 \cdot IC \quad (3.1)$$

$$\partial_t pSMAD2 = pSMAD2\_diffusionRate \cdot \triangle\, pSMAD2 - 2ktrim \cdot pSMAD2^2 \cdot SMAD4 + 2kdiss \cdot TRIMER + kcat2 \cdot SMAD2 \cdot C - kr \cdot pSMAD2 + kcat4 \cdot SMAD2 \cdot IC \quad (3.2)$$

In which $kcat2$ and $kcat4$ are functions:

$$\partial_t kcat2 = kcat \cdot \left(1.0 + \frac{k\_extra2 \cdot I}{Imax}\right) \tag{3.3}$$

$$\partial_t kcat4 = kcat \cdot \left(1.0 + \frac{k\_extra4 \cdot I}{Imax}\right) \tag{3.4}$$

$$\partial_t TRIMER = TRIMER\_diffusionRate \cdot \triangle TRIMER + ktrim \cdot pSMAD2^2 \cdot SMAD4 - kdiss \cdot TRIMER \tag{3.5}$$

$$\partial_t SMAD4 = SMAD4\_diffusionRate \cdot \triangle SMAD4 - ktrim \cdot pSMAD2^2 \cdot SMAD4 + kdiss \cdot TRIMER \tag{3.6}$$

$$\partial_t I = I\_diffusionRate \cdot \triangle I + ksyn\_intg - kdeg\_intg \cdot I + kdiss\_IC \cdot IC - kform\_IC \cdot I \cdot C \tag{3.7}$$

$$\partial_t C = C\_diffusionRate \cdot \triangle C + ksyn2 - kdeg2 \cdot C + kdiss\_IC \cdot IC - kform\_IC \cdot I \cdot C \tag{3.8}$$

In which $ksyn2$ and $kdeg2$ are functions:

$$\partial_t ksyn2 = ksyn \cdot \left(1.0 + \frac{k\_extra1 \cdot I}{Imax}\right) \tag{3.9}$$

$$\partial_t kdeg2 = kdeg \cdot \left(1.0 - \frac{k\_extra3 \cdot I}{Imax}\right) \tag{3.10}$$

$$\partial_t IC = IC\_diffusionRate \cdot \triangle IC + kform\_IC \cdot I \cdot C - kdiss\_IC \cdot IC \tag{3.11}$$

The boundary conditions applicable to the cell membrane, $x \in \Omega_{cyt}$, are represented by Equations (3.12) – (3.18).

$$SMAD2\_diffusionRate \cdot \partial_n SMAD2 = 0 \tag{3.12}$$

$$pSMAD2\_diffusionRate \cdot \partial_n pSMAD2 = 0 \tag{3.13}$$

$$TRIMER\_diffusionRate \cdot \partial_n TRIMER = 0 \tag{3.14}$$

$$SMAD4\_diffusionRate \cdot \partial_n SMAD4 = 0 \tag{3.15}$$

$$C\_diffusionRate \cdot \partial_n C = 0 \tag{3.16}$$

$$I\_diffusionRate \cdot \partial_n I = 0 \tag{3.17}$$

$$IC\_diffusionRate \cdot \partial_n IC = 0 \tag{3.18}$$

In what concerns the reactions which take place in the nucleus and boundary conditions at the nuclear membrane, Equations (3.19) – (3.26) are applicable, $x \in \Omega_{nuc}$, $t \in (0, T]$. The majority of Equations (3.19) – (3.22) are identical to Equations (3.1), (3.2), (3.5) and (3.6), respectively. Equations (3.1) and (3.19) and Equations (3.2) and (3.20) differ among each other since the phosphorylation of the SMAD2 only occurs in the cytoplasm and its dephosphorylation only occurs in the nucleus (the dephosphorylation constant for the pSMAD2 in the cytoplasm, kr, assumes value 0.0 s^{-1}, see Section **3.2.2 – Parameters**, Table 3.3) being this difference represented by the term $kdephos \cdot pSMAD2n$ in Equations (3.19) and (3.20).

$$\partial_t SMAD2n = SMAD2n\_diffusionRate \cdot \triangle SMAD2n + kdephos \cdot pSMAD2n \tag{3.19}$$

$$\begin{aligned} \partial_t pSMAD2n = {} & pSMAD2n\_diffusionRate \cdot \triangle pSMAD2n \\ & - 2ktrim \cdot pSMAD2n^2 \cdot SMAD4n + 2kdiss \cdot TRIMER\_n \\ & - kdephos \cdot pSMAD2n \end{aligned} \tag{3.20}$$

$$\partial_t TRIMER\_n = TRIMERn\_diffusionRate \cdot \triangle TRIMER\_n + ktrim \cdot pSMAD2n^2 \cdot SMAD4n - kdiss \cdot TRIMER\_n \quad (3.21)$$

$$\partial_t SMAD4n = SMAD4n\_diffusionRate \cdot \triangle SMAD4n - ktrim \cdot pSMAD2n^2 \cdot SMAD4 + kdiss \cdot TRIMER\_n \quad (3.22)$$

The boundary conditions for the nuclear membrane $x \in \Omega_{nuc}$ are represented by Equations (3.23) – (3.26). These are also the boundary conditions used in [11].

$$SMAD2\_diffusionRate \cdot \partial_n SMAD2 = -kcn\_SMAD2 \cdot SMAD2 + knc\_SMAD2 \cdot SMAD2n \quad (3.23)$$

$$pSMAD2\_diffusionRate \cdot \partial_n pSMAD2 = -kcn\_pSMAD2 \cdot pSMAD2 \quad (3.24)$$

$$TRIMER\_diffusionRate \cdot \partial_n TRIMER = -kcn\_TRIMER \cdot TRIMER \quad (3.25)$$

$$SMAD4\_diffusionRate \cdot \partial_n SMAD4 = -kcn\_SMAD4 \cdot SMAD4 + knc\_SMAD4 \cdot SMAD4n \quad (3.26)$$

3.2.4 – Geometry

The model presented in this work is constructed in a 2D domain and there are three implemented subdomains: *EC*, *CYT* and *NUC* (Figure 3.5).

The domain of the model has dimensions 40 μm x 40 μm.

The cell used in our standard model (subdomains *CYT* + *NUC*) has a circular geometry, with an area of 707.36 μm^2 (Standard Size) and a radius of approximately 15 μm. The previous dimensions are of the same magnitude order found in [2], [11] and [18].

The nucleus (subdomain *NUC*) is represented by a circle concentric with the circle which delimits the cell, has an area of 78.24 μm^2 and has a radius of approximately 5 μm. The selected value for the area of the nucleus matches the magnitude order of its analogous in [2], [11] and [18].

The activation area, which is the site of the cell in which the SMAD2 can be phosphorylated, is 254.469 μm^2. This area is delimited, on its inner part, by a circle of radius 12 μm, and on its outer part by the cytoplasmic membrane.

The subdomain *EC* represents the extracellular matrix.

The subdomain *CYT r*epresents the cytoplasm in this model.

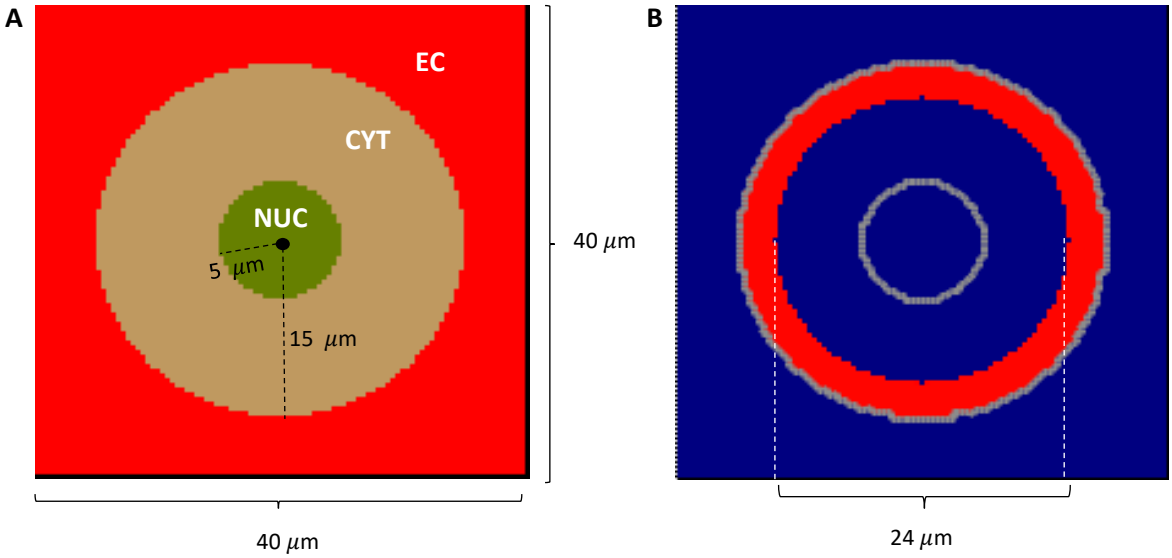

Figure 3.5 – Subdomains of the model and activation ring. A – Geometry of the subdomains; each subdomain is identified by its name in the corresponding region and by its dimensions; B – Activation area for the signaling pathway; the region in red corresponds to the site of the cell in which the SMAD2 can be phosphorylated.

CHAPTER 4 – Results and Discussion

4.1 – Parameter Fitting

In order to obtain the curves for the average concentration of the SMAD2, pSMAD2, TRIMER and SMAD4 (in the nucleus and in the cytoplasm), matching the information presented in the literature, certain parameters were fitted to the available data, obtained via simulations. These parameters are: kcn_SMAD2, knc_SMAD2, kcn_pSMAD2, kcn_TRIMER, kcn_SMAD4, knc_SMAD4, $kcat$ and $kdephos$ (see Table 3.2 and Table 3.3 in Section **3.2.2 – Parameters**, for their final, fitted value). It is important to mention that all the parameters which were introduced for Module II, crosstalk with mechanotransduction, such as $ksyn\_intg$, $kdeg\_intg$, $kform\_IC$, $kdiss\_IC$ and k_{extra_i}, $i = 1, 2, 3, 4$, were all considered to assume value 0.0 at this stage of the project.

In order to perform this process, intracompartmental ratios were determined, depending on the considered species. For the SMAD2, SMAD2n, SMAD4 and SMAD4n, the ratio between the steady-concentration and the initial concentration was calculated, while for the pSMAD2, pSMAD2n, the TRIMER and the TRIMER_n, the ratio between the steady-state concentration and the peak concentration was determined (since the initial value for these species is 0.0 μM). The ratios previously mentioned were compared to their analogous ratios in *Claus et al.*, 2013 [11] and *Nicklas & Saiz*, 2013 [2]. The results are presented in Figure 4.1. The grey color plot is not represented for the TRIMER, SMAD4, SMAD2n and TRIMER_n because there is not data in *Nicklas & Saiz*, 2013 [2] which allows to determine the value of the intracompartmental ratios.

In a second stage of the fitting process, the intercompartmental ratios between the steady-state concentrations for each species (e.g.: $TRIMER/TRIMER\_n$) were determined for the implemented model and for the results presented in *Claus et al.*, 2013 [11] and *Nicklas & Saiz*, 2013 [2]. The bar plots for these ratios are presented in Figure 4.2. The grey color plot is only represented for the SMAD2 because there is not data in *Nicklas & Saiz*, 2013 [2] which allows to determine the intercompartmental ratio values for the pSMAD2, the SMAD4 and the TRIMER.

As it can be inferred by analyzing Figure 4.1 and Figure 4.2, the ratios determined for each species, for the literature and the simulations, remain within the same order of magnitude, which reveals that the estimated values for the parameters are acceptable approximations. However, the bar plots in Figure 4.1 and Figure 4.2 indicate that there is a greater discrepancy between the ratios obtained via simulations and the ratios calculated for the results in *Nicklas & Saiz*, 2013 [2] than between the ratios for the results of this model and the values determined for *Claus et al.*, 2013 [11]. This finding is most likely due to the fact that the model in *Nicklas & Saiz*, 2013 [2] is more intricate than the one presented in *Claus et al.*, 2013 [11], since it includes various mechanisms of the pathway which are absent in the proposed computational model (e.g. negative feedback loop).

It is worth noting that besides comparing the model ratios with data from *Claus et al*., 2013 [11] and *Nicklas & Saiz*, 2013 [2], which are the articles that inspired the basis of the model presented in this document, data from other sources of literature were consulted in order to make sure that these ratios are indeed acceptable approximations, namely [18] and [82].

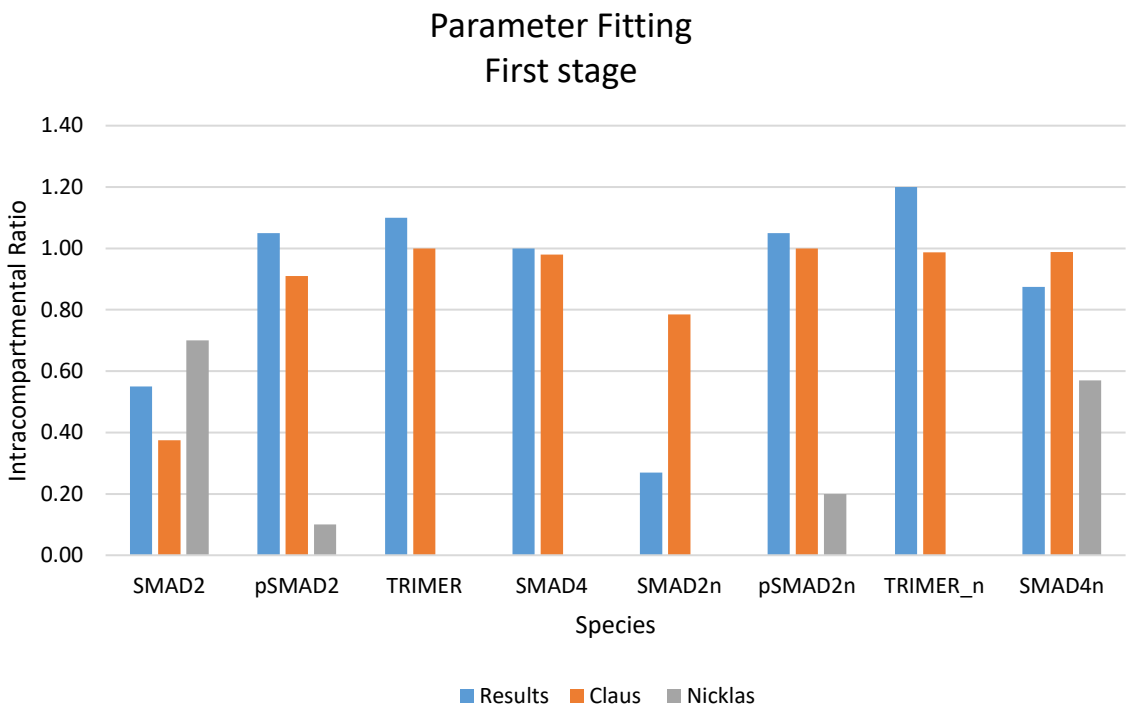

Figure 4.1 – Outcomes of the first stage of parameter fitting – intracompartmental ratios. The legend label "Results" stands for the determined ratios for each species based on simulations performed during the course of this project – ratio between the steady-concentration and the initial concentration for the SMAD2, SMAD2n, SMAD4 and SMAD4n and ratio between the steady-state concentration and the peak concentration for the pSMAD2, pSMAD2n, the TRIMER and the TRIMER_n. The legend designations Nicklas" and "Claus" are attributed to the color plots which translate the value of these same ratios for the results presented in literature – *Nicklas & Saiz,* 2013 *[2]*, and *Claus et al*, 2013 *[11]*, respectively. The bar plots for "Results" are presented in blue, the bar plots for "Claus" are presented in orange and the bar plots for "Nicklas" are presented in grey. For some species, there is not a grey bar – ratios for the results in *Nicklas & Saiz*, 2013 *[2]* – because there is no available data to determine those values in this reference.

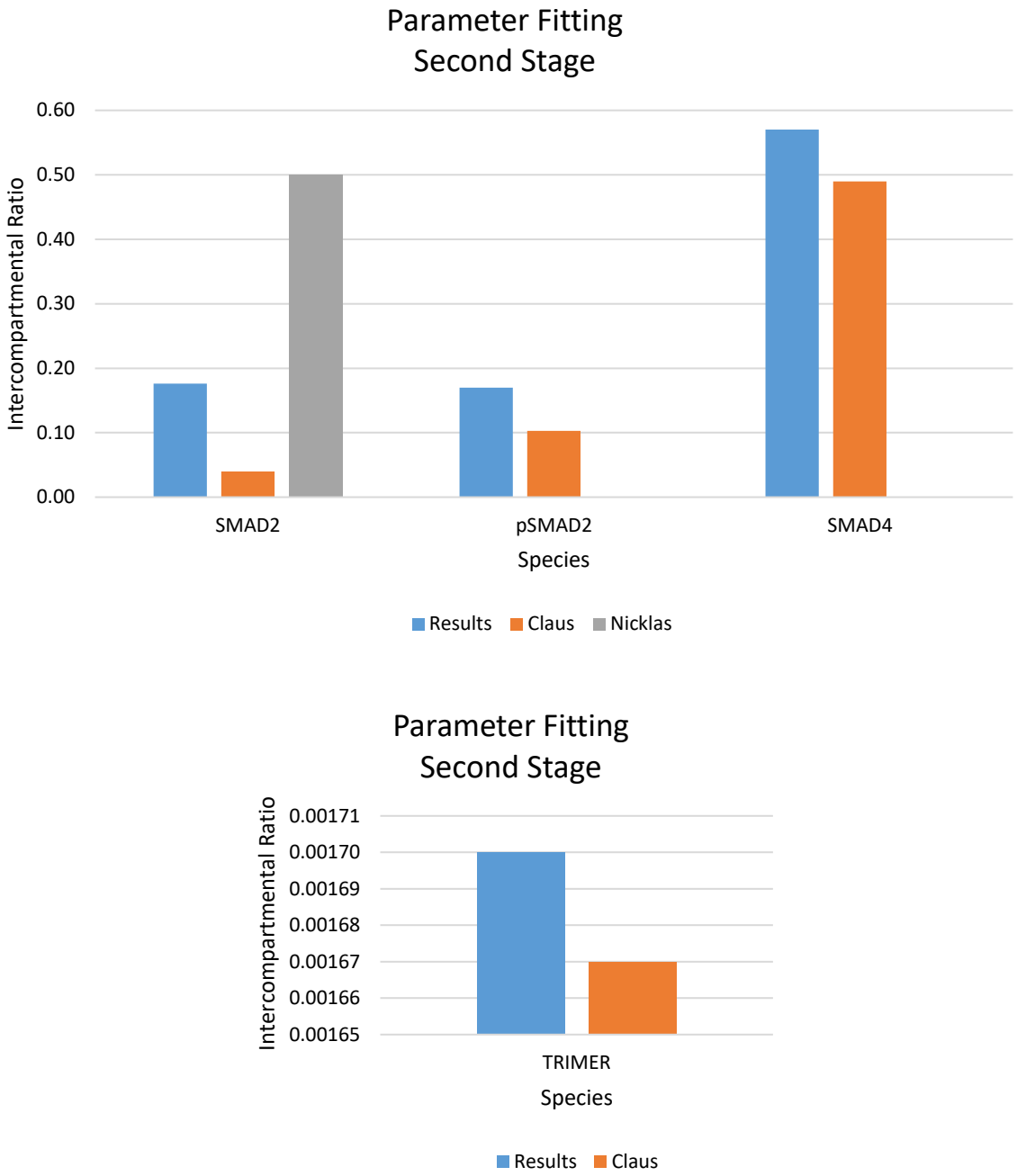

Figure 4.2 – Outcomes of the second stage of parameter fitting – intercompartmental ratios. The legend label "Results" stands for the determined intercompartmental ratios for the SMAD2, pSMAD2, SMAD4 and TRIMER. The legend designations "Nicklas" and "Claus" are attributed to the color plots which translate the value of these same ratios for the results presented in literature – *Nicklas & Saiz,* 2013 *[2]*, and *Claus et al*, 2013 *[11]*, respectively. The bar plots for "Results" are presented in blue, the bar plots for "Claus" are presented in orange and the bar plots for "Nicklas" are presented in grey. For some species, there is not a grey bar – ratios for the results in *Nicklas & Saiz,* 2013 *[2]* – because there is no available data to determine those values in this reference.

For Module II of this work, the parameters $ksyn\_intg$, $kdeg\_intg$, $kform\_IC$, $kdiss\_IC$ and k_{extra_i}, $i = 1, 2, 3, 4$, were fitted to the data provided by simulations. Their final values are included in Section **3.2.2 – Parameters**, Table 3.2 and Table 3.3. It is worth mentioning that most of these parameters do not have value attributed to them in literature. Therefore, the parameters $ksyn\_intg$, $kdeg\_intg$, $kform\_IC$ and $kdiss\_IC$ initially assumed similar values to other parameters pre-included

in the model, namely in Module I: $ksyn\_intg$ assumed a similar value to $ksyn$, $kdeg\_intg$ assumed an identical value to $kdeg$, $kform\_IC$ assumed a similar value to $ktrim$ and $kdiss\_IC$ assumed an identical value to $kdiss$. From there, the value of these parameters was adapted via different simulations based on restrictions placed by the remaining parameters: $Imax$ and k_{extra_i}, $i = 1, 2, 3, 4$. The value of $Imax$ was selected based on the maximum value of *I* for different simulations, with different values of k_{extra_i}, $i = 1, 2, 3, 4$. This value had to be high enough to always top the highest reached concentrations of *I* (see Section **3.2.3 – Equations and Boundary Conditions**, Equations (3.3), (3.4), (3.9) and (3.10)). Lastly, for all k_{extra_i}, $i = 1, 2, 3, 4$, different simulations ran in order to obtain a relatively different outcome for the TRIMER_n in relation to the model results concerning Module I alone. These parameters had to necessarily assume a value between 0 and 1 (see Section **3.2.3 – Equations and Boundary Conditions**, Equations (3.3), (3.4), (3.9) and (3.10)).

