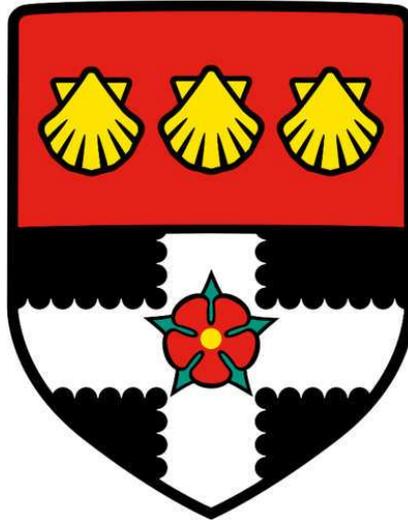


UNIVERSITY OF READING



**Modelling of collagen receptor
clustering and signalling**

A thesis submitted for the degree of Doctor of Philosophy

Chukiat Tantiwong

School of Biological Sciences

April 2023

Declaration

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

Signed: Chukiat Tantiwong

Date: 24 April 2023

Acknowledgement

I am deeply grateful to my advisors, Prof. Jonathan M Gibbins, Dr Joanne Dunster, Assist. Prof. Rachel Cavill, and Prof. Johan Heemskerk, for their unwavering support and guidance throughout my PhD research, and Dr Christoph Wierling and Prof. Marcus Tindall for their helpful suggestion. I also extend my thanks to the funding, Marie Skłodowska-Curie Innovative Training Network (ITN), and institutions for their contributions to my research. Special thanks to the participants, staff, and facilities at the University of Reading, Maastricht University, and Alacris Theranostics GmbH. I also appreciate my colleagues, Zahra Maqsood, Hilaire Cheung, and other TAPAS ESRs, for their kind support, both academically and emotionally. I am thankful to my family and friends for their love and encouragement. Thank you all for your invaluable contributions to my thesis.

Abstract

Platelets play a crucial role in haemostasis and thrombosis, and their activation is regulated by various agonists and intracellular signalling pathways. This PhD thesis aims to investigate the complex interplay between platelet receptors, signalling pathways, and intracellular calcium dynamics using mathematical modelling approaches. We present an agent-based model (ABM) to investigate the dimerisation of GPVI and the role of glycolipid-enriched raft-like domains in regulating receptor diffusion. We introduce an ordinary differential equation (ODE) model to unravel the regulatory mechanisms of GPVI signalling from Syk phosphorylation to releasing inositol trisphosphate (IP₃) into the platelet cytosol via LAT signalosome. To fulfil the signalling complexity more completely than has previously been possible, a mathematical model of phosphoinositide (PI) metabolism in human platelets in response to GPVI activation was developed and calibrated against experimental data capturing transient time-course changes in phosphoinositide positional isomers. Finally, two machine learning models, a non-linear autoregressive network with exogenous inputs (NARX) and a partial least square (PLS) regression model, were developed to investigate how different agonists and inhibitors impact on intracellular calcium dynamics in platelets.

The findings from this thesis provide valuable insights into the complex regulatory mechanisms of platelet activation and intracellular calcium dynamics and how they are influenced by various agonists and inhibitors. These mathematical modelling approaches have the potential to be used in the development of simulation frameworks for studying spatiotemporal concentrations of ligands and inhibitors in platelets. The results of this thesis may have implications for the development of targeted therapies for platelet-related diseases, such as thrombosis and haemostasis disorders. The models also have wide applicability to other cell systems. Overall, this PhD thesis contributes to our understanding of platelet biology and provides novel insights into platelet signalling and calcium dynamics using mathematical modelling approaches.

List of abbreviations

2-APB	2-Aminoethyldiphenyl borate
ABC	Approximate Bayesian Computation
ABM	Agent-Based Modelling
absdev	Average absolute deviation from straight line
ACD	Acid citrate dextrose
ADP	Adenosine diphosphate
AI	Apyrase and Indomethacin
AIC	Akaike Information Criteria
AICc	Akaike Information Criteria with a correction of small sample size
ATP	Adenosine triphosphate
BCA	Bicinchoninic acid
Btk	Bruton's tyrosine kinase
c-Cbl	Ubiquitin E3 ligase
CD	Confined Domain
Col	Collagen
CRP	Collagen-related peptide
CRP-XL	Cross-linked collagen-related peptide
Cyt	Cytosol
DAG	Diacylglycerol
DMSO	Dimethyl sulfoxide
DTS	Dense Tubular System
EGFR	Epidermal growth factor receptor