4.2 – "Sanity Checks"

In order to check the model for mass conservation and consistency, some simulations were run. These sanity checks act as tool to understand if each parameter is acting according to what is known in literature and serves its purpose in the model, therefore displaying if the model is well constructed or not.

In order to assess the validity of the implemented model for the TGF-β signaling pathway, Equations (4.1) and (4.2) had to be satisfied. Equations (4.1) and (4.2) are mass-conservation equations based on the scheme of the TGF-β signaling pathway (see Section **3.2.1 – Diagrams and Components**, Figure 3.3).

$$SMAD2 + SMAD2n + pSMAD2 + pSMAD2n + 2 \cdot TRIMER + 2 \cdot TRIMER\_n = constant\ 1 \quad (4.1)$$

$$SMAD4 + SMAD4n + TRIMER + TRIMER\_n = constant\ 2 \quad (4.2)$$

When substituting each species in Equations (4.1) and (4.2) by their respective number of molecules in each time point of the simulation, the value in the right member of the equations should remain the same. This requirement was verified with the implemented model. The value for $constant\ 1$ remained equal to 7410.19 throughout all the time points, while the value for $constant\ 2$ remained equal to 3783.45. Therefore, it can be concluded that the model respects mass-conservation.

4.3 – Module I Results

The results of the different set of experiments performed to assess the influence of cell morphology upon the TGF-β signaling pathway. The generated basis script for this module of the project is displayed in Section **Module I Script** from **APPENDIX C**. It is noteworthy that according to the objective of the experiment, pertinent adaptations had to be made to the basis code for each simulation, such as setting a different geometry for the cell or different cell dimensions from the default values stated in Section **3.2.4 – Geometry**.

One should bear in mind that in Module I of the model, all parameters related to the crosstalk between the TGF-β signaling pathway and the integrins are set to 0.0, namely: $Imax$, $ksyn\_intg$, $kdeg\_intg$, $kform\_IC$, $kdiss\_IC$ and k_{extra_i}, $i = 1, 2, 3, 4$.

4.3.1 – The Influence of Cell Shape on the Nuclear Concentration of the Species

Numerous spatial and non-spatial computational models of TGF-β signaling have been presented by different authors but almost all of them do not consider cell morphology, which is typically considered to be a very relevant measure of how cells respond to external mechanical cues. Therefore, some experiments were set in order to assess the impact of cell shape on the nuclear concentration of species and understand the influence of this morphologic feature at a downstream level of the signaling cascade. The most relevant species to assess the effects of cell morphology on the TGF-β pathway is the nuclear TRIMER, TRIMER_n, since this species is able to bind to DNA and, thus, induce genetic expression. However, the other species are also included in this experiment to constitute a term of comparison and aid in the comprehension of the results.

Simulations ran for different cell geometries, maintaining the same area for the cell, nucleus and area of activation of the standard cell (see Section **3.2.4 – Geometry**) to assess the influence of cell shape upon the concentration of the species in the nucleus. The different geometries used in the simulations are presented in Figure 4.3. The standard geometry is the circle. The added geometries include two ellipses, ellipse 1 and ellipse 2, one rectangle and one square. The difference between ellipse 1 and ellipse 2 relies on the width/height ratio – the ellipse 2 has a higher value for this ratio.

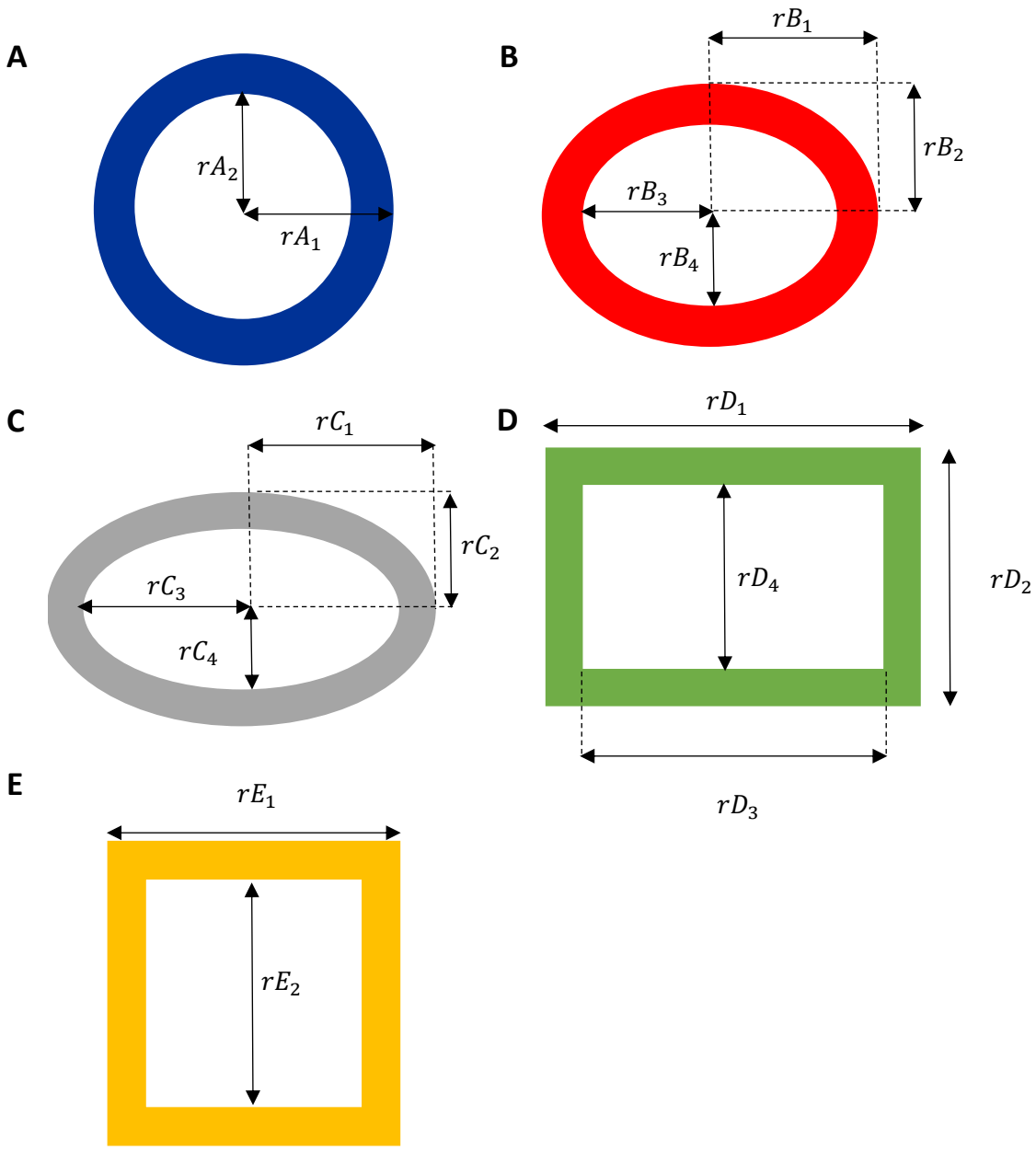

Figure 4.3 – Geometries of the cell used in the experiments to assess the influence of cell shape on the nuclear concentration of the species. The area in which the SMAD2 is phosphorylated, activation area, is highlighted in color. A – Circle geometry; this is the standard geometry of the cell; radius rA_1 is the cell radius and rA_2 is the radius of the circumference which delimits the area without activation of SMAD2. B – Ellipse 1 geometry; $rB_1$and $rB_2$are the half-width and half-height of the cell, while $rB_3$and rB_4 are their analogous for the ellipse which delimits the non-activation area; when comparing the two ellipses established as cell shape, aside from the circle, this is the ellipse with a lower width/height ratio. C – Ellipse 2 geometry; rC_1 and rC_2 are the half-width and half-height of the cell, while rC_3 and rC_4 are their analogous for the ellipse which delimits the non-activation area; when comparing the two ellipses established as cell shape, aside from the circle, this is the ellipse with a higher width/height ratio. D – Rectangle geometry; rD_1 and rD_2 are the width and height of the cell, while rD_3 and rD_4 are their analogous for the rectangle which delimits the non-activation area. E – Square geometry; $rE_1$is the width of each side of the cell, while rE_2 is its analogous for the non-activation area.

The concentration of each species in the nucleus was compared for the different cell geometries after running simulations. The outcomes are presented in Figure 4.4.

In order to easily compare the results for the average concentration of each species in the nucleus across the different cell shapes, the results are presented in the form of a ratio: (Species steady-state

concentration in a given **S**hape)/(Species steady-state concentration in the **C**ircle) – **SC** ratio. As we can see in Figure 4.4, the differences between the values of the **SC r**atio for each species, across cell shapes, are smaller for the SMAD2n and SMAD4n than for the pSMAD2n and TRIMER_n. It is worth noting that the fact that each of the bar plots presented in Figure 4.4 corresponds to a simulation belonging to a different MathModel (see Section **3.1.1 – Virtual Cell Software**) might have contributed to a certain extent to these differences. The establishment of a new geometry for the cell in each MathModel resulted in slight variations for the area of the nucleus due to resolution limitations of the software. In what concerns the pSMAD2n, the bigger differences in the value of the SC ratio of this species throughout the different cell shapes, in comparison to the SMAD2n and the SMAD4n, can possibly be related with the fact that the equation which translates the variation of concentration of the pSMAD2n, Equation (3.20), is the equation that includes more terms. Therefore, any slight variation in the value of the concentration for the other species, contributed to the exacerbation of the variation of the pSMAD2n concentration across distinct geometries.

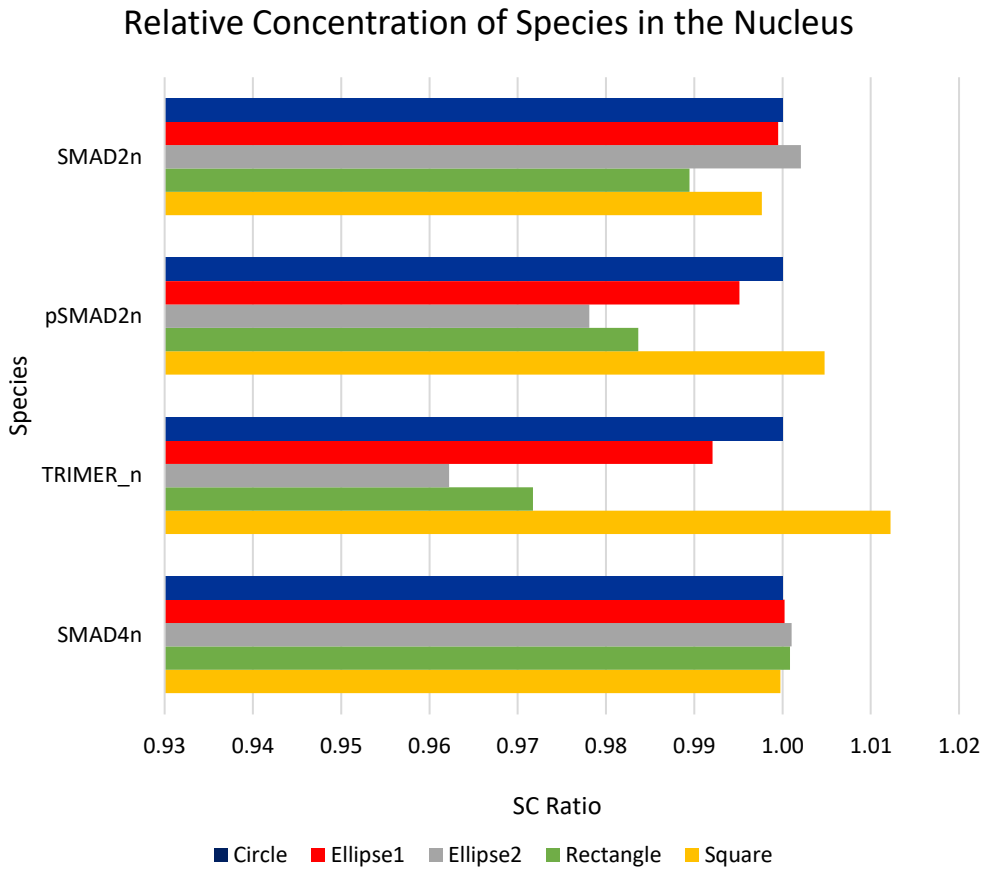

Figure 4.4 – Relative concentration of species in the nucleus, for each cell geometry. The bar plots correspond to the ratio between the concentration of a species in a given shape and its concentration in the circle – SC ratio – for different cell geometries: Circle, represented in blue; Ellipse 1, represented in red; Ellipse 2, represented in grey; Rectangle, represented in green; Square, represented in yellow. The different species, namely the SMAD2n, pSMAD2n, TRIMER_n and SMAD4n, are presented in the vertical axis.

4.3.2 – The Influence of the Width/Height Ratio on the Nuclear Concentration of Species

In order to study more in depth the influence of the cell shape on the concentration of the nuclear species of the TGF-β (briefly assessed in the previous section of results, Section **4.3.1 – The Influence of Cell Shape on the Nuclear Concentration of the Species**) an experiment was carried for different width/height ratio values of the cell – within a range from 1 to 5. The experiment encompasses both rectangular and elliptic shapes of cells. The results are presented in Figure 4.5, Figure 4.6, Figure 4.7 and Figure 4.8, for the SMAD2n, pSMAD2n, TRIMER_n and SMAD4n, respectively. Each figure shows a scatter plot for the SC ratio dependent on the width/height ratio of the cell, of both elliptical and rectangular cell shapes. In addition, a simple linear regression model was applied in each case; therefore, each figure displays the trend line of each scatter plot and the adjusted coefficient of determination, denoted by R^2_{adj}. The R^2_{adj} was selected to measure the quality of each linear regression and aid in the evaluation of the results. The coefficient of determination R^2 is sometimes used to assess if the model in question is of good quality, defined as the percentage of the variation that can be explained by the regression equation. However, the value of the variable R^2 always increases as new variables are added to the model, which does not happen with R^2_{adj}. The value of R^2_{adj} only increases if the new variable provides, in fact, a better fit to data. Bearing in mind that additional experiments can be performed following the work presented in this document, it is pertinent to use the R^2_{adj} instead of the R^2 to evaluate the quality of the linear regressions since it supplies information concerning the new variable to add to the model, indicating if it provides a better fit to the data. The value of the R^2_{adj} is comprised between 0 and 1 and if it is close to 1, that means that the independent variable is a good linear predictor of the dependent variable [83].

R^2_{adj} is obtained through Equation (4.3):

$$R^2_{adj} = 1 - \left[(1 - R^2)\left(\frac{n-1}{n-p-1}\right)\right] \tag{4.3}$$

In which n denotes the sample size, p denotes the number of independent variables excluding the constant and R^2 represents the coefficient of determination, which is obtained with Equation (4.4):

$$R^2 = \frac{SSR}{SST} \tag{4.4}$$

In which SSR represents the sum of squares $\Sigma\left(\widehat{y_i} - \bar{y}\right)^2$, SST is the total sum of squares $\Sigma(y_i - \bar{y})^2$, i is the number of the measurement and Y is the dependent variable. In this particular case, i is the width/height ratio number i, n =8, p=1, the width/height ratio is the independent variable and the dependent variable is the SC ratio, substituting variable Y in Equations (4.3) and (4.4).

By looking at Figure 4.5, it is noticeable that, in general, when taking into account a certain value of the width/height ratio, the SC ratio has a higher value for an elliptical cell shape than for a rectangular one. However, when the width/height ratio is approximately 5.2, the regression lines intersect and the SC ratio becomes higher for the rectangular cell shape. It is also immediately noticeable that the slopes of the regression lines for both cell shapes are negative, namely -0.0029 for the ellipse and -0.0012 for the rectangle. These slope values indicate that as the width/height ratio increases, the SC ratio for the SMAD2n decreases. In addition, since the absolute value of the regression line slope for the ellipse is higher than the absolute value of the regression line slope for the rectangle, that means that for each additional unit of the width/height ratio there is a higher decrease of the SC ratio for the ellipse, indicating that an elliptical cell shape induces less concentration of the SMAD2n as the width/height ratio increases. It is worth noting that the R^2_{adj} for both cases, for the ellipse and for the rectangle, is considerably below 1 – 0.537 and 0.102, respectively. These values show that for the SMAD2n, the width/height ratio is a mediocre linear predictor to the SC ratio, in the case of the elliptical cell shape, and not a good linear predictor to the SC ratio, in the case of the rectangle. This means that there is a big dispersion of the observations for the SC ratio, especially for the rectangular cell shape.

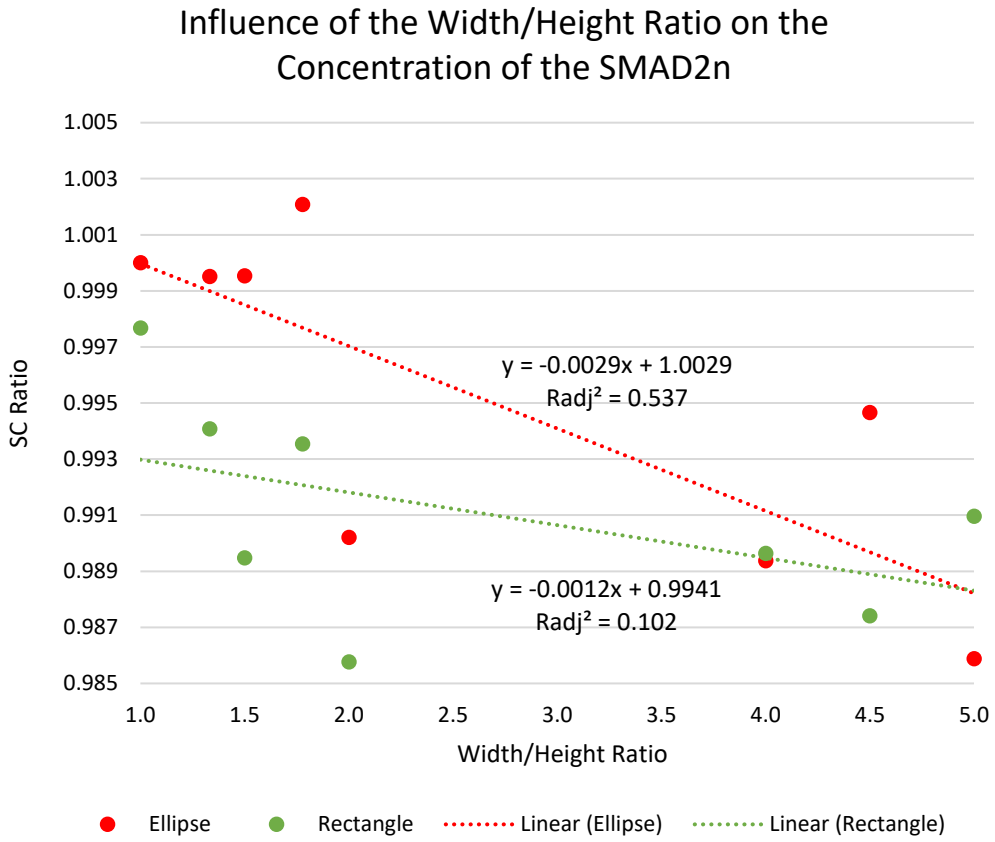

Figure 4.5 – Influence of the width/height ratio on the concentration of the SMAD2n measured by the SC ratio. The figure displays a scatter plot for different cell shapes – for the ellipse, represented by red dots, and for the rectangle, represented by green dots. The respective scatter plot regression lines are displayed by dashed lines of the same color. The trend line equations and respective adjusted R-squared values are also shown.

When analyzing Figure 4.6 and Figure 4.7, concerning the pSMAD2 and the TRIMER_n, respectively, it can be observed that for both sets of results there is a decrease in value for the SC ratio as the width/height ratio increases, which is corroborated by the negative slope of every regression line. In the case of the pSMAD2n, the slope of the regression line for the elliptical cell shape is -0.0332 and for the rectangular cell shape it is -0.0451, as seen in Figure 4.6. Initially, when considering the same width/height ratio for both cell shapes, the SC ratio is superior for the rectangular cell. However, as the width/height ratio approaches a value of 2.4, the regression lines of both cell shapes intersect and the SC ratio becomes higher for the elliptical cell shape. In the case of the TRIMER_n, the slope of the regression line for the elliptical cell shape is -0.0501 and for the rectangular shape it is -0.0685, as depicted in Figure 4.7. Similar to the results concerning the pSMAD2n, for the TRIMER_n, when considering the same width/height ratio for both cell shapes, the SC ratio is superior for the rectangular cell at first but around a width/height ratio of 2.6, the SC ratio becomes higher for the elliptical cell. The slope values for the regression lines concerning the pSMAD2n and the TRIMER_n (previously mentioned in this paragraph) also indicate that for the pSMAD2n and the TRIMER_n, there is a higher decrease of the SC ratio by additional unit of the width/height ratio for a rectangular cell shape than for an elliptical shape, since the absolute value of the slopes is superior when considering a rectangular cell shape. In what concerns the quality of the model, the values assumed by R^2_{adj} are near 1 for the pSMAD2n and the TRIMER_n. For the pSMAD2n, the variable assumes values of 0.896 and 0.985,

when considering an elliptical cell shape and a rectangular cell shape, respectively. For the TRIMER_n, R^2_{adj} has values of 0.877 and 0.982 for an elliptical cell shape and a rectangular cell shape, respectively. Since these values of R^2_{adj} are near 1, the linear regression models are shown to be good models for the considered dependent and independent variables. It can be concluded that for the pSMAD2n and the TRIMER_n, the width/height ratio is a good linear predictor of the SC ratio. Lastly, the similar slopes and R^2_{adj} values of the regression lines of each cell shape for both species, pSMAD2n and TRIMER_n, display the similarity in the Equations which represent the variation of the concentration of the pSMAD2n and the TRIMER_n, Equations (3.20) and (3.21) since both of these equations have the same number of quadratic terms.

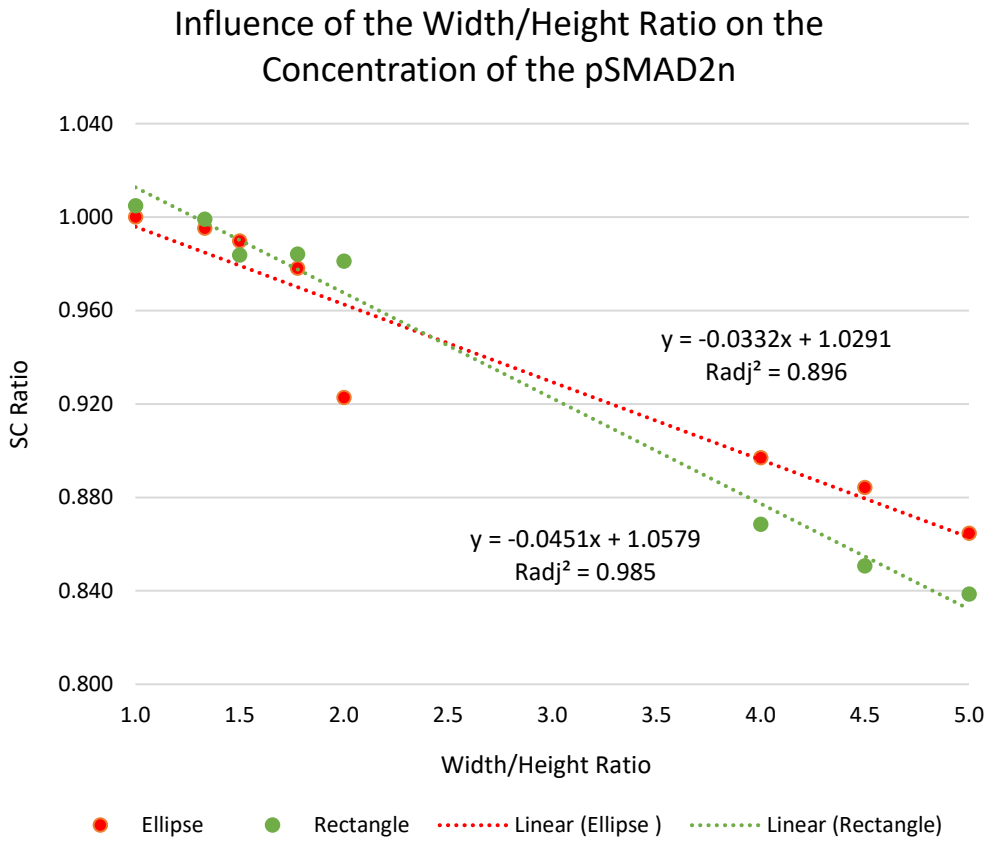

Figure 4.6 – Influence of the width/height ratio on the concentration of the pSMAD2n measured by the SC ratio. The figure displays a scatter plot for different cell shapes – for the ellipse, represented by red dots, and for the rectangle, represented by green dots. The respective scatter plot regression lines are displayed by dashed lines of the same color. The trend line equations and respective adjusted R-squared values are also shown.

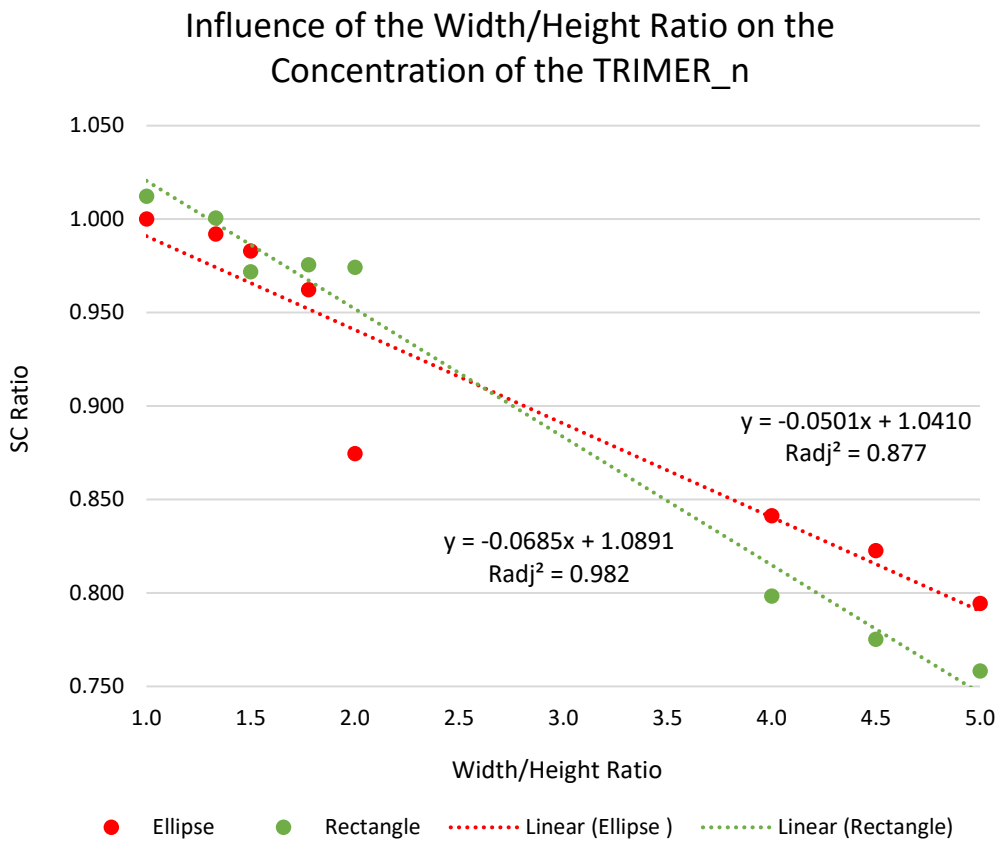

Figure 4.7 – Influence of the width/height ratio on the concentration of the TRIMER_n measured by the SC ratio. The figure displays a scatter plot for different cell shapes – for the ellipse, represented by red dots, and for the rectangle, represented by green dots. The respective scatter plot regression lines are displayed by dashed lines of the same color. The trend line equations and respective adjusted R-squared values are also shown.

In Figure 4.8, which concerns the SMAD4n, there is an increase in the SC ratio as the width/height ratio increases, for both cell shapes: elliptical and rectangular. This finding is corroborated by the positive slopes of the regression lines for each cell shape, assuming a value of 0.0014 for an elliptical cell shape and a value of 0.0019 for a rectangular cell shape. Initially, the SC ratio is higher for the elliptical cell shape. However, as the width/height ratio approaches a value of 2.6, the regression lines intersect and, then, the SC ratio becomes higher when considering the rectangular cell shape. Considering the SMAD4n, the slope values of the regression lines for each cell shape indicate that there is an increase by 0.0014 in the SC ratio by unit of the width/height ratio for an elliptical cell shape, while there is an increase by 0.0019 in the SC ratio by unit of the width/height ratio for a rectangular cell shape. This finding indicates that a rectangular cell shape promotes higher concentration of the SMAD4n as the width/height ratio increases, in comparison with an elliptical cell shape. The R^2_{adj} values obtained for the SMAD4 are 0.982 for the rectangle and 0.840 for the ellipse. Since these values are high values of R^2_{adj}, the linear regression models are shown to be of good quality – the width/height ratio is a good linear predictor for the SC ratio, when considering the SMAD4n. It is worth noting that the determined values of R^2_{adj} for both cell shapes in the case of the SMAD4n, are similar to the values of R^2_{adj} for the pSMAD2n and the TRIMER_n (see Figure 4.6 and Figure 4.7). The similarities between these values are most likely due to the fact that the equation which translates the change in concentration of the

SMAD4n, Equation (3.22), has the same number of quadratic terms as the equations which translate the variation in concentration of the pSMAD2n and the TRIMER_n – Equations (3.20) and (3.21), respectively. In addition, it is worth noting that the increase of the SC ratio as the width/height ratio grows in both cell shapes, considering the SMAD4n, was expected. This expectation resides in the fact that for the TRIMER_n, there is a decrease in the SC ratio as the width/height ratio increases and the Equations which translate the change in concentration of both of this species – Equation (3.21) for the TRIMER_n and Equation (3.22) for the SMAD4n – are symmetrical (apart for their first term).

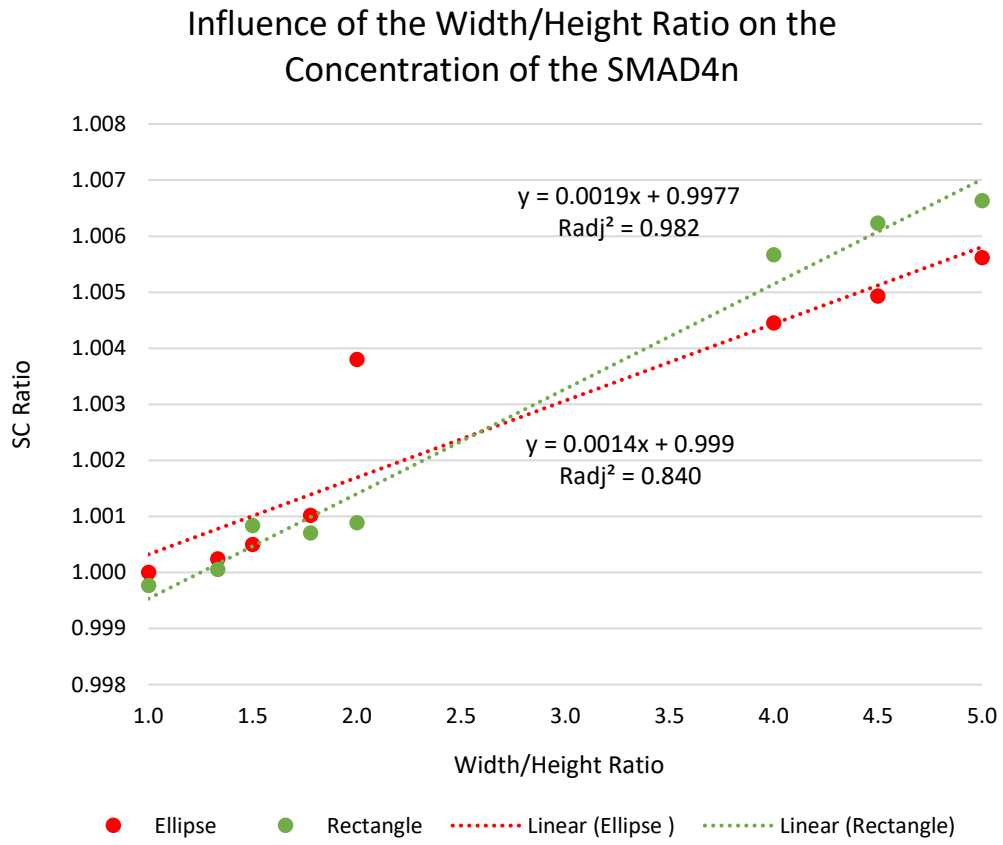

Figure 4.8 – Influence of the width/height ratio on the concentration of the SMAD4n measured by the SC ratio. The figure displays a scatter plot for different cell shapes – for the ellipse, represented by red dots, and for the rectangle, represented by green dots. The respective scatter plot regression lines are displayed by dashed lines of the same color. The trend line equations and respective adjusted R-squared values are also shown.

When comparing the presented results for every species among each other, it can be concluded that for an elliptical cell shape, the lowest absolute value for a regression line slope is obtained for the SMAD4n, which equals 0.0014, followed by the determined value when considering the SMAD2n, 0.0029. Also for an elliptical cell shape, the highest absolute value of a regression line slope is obtained for the case of the TRIMER_n, in which it assumes a value of 0.0501, followed by the pSMAD2n, where it equals 0.0332. This means that considering the same cell size and an elliptical cell shape, there is less change in concentration of the SMAD2n and the SMAD4n by additional unit of the width/height ratio than for the concentration of the pSMAD2n and the TRIMER_n (which is corroborated by the results presented in Figure 4.4, Section **4.3.1 – The Influence of Cell Shape on the Nuclear Concentration**

of the Species). For a rectangular cell shape, the lowest absolute value determined for the regression line slope is observed for the SMAD2n, assuming a value of 0.0012, followed by the observed value for the SMAD4n, which is 0.0019. On the other hand, the highest absolute value obtained for the regression line of the rectangular cell shape is observed for the TRIMER_n, assuming a value of 0.0685, succeeded by the pSMAD2n, where it equals 0.0451.

In summary, when considering a simple linear regression model to evaluate the influence of the width/height ratio on the SC ratio, it can be concluded that a bigger width/height ratio of the cell promotes less concentration of the SMAD2n, the pSMAD2n and of the TRIMER_n for elliptical and rectangular cell shapes. Conversely, it was found that as the width/height ratio of the cell increases, the concentration of the SMAD4n also increases, for elliptical and rectangular cell shapes. Still considering elliptical and rectangular cell shapes, the species whose concentration are less affected by a rise in the width/height ratio are the SMAD2n and the SMAD4n and the species which suffer more variation in their concentration as the width/height ratio becomes higher are the pSMAD2n and the TRIMER_n. In particular, for an elliptical cell shape, the highest variation in the SC ratio by width/height ratio unit is detected for the TRIMER_n and the lowest is observed for the SMAD4n. For a rectangular cell shape, the species whose concentration suffer more variation by width/height unit is the TRIMER_n and the species concentration less affected by the increase of the width/height ratio is the concentration of the SMAD2n. Out of all the presented species, the SMAD2n represents the higher discrepancy in its concentration value across cell shapes – the ratio between the slope of the regression line of the ellipse and the slope of the regression line of the rectangle equals 2.417, meaning that the concentration of the SMAD2n is 142% higher for an elliptical cell than for a rectangular one, by width/height ratio unit. For the other species, the inverse ratio was determined since the slope is higher for the rectangular shape (the variation of the species concentration across cell shapes is what is being considered, not the variation from one specific shape to another) and the respective percentage of increase is presented. The pSMAD2n, the TRIMER_n and the SMAD4n have similar variations across cell shapes. Following the SMAD2n, the second highest discrepancy for the concentration value across cell shapes is determined for the TRIMER_n, for which the concentration is 36.7% higher when the cell shape is rectangular, in terms of width/height ratio unit. Following the TRIMER_n, the pSMAD2n holds the third highest discrepancy in its concentration value across cell shapes, for which a rectangular cell shape induces a concentration 35.8% superior to an elliptical cell shape by width/height ratio unit. Lastly, the lowest discrepancy in the concentration of a species across shapes is determined for the SMAD4n, for which a concentration 35.7% superior is determined for a rectangular cell shape in relation to an elliptical cell.

However, it is worth noting that the linear regression model was shown to not be of good quality for the SMAD2n. The prominent dispersion of values of the SC ratio found for the SMAD2 might be related with the fact that the equation which translates the variation in the concentration of the SMAD2n – Equation (3.19) – has less terms than the equations which represent the variation of the concentration of the other species in the nucleus – Equations (3.20) – (3.22) – which would normally result in less computational miscalculation, thus, less deviation from real values.

4.3.3 – The Influence of the Cell Size on the Concentration of the TRIMER_n

This set of experiments aims to assess the influence of the area of the cell (which is the other morphological characteristic in study besides the shape of the cell) upon the concentration of the TRIMER_n. In this experiment, the only outcome taken into account is the average concentration of the TRIMER in the nucleus, the TRIMER_n, due to the fact that this is the only species capable of binding to the DNA, thus allowing the occurrence of genetic expression and activating the signaling pathway downstream.

Simulations ran to assess the influence of the cell size on the TRIMER_n concentration for different cell geometries, with different areas for the cell, while maintaining the area of the nucleus and the area of the activation ring (see Section **3.2.4 – Geometry**). The different values of the cell size used in these simulations are the Standard Size (see Section **3.2.4 – Geometry**), 0.6·Standard Size; 0.8·Standard Size; 2.0·Standard Size; 6.0·Standard Size; 8.0·Standard Size; 10.0·Standard Size. The results are presented in Figure 4.9, in which the TRIMER_n concentration is represented by the SC ratio. This figure displays scatter plots, for the different cell geometries, of the SC ratio depending on the cell size. In order to interpret the results obtained via simulations, a trend line was fit to the data for all different cell shapes: circle, ellipse 1, ellipse 2, rectangle and square. The type of trend line that fitted the best the available data is a power trend line. The power trend lines of each scatter plot and the respective power functions are also depicted in Figure 4.9, along with the R^2_{adj}, which indicates the reliability of the trend line (see Section **4.3.2 – The Influence of the Width/Height Ratio on the Nuclear Concentration of Species**).

Firstly, by observing the values attributed to R^2_{adj} for the circle (Figure 4.9A), ellipse 1 (Figure 4.9B), ellipse 2 (Figure 4.9C), rectangle (Figure 4.9D) and square (Figure 4.9E), which are 0.983, 0.985, 0.986, 0.978 and 0.978, respectively, it can be concluded that all R^2_{adj} are close to 1, thus indicating that the trend lines used to fit the data hold a high value of reliability. This suggests that the influence of the cell size on the TRIMER_n concentration can be modeled by a power function, regardless of the cell geometry. In addition, a power function implies a decrease or increase in the value of the dependent variable, at a decreasing rate or increasing rate, respectively, as the independent variable increases. In the specific case of the results depicted in Figure 4.9, there is a decrease in the SC ratio, thus in the TRIMER_n concentration, as the size of the cell increases. The decrease in value of the SC ratio by area unit becomes less accentuated as the cell size increases.

Considering the Standard Size of the cell (707.36 μm^2) as the reference cell size and the cell size as the independent variable in the functions which translate the trend lines presented for the circle, ellipse 1 and ellipse 2 (Figure 4.9A, Figure 4.9B and Figure 4.9C, respectively), calculations were performed to assess the variation in the SC ratio value for the following cell sizes: 2.0·Standard Size (1414.72 μm^2), 5.0·Standard Size (3536.80 μm^2) and 10.0 (7073.60 μm^2) ·Standard Size. All three cell shapes display the same percentage of drop in the TRIMER_n concentration for the considered cell sizes. In particular, for a cell size correspondent to 2.0·Standard Size, there is a decrease of 18% in the absolute value of the SC ratio (in comparison to a cell of Standard Size) while for cell sizes of 5.0·Standard Size and 10·Standard Size, there is a decrease of 38 % and of 49%, respectively. When applying the same calculations to a rectangular cell, for which the results are displayed in Figure 4.9D, it can be concluded that there is a drop of 19% in the SC ratio value for a cell size of 2.0·Standard Size (in relation to a cell of Standard Size), while there is a decrease in value of the SC ratio equal to 39% and 51% for cell sizes of 5.0·Standard Size and 10.0·Standard Size, respectively. Lastly, when applying the same calculations to a square cell, for which the results are displayed in Figure 4.9E, it can be concluded that there is a drop of 18% in the SC ratio value for a cell size of 2.0·Standard Size, while there is a decrease in the SC ratio of 36% and 48% for cell sizes of 5.0·Standard Size and 10.0·Standard Size (when comparing

to a cell of Standard Size). However, it is noteworthy that the discrepancy between the observed percentages of decrease of the SC ratio across cell shapes might be related with minor alterations in the area of the activation ring (see **3.2.4 – Geometry**) which automatically took place in VCell when modifying the cytoplasmic area of the cell from one MathModel to another.

In summary, when analyzing the results from Figure 4.9, it can be inferred that as the area of the cell increases, the SC ratio decreases for all the different shapes. For a range of cell sizes which extends from the Standard Size up until 10·Standard Size, there is a drop of approximately 50% in the concentration of the TRIMER_n for all cell sizes. In particular, the biggest drop of concentration for the considered cell sizes was observed for the rectangular cell, followed by the circular and two elliptical geometries, and lastly, the SC ratio decreases the least for the square cell.

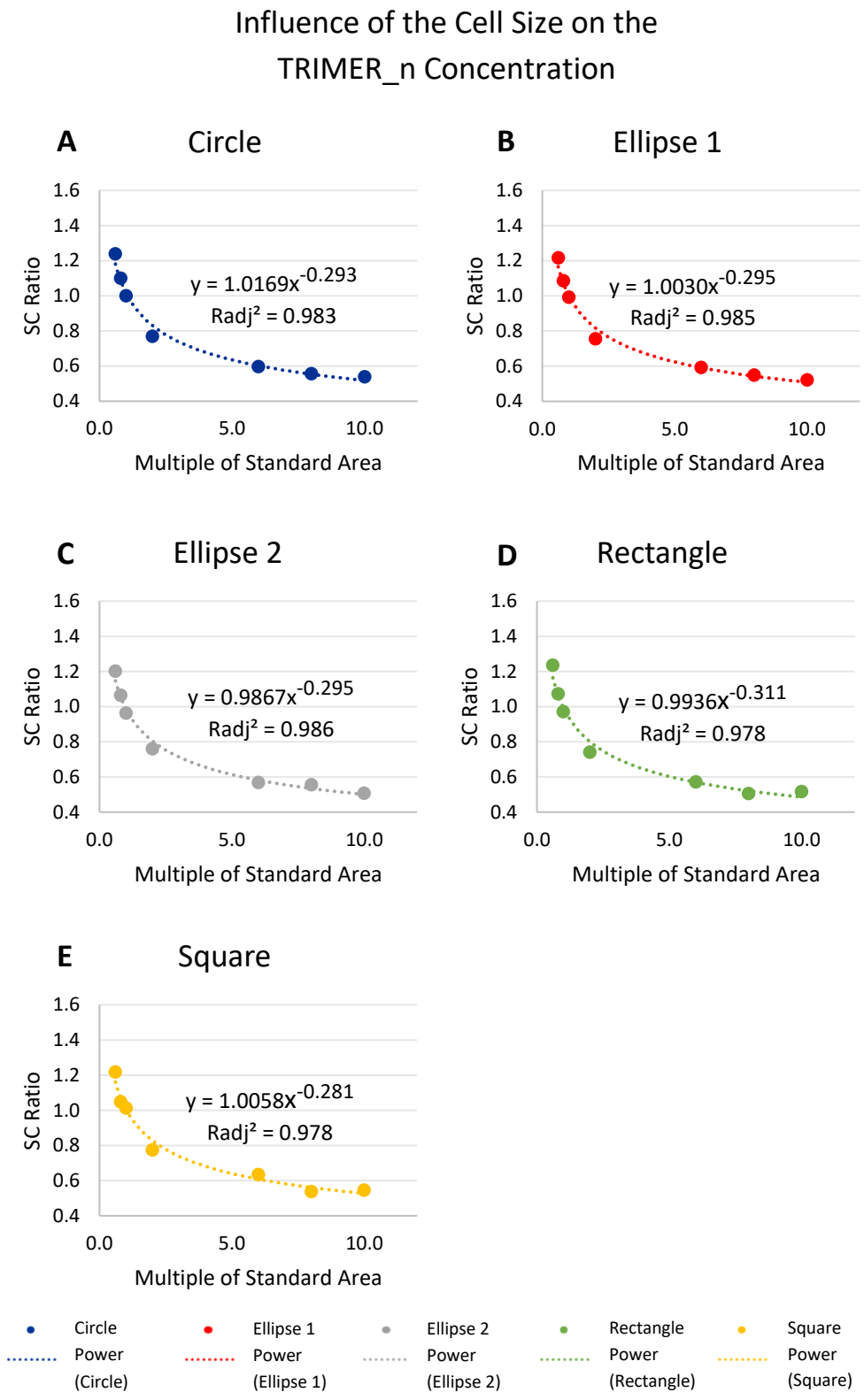

Figure 4.9 – Influence of the cell size on the concentration of the TRIMER_n measured by the SC ratio. The cell size range extends from the standard cell area up to 10 times the standard area. The figure displays a scatter plot for different cell shapes, power trend line equations and respective adjusted R-squared values. A –Circle geometry; the scatter plot is represented by blue dots and the trend line is represented by a blue dashed curve; B –Ellipse 1 geometry; the scatter plot is represented by red dots and the trend line is represented by a red dashed curve; C –Ellipse 2 geometry; the scatter plot is represented by grey dots and the trend line is represented by a grey dashed curve; D –Rectangle geometry; the scatter plot is represented by green dots and the trend line is represented by a green dashed curve; E –Square geometry; the scatter plot is represented by yellow dots and the trend line is represented by a yellow dashed curve.