EGTA / EG	Ethylene-bis(oxyethylenitrilo)tetraacetic acid / egtazic acid
EHR	electronic health record
ER	Endoplasmic Reticulum
ErbB	Epidermal growth factor
FcR γ	Fc receptor γ -chain
GPCR	G protein-coupled receptor
GPVI	collagen receptor glycoprotein VI
GSK-A1	PI4KA inhibitor
HER2	Human epidermal growth factor receptor 2
IP1 / InsP1	Inositol monophosphate
IP3 / InsP3	Inositol trisphosphate
IP3R	Inositol trisphosphate (IP ₃) Receptor
IPA	Isopropanol
ITAM	Immunoreceptor tyrosine-based activation motif
LAT	Linker for activation of T cells
MC	Monte-Carlo
MCS	Membrane Contact Site
MD	Molecular Dynamics
MLP	Multilayer perceptron
MSD	Mean Square Distance
MSE	Mean Square Error
NARX	Non-linear AutoRegressive network with eXogenous inputs
NF- κ B	Nuclear factor κ -light-chain-enhancer of activated B cells

NN	Neural Network
NO	Nitric Oxide
OAT	One-factor-At-a-Time
OCRL	Inositol polyphosphate-5-phosphatase
ODE	Ordinary Differential Equation
PAR	Protease-activated receptors
PCA	Principal Component Analysis
PDE	Partial Differential Equation
PECAM1	Platelet and endothelial cell adhesion molecule 1
PGI ₂	Prostaglandin I ₂ (Prostacyclin)
PH	Pleckstrin homology
PI	Phosphoinositide
PI3K	Phosphoinositide 3-kinase
PI4KA	Phosphatidylinositol 4-kinase A
PI4P / PtdIns4P	Phosphatidylinositol-4-phosphate
PIP ₂ / PtdInsP ₂	Phosphatidylinositol bisphosphate
PIP ₃ / PtdInsP ₃	Phosphatidylinositol trisphosphate
PITP α	Phosphatidyl-inositol transfer protein type α
PKC	Protein kinase C
PLC β	Phospholipase C β
PLC γ 2	Phospholipase C γ 2
PLS	Partial Least Square
PLSR	Partial Least Square Regression

PM	Plasma Membrane
PMCA	Plasma membrane Ca^{2+} ATPase
PRP	Platelet-rich plasma
PS	Phosphatidylserine
PtdIns(4,5) P_2 -FP	PtdIns(4,5) P_2 α -fluorovinylphosphonate
PTEN	Protein phosphatase and tensin homolog
ReLU	Rectified Linear Unit
SDS	Sodium dodecyl sulfate
SERCA	Sarcoendoplasmic reticulum (SR) calcium transport ATPase
SFKs	Src-family kinases
siRNA	Small interfering ribonucleic acid
SMOTE	Synthetic Minority Oversampling Technique
SOCE	Store-operated calcium entry
SRM	Selected reaction monitoring
SSE	Sum Squared due to Error
STIM	Stromal interaction molecule
Syk	Spleen tyrosine kinase
tansig	Hyperbolic tangent sigmoid
Thap	Thapsigargin
Thr	Thrombin
TKLR	Tyrosine kinase-linked receptor
tmax	Time point that calcium level reaches the maximal value
TNF	Tumor necrosis factor

TP	Thromboxane receptor
Tris	Tris(hydroxymethyl)-aminomethane
TULA-2	T-cell ubiquitin ligand-2
TXA2	Thromboxane A ₂
Vcell	Virtual Cell computational modelling and simulation software
VWF	von Willebrand factor
ylast	Terminal value of calcium level
YU142670	OCRL inhibitor

Table of Contents

	Page
Declaration	i
Acknowledgement	ii
Abstract	iii
List of abbreviations	iv
Chapter 1 General Introduction	1
Chapter 2 An agent-based approach for modelling and simulation of glycoprotein VI receptor diffusion, localisation and dimerisation in platelet lipid rafts	35
Chapter 3 The importance of regulation on controlling GPVI signal transduction: a systems biology approach	74
Chapter 4 Development of a mathematical model of the platelet phosphoinositide metabolism	131
Chapter 5 Additive benefit of modelling of platelet calcium signalling by combined recurrent neural network and partial least squares analyses	173
Chapter 6 General Discussion	214

Chapter 1

General Introduction

In this chapter, the general understanding of platelet biology will be introduced, from platelet formation and lifetime to the control of activation, thrombus formation, and coagulation. Then we will present our main focus in this thesis, platelet collagen receptor, GPVI, its structure and signalling. Some techniques that can be used to model GPVI movement, interaction, and signalling will be reviewed. The problem of platelet cytosolic calcium and proposed machine-learning approaches will be discussed. Finally, the aims and outline of this thesis will be summarised at the end of this chapter.