4.3.4 – Sensitivity Analysis

Every parameter of the model was subjected to a sensitivity analysis procedure in order to understand which parameters affect the TRIMER_n concentration the most. The sensitivity of each parameter was evaluated through the following equation:

$$Parameter\ sensitivity = \left| \frac{|Concentration\ of\ the\ TRIMER\_n(k + \Delta k) - Concentration\ of\ the\ TRIMER\_n(k)|}{Concentration\ of\ the\ TRIMER\_n(k)} \Big/ \left(\frac{\Delta k}{k}\right) \right| \quad (4.5)$$

Which was adapted from [84]. It is noteworthy that the concentration of the TRIMER_n mentioned in Equation (4.5) is the steady-state concentration. The ratio $\frac{\Delta k}{k}$ assumes the following values per simulation: +10%, +50%, +100%, -10%, -50% and -100%. Equation (4.5) was selected to evaluate the parameter sensitivity because it was also used for sensitivity analysis purposes in a different pathway, the YAP/TAZ pathway, included in the PhD project related to this project (see Section **1.1 – Contextualization and Motivation**). The achieved results from Equation (4.5) place the dependence of the TRIMER_n on the considered parameter in one of the three categories below:

- $Parameter\ sensitivity = 1$: the dependence on the parameter is linear;
- $Parameter\ sensitivity < 1$: the dependence on the parameter is small;
- $Parameter\ sensitivity > 1$: The dependence on the parameter is high.

The results obtained through the sensitivity analysis process are presented in Figure 4.10, Figure 4.11 and Figure 4.12.

When comparing the results obtained for the 10% fold change, Figure 4.10, with the results determined for the 50% and 100% fold change, Figure 4.11 and Figure 4.12, respectively, it can be seen that the difference between the parameter sensitivity value for the fold increase and fold decrease, for every parameter, is smaller for the 10% fold change than for the latter fold changes, which is accordingly to what is expected since the values for the parameters remain closer to their original values when the fold change is smaller, resulting in a more subtle change of outcomes.

The results from Figure 4.10, Figure 4.11 and Figure 4.12 suggest that the parameter sensitivity is higher when there is a decrease in the value of each parameter than when it is increased. This finding is most likely due to saturation in the outcome. In the case of diffusive processes, at a given instant, the molecules are already well-mixed so a rise in the value of the diffusion coefficients will not change the outcome. In the case of reactions, by greatly increasing a reaction rate, the reaction will mostly move in

one direction; therefore, a high value for the reaction rate will not have a big effect on the outcome from a certain point. Decreasing the value of these parameters will prompt larger sensitivity. Furthermore, the graphs support that the reactions which take place in the cell have in general, particularly for 10% and 50% fold change of the parameters, a stronger impact upon the outcome – concentration of the TRIMER_n – than the diffusion of the species. In particular, there is a high dependence of the outcome on the dissociation rate of the TRIMER $kdiss$ and the dephosphorylation rate of the pSMAD2n $kdephos$.

In the case of $kdiss$, this dependence might be due to the fact that it directly influences the concentration of the TRIMER_n, which is the outcome used for sensitivity analysis, since it is its dissociation rate and it takes part in every equation which states the change of concentration of species of the pSMAD2, the TRIMER and the SMAD4 (both in the cytoplasm and in the nucleus) – Equations (3.2), (3.5), (3.6) and (3.20) – (3.22). In [18], it was also concluded through sensitivity analysis that the dissociation rate of the TRIMER had a strong impact upon the concentration of the TRIMER_n, even though this finding only applies to an intermediate to longer period of time following ligand stimulation. However, it is worth noting that in [18] the dissociation rate of the TRIMER is represented by distinctive variables in the cytoplasm and the nucleus and this separation was not made in this project. In [15], the concentration of the TRIMER_n was also shown to be very sensitive to the analogous of $kdiss$ in their work. Regarding the sensitivity analysis for $kdiss$ performed in this work, it is worth noting that the maximum end time for the simulations was of 1,000,000s and this limiting factor did not allow to fully reach the steady-state concentration of the TRIMER_n. However, the graph obtained via simulation suggests that the steady-state concentration is approximately 0.070 μM (the value used for the sensitivity analysis is 0.0717 μM).

In what concerns the parameter $kdephos$, its sensitivity might be explained by the fact that this rate only multiplies by the pSMAD2n in the term in which it is inserted in Equations (3.19) and (3.20), it does not multiply by the concentration of the other species. Therefore, taking into account that almost all the rates and concentrations have an absolute value below 1, this makes the term in which it is inserted to have higher magnitude than many of the other terms present in all the equations, culminating in a larger weight on the signaling pathway outcome. In addition, the parameter $kdephos$ has a direct effect on the concentration of the pSMAD2n, since it consists in its dephosphorylation rate. This fact consequently has a strong effect on the TRIMER_n concentration as the pSMAD2n integrates the TRIMER_n complex molecules. The dephosphorylation rate of the nuclear phosphorylated SMAD2 is also identified in literature as one of the parameters that the TRIMER_n has shown more sensitivity to, particularly in [18].

The results also show that the parameter $ktrim$, which is the formation rate of the TRIMER, is also one of the parameters to which the outcome displays more sensitivity. This is most likely due to the fact that this parameter is directly connected to the TRIMER, regardless of it being located in the cytoplasm and in the nucleus, translating the rate at which the pSMAD2 and the SMAD4 molecules associate with each other in order to form the TRIMER complex. In [15], this parameter has also been identified has one which carries the most influence upon the nuclear TRIMER complex concentration.

The import and export rates of the SMAD4 molecules, kcn_SMAD4 and knc_SMAD4, respectively, have also been identified through sensitivity analysis as two of the parameters to which the nuclear concentration of the TRIMER is the most sensitive to. This goes accordingly to what was found in [18]. Taking into account that the initial concentration of the species was the same across simulations, the most plausible explanation for the outcome sensitivity in relation to the previously stated rates is the fact that these rates directly influence the amount of SMAD4 available in the nucleus and these SMAD4n molecules associate with the SMAD2n molecules to constitute the TRIMER_n. However, it is noteworthy that in the case of the pSMAD2n, even though it also takes part in the composition of the TRIMER_n, the outcome did not reveal as much sensitivity to its import rate to the nucleus. This is probably due to the fact that the pSMAD2 is able to translocate from the cytoplasm into the nucleus but cannot leave the

nucleus in order to return to the cytoplasm, in opposition to the SMAD4n, which is able to undergo translocation both ways. This means that in the case of the pSMAD2n, a change in value of its import rate played a less important role on the concentration of the TRIMER_n since any of its molecules which migrate into the nucleus end up staying in the nucleus regardless, which means that the sole direct factor which plays a pivotal role in the determination of the number of pSMAD2n molecules that participate in the formation of the TRIMER_n is the dephosphorylation rate of the pSMAD2n, represented by $kdephos$. The latter had previously been identified as one of the parameters to which the outcome displayed more sensitivity.

Even though these results suggest that diffusion processes do not have as much effect upon the nuclear TRIMER concentration as the reactions, it seems that their absence in the model would have a big effect on the outcome, as it can be seen in Figure 4.12. When the diffusion coefficients of the species are equal to 0.0 $\mu m^2 s^{-1}$, the outcome displays approximately linear to high sensitivity to these parameters (the exceptions are the $SMAD4n\_diffusionRate$, the $TRIMERn\_diffusionRate$ and the $pSMAD2\_diffusionRate$).

In summary, it can be concluded that the model shows that the dissociation and formation of the TRIMER (both in nucleus and in the cytoplasm), the export and import of the nuclear SMAD4 and the dephosphorylation of the nuclear pSMAD2 are the cell processes which have more impact on the concentration of the nuclear TRIMER. In addition, the diffusion coefficients do not contribute as much for the concentration of the nuclear TRIMER as the different rates associated to the species included in this model. However, these parameters should still be included in models build to mimic and study the dynamics of the TGF-β pathway since, in general, a decrease of 100% in their value displays that there is a linear to high sensitivity on the outcome.

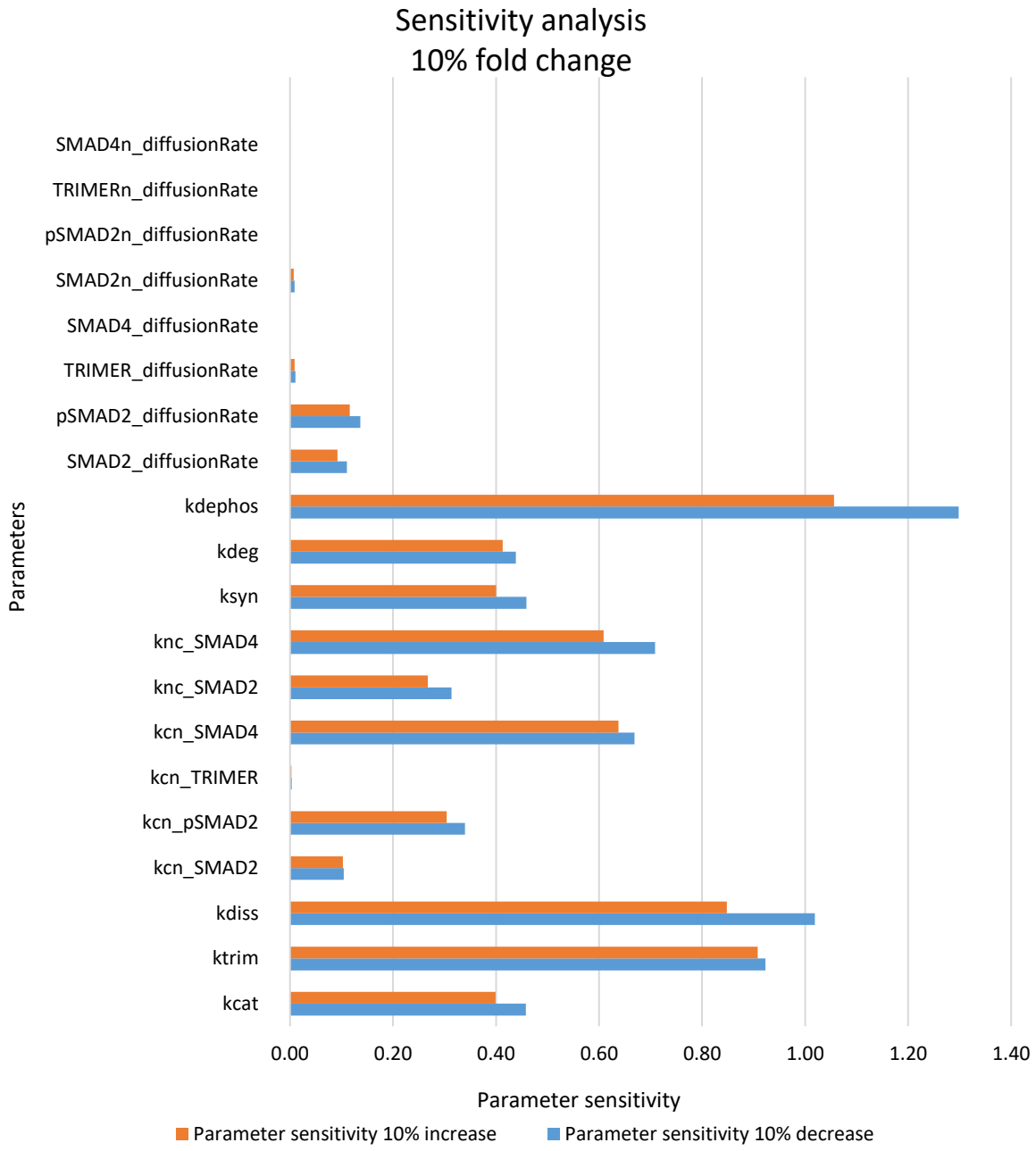

Figure 4.10 – Parameter sensitivity for 10% fold change of the parameters. The bars in blue represent the parameter sensitivity due to a decrease of 10% of the value of the parameter, while the bars in orange represent an increase of 10% of the value of the parameter.

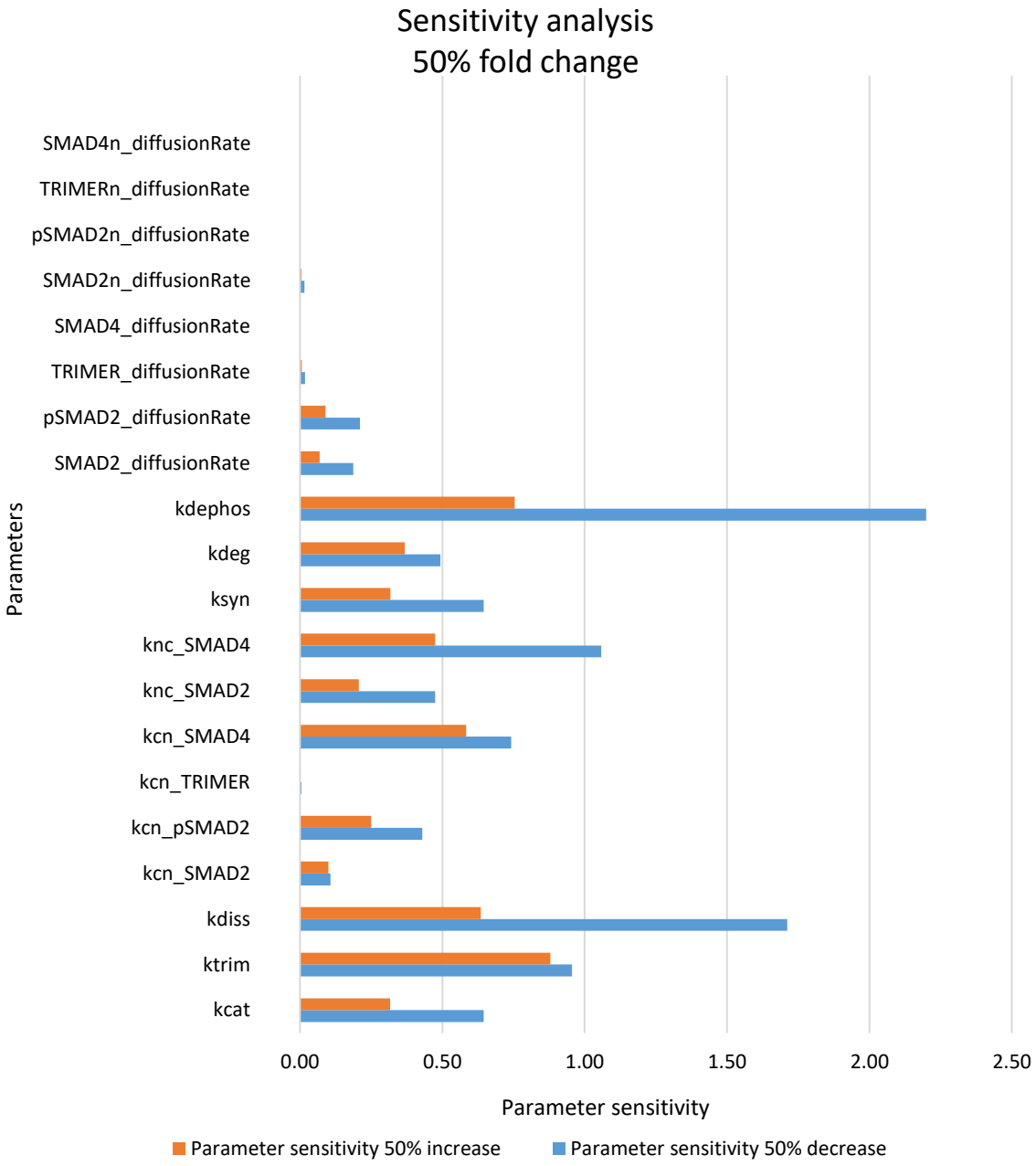

Figure 4.11 – Parameter sensitivity for 50% fold change of the parameters. The bars in blue represent the parameter sensitivity due to a decrease of 50% of the value of the parameter, while the bars in orange represent an increase of 50% of the value of the parameter.

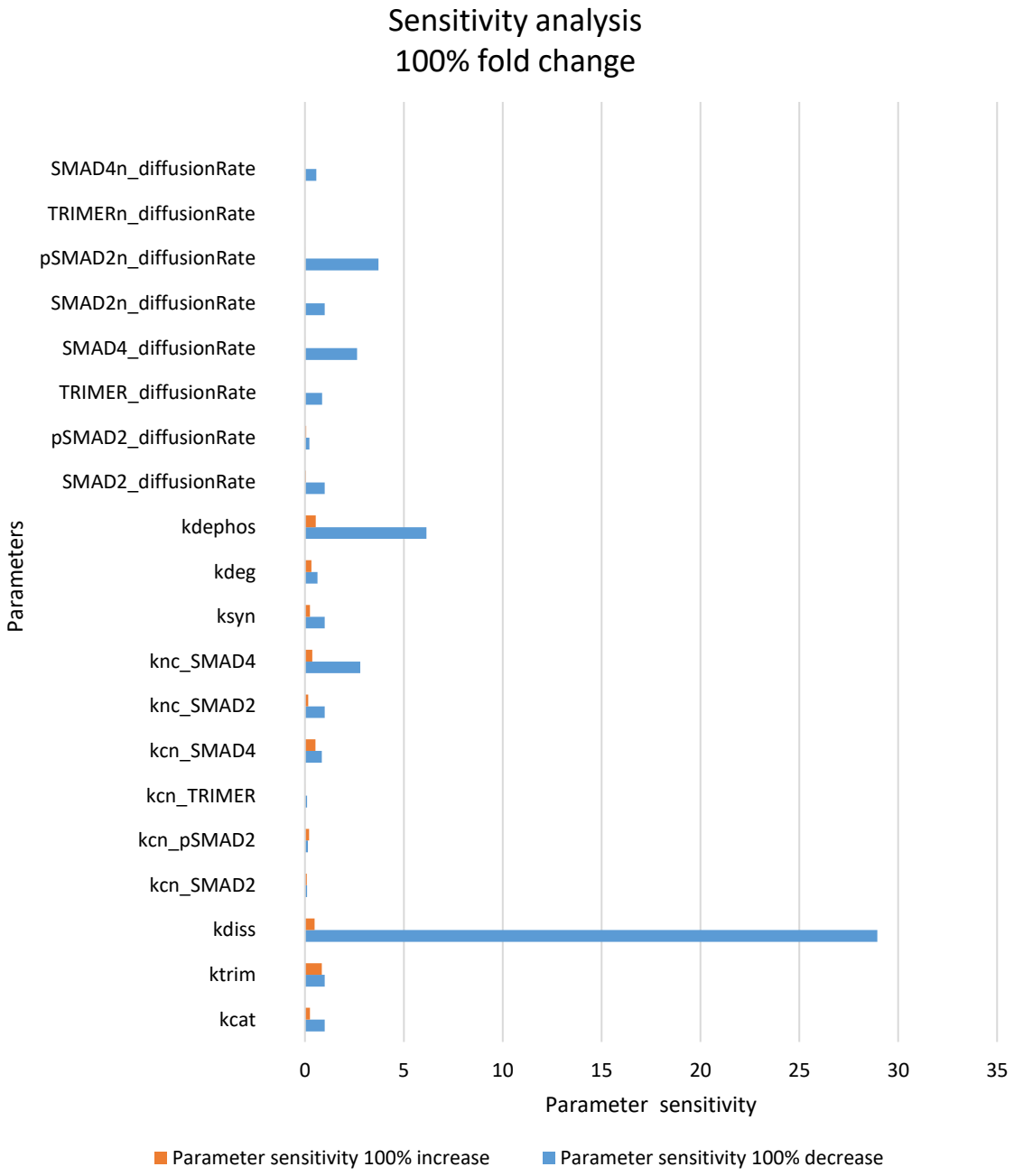

Figure 4.12 – Parameter sensitivity for 100% fold change of the parameters. The bars in blue represent the parameter sensitivity due to a decrease of 100% of the value of the parameter, while the bars in orange represent an increase of 100% of the value of the parameter.

4.4 – Module II Results

As previously mentioned in Section **1.2 – Objectives**, one of the aims of this work is to assess the crosstalk between mechanotransduction, in specific between integrins, and the TGF-β signaling pathway. The generated basis script for this module of the work is shown in Section **Module II Script** from

APPENDIX C. It is worth noting that all the applicable equations and boundary conditions for this module are present in Section **3.2.3 – Equations and Boundary Conditions** and that every parameter associated to the integrin module, namely $ksyn\_intg$, $kdeg\_intg$, $kdiss\_IC$, $kform\_IC$ and $Imax$, now assume the value stated in Section **3.2.2 – Parameters**, Table 3.2 and Table 3.3. For some simulations, appropriate modifications were made to $k_{extra_i}, i = 1,2,3,4$. Therefore, all $k_{extra_i}, i = 1,2,3,4$, assume the value stated in Table 3.3, which is 0.1, unless stated otherwise.

4.4.1 – Assessment of the Interactions between the TGF-β Signaling Pathway and the Integrins

As stated in Section **1.1 – Contextualization and Motivation**, a deeper understanding regarding the crosstalk between the TGF-β and mechanotransduction can potentially provide valuable information to achieve control of TGF-β expression and this knowledge can subsequently be used to uncover cell-biomaterial interactions employed in numerous therapies, such as therapies aimed to control the progression of tumors. Therefore, a series of experiments were performed to understand the influence of the crosstalk interactions with the integrins upon the signaling cascade. These simulations were created to evaluate the influence of each type of crosstalk interaction, either in separate or together with another type of interaction, on the concentration of the TRIMER_n, since this species binds to the nucleus, thus allowing genetic expression.

The four types of interaction which were considered in these experiments are:

1) The upregulation of the TGF-β receptors elicited by the integrins, which is translated into an increase of the synthesis rate of the *C* complex;
2) The *C* complex, whose synthesis is elicited by the integrins, binds to the SMAD2, prompting an increment in the catalytic constant of the phosphorylation reaction of the SMAD2 mediated by the *C* complex;
3) The stabilization of the TGF-β receptors elicited by the integrins, resulting in the decrease of the degradation rate of the *C* complex;
4) The formation of the *IC* complex, a complex made up by the integrins, *I*, and the TGF-β receptors, represented by the *C* complex; the *IC* complex binds to the SMAD2, eliciting the increase of the catalytic constant of the phosphorylation reaction of the SMAD2 mediated by the *IC* complex.

The results obtained via simulation concerning the influence of one type, two types and three types of crosstalk interactions upon the TGF-β signaling pathway are presented in Figure 4.13, Figure 4.14 and Figure 4.15, respectively. All these figures depict the value of the relative nuclear different bar plots. The relative concentration for each interaction was calculated as a ratio between the absolute value of the nuclear TRIMER_n concentration for the interaction or simultaneous interactions taken into account (1, 2, 3 or 4) and the TRIMER_n concentration for all k_{extra_i} =0.1. One should bear in mind that $k_{extra_i}, i = 1,2,3,4$, are parameters included in functions created to evaluate different types of crosstalk interactions (see Section **3.2.3 – Equations and Boundary Conditions**). Each number is associated to each type of crosstalk in the following manner: 1) k_{extra_1} is associated to an increase of the synthesis

rate of the *C* complex; 2) k_{extra_2} is associated to an increase in the catalytic constant of the phosphorylation of the SMAD2 mediated by the *C* complex; 3) k_{extra_3} is associated to a decrease in the degradation rate of the *C* complex; 4) k_{extra_4} is associated to an increase of the catalytic constant of the phosphorylation of the SMAD2 mediated by the *IC* complex. The bar plot which concerns the simulation for all k_{extra_i} =0.1 (all k_{extra_i} assume their default value) is denoted as *Control*. In order to identify if a certain type of crosstalk interaction suffers any modification from the default value of 0.1, there are symbols associated to each bar plot: the symbol "+" is used to denote that there is a modification of k_{extra_i} in relation to its default value; if there is a modification, k_{extra_i} assumes the value identified in the legend of each figure, placed on their right side, as opposed to the "- " symbol, which is used when the default value of these parameters are maintained. Therefore, a certain type of interaction is never considered to be completely "off"; the assessment of the influence of crosstalk interactions with the integrins is made by comparing the concentration of the TRIMER_n when there is an increase of the absolute value of one, two or three different k_{extra_i}, to the concentration of the TRIMER_n when there is not an increase in the absolute value of any k_{extra_i} (represented by the *Control* bar). It is worth noting that the values for the concentration of the TRIMER_n which were considered to determine the relative concentration displayed by the different bar plots are steady-state concentrations and were obtained via simulation.

By analyzing Figure 4.13, which portrays the influence that one way of crosstalk of the TGF-β with the integrins has upon the concentration of the TRIMER_n, it can be concluded that in general, when there is an increase in the absolute value of $k_{extra_i}, i = 1,2,3$, there is an increase in the outcome, which is the concentration of the TRIMER_n. This finding indicates more genetic expression. The results presented in Figure 4.13 suggest that a higher synthesis rate of the TGF-β receptors, a higher catalytic constant of the phosphorylation of the SMAD2 mediated by the *C* complex and a lower degradation rate of the TGF-β receptors, individually result in a higher concentration of the TRIMER_n. On the other hand, when k_{extra_4} is increased, the absolute value of the TRIMER_n concentration practically remains the same and is identical to the concentration in default conditions, which suggests that a higher catalytic constant of the phosphorylation of the SMAD2 due to *IC* binding is not relevant to increase the concentration of the TRIMER_n. Another conclusion that can be drawn by looking at Figure 4.13 is that the crosstalk between the TGF-β pathway and the integrins via decrease of the degradation rate of the receptors promotes more concentration of the TRIMER_n than the other ways of crosstalk. When k_{extra_3} assumes a value of 0.5, the outcome is about 1.15 times the value of TRIMER_n concentration in default conditions and when the value of k_{extra_i} increases to 0.7 and 1.0, the outcome changes to 1.2 times and 1.3 times the *Control* outcome, respectively. These values are the highest for every set of $k_{extra_i} i = 1,2,3,4$.

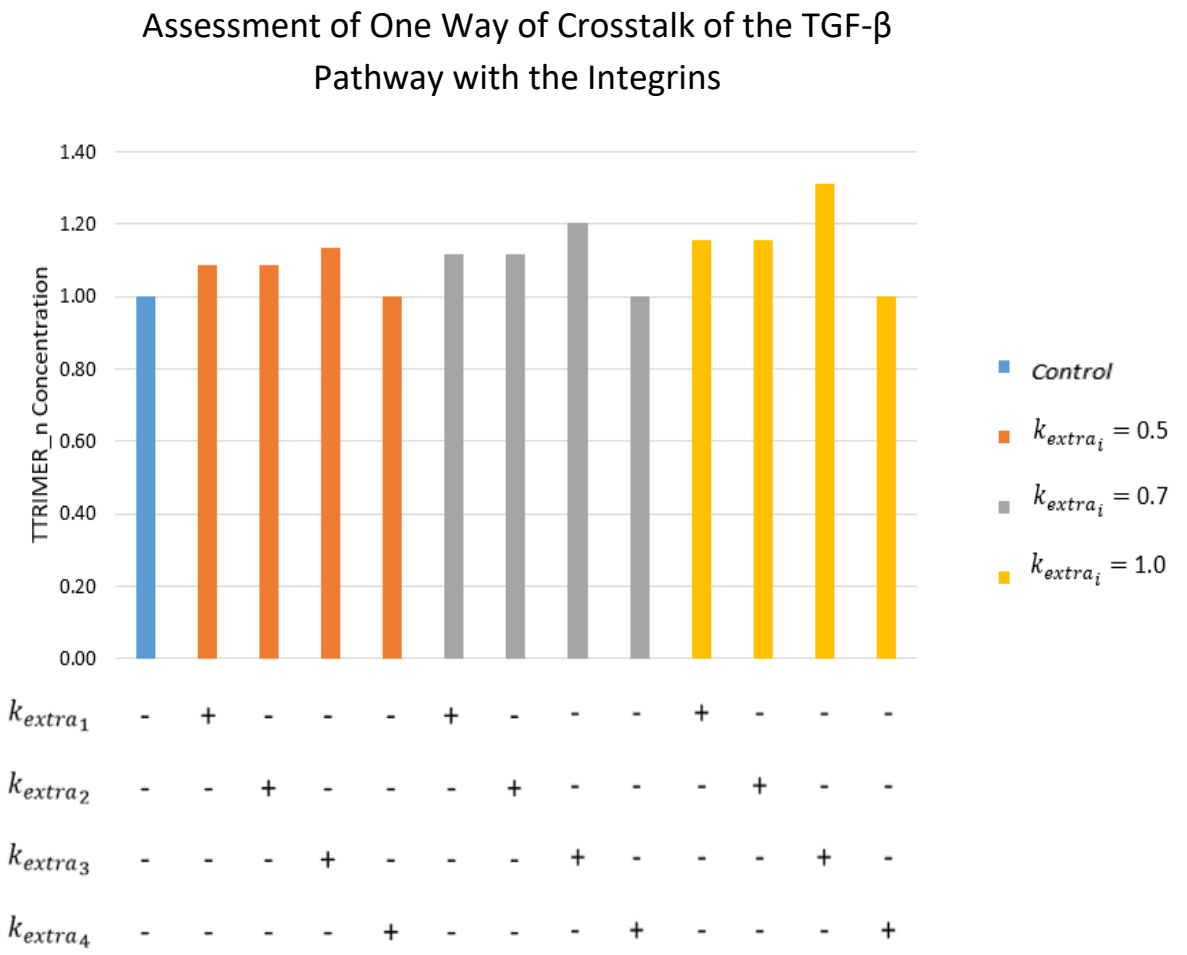

Figure 4.13 – Assessment of one way of crosstalk of the TGF-β pathway with the integrins. The bar plots display the relative concentration of the TRIMER_n due to one way of crosstalk with the integrins, except for the first bar plot, which is the control bar for the rest of the results. Each $k_{extra_i}, i = 1,2,3,4,$ is associated to a specific type of crosstalk between the integrins and the TGF-β pathway represented by i: 1 – Upregulation of the TGF-β receptors elicited by the integrins; 2 – Increase of the catalytic constant of the phosphorylation of the SMAD2 through binding to the *C* complex; 3 – Decrease of the degradation rate of the receptors; 4 – Increase of the catalytic constant of the phosphorylation of the SMAD2 through binding to the *IC* complex. When the "+" symbol is spotted in the line i, that means that the parameter k_{extra_i} suffers a modification and the new value assumed is shown by the legend on the right side of the color plots. The "-" symbol means that changes are not made to the default value.

When analyzing Figure 4.14, which assesses the effects of two ways of crosstalk of the TGFβ with the integrins, it can be seen that for every possible combination of two types of crosstalk, considering all different values assumed by k_{extra_i}, $i = 1,2,3,4$, there is a higher value for the outcome than the one determined in default conditions, which means that when the effects of two different types of crosstalk are exacerbated simultaneously, more genetic expression is induced. When considering each possible value for k_{extra_i}, for almost all the combinations which include an increase in the value of k_{extra_3} with a simultaneous increase of other $k_{extra_i}, i = 1,2,4$, the outcome assumes a higher value than in comparison with other combinations which exclude an increase of k_{extra_3}. This finding suggests that the crosstalk of the integrins with the TGF-β via decrease of the degradation rate of the receptors is the type of interaction which promotes a higher value of the outcome, similar to what is implied by the results presented in Figure 4.13. The only exception to this pattern in Figure 4.14 is detected in the joint effect of the decrease in the receptor degradation rate and the increase of the catalytic constant of the

phosphorylation of the SMAD2 due to *IC* binding, associated to parameters k_{extra_3} and k_{extra_4}. This combination of interactions is not part of the three bar plots for k_{extra_i}=0.5 and k_{extra_i}=0.7 which indicate higher concentration values of the TRIMER_n. Similar to findings from Figure 4.13, which assess the effects of one way of crosstalk on the TRIMER_n concentration, it is suggested that the phosphorylation of the SMAD2 mediated by the *IC* complex does not promote as much formation of the TRIMER_n in comparison with the other types of crosstalk. In Figure 4.14, for k_{extra_i}=0.5, the highest concentration for the outcome is approximately 1.2 times the TRIMER_n concentration in default conditions. When comparing it to the highest concentration presented in Figure 4.13, also for k_{extra_i}=0.5, the highest value is detected when assessing two ways of crosstalk of the TGF-β with the integrins than just one. For the case in which k_{extra_i}=0.7, the greater concentration achieved for the outcome when considering the joint effect of two types of crosstalk, seen in Figure 4.14, is approximately 1.28 times the concentration for the TRIMER_n in control conditions. This value is above the highest value seen in Figure 4.13, which concerns the assessment of one way of crosstalk; in Figure 4.13, the outcome for k_{extra_i}=0.7 is around 1.2 times the outcome in default conditions. This finding is also verified for k_{extra_i}=1.0, since the highest value for the outcome in this group of results, in Figure 4.14, is around 1.38 times the *Control* outcome versus 1.3 times the outcome concentration in Figure 4.13, which concerns the change in value of only one rate, associated to one type of crosstalk.

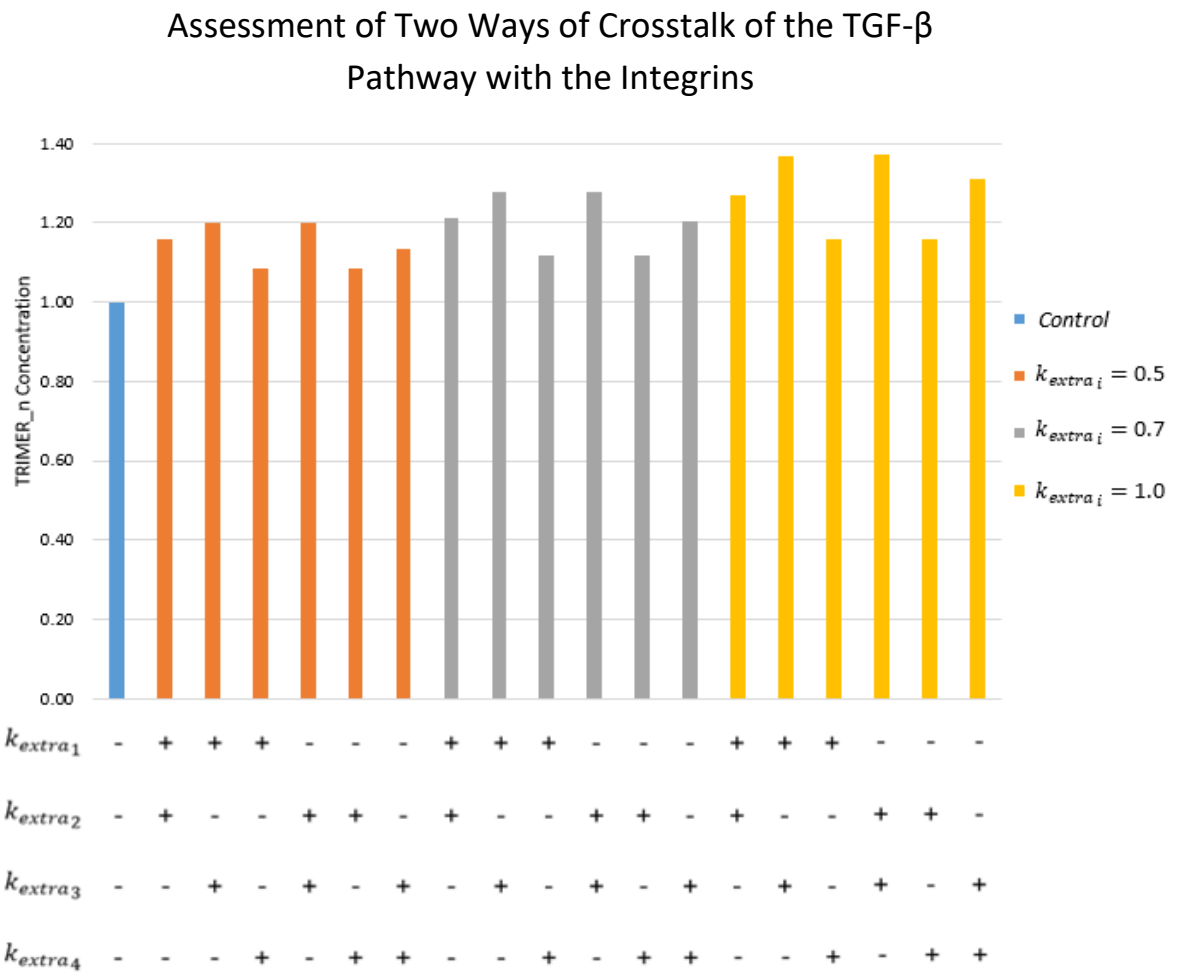

Figure 4.14 – Assessment of two ways of crosstalk of the TGF-β pathway with the integrins. The bar plots display the relative concentration of the TRIMER_n due to two ways of crosstalk with the integrins, except for the first bar plot, which is the control bar for the rest of the results. Each $k_{extra_i}, i = 1,2,3,4,$ is associated to a specific type of crosstalk between the integrins and the TGF-β pathway represented by i: 1 – Upregulation of the TGF-β receptors elicited by the integrins; 2 – Increase of the catalytic constant of the phosphorylation of the SMAD2 through binding to the *C* complex; 3 – Decrease of the degradation rate of the receptors; 4 – Increase of the catalytic constant of the phosphorylation of the SMAD2 through binding to the *IC* complex. When the "+" symbol is spotted in the line i, that means that the parameter k_{extra_i} suffers a modification and the new value assumed is shown by the legend on the right side of the color plots. The "-" symbol means that changes are not made to the default value.

By looking at Figure 4.15, when assessing the impact of three ways of crosstalk upon the outcome, there is not a general trend when k_{extra_i} increases. Therefore, it is pertinent to comment on the influence of each type of crosstalk when k_{extra_i} changes, that is, in a case-by-case manner. For the combination of k_{extra_1}, k_{extra_2} and k_{extra_3} (which corresponds to the joint effect of an increase in the synthesis rate of the receptors, in the catalytic constant of the phosphorylation of the SMAD2 mediated by the *C* complex and of a decrease in the degradation rate of the receptors), as the value of k_{extra_i} is incremented, there is a decline in the concentration of the TRIMER_n. The combination made up by k_{extra_1}, k_{extra_2} and k_{extra_4} (which corresponds to the joint effect of an increase in the synthesis of the receptors, in the catalytic constant of the phosphorylation of the SMAD2 mediated by the *C* complex and in the catalytic constant of the phosphorylation of the SMAD2 mediated by the *IC* complex) leads to less concentration of the outcome with higher k_{extra_i}. When considering the combination of k_{extra_1}, k_{extra_3} and k_{extra_4}

(which corresponds to the joint effect of an increase in the synthesis of the receptors, decrease in the degradation rate of the receptors and increase in the catalytic constant of the phosphorylation of the SMAD2 mediated by the *IC* complex), there is a slight increase in the outcome as k_{extra_i} gets higher. For the last combination of crosstalk interactions, associated to k_{extra_2}, k_{extra_3} and k_{extra_4}, (which corresponds to the joint effect of an increase in the phosphorylation rate of the SMAD2 mediated by the *C* complex, decrease in the degradation rate of the receptors and increase in the catalytic constant of the phosphorylation of the SMAD2 mediated by the *IC* complex), the concentration of the TRIMER_n practically remains the same as k_{extra_i} is incremented. Still in Figure 4.15, the highest concentration for the TRIMER_n is found for the combination of k_{extra_1}, k_{extra_2} and k_{extra_3}. This value is about 1.25 times the outcome in *Control* circumstances when $k_{extra_i} = 0.5$. In addition, for the same combination of crosstalk interactions, when k_{extra_i}= 0.7 and k_{extra_i}=1.0, the TRIMER_n concentration is around 1.1 times the *Control* outcome. Once again, these values are the highest among every combination of different interactions. In comparison with the results concerning one way and two ways of crosstalk, the highest detected value for the outcome when assessing three ways of crosstalk is identical to its analogous presented for the results which assess the influence of one way of crosstalk but it is below the highest concentration of the TRIMER_n found for two ways of crosstalk. On the other hand, the lowest values of the outcome for every value of k_{extra_i} , in Figure 4.15, are found for the combination of k_{extra_1} , k_{extra_2} and k_{extra_4}. The outcome for the combination made up by k_{extra_1} , k_{extra_2} and k_{extra_4} lies, approximately, between 0.9 and 0.95 times the *Control* outcome, for every k_{extra_i}. Therefore, it can be concluded that the lowest values for the outcome in Figure 4.15 correspond to crosstalk combinations which exclude the decline of the degradation rate of the receptors as one of the interactions modified from the *Control* state (k_{extra_3} is "−") and, simultaneously, include the increase of the catalytic constant of the phosphorylation of the SMAD2 mediated by the *IC* complex (k_{extra_4} is "+"). From the analysis of the results concerning the effects of three ways of crosstalk, it can be concluded that the crosstalk with the integrins which results in less outcome is the increase of the catalytic constant of the phosphorylation of the SMAD2 mediated by the *IC* complex and the interaction type which promotes a higher level of the outcome is the decrease of the degradation rate of the receptors due to stabilization of the integrins. These findings are also corroborated by the results from Figure 4.13 and Figure 4.14, concerning one and two types of interaction.