General Introduction

Blood platelets are required for haemostasis and do also play an important role in arterial thrombosis¹. Injury of a blood vessel or rupture of an atherosclerotic plaque leads to platelet adhesion to matrix proteins such as collagens, which lie underneath the endothelial layer². Subsequent receptor-induced intracellular signalling, followed by secretion of bioactive molecules, attracts new platelets which results in the formation of platelet aggregates. Simultaneously with the formation of this platelet plug or thrombus, activation of the coagulation cascade leads to generation of a fibrin mesh, which wraps around and supports the thrombus. These events prevent blood from flowing out of the vessel, thus stopping a bleeding, or in case of thrombosis can obstruct the vessel with tissue ischemia as a result.

Current antiplatelet drugs, aiming to prevent a next thrombotic event (transient ischemic attack, stroke or heart infarction), all have a bleeding side effect, likely because of over-inhibition of the platelets³. The collagen receptor glycoprotein VI (GPVI), only expressed on platelets, is a new interesting target for antithrombotic therapy, given that mouse and patient data report no severe bleeding in the absence of this receptor⁴. For this thesis, using a variety of modelling approaches, I have been interested in the potential of different ways to modulate the activation of platelet GPVI, *i.e.*, by interfering with receptor clustering, by blockage of GPVI-dependent protein or lipid kinases, or by inhibiting Ca²⁺ entry channels.

My work on GPVI-dependent platelet activation pathways thereby aims to fill a gap between physiological and modelling interfering approaches. Mathematical and neural network modelling are to be used to predict effects of regulatory perturbations, to identify promising therapeutic targets, and to design optimised experiments. In the present chapter, I introduce for this thesis relevant background information regarding platelet and GPVI

functions and on potential ways of modelling these.

Platelets formation and life time

In comparison to red and white blood cells, platelets are smallest anucleate cells in the blood, with a diameter of 2-3 μm and a volume of 4-8 fL ⁵. For a healthy subject, the platelet count is 150,000-400,000 per microliter of blood⁶, meaning that platelets given their small size comprise only a small fraction of the blood volume⁷. Upon activation, platelets change in shape from a biconvex discoid structure (maintained by cytoskeleton elements)⁸ to a spread appearance with filopodia and lamellipodia⁹.

Multiple signalling and adhesive glycoprotein receptors are present on the platelet plasma membrane, the most relevant of which, for this thesis are the collagen receptors, GPVI and integrin $\alpha 2\beta 1$, the shear-sensitive receptor GPIb-V-IX, and the fibrinogen receptor integrin $\alpha \text{IIb}\beta 3$ ¹⁰. Beneath the plasma membrane, the actin-myosin filament system controls platelet shape^{11,12}. Several organelles are present, including mitochondria,¹³ storage granules for secretion¹⁴, and the dense tubular system (DTS) with protein processing and Ca^{2+} storage functions¹⁵⁻¹⁷. The DTS is derived from the endoplasmic reticulum of megakaryocytes¹⁸.

Megakaryocytes in the bone marrow¹⁹ and lungs²⁰ are able to produce multiple proplatelets during their life time²¹⁻²³. As a megakaryocyte matures, it undergoes the process of endomitosis²⁴, in which the nucleus but not the cytoplasm divides, resulting in a large multinucleated cell (Figure 1A). So-called proplatelets, as protrusions from a megakaryocyte, enter into blood stream^{25,26} in a shear stress-dependent way²⁷. These proplatelets divide into platelets by a microtubular twist mechanism²⁸. After an average lifetime of 8-9 days, platelets in the circulation undergo an apoptotic pathway or become desialylated, resulting in their removal by liver hepatocytes or by splenic macrophages (Figure 1C)^{7,19}.