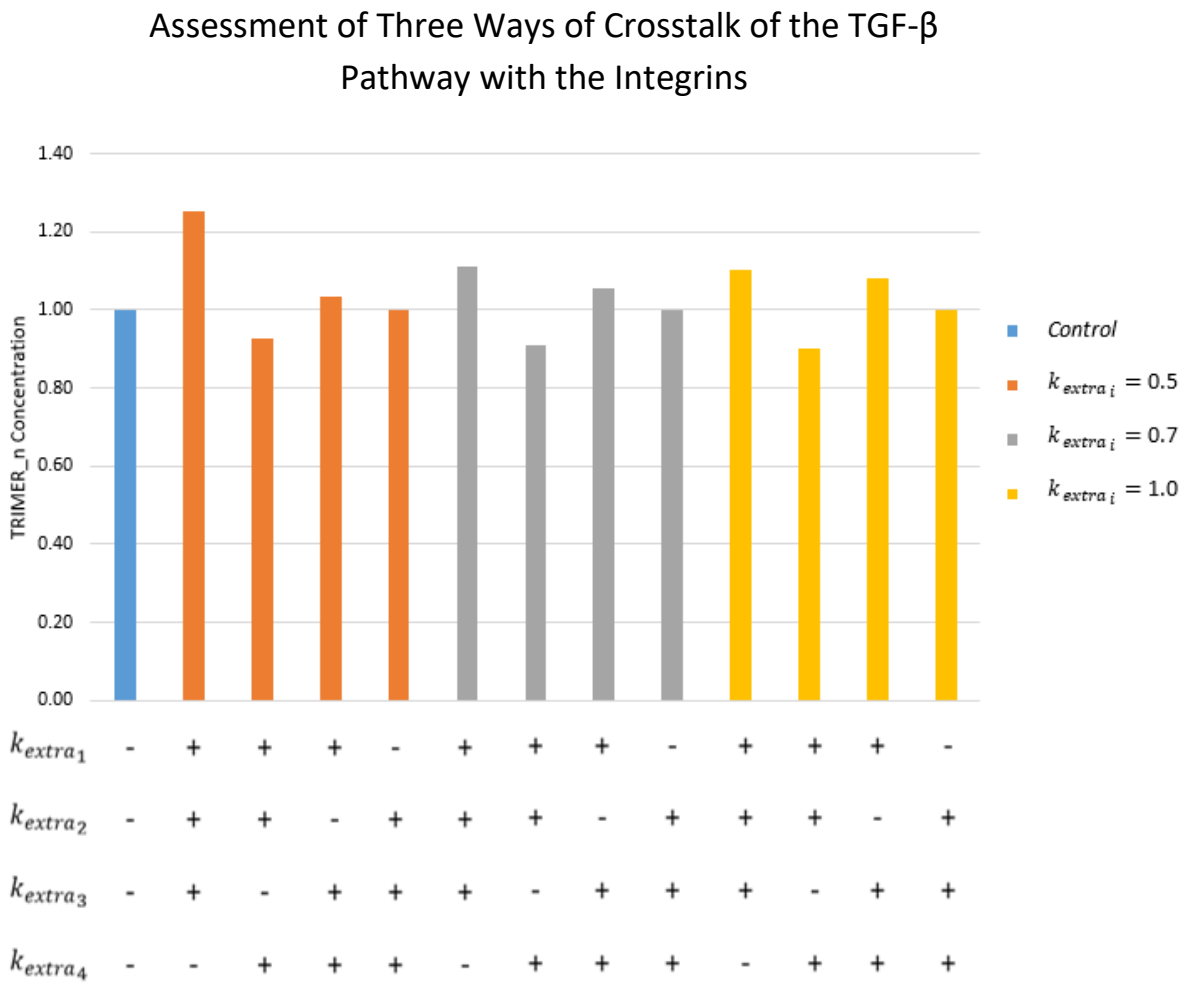

Figure 4.15 – Assessment of three ways of crosstalk of the TGF-β pathway with the integrins. The bar plots display the relative concentration of the TRIMER_n due to three ways of crosstalk with the integrins, except for the first bar plot, which is the control bar for the rest of the results. Each $k_{extra_i}, i = 1,2,3,4,$ is associated to a specific type of crosstalk between the integrins and the TGF-β pathway represented by i: 1 – Upregulation of the TGF-β receptors elicited by the integrins; 2 – Increase of the catalytic constant of the phosphorylation of the SMAD2 through binding to the C complex; 3 – Decrease of the degradation rate of the receptors; 4 – Increase of the catalytic constant of the phosphorylation of the SMAD2 through binding to the IC complex. When the "+" symbol is spotted in the line i, that means that the parameter k_{extra_i} suffers a modification and the new value assumed is shown by the legend on the right side of the color plots. The "-" symbol means that changes are not made to the default value.

In sumary, the results concerning Figure 4.13, Figure 4.14 and Figure 4.15 indicate that the type of crosstalk interaction of the TGF-β with the integrins which induces less genetic expression is the increase of the catalytic constant of the phosphorylation of the SMAD2 mediated by the *IC* complex. On the other hand, the type of crosstalk interaction which promotes a higher level of the outcome is the decrease in the degradation rate of the receptors due to stabilization of the integrins.

Another conclusion that can be drawn from Figure 4.13, Figure 4.14 and Figure 4.15 is that in general, genetic expression is stimulated more strongly by the joint behavior of two ways of crosstalk than by the effects of one way of crosstalk independently and the combination of the influence of three types of crosstalk. This finding becomes more evident as the value of k_{extra_i} increases, which is associated to a more intense effect of each type of crosstalk.

CHAPTER 5 – Conclusions

The present work was performed with the intent of creating a new spatial model for the TGF-β pathway, built on previously developed models, in order to reach a deeper understanding concerning the effect that cell morphology, particularly the shape and the size of the cell, has upon the TGF-β cascade. In addition, this work aimed to analyze the crosstalk between the TGF-β pathway and mechanotransduction, namely with the integrins. The relevance of this computational study lies in the potential of having a new insight into the cell-biomaterial interactions and subsequently applying this knowledge in various therapies, such as cancer therapy.

Following the experimental procedures and analysis of the information, the results suggest that a bigger width/height ratio of the cell induces less concentration of the SMAD2n, pSMAD2n and TRIMER_n, for elliptical and rectangular cell shapes. Conversely, the concentration of the SMAD4n increases as the width/height ratio of the cell increases, for elliptical and rectangular cell shapes. When considering different cell shapes, particularly elliptical and rectangular cell shapes, but considering the same cell area with increasing width/height ratio of the cell, the species in the nucleus whose concentration suffers the highest percentage of variation across different cell shapes is the SMAD2n. Further experimental procedures also suggest that the bigger the size of the cell, the less TRIMER_n concentration will be obtained downstream – less genetic expression. Particularly, when considering a range of cell areas which extends from the standard cell area until 10 times the standard cell area, there is a decrease of approximately 50% in the concentration of the TRIMER_n, for different cell shapes. The results also suggest that the reactions which take place in the TGF-β pathway play a more important role on downstream signaling than the diffusion of species. However, sensitivity analysis of the model parameters indicate that genetic expression prompted by the TGF-β when considering diffusive processes differs considerably from results without considering this phenomenon. Therefore, the diffusion of species should be included in models build to mimic and study the dynamics of the TGF-β pathway.

In what concerns the crosstalk of the integrins with the TGF-β signaling pathway, it seems that the way of crosstalk which increases more heavily genetic expression is the decrease of degradation rate of the receptors due to stabilization from the integrins. Conversely, out of the processes of crosstalk with the integrins included in this work, the increase of the catalytic constant of the phosphorylation of the SMAD2 mediated by the *IC* complex, complex made up by the receptors and the integrins, seems to contribute the least for the previously stated outcome. It was also suggested from the experimental procedure that in general, the TRIMER_n complex undergoes more upregulation when influenced by the exacerbated joint behavior of two ways of crosstalk instead of one way of crosstalk independently or the combination of three types of crosstalk. This finding becomes more apparent the more intense the effect of each type of crosstalk is (e.g. the higher the synthesis rate of the receptors elicited by the integrins and the smaller the degradation rate of the receptors due to stabilization of the receptors, then higher the final concentration of the TRIMER will be when these two types of crosstalk are working simultaneously).

It is noteworthy that this model can be subjected to some improvements, which is particularly supported by the sensitivity analysis to which the parameters were subjected to in this work, concretely in Module I. In order to draw more accurate conclusions regarding the parameters of the model to which the nuclear TRIMER complex is more sensitive to, the sensitivity analysis methodology would benefit from the analysis of the nuclear TRIMER average concentration at different time points until the end of each simulation instead of solely recurring to the steady-state concentration. In addition, the formation rate of the TRIMER, $ktrim$, and the dissociation rate of the TRIMER, $kdiss$, could both be divided into

two different parameters: one rate applicable in the cytoplasm and one rate applicable in the nucleus. These modifications would allow to assess if there is a difference between both compartments in what concerns the sensitivity of the outcome to the formation and dissociation of the nuclear TRIMER.

Following the work presented in this document, it is pertinent to take as next steps the establishment and performance of a set of experiments to assess the crosstalk of the TGF-β pathway with the integrins for different cell shapes and, therefore, analyze simultaneously the effect of cell morphology and mechanotransduction on the signaling cascade. In addition, since there is evidence of an imbalance in released G/F-actin–regulated proteins due to cytoskeletal remodeling, which results in downstream genetic expression of TGF-β and BMP receptors, the model would benefit from the insertion of a module concerning the interactions with G/F-actin-regulated proteins, thus evaluating the effects of motile functions of the cell on the concentration of the TRIMER_n.

Lastly, some *in vitro* experiments with the TGF-β pathway should be carried in order to compare the *in vitro* results with the information obtained *in silico* for further validation and corroboration of this model.

REFERENCES

[1] D. A. Clark and R. Coker, "Molecules in focus - Transforming growth factor-beta (TGF-β)," *Int. J. Biochem. Cell Biol.*, vol. 30, no. 3, pp. 293–298, 1998.

[2] D. Nicklas and L. Saiz, "Computational modelling of Smad- mediated negative feedback and crosstalk in the TGF-β superfamily network," *J. R. Soc. Interface*, vol. 10, no. 86, pp. 1–17, 2013.

[3] J. C. Arciero, T. L. Jackson, and D. E. Kirschner, "A Mathematical Model Of Tumor-Immune Evasion And siRNA Treatment," *Discret. Contin. Dyn. Syst. B*, vol. 4, no. 1, pp. 39–58, 2004.

[4] S. Chung, C. R. Cooper, and B. A. Ogunnaike, "A control engineering approach to understanding the TGF-β paradox in cancer," *J. R. Soc. Interface*, vol. 9, pp. 1389–1397, 2011.

[5] Y. Kim and H. G. Othmer, "A Hybrid Model of Tumor – Stromal Interactions in Breast Cancer," *Bull Math Biol.*, vol. 75, no. 8, pp. 1304–1350, 2012.

[6] S. E. Wang *et al.*, "A mathematical model quantifies proliferation and motility effects of TGF-β on cancer cells," *Comput. Math. Methods Med.*, vol. 10, no. 1, pp. 71–83, 2009.

[7] G. Ascolani and P. Lio, "Modeling TGF-β in Early Stages of Cancer Tissue Dynamics," *PLoS One*, vol. 9, no. 2, p. e88533, 2014.

[8] R. Eftimie, M. Perez, and P. Buono, "Pattern formation in a nonlocal mathematical model for the multiple roles of the TGF-β pathway in tumour dynamics," *Math. Biosci.*, vol. 289, pp. 96–115, 2017.

[9] J. M. G. Vilar, R. Jansen, and C. Sander, "Signal Processing in the TGF-β Superfamily Ligand-Receptor Network," *PLoS Comput. Biol.*, vol. 2, no. 1, p. e3, 2006.

[10] K. Wegner *et al.*, "Biophysical Chemistry Dynamics and feedback loops in the transforming growth factor β signaling pathway," *Biophys. Chem.*, vol. 162, pp. 22–34, 2012.

[11] J. Claus, E. Friedmann, U. Klingmuller, R. Rannacher, and T. Szekeres, "Spatial aspects in the SMAD signaling pathway," *J. Math. Biol.*, vol. 67, pp. 1171–1197, 2012.

[12] S. Khatibi, H. Zhu, J. Wagner, C. W. Tan, J. H. Manton, and A. W. Burgess, "Mathematical model of TGF-β signalling : feedback coupling is consistent with signal switching," *BMC Syst. Biol.*, vol. 11, no. 48, pp. 1–15, 2017.

[13] A. Morshed, P. Dutta, and R. H. Dillon, "Mathematical modeling and numerical simulation of the TGF-β / Smad signaling pathway in tumor microenvironments," *Appl. Numer. Math.*, vol. 133, pp. 41–51, 2018.

[14] D. C. Clarke, M. D. Betterton, and X. Liu, "Systems theory of Smad signalling," *Syst. Biol.*, vol. 153, no. 6, pp. 412–424, 2006.

[15] Z. Zi and E. Klipp, "Constraint-Based Modeling and Kinetic Analysis of the Smad Dependent TGF-β Signaling Pathway," *PLoS One*, vol. 2, no. 9, p. e936, 2007.

[16] P. Melke, H. Jonsson, E. Pardali, P. Ten Dijke, and C. Peterson, "A Rate Equation Approach to Elucidate the Kinetics and Robustness of the TGF-β Pathway," *Biophys. J.*, vol. 91, no. 12, pp. 4368–4380, 2006.

[17] B. Schmierer, A. L. Tournier, P. A. Bates, and C. S. Hill, "Mathematical modeling identifies

Smad nucleocytoplasmic shuttling as a dynamic signal-interpreting system," *PNAS*, vol. 105, no. 18, pp. 6608–6613, 2008.

[18] S. Chung, F. L. Miles, R. A. Sikes, C. R. Cooper, M. C. Farach-carson, and B. A. Ogunnaike, "Quantitative Modeling and Analysis of the Transforming Growth Factor β Signaling Pathway," *Biophysj*, vol. 96, no. 5, pp. 1733–1750, 2009.

[19] Z. Zi *et al.*, "Quantitative analysis of transient and sustained transforming growth factor-β signaling dynamics," *Mol. Syst. Biol.*, vol. 7, no. 492, pp. 1–12, 2011.

[20] J. M. G. Vilar and L. Saiz, "Trafficking Coordinate Description of Intracellular Transport Control of Signaling Networks," *Biophysj*, vol. 101, no. 10, pp. 2315–2323, 2011.

[21] G. Cellière, G. Fengos, M. Hervé, and D. Iber, "The plasticity of TGF-β signaling," *BMC Syst. Biol.*, vol. 5, no. 184, pp. 1–13, 2011.

[22] M. J. Dalby, "Topographically induced direct cell mechanotransduction," *Med. Eng. Phys.*, vol. 27, no. 9, pp. 730–742, 2005.

[23] M. J. Dalby, N. Gadegaard, and R. O. C. Oreffo, "Harnessing nanotopography and integrin–matrix interactions to influence stem cell fate," *Nat. Publ. Gr.*, vol. 13, pp. 558–569, 2014.

[24] Y. K. Park and Y. Goda, "Integrins in synapse regulation," *Nat. Rev. Neurosci.*, vol. 17, no. 12, pp. 745–756, 2016.

[25] Y. A. Kadry and D. A. Calderwood, "Chapter 22: Structural and signaling functions of integrins," *Biochim. Biophys. Acta - Biomembr.*, vol. 1862, no. 5, pp. 1–15, 2020.

[26] M. Seo, T. J. Kang, C. H. Lee, A. Lee, and M. Noh, "HaCaT Keratinocytes and Primary Epidermal Keratinocytes Have Different Transcriptional Profiles of Cornified Envelope-Associated Genes to T Helper Cell Cytokines," *Biomol Ther*, vol. 20, no. 2, pp. 171–176, 2012.

[27] K. Grannas *et al.*, "Crosstalk between Hippo and TGF β : Subcellular Localization of YAP / TAZ / Smad Complexes," *J. Mol. Biol.*, vol. 427, no. 21, pp. 3407–3415, 2015.

[28] J. A. Hubbell, "Biomaterials in Tissue Engineering," *Nat. Biotechnol.*, vol. 13, no. 6, pp. 565–576, 1995.

[29] R. Lanza, R. Langer, J. Vacanti, and A. Atala, *Principles of Tissue Engineering*, 5th ed. Elsevier Science Publishing Co Inc, 2020.

[30] L. G. Griffith and G. Naughton, "Tissue engineering - Current challenges and expanding opportunities," *Science*, vol. 295, no. 5557. pp. 1009–1016, 2002.

[31] R. Langer and J. Vacanti, "Advances in Tissue Engineering," *J. Pediatr. Surg.*, vol. 51, no. 1, pp. 8–12, 2016.

[32] J. P. K. Armstrong *et al.*, "Artificial membrane-binding proteins stimulate oxygenation of stem cells during engineering of large cartilage tissue," *Nat. Commun.*, vol. 6, no. 7405, pp. 1–6, 2015.

[33] N. C. Institute, "Allogeneic Stem Cell Transplant." [Online]. Available: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/allogeneic-stem-cell-transplant. [Accessed: 05-Jan-2020].

[34] J. P. K. Armstrong and M. M. Stevens, "Using Remote Fields for Complex Tissue Engineering," *Trends Biotechnol.*, vol. 38, no. 3, pp. 254–263, 2019.

[35] J. Kopf, A. Petersen, G. N. Duda, and P. Knaus, "BMP2 and mechanical loading cooperatively regulate immediate early signalling events in the BMP pathway," *BMC Biol.*, vol. 10, no. 37, pp.

1–12, 2012.

[36] A. H. Reddi, “Bone morphogenetic proteins: From basic science to clinical applications,” *J. Bone Jt. Surg. - Ser. A*, vol. 83, pp. 1–6, 2001.

[37] P. De Biase and R. Capanna, “Clinical applications of BMPs.,” *Injury*, vol. 36 Suppl 3, no. 3, pp. S43–S46, 2005.

[38] G. S. Krishnakumar, A. Roffi, D. Reale, E. Kon, and G. Filardo, “Clinical application of bone morphogenetic proteins for bone healing: a systematic review,” *Int. Orthop.*, vol. 41, no. 6, pp. 1073–1083, 2017.

[39] O. P. Gautschi, S. P. Frey, and R. Zellweger, “Bone morphogenetic proteins in clinical applications,” *ANZ J. Surg.*, vol. 77, no. 8, pp. 626–631, 2007.

[40] Q. Wang, H. Zhang, H. Gan, H. Wang, Q. Li, and Z. Wang, “Application of combined porous tantalum scaffolds loaded with bone morphogenetic protein 7 to repair of osteochondral defect in rabbits *,” *Int. Orthop.*, vol. 42, no. 7, pp. 1437–1448, 2018.

[41] I. Marquetti and S. Desai, “Molecular modeling of bone morphogenetic protein for tissue engineering applications,” *IISE Annu. Conf. Expo 2018*, pp. 1759–1764, 2018.

[42] A. Moghaddam, C. Elleser, B. Biglari, A. Wentzensen, and G. Zimmermann, “Clinical application of BMP 7 in long bone non-unions,” *Arch Orthop Trauma Surg*, vol. 130, pp. 71–76, 2010.

[43] J. J. Worthington, J. E. Klementowicz, and M. A. Travis, “TGFβ : a sleeping giant awoken by integrins,” *Cell Press*, vol. 36, no. 1, pp. 47–54, 2011.

[44] R. Derynck *et al.*, “Nomenclature: vertebrate mediators of TGFβ family signals.,” *Cell*, vol. 87, no. 2, p. 173, 1996.

[45] S. Itoh, F. Itoh, M. J. Goumans, and P. Ten Dijke, “Signaling of transforming growth factor-β family members through Smad proteins,” *Eur. J. Biochem.*, vol. 267, no. 24, pp. 6954–6967, 2000.

[46] Q. Wei, T. L. M. Pohl, A. Seckinger, J. P. Spatz, and E. A. Cavalcanti-adam, “Regulation of integrin and growth factor signaling in biomaterials for osteodifferentiation,” *Beilstein J. Org. Chem.*, vol. 11, pp. 773–783, 2015.

[47] B. Hallberg and R. H. Palmer, “The role of the ALK receptor in cancer biology,” *Ann. Oncol.*, vol. 27, no. Supplement 3, pp. iii4–iii15, 2016.

[48] N. H. G. R. Institute, “Deoxyribonucleic Acid (DNA).” [Online]. Available: https://www.genome.gov/genetics-glossary/Deoxyribonucleic-Acid. [Accessed: 07-Nov-2020].

[49] R. Kamada *et al.*, “Metal-dependent Ser/Thr protein phosphatase PPM family: Evolution, structures, diseases and inhibitors,” *Pharmacol. Ther.*, vol. 215, p. 107622, 2020.

[50] Z. Zi, D. A. Chapnick, and X. Liu, “Dynamics of TGF-β / Smad signaling,” *FEBS Lett.*, vol. 586, no. 14, pp. 1921–1928, 2012.

[51] S. N. Sreenath, K. H. Cho, and P. Wellstead, “Modelling the dynamics of signalling pathways,” *Essays Biochem.*, vol. 45, pp. 1–28, 2008.

[52] J. J. Hughey and T. K. Lee, “Computational Modeling of Mammalian Signaling Networks,” *Wiley Interdiscip Rev Syst Biol Med*, vol. 2, no. 2, pp. 194–209, 2010.

[53] J. Scholma, S. Schivo, R. A. Urquidi Camacho, J. van de Pol, M. Karperien, and J. N. Post,

"Biological networks 101: Computational modeling for molecular biologists," *Gene*, vol. 533, no. 1, pp. 379–384, 2014.

[54] J. Zhang, X. J. Tian, Y. J. Chen, W. Wang, S. Watkins, and J. Xing, "Pathway crosstalk enables cells to interpret TGF-β duration," *npj Syst. Biol. Appl.*, vol. 4, no. 1, pp. 1–12, 2018.

[55] J. Cooper and F. G. Giancotti, "Integrin Signaling in Cancer: Mechanotransduction, Stemness, Epithelial Plasticity, and Therapeutic Resistance," *Cancer Cell*, vol. 35, no. 3, pp. 347–367, 2019.

[56] D. S. Harburger and D. A. Calderwood, "Erratum: Integrin signalling at a glance," *J. Cell Sci.*, vol. 122, no. 9, pp. 159–163, 2009.

[57] B. P. Eliceiri, "Integrin and growth factor receptor crosstalk," *Circ. Res.*, vol. 89, no. 12, pp. 1104–1110, 2001.

[58] M. Bachmann, S. Kukkurainen, V. P. Hytönen, and B. Wehrle-Haller, "Cell adhesion by integrins," *Physiol. Rev.*, vol. 99, no. 4, pp. 1655–1699, 2019.

[59] Nat Center for Biotechnology Information, "Calcium Ion." [Online]. Available: https://pubchem.ncbi.nlm.nih.gov/compound/Calcium-ion. [Accessed: 07-Nov-2020].

[60] S. R. Neves, P. T. Ram, and R. Iyenga, "G Protein Pathways," *Science.*, vol. 296, no. 5573, pp. 1636–1639, 2002.

[61] M. J. Dalby, "Cellular response to low adhesion nanotopographies," *Int. J. Nanomedicine*, vol. 2, no. 3, pp. 373–381, 2007.

[62] M. A. Schwartz and M. H. Ginsberg, "Networks and crosstalk: Integrin signalling spreads," *Nat. Cell Biol.*, vol. 4, no. 4, pp. 65–68, 2002.

[63] B. Cherksey, R. Durrie, P. E. Braun, and V. S. Sapirstein, "In vitro analysis of ion channels in periaxolemmal-myelin and white matter clathrin coated vesicles: Modulation by calcium and GTPγS," *Neurochem. Res.*, vol. 19, no. 8, pp. 1101–1106, 1994.

[64] S. Song *et al.*, "Small GTPases: Structure, biological function and its interaction with nanoparticles," *Asian J. Pharm. Sci.*, vol. 14, no. 1, pp. 30–39, 2019.

[65] J. M.Koenig, M. B. Joseph, and M. M. Mariscalco, "126-Normal and Abnormal Neutrophil Physiology in the Newborn," *Fetal Neonatal Physiol.*, vol. 2, pp. 1216–1229, 2017.

[66] UniProt, "UniProtKB - P12931 (SRC_HUMAN)." [Online]. Available: https://www.uniprot.org/uniprot/P12931. [Accessed: 07-Nov-2020].

[67] C. Margadant and A. Sonnenberg, "Integrin-TGF-β crosstalk in fibrosis, cancer and wound healing," *EMBO Rep.*, vol. 11, no. 2, pp. 97–105, 2010.

[68] C. Chandra Kumar, "Signaling by integrin receptors," *Oncogene*, vol. 17, pp. 1365–1373, 1998.

[69] H. Xu *et al.*, "Actin cytoskeleton mediates BMP2-Smad signaling via calponin 1 in preosteoblast under simulated microgravity," *Biochimie*, vol. 138, pp. 184–193, 2017.

[70] UniProt, "UniProtKB - P51911 (CNN1_HUMAN)." [Online]. Available: https://www.uniprot.org/uniprot/P51911. [Accessed: 05-Nov-2020].

[71] A. Sasaki, Y. Masuda, Y. Ohta, K. Ikeda, and K. Watanabe, "Filamin Associates with Smads and Regulates Transforming Growth Factor-β Signaling *," *J Biol Chem.*, vol. 276, no. 21, pp. 17871–17877, 2001.

[72] S. J. Winder, R. Kathryn, S. J. Winder, and K. R. Ayscough, “Actin-binding proteins,” *J. Cell Sci.*, vol. 118, no. 4, pp. 651–654, 2005.

[73] E. N. Olson and A. Nordheim, “Linking actin dynamics and gene transcription to drive cellular motile functions,” *Mol Cell Biol*, vol. 11, no. 5, pp. 353–365, 2010.

[74] P. Lenzi, G. Bocci, and G. Natale, “John Hunter and the origin of the the term ‘angiogenesis,’” *Angiogenesis*, vol. 19, pp. 255–256, 2016.

[75] A. L. Bauer, T. L. Jackson, Y. Jiang, and T. Rohlf, “Receptor cross-talk in angiogenesis : Mapping environmental cues to cell phenotype using a stochastic , Boolean signaling network model,” *J. Theor. Biol.*, vol. 264, pp. 838–846, 2010.

[76] G. W. Brodland, “How computational models can help unlock biological systems,” *Semin. Cell Dev. Biol.*, vol. 48, pp. 62–73, 2015.

[77] C. for C. A. and Modeling and Uc. Health, “VCell- Modeling & Analysis Software.” [Online]. Available: https://vcell.org/. [Accessed: 05-Jan-2020].

[78] I. I. Moraru *et al.*, “The Virtual Cell Modeling and Simulation Software Envirnoment,” *IET Syst. Biol.*, vol. 2, no. 5, pp. 352–362, 2009.

[79] X. Ma, Q. Wang, Y. Jiang, Z. Xiao, X. Fang, and Y. G. Chen, “Lateral diffusion of TGF-β type I receptor studied by single-molecule imaging,” *Biochem. Biophys. Res. Commun.*, vol. 356, no. 1, pp. 67–71, 2007.

[80] E. S. Welf, U. P. Naik, and B. A. Ogunnaike, “A spatial model for integrin clustering as a result of feedback between integrin activation and integrin binding,” *Biophys. J.*, vol. 103, no. 6, pp. 1379–1389, 2012.

[81] D. Lepzelter and M. H. Zaman, “Clustered diffusion of integrins,” *Biophys. J.*, vol. 99, no. 12, pp. L106–L108, 2010.

[82] C. E. Pierreux, F. J. Nicolás, and C. S. Hill, “Transforming Growth Factor β-Independent Shuttling of Smad4 between the Cytoplasm and Nucleus,” *Mol. Cell. Biol.*, vol. 20, no. 23, pp. 9041–9054, 2000.

[83] O. Harel, “The estimation of R2 and adjusted R2 in incomplete data sets using multiple imputation,” *J. Appl. Stat.*, vol. 36, no. 10, pp. 1109–1118, 2009.

[84] M. Sun, F. Spill, and M. H. Zaman, “A Computational Model of YAP/TAZ Mechanosensing,” *Biophys. J.*, vol. 110, no. 11, pp. 2540–2550, 2016.

APPENDIX A

Table A. 1 – Proposed mathematical models for the canonical TGF-β signaling pathway- part 1. Each model is classified according to model type, aims and conclusion, topics, dimension and length scale of the implemented model.

Article	Vilar et al. January 2006 [9]	Clarke et al. November 2006 [14]	Melke et al. December 2006 [16]	Zi & Klipp September 2007 [15]	Schmierer et al. May 2008 [17]
Model type	ODEs	ODEs	ODEs	ODEs	ODEs
Aims and Conclusion	Display versatile signal-processing unit behavior from coupling of signaling with receptor trafficking Alterations in receptor trafficking patterns seemingly promote tumor growth	Study mechanisms of Smad accumulation Unevenness of R-Smad phosphorylation and dephosphorylation rates is probably a key factor in the nuclear accumulation process	Build a rate equation model for the TGF-β pathway to assess its kinetics Model outcomes match experimental data Inhibitory feedback provides robustness to the TGF-β signaling system	Assess the dynamic relationship between receptor trafficking and activation of phosphorylated Smad TGF-β response regulated by the balance between clathrin dependent endocytosis and non-clathrin mediated endocytosis	Assess nature of Smad accumulation – static or dynamic Determine which reactions are rate-limiting and potential points for therapeutic intervention Smad nucleocytoplasmic shuttling is established as a dynamic network
Topics	Formation of ligand-receptor complexes Receptor trafficking – internalization and recycling Degradation of receptors and ligand-receptor complexes Receptors continuously produced via gene expression	Formation of ligand-receptor complexes Formation, phosphorylation, and dephosphorylation of R-Smad complexes Smad nucleocytoplasmic shuttling	Formation of ligand-receptor complexes Phosphorylation and formation of R-Smad complexes Smad nucleocytoplasmic shuttling Inhibitory feedback	Receptor trafficking – internalization and recycling Degradation of receptor and ligand-receptor complexes Phosphorylation and formation of R-Smad complexes Smad nucleocytoplasmic shuttling	Smad nucleocytoplasmic shuttling
Dimension	1D	1D	1D	1D	1D
Length Scale	Receptor level	Cell level	Cell level	Cell level	Cell level – specifically nucleus and cytoplasm

Table A. 2 – Proposed mathematical models for the canonical TGF-β signaling pathway- part 2. Each model is classified according to model type, aims and conclusion, topics, dimension and length scale of the implemented model.

Article	Chung et al. March 2009 [18]	Zi et al. May 2011 [19]	Vilar & Saiz November 2011 [20]	Cellière et al. November 2011 [21]
Model type	ODEs	ODEs	ODEs and PDEs	ODEs
Aims and Conclusion	Elucidate how modulation and regulation of TGF-β responses occur Study TGF-β role as a suppressor for tumors TGF-β as a promoter and suppressor of tumor growth is explained by four hypotheses: distinct thresholds for gene expression; fast degradation of signaling components; competitive inhibition by nuclear binding partner of phosphorylated Smad; presence of a TGF-β control system	Assess the effect of TGF-β ligand doses upon signaling dynamics One TGF-β stimulation pulse prompts a transient signaling response TGF-β pathway shows different levels of sensitivity to ligand doses at distinct time scales	Capture continuous change in properties of compartments using a trafficking coordinate Different TGF-β dynamics for cancerous and non-cancerous cells	Assess the least required changes in parameters to obtain a different type of signal processing for the TGF-β pathway Transient responses are promoted by inhibitory feedback and fast shuttling Alteration of single parameters is enough to alter signaling response
Topics	Formation of ligand-receptor complexes Receptor trafficking – internalization and recycling Phosphorylation, formation, dephosphorylation and dissociation of Smad complexes Smad nucleocytoplasmic shuttling Inhibitory feedback Degradation of receptor and ligand-receptor complexes Protein synthesis	Formation of ligand-receptor complexes Receptor trafficking – internalization and recycling Phosphorylation, formation, dephosphorylation and dissociation of Smad complexes Smad nucleocytoplasmic shuttling Inhibitory feedback Degradation of receptors	Formation of ligand-receptor complexes Receptors continuously produced via gene expression Receptor trafficking, internalization and recycling Phosphorylation, formation, dephosphorylation and dissociation of Smad complexes Smad nucleocytoplasmic shuttling Degradation of receptors	Formation of ligand-receptor complexes Phosphorylation and formation Smad complexes Smad nucleocytoplasmic shuttling Inhibitory feedback
Dimension	1D	1D	2D	1D
Length Scale	Cell level	Cell level	Cell level	Cell level

Table A. 3 – Proposed mathematical models for the canonical TGF-β signaling pathway- part 3. Each model is classified according to model type, aims and conclusion, topics, dimension and length scale of the implemented model.

Article	Wegner et al. January 2012 [10]	Claus et al. September 2012 [11]	Nicklas & Saiz May 2013 [2]
Model type	ODEs	ODEs and PDEs	ODEs
Aims and Conclusion	Pinpoint the importance of different feedback regulators Smad7 and Smurf2 possess the most potential to intervene in negative feedback	Compute quantitatively the effect of diffusion for signaling pathway Compare results obtained by ODEs and PDEs Effect of cell shape relevant for dendritic cells/ cells with extensions Effect of cell shape to regulate genes in ordinary shaped cells with small cytoplasmic domain not significant for gene expression	Assess distinct signaling patterns in TGF-β upon similar environmental conditions Negative feedback, both R-Smad channels (Smad 1/5/8 and Smad 2/3 channels) and stimulation by more than one ligand prompt peculiar signaling patterns
Topics	Formation of ligand-receptor complexes Receptor trafficking – internalization and recycling Phosphorylation, formation, dephosphorylation and dissociation of Smad complexes Smad nucleocytoplasmic shuttling Inhibitory feedback Degradation of receptors Protein synthesis	Formation of ligand-receptor complexes Phosphorylation, formation, dephosphorylation and dissociation of Smad complexes Smad nucleocytoplasmic shuttling Inhibitory feedback	Formation of ligand-receptor complexes Receptor trafficking – internalization Smad nucleocytoplasmic shuttling Phosphorylation and formation of Smad complexes Degradation of receptor, ligand-receptor complexes and R-Smads Protein synthesis Crosstalk between Smad 1/5/8 and Smad 2/3 channels mediated by ligand-induced negative feedback loop Inhibitory feedback for different cell lines- HaCaT, BAEC and C2C12
Dimension	1D	1D and 4D	1D
Length Scale	Cell level	Cell level	Cell level

Table A. 4 – Proposed mathematical models for the canonical TGF-β signaling pathway- part 4. Each model is classified according to model type, aims and conclusion, topics, dimension and length scale of the implemented model.

Article	Khatibi et al. April 2017 [12]	Morshed et al. November 2018 [13]
Model type	ODEs	ODEs
Aims and Conclusion	Elucidate outcomes from the TGF-β pathway TGF-β signaling pathway outcomes vary according to switching- long-term or short-term	Assess variety of TGF-β pathway outcomes according with different concentrations of TGF-β in the extracellular medium TGF-β promotes cancer cell growth and proliferation of tumors for a wide range of values of TGF-β concentration
Topics	Formation of ligand-receptor complexes Phosphorylation and formation of Smad complexes Inhibitory and positive feedback Degradation of receptors Protein synthesis	Formation of ligand-receptor complexes Receptor trafficking – internalization and recycling Phosphorylation, formation, dephosphorylation and dissociation of Smad complexes Smad nucleocytoplasmic shuttling Inhibitory feedback Degradation of receptors Protein synthesis
Dimension	1D	1D
Length Scale	Cell level	Cell level

APPENDIX B

Table B. 1 – Assessment of which species migrate from the cytoplasm into the nucleus according with different literature. A cross displayed below a certain species indicates that the species translocates from the cytoplasm into the nucleus in the model presented on the reference on the left side of the table.

Article	SMAD2	pSMAD2	TRIMER	SMAD4
Clarke et al. 2006 [14]	X	X	X	X
Schmierer et al. 2008 [17]			X	
Chung et al. 2009 [18]	X	X	X	X
Zi et al. 2011 [50]	X	X	X	X
Cellière et al. 2011 [21]	X	X	X	X
Claus et al. 2013 [11]	X	X	X	X
Nicklas & Saiz 2013 [2]	X	X	X	X
Khatibi et al. 2017 [12]	X	X	X	X
Morshed et al. 2018 [13]		X		

Table B. 2 – Assessment of which species migrate from the nucleus into the cytoplasm according with different literature. A cross displayed below a certain species indicates that the considered species translocates from the nucleus into the cytoplasm in the model presented on the reference on the left side of the table.

Article	SMAD2n	pSMAD2n	TRIMER_n	SMAD4n
Clarke et al. 2006 [14]	X			X
Schmierer et al. 2008 [17]				
Chung et al. 2009 [18]	X			X
Zi et al. 2011 [50]	X	X		X
Cellière et al. 2011 [21]	X	X		X
Claus et al. 2013 [11]	X			X
Nicklas & Saiz 2013 [2]	X			X
Khatibi et al. 2017 [12]	X	X		X
Morshed et al. 2018 [13]				

APPENDIX C

I) Module I Script

```
MathDescription {

 /* Automatically established by VCell ;                    */
Constant  _F_        96485.3321;
Constant  _F_nmol_          9.64853321E-5;
Constant  _K_GHK_1.0E-9;
Constant  _N_pmol_          6.02214179E11;
Constant  _PI_       3.141592653589793;
Constant  _R_        8314.46261815;
Constant  _T_        300.0;
Constant  Area_cyt 514.626;
Constant  Area_nuc78.24;
Constant  AreaPerUnitArea_Nuclear_membrane  1.0;
Constant  AreaPerUnitArea_Plasma_membrane   1.0;
Constant  AreaPerUnitVolume_m0       1.0;
 /* Diffusion rate and initial concentration of the C complex;                    */
Constant  C_diffusionRate     0.0;
Constant  C_init_uM           0.0;
 /* Automatically established by VCell ;                    */
Constant  K_millivolts_per_volt          1000.0;
 /* Catalytic constant of the phosphorylation of the SMAD2 mediated by the C complex;
*/
Constant  Kcat       800.0;
 /* Nuclear import rate of the pSMAD2, SMAD2, SMAD4 and TRIMER complex, respectively;
*/
Constant  kcn_pSMAD2        130.0;
Constant  kcn_SMAD2         55.0;
Constant  kcn_SMAD4         3.5;
Constant  kcn_Trimer        90.0;
 /* Degradation rate of the C complex ;                    */
Constant  kdeg       4.63E-4;
```

/* Dephosphorylation rate of the pSMAD2 in the nucleus; */

Constant Kdephos 0.08;

/* Dissociation rate of the TRIMER ; */

Constant Kdiss_r1_DirectHalf 0.00167;

/* Automatically established by VCell ; */

Constant Kdiss_r1_InverseHalf 0.0;

/* Automatically established by VCell ; */

Constant Km 2.0E-5;

Constant KMOLE 0.001660538783162726;

/* Nuclear export rates of the SMAD2 and SMAD4 ; */

Constant knc_SMAD2 40.0;

Constant knc_SMAD4 2.0;

/* Dephosphorylation rate of the pSMAD2 in the cytoplasm; */

Constant Kr 0.0;

/* Synthesis rate of the C complex ; */

Constant ksyn 1.467E-7;

/* Formation rate of the TRIMER; */

Constant Ktrim_r1_DirectHalf 0.167;

Constant Ktrim_r1_InverseHalf 0.0;

/* Diffusion rate and initial concentration of the pSMAD2 and pSMAD2n, respectively; */

Constant pSMAD2_diffusionRate 15.0;

Constant pSMAD2_init_uM 0.0;

Constant pSMAD2n_diffusionRate 15.0;

Constant pSMAD2n_init_uM 0.0;

/* Automatically established by VCell ; */

Constant R_beta 0.001;

/* Diffusion rate and initial concentration of the SMAD2, SMAD2n, SMAD4, SMAD4n, TRIMER and TRIMER_n, respectively; */

Constant SMAD2_diffusionRate 15.0;

Constant SMAD2_init_uM 0.008;

Constant SMAD2n_diffusionRate 15.0;

Constant SMAD2n_init_uM 0.093;

Constant SMAD4_diffusionRate 15.0;

Constant SMAD4_init_uM 0.008;

Constant SMAD4n_diffusionRate 15.0;

Constant SMAD4n_init_uM 0.016;

```
Constant  Trimer_diffusionRate          1.0;

Constant  Trimer_init_uM      0.0;

Constant  Trimern_diffusionRate          1.0;

Constant  Trimern_init_uM    0.0;

 /* Automatically established by VCell ;                                    */

Constant  V_cyt      310.56;

Constant  Voltage_m0          0.0;

Constant  VolumePerUnitVolume_cyt    1.0;

Constant  VolumePerUnitVolume_Nucleus          1.0;


 /* Species defined in the respective compartments: cytoplasm or nucleus ;                                    */

VolumeVariable   cell::C

VolumeVariable   cell::pSMAD2

VolumeVariable   nucleus::pSMAD2n

VolumeVariable   cell::SMAD2

VolumeVariable   nucleus::SMAD2n

VolumeVariable   cell::SMAD4

VolumeVariable   nucleus::SMAD4n

VolumeVariable   cell::Trimer

VolumeVariable   nucleus::Trimern


 /* Restriction of the C complex to an activation ring ;                                    */

Function  A            (C * !((((((x - 20.0) ^ 2.0) / (12.0 ^ 2.0)) + (((y - 20.0) ^ 2.0) / (12.0 ^ 2.0))) <= 1.0)));

 /* Functions for the boundary conditions in the nuclear membrane for the pSMAD2, SMAD2, SMAD4 and Trimer, respectively ;                                    */

Function  cell_nucleus_membrane::J_flux_pSMAD2          ( - (kcn_pSMAD2 / Area_nuc) * pSMAD2);

Function  cell_nucleus_membrane::J_flux_SMAD2          ((( - kcn_SMAD2 / Area_nuc) * SMAD2) + ((knc_SMAD2 /
Area_nuc) * SMAD2n));

Function  cell_nucleus_membrane::J_flux_SMAD4          ((( - kcn_SMAD4 / Area_nuc) * SMAD4) + ((knc_SMAD4 /
Area_nuc) * SMAD4n));

Function  cell_nucleus_membrane::J_flux_Trimer          ( - (kcn_Trimer / Area_nuc) * Trimer);

 /* Equations/terms of the equations which translate the variation of the concentration of the SMAD2, pSMAD2, TRIMER and SMAD4 in the cytoplasm, stated in the section "Rate" of "CompartmentSubDomain cell"  ;                                    */

Function  cell::J_r0   (( - Kcat * SMAD2 * A) + (Kr * pSMAD2));

Function  nucleus::J_r0_2       (Kdephos * pSMAD2n);

Function  cell::J_r1_DirectHalf          (((Ktrim_r1_DirectHalf * SMAD4) * pow(pSMAD2,2.0)) - (Kdiss_r1_DirectHalf * Trimer));
```

```
Function nucleus::J_r1_DirectHalf2 (((Ktrim_r1_DirectHalf * SMAD4n) * pow(pSMAD2n,2.0)) - (Kdiss_r1_DirectHalf * Trimern));

Function cell::J_r1_InverseHalf ((Kdiss_r1_DirectHalf * Trimer) - ((Ktrim_r1_DirectHalf * pow(pSMAD2,2.0)) * SMAD4));

Function nucleus::J_r1_InverseHalf2 ((Kdiss_r1_DirectHalf * Trimern) - ((Ktrim_r1_DirectHalf * pow(pSMAD2n,2.0)) * SMAD4n));

/* Automatically established by VCell ; */

Function cell::O0_SMAD2_tot SMAD2;

Function cell::O0_SMAD4_tot SMAD4;

Function cell::Size_cyt (VolumePerUnitVolume_cyt * vcRegionVolume('cell'));

Function EC::Size_m0 (AreaPerUnitVolume_m0 * vcRegionVolume('EC'));

Function nucleus::Size_Nucleus (VolumePerUnitVolume_Nucleus * vcRegionVolume('Nucleus'));

Function EC_cell_membrane::Size_Plasma_membrane (AreaPerUnitArea_Plasma_membrane * vcRegion-Area('EC_cell_membrane'));

Function EC_cell_membrane::sobj_cell1_EC0_size vcRegionArea('EC_cell_membrane');

Function EC_cell_membrane::sobj_Cell1_EC0_size vcRegionArea('EC_cell_membrane');

/* Function for the variation of the concentration of the C complex; */

Function cell::Var_complex (ksyn - (kdeg * A));

/* Automatically established by VCell ; */

Function cell::vobj_cell1_size vcRegionVolume('cell');

Function EC::vobj_EC0_size vcRegionVolume('EC');

Function nucleus::vobj_Nucleus2_size vcRegionVolume('Nucleus');


/* Definition of the PDE's which translate the variation of the concentration of species in the cytoplasm
*/

CompartmentSubDomain cell {

        BoundaryXm      Flux

        BoundaryXp      Flux

        BoundaryYm      Flux

        BoundaryYp      Flux

        PdeEquation SMAD4 {

                Rate    ((J_r1_InverseHalf - J_r1_DirectHalf) / 2.0);

                Diffusion SMAD4_diffusionRate;

                Initial SMAD4_init_uM;

        }

        PdeEquation SMAD2 {

                Rate    J_r0;
```

```
		Diffusion	SMAD2_diffusionRate;
		Initial		SMAD2_init_uM;
	}
	PdeEquation pSMAD2 {
		Rate		((J_r1_InverseHalf - J_r1_DirectHalf) - J_r0);
		Diffusion	pSMAD2_diffusionRate;
		Initial		pSMAD2_init_uM;
	}
	PdeEquation Trimer {
		Rate		(( - J_r1_InverseHalf + J_r1_DirectHalf) / 2.0);
		Diffusion	Trimer_diffusionRate;
		Initial		Trimer_init_uM;
	}
	PdeEquation C {
		Rate		Var_complex;
		Diffusion	C_diffusionRate;
		Initial		C_init_uM;
	}
}

 /*  Definition of the PDE's which translate the variation of concentration of species in the nucleus
*/

CompartmentSubDomain nucleus {
	BoundaryXm		Flux
	BoundaryXp		Flux
	BoundaryYm		Flux
	BoundaryYp		Flux
	PdeEquation Trimern {
		Rate		(( - J_r1_InverseHalf2 + J_r1_DirectHalf2) / 2.0);
		Diffusion	Trimern_diffusionRate;
		Initial		Trimern_init_uM;
	}
	PdeEquation SMAD2n {
		Rate		J_r0_2;
		Diffusion	SMAD2n_diffusionRate;
		Initial		SMAD2n_init_uM;
```

```
        }
        PdeEquation pSMAD2n {
                Rate        ((J_r1_InverseHalf2 - J_r1_DirectHalf2) - J_r0_2);
                Diffusion   pSMAD2n_diffusionRate;
                Initial     pSMAD2n_init_uM;
        }
        PdeEquation SMAD4n {
                Rate        ((J_r1_InverseHalf2 - J_r1_DirectHalf2) / 2.0);
                Diffusion   SMAD4n_diffusionRate;
                Initial     SMAD4n_init_uM;
        }
}

CompartmentSubDomain EC {
        BoundaryXm          Flux
        BoundaryXp          Flux
        BoundaryYm          Flux
        BoundaryYp          Flux
}

MembraneSubDomain EC cell {
        Name        EC_cell_membrane
        BoundaryXm          Flux
        BoundaryXp          Flux
        BoundaryYm          Flux
        BoundaryYp          Flux
        JumpCondition SMAD4 {
                InFlux      0.0;
                OutFlux     0.0;
        }
        JumpCondition SMAD2 {
                InFlux      0.0;
                OutFlux     0.0;
        }
        JumpCondition pSMAD2 {
```

```
                InFlux      0.0;
                OutFlux     0.0;
        }
        JumpCondition Trimer {
                InFlux      0.0;
                OutFlux     0.0;
        }
        JumpCondition C {
                InFlux      0.0;
                OutFlux     0.0;
        }
}

 /*  Boundary conditions at the nuclear membrane
*/

MembraneSubDomain cell nucleus {
        Name        cell_nucleus_membrane
        BoundaryXm          Flux
        BoundaryXp          Flux
        BoundaryYm          Flux
        BoundaryYp          Flux
        JumpCondition SMAD4 {
                InFlux      J_flux_SMAD4;
                OutFlux     0.0;
        }
        JumpCondition SMAD2 {
                InFlux      J_flux_SMAD2;
                OutFlux     0.0;
        }
        JumpCondition pSMAD2 {
                InFlux      J_flux_pSMAD2;
                OutFlux     0.0;
        }
        JumpCondition Trimer {
                InFlux      J_flux_Trimer;
                OutFlux     0.0;
```

```
        }
        JumpCondition SMAD2n {
                InFlux      0.0;
                OutFlux     - J_flux_SMAD2;
        }
        JumpCondition pSMAD2n {
                InFlux      0.0;
                OutFlux     - J_flux_pSMAD2;
        }
        JumpCondition Trimern {
                InFlux      0.0;
                OutFlux     - J_flux_Trimer;
        }
        JumpCondition SMAD4n {
                InFlux      0.0;
                OutFlux     - J_flux_SMAD4;
        }
        JumpCondition C {
                InFlux      0.0;
                OutFlux   0.0;
        }
}


}
```