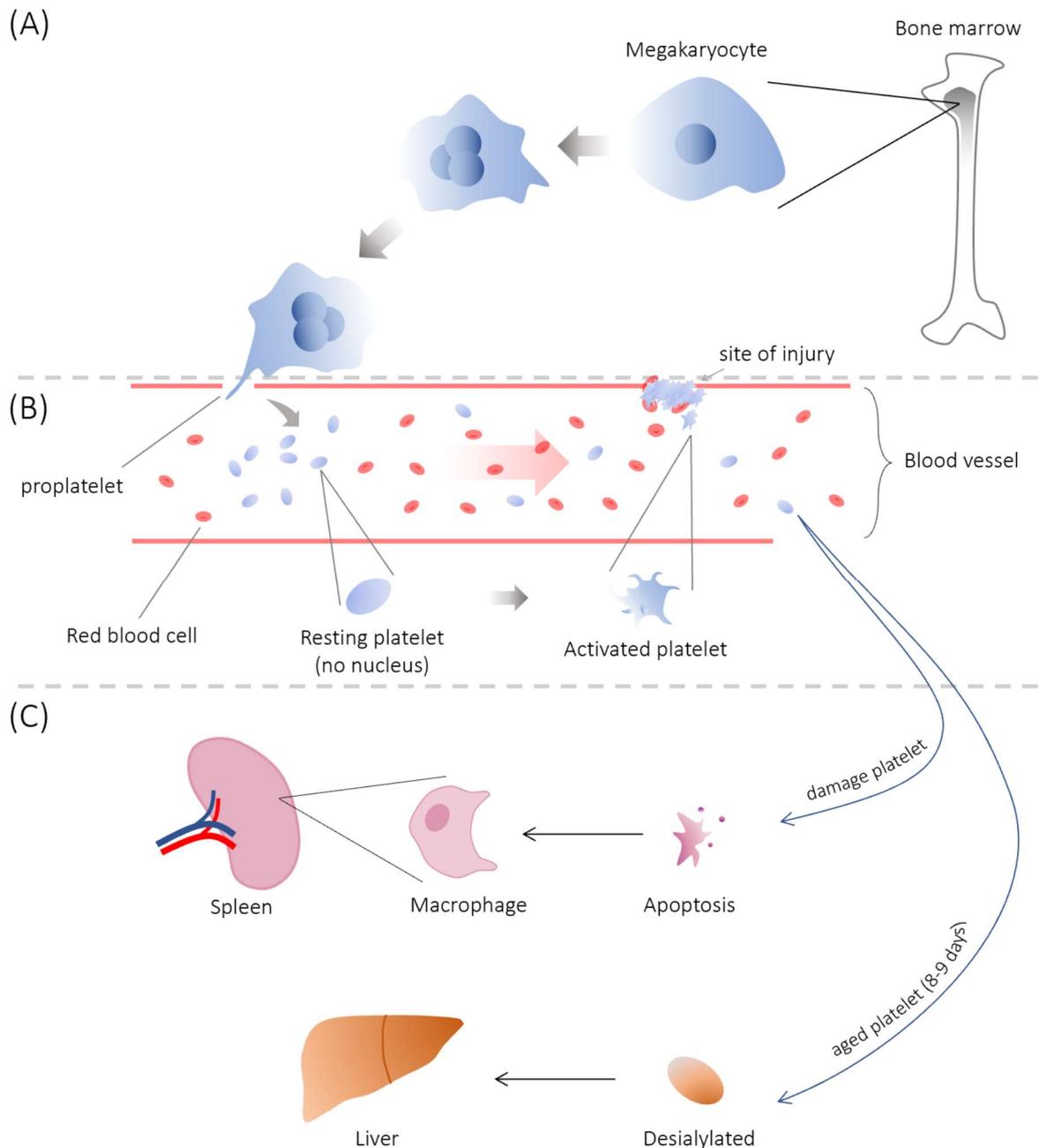


Figure 1. A life cycle of platelets. (A) Mature, multinucleated megakaryocytes in the bone marrow cell develop protrusions which separate into the blood stream as proplatelets, then dividing into platelets. (B) Adhesion and activation of platelets onto sites of vessel wall injury for instance by subendothelial collagen. (C) Clearance of apoptosis and desialylated platelets by the spleen and liver, respectively.

Under the dynamic flow conditions of the circulation, platelets are pushed by red blood cells out of the centre of the blood stream²⁹, which effect is favoured by the convex shape of

the red cells^{30,31}. After vascular injury and subendothelial matrix exposure^{32,33}, platelets adhere to the exposed collagen fibres, causing activation, change in shape from discoid to spherical, and filipodia formation to enlarge the surface (Figure 1B)³⁴. Increased adhesiveness is mediated by several integrin receptors^{35,36}, while granular release of ADP and polyphosphates leads to the activation of other platelets as well as the coagulation process³⁷. All of this results in a gradual formation of platelet aggregates (so-called platelet plugs) in primary haemostasis. Consolidation of the plug occurs by coagulation-induced fibrin formation to result in a stable clot³⁸.