II) Module II Script

MathDescription {

/* Automatically established by VCell ; */

Constant _F_ 96485.3321;

Constant _F_nmol_ 9.64853321E-5;

Constant _K_GHK_ 1.0E-9;

Constant _N_pmol_ 6.02214179E11;

Constant _PI_ 3.141592653589793;

Constant _R_ 8314.46261815;

Constant _T_ 300.0;

Constant Area_cyt 514.626;

Constant Area_nuc78.24;

Constant AreaPerUnitArea_Nuclear_membrane 1.0;

Constant AreaPerUnitArea_Plasma_membrane 1.0;

Constant AreaPerUnitVolume_m0 1.0;

/* Diffusion rate and initial concentration of the C complex, IC complex and integrins, respectively; */

Constant C_diffusionRate 0.0;

Constant C_init_uM 0.0;

Constant I_diffusionRate 0.0;

Constant I_init_uM 0.0;

Constant IC_diffusionRate 0.0;

Constant IC_init_uM 0.0;

/* Limit concentration of the integrins; */

Constant Imax 0.004;

/* Steps included in the functions for the increase of the synthesis rate of the C complex, for the increase of the catalytic constant of the phosphorylation of the SMAD2 mediated by the C complex, for the decrease of the degradation of the C complex and for the increase of the catalytic constant of the phosphorylation of the SMAD2 mediated by the IC complex, respectively ; */

Constant k_extra1 0.1;

Constant k_extra2 0.1;

Constant k_extra3 0.1;

Constant k_extra4 0.1;

/* Automatically established by VCell ; */

Constant K_millivolts_per_volt 1000.0;

/* Catalytic constant of the phosphorylation of the SMAD2 mediated by the C complex; */

Constant Kcat 800.0;

/* Nuclear import rate of the pSMAD2, SMAD2, SMAD4 and TRIMER complex, respectively; */

Constant kcn_pSMAD2 130.0;

Constant kcn_SMAD2 55.0;

Constant kcn_SMAD4 3.5;

Constant kcn_Trimer 90.0;

/* Degradation rate of the C complex ; */

Constant kdeg 4.63E-4;

/* Degradation rate of the integrins ; */

Constant kdeg_intg 4.6E-4;

/* Dephosphorylation rate of the pSMAD2 in the nucleus; */

Constant Kdephos 0.08;

/* Dissociation rate of the IC complex ; */

Constant kdiss_IC 0.016;

/* Dissociation rate of the TRIMER ; */

Constant Kdiss_r1_DirectHalf 0.00167;

/* Automatically established by VCell ; */

Constant Kdiss_r1_InverseHalf 0.0;

/* Formation rate of the IC complex ; */

Constant kform_IC 0.00146;

/* Automatically established by VCell ; */

Constant Km 2.0E-5;

Constant KMOLE 0.001660538783162726;

/* Nuclear export rates of the SMAD2 and the SMAD4 ; */

Constant knc_SMAD2 40.0;

Constant knc_SMAD4 2.0;

/* Dephosphorylation rate of the pSMAD2 in the cytoplasm; */

Constant Kr 0.0;

/* Synthesis rate of the C complex ; */

Constant ksyn 1.467E-7;

/* Synthesis rate of the integrins ; */

Constant ksyn_intg 1.5E-6;

/* Formation rate of the TRIMER; */

Constant Ktrim_r1_DirectHalf 0.167;

```
Constant  Ktrim_r1_InverseHalf          0.0;
 /* Diffusion rate and initial concentration of the pSMAD2 and pSMAD2n, respectively;                  */
Constant  pSMAD2_diffusionRate        15.0;
Constant  pSMAD2_init_uM   0.0;
Constant  pSMAD2n_diffusionRate       15.0;
Constant  pSMAD2n_init_uM 0.0;
 /* Automatically established by VCell ;                  */
Constant  R_beta    0.001;
 /* Diffusion rate and initial concentration of the SMAD2, SMAD2n, SMAD4, SMAD4n, TRIMER and TRIMER_n, respectively;
*/
Constant  SMAD2_diffusionRate         15.0;
Constant  SMAD2_init_uM     0.008;
Constant  SMAD2n_diffusionRate        15.0;
Constant  SMAD2n_init_uM   0.093;
Constant  SMAD4_diffusionRate         15.0;
Constant  SMAD4_init_uM     0.008;
Constant  SMAD4n_diffusionRate        15.0;
Constant  SMAD4n_init_uM   0.016;
Constant  Trimer_diffusionRate          1.0;
Constant  Trimer_init_uM      0.0;
Constant  Trimern_diffusionRate         1.0;
Constant  Trimern_init_uM    0.0;
 /* Automatically established by VCell ;                  */
Constant  V_cyt      310.56;
Constant  Voltage_m0          0.0;
Constant  VolumePerUnitVolume_cyt   1.0;
Constant  VolumePerUnitVolume_Nucleus          1.0;

 /* Species defined in the respective compartments: cytoplasm or nucleus ;                  */
VolumeVariable   cell::C
VolumeVariable   cell::I
VolumeVariable   cell::IC
VolumeVariable   cell::pSMAD2
VolumeVariable   nucleus::pSMAD2n
VolumeVariable   cell::SMAD2
VolumeVariable   nucleus::SMAD2n
```

```
VolumeVariable   cell::SMAD4

VolumeVariable   nucleus::SMAD4n

VolumeVariable   cell::Trimer

VolumeVariable   nucleus::Trimern


 /* Restriction of the C complex, integrins and IC complex to an activation ring ;                 */

Function  A          (C * !((((((x - 20.0) ^ 2.0) / (12.0 ^ 2.0)) + (((y - 20.0) ^ 2.0) / (12.0 ^ 2.0))) <= 1.0)));

Function  A2         (I * !((((((x - 20.0) ^ 2.0) / (12.0 ^ 2.0)) + (((y - 20.0) ^ 2.0) / (12.0 ^ 2.0))) <= 1.0)));

Function  A3         (IC * !((((((x - 20.0) ^ 2.0) / (12.0 ^ 2.0)) + (((y - 20.0) ^ 2.0) / (12.0 ^ 2.0))) <= 1.0)));

 /* Functions for the boundary conditions in the nuclear membrane for the pSMAD2, SMAD2, SMAD4 and TRIMER, respectively ;                 */

Function  cell_nucleus_membrane::J_flux_pSMAD2          ( - (kcn_pSMAD2 / Area_nuc) * pSMAD2);

Function  cell_nucleus_membrane::J_flux_SMAD2           ((( - kcn_SMAD2 / Area_nuc) * SMAD2) + ((knc_SMAD2 / Area_nuc) * SMAD2n));

Function  cell_nucleus_membrane::J_flux_SMAD4           ((( - kcn_SMAD4 / Area_nuc) * SMAD4) + ((knc_SMAD4 / Area_nuc) * SMAD4n));

Function  cell_nucleus_membrane::J_flux_Trimer          ( - (kcn_Trimer / Area_nuc) * Trimer);

 /* Equations/terms of the equations which translate the variation of the concentration of the SMAD2, pSMAD2, TRIMER and the SMAD4 in the cytoplasm, stated in the section "Rate" of "CompartmentSubDomain cell" ;                 */

Function  cell::J_r0   (( - kcat3 * SMAD2 * A) + (Kr * pSMAD2) + ( - kcat5 * A3 * SMAD2));

Function  nucleus::J_r0_2       (Kdephos * pSMAD2n);

Function  cell::J_r1_DirectHalf             (((Ktrim_r1_DirectHalf * SMAD4) * pow(pSMAD2,2.0)) - (Kdiss_r1_DirectHalf * Trimer));

Function  nucleus::J_r1_DirectHalf2         (((Ktrim_r1_DirectHalf * SMAD4n) * pow(pSMAD2n,2.0)) - (Kdiss_r1_DirectHalf * Trimern));

Function  cell::J_r1_InverseHalf            ((Kdiss_r1_DirectHalf * Trimer) - ((Ktrim_r1_DirectHalf * pow(pSMAD2,2.0)) * SMAD4));

Function  nucleus::J_r1_InverseHalf2        ((Kdiss_r1_DirectHalf * Trimern) - ((Ktrim_r1_DirectHalf * pow(pSMAD2n,2.0)) * SMAD4n));

 /* Function for the catalytic constant of the phosphorylation of the SMAD2 mediated by the C complex with an increase step of k_extra2 ;                 */

Function  kcat2       (Kcat * (1.0 + ((k_extra2 * A2) / Imax)));

 /* Restriction of the function for the catalytic constant of the phosphorylation of the SMAD2 mediated by the C complex to an activation ring ;                 */

Function  kcat3       (kcat2 * !((((((x - 20.0) ^ 2.0) / (12.0 ^ 2.0)) + (((y - 20.0) ^ 2.0) / (12.0 ^ 2.0))) <= 1.0)));

 /* Function for the catalytic constant of the phosphorylation of the SMAD2 mediated by the IC complex with an increase step of k_extra4 ;                 */

Function  kcat4       (Kcat * (1.0 + ((k_extra4 * A2) / Imax)));

 /* Restriction of the function for the catalytic constant of the phosphorylation of the SMAD2 mediated by the IC complex to an activation ring ;                 */
```

```
Function kcat5 (kcat4 * !((((((x - 20.0) ^ 2.0) / (12.0 ^ 2.0)) + (((y - 20.0) ^ 2.0) / (12.0 ^ 2.0))) <= 1.0)));
/* Function for the degradation rate of the C complex with a decrease step of k_extra3 ; */
Function kdeg2 (kdeg * (1.0 - ((k_extra3 * A2) / Imax)));
/* Restriction of the function for the degradation rate of the C complex to an activation ring ; */
Function kdeg3 (kdeg2 * !((((((x - 20.0) ^ 2.0) / (12.0 ^ 2.0)) + (((y - 20.0) ^ 2.0) / (12.0 ^ 2.0))) <= 1.0)));
/* Restriction of the degradation rate of the integrins to an activation ring ; */
Function kdeg_intg2 (kdeg_intg * !((((((x - 20.0) ^ 2.0) / (12.0 ^ 2.0)) + (((y - 20.0) ^ 2.0) / (12.0 ^ 2.0))) <= 1.0)));
/* Restriction of the dissociation rate of the IC complex to an activation ring ; */
Function kdiss_IC2 (kdiss_IC * !((((((x - 20.0) ^ 2.0) / (12.0 ^ 2.0)) + (((y - 20.0) ^ 2.0) / (12.0 ^ 2.0))) <= 1.0)));
/* Restriction of the formation rate of the IC complex to an activation ring ; */
Function kform_IC2 (kform_IC * !((((((x - 20.0) ^ 2.0) / (12.0 ^ 2.0)) + (((y - 20.0) ^ 2.0) / (12.0 ^ 2.0))) <= 1.0)));
/* Function for the synthesis rate of the C complex with an increase step of k_extra1 ; */
Function ksyn2 (ksyn * (1.0 + ((k_extra1 * A2) / Imax)));
/* Restriction of the function for the synthesis rate of the C complex to an activation ring ; */
Function ksyn3 (ksyn2 * !((((((x - 20.0) ^ 2.0) / (12.0 ^ 2.0)) + (((y - 20.0) ^ 2.0) / (12.0 ^ 2.0))) <= 1.0)));
/* Restriction of the synthesis rate of the integrins to an activation ring ; */
Function ksyn_intg2 (ksyn_intg * !((((((x - 20.0) ^ 2.0) / (12.0 ^ 2.0)) + (((y - 20.0) ^ 2.0) / (12.0 ^ 2.0))) <= 1.0)));
/* Automatically established by VCell ; */
Function cell::O0_SMAD2_tot SMAD2;
Function cell::O0_SMAD4_tot SMAD4;
Function cell::Size_cyt (VolumePerUnitVolume_cyt * vcRegionVolume('cell'));
Function EC::Size_m0 (AreaPerUnitVolume_m0 * vcRegionVolume('EC'));
Function nucleus::Size_Nucleus (VolumePerUnitVolume_Nucleus * vcRegionVolume('Nucleus'));
Function EC_cell_membrane::Size_Plasma_membrane (AreaPerUnitArea_Plasma_membrane * vcRegion-
Area('EC_cell_membrane'));
Function EC_cell_membrane::sobj_cell1_EC0_size vcRegionArea('EC_cell_membrane');
Function EC_cell_membrane::sobj_Cell1_EC0_size vcRegionArea('EC_cell_membrane');
/* Functions for the variation of the concentration of the C complex, IC complex and integrins, respectively; */
Function cell::Var_complex (ksyn3 - (kdeg3 * A) + (kdiss_IC2 * A3) - (kform_IC2 * A2 * A));
Function cell::Var_IC ((kform_IC2 * A2 * A) - (kdiss_IC2 * A3));
Function cell::Var_intg (ksyn_intg2 - (kdeg_intg2 * A2) + (kdiss_IC2 * A3) - (kform_IC2 * A2 * A));
/* Automatically established by VCell ; */
Function cell::vobj_cell1_size vcRegionVolume('cell');
Function EC::vobj_EC0_size vcRegionVolume('EC');
Function nucleus::vobj_Nucleus2_size vcRegionVolume('Nucleus');
```

```
/*  Definition of the PDE's which translate the variation of the concentration of species in the cytoplasm
*/

CompartmentSubDomain cell {

        BoundaryXm      Flux

        BoundaryXp      Flux

        BoundaryYm      Flux

        BoundaryYp      Flux

        PdeEquation SMAD4 {

                Rate        ((J_r1_InverseHalf - J_r1_DirectHalf) / 2.0);

                Diffusion   SMAD4_diffusionRate;

                Initial     SMAD4_init_uM;

        }

        PdeEquation SMAD2 {

                Rate        J_r0;

                Diffusion   SMAD2_diffusionRate;

                Initial     SMAD2_init_uM;

        }

        PdeEquation pSMAD2 {

                Rate        ((J_r1_InverseHalf - J_r1_DirectHalf) - J_r0);

                Diffusion   pSMAD2_diffusionRate;

                Initial     pSMAD2_init_uM;

        }

        PdeEquation Trimer {

                Rate        (( - J_r1_InverseHalf + J_r1_DirectHalf) / 2.0);

                Diffusion   Trimer_diffusionRate;

                Initial     Trimer_init_uM;

        }

        PdeEquation C {

                Rate        Var_complex;

                Diffusion   C_diffusionRate;

                Initial     C_init_uM;

        }

        PdeEquation I {

                Rate        Var_intg;

                Diffusion   I_diffusionRate;

                Initial     I_init_uM;
```

```
        }
        PdeEquation IC {
                Rate        Var_IC;
                Diffusion   IC_diffusionRate;
                Initial     IC_init_uM;
        }
}

/*  Definition of the PDE's which translate the variation of the concentration of species in the nucleus */
CompartmentSubDomain nucleus {
        BoundaryXm          Flux
        BoundaryXp          Flux
        BoundaryYm          Flux
        BoundaryYp          Flux
        PdeEquation Trimern {
                Rate        (( - J_r1_InverseHalf2 + J_r1_DirectHalf2) / 2.0);
                Diffusion   Trimern_diffusionRate;
                Initial     Trimern_init_uM;
        }
        PdeEquation SMAD2n {
                Rate        J_r0_2;
                Diffusion   SMAD2n_diffusionRate;
                Initial     SMAD2n_init_uM;
        }
        PdeEquation pSMAD2n {
                Rate        ((J_r1_InverseHalf2 - J_r1_DirectHalf2) - J_r0_2);
                Diffusion   pSMAD2n_diffusionRate;
                Initial     pSMAD2n_init_uM;
        }
        PdeEquation SMAD4n {
                Rate        ((J_r1_InverseHalf2 - J_r1_DirectHalf2) / 2.0);
                Diffusion   SMAD4n_diffusionRate;
                Initial     SMAD4n_init_uM;
        }
}
```

```
CompartmentSubDomain EC {
        BoundaryXm      Flux
        BoundaryXp      Flux
        BoundaryYm      Flux
        BoundaryYp      Flux
}

MembraneSubDomain EC cell {
        Name    EC_cell_membrane
        BoundaryXm      Flux
        BoundaryXp      Flux
        BoundaryYm      Flux
        BoundaryYp      Flux
        JumpCondition SMAD4 {
                InFlux  0.0;
                OutFlux 0.0;
        }
        JumpCondition SMAD2 {
                InFlux  0.0;
                OutFlux 0.0;
        }
        JumpCondition pSMAD2 {
                InFlux  0.0;
                OutFlux 0.0;
        }
        JumpCondition Trimer {
                InFlux  0.0;
                OutFlux 0.0;
        }
        JumpCondition C {
                InFlux  0.0;
                OutFlux 0.0;
        }
        JumpCondition I {
```

```
                InFlux      0.0;
                OutFlux     0.0;
        }
        JumpCondition IC {
                InFlux      0.0;
                OutFlux     0.0;
        }
}

 /*  Boundary conditions at the nuclear membrane                                    */
MembraneSubDomain cell nucleus {
        Name        cell_nucleus_membrane
        BoundaryXm          Flux
        BoundaryXp          Flux
        BoundaryYm          Flux
        BoundaryYp          Flux
        JumpCondition SMAD4 {
                InFlux      J_flux_SMAD4;
                OutFlux     0.0;
        }
        JumpCondition SMAD2 {
                InFlux      J_flux_SMAD2;
                OutFlux     0.0;
        }
        JumpCondition pSMAD2 {
                InFlux      J_flux_pSMAD2;
                OutFlux     0.0;
        }
        JumpCondition Trimer {
                InFlux      J_flux_Trimer;
                OutFlux     0.0;
        }
        JumpCondition SMAD2n {
                InFlux      0.0;
                OutFlux     - J_flux_SMAD2;
```

```
        }
        JumpCondition pSMAD2n {
                InFlux      0.0;
                OutFlux     - J_flux_pSMAD2;
        }
        JumpCondition Trimern {
                InFlux      0.0;
                OutFlux     - J_flux_Trimer;
        }
        JumpCondition SMAD4n {
                InFlux      0.0;
                OutFlux     - J_flux_SMAD4;
        }
        JumpCondition C {
                InFlux      0.0;
                OutFlux   0.0;
        }
        JumpCondition I {
                InFlux      0.0;
                OutFlux   0.0;
        }
        JumpCondition IC {
                InFlux      0.0;
                OutFlux   0.0;
        }
}


}
```