Control of platelet activation and thrombus formation

In the intact vessel wall, platelets are inhibited by nitric oxide and prostaglandin I₂ (prostacyclin)³⁹, produced by vascular endothelial cells⁴⁰. Both nitric oxide⁴¹ and prostaglandin I₂⁴² suppress the adhesion and activation processes of platelets. When the tissue and matrix underneath the endothelium becomes exposed, platelets can adhere to collagen fibres⁴³. The commonly accepted event sequence of events of platelets in contact with collagen and von Willebrand factor (VWF) is illustrated in Figure 2. Initial adhesion under shear occurs by the interaction of GPIb-V-IX with VWF⁴⁴, which is produced by endothelial cells and megakaryocytes⁴⁵. The collagen receptors integrin $\alpha 2\beta 1$ and GPVI then become engaged, accomplishing stable platelet adhesion and activation⁴⁶.

The mechanism of collagen-induced activation via GPVI is well understood (for details see below). Principal steps involve activation of the protein tyrosine kinase Syk⁴⁷ and the phospholipase Cy2 (PLCy2), which mediates in a rise in cytosolic Ca²⁺.⁴⁸ Feedforward processes are through the release of autocrine mediators ADP³⁷ and thromboxane A₂^{49,50}.

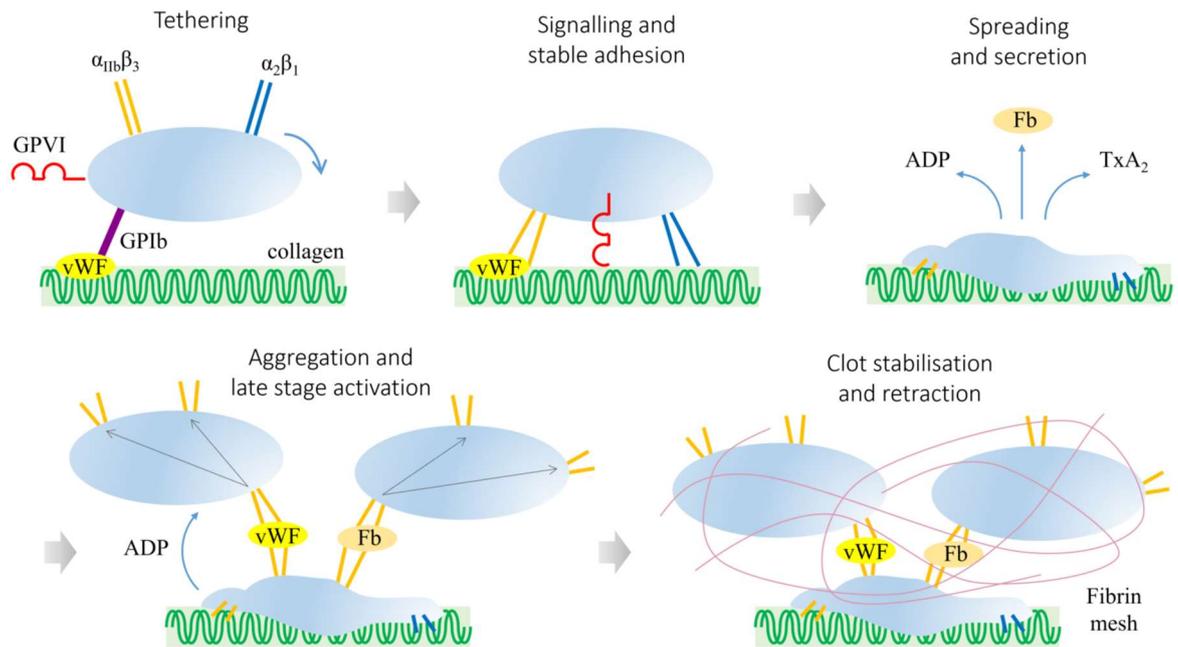


Figure 2. Sequential steps of collagen-induced platelet adhesion, activation and thrombus formation. Figure adapted from Ref^{A6}. At a damaged vessel exposing collagen and VWF, initial platelet adhesion occurs by shear-dependent VWF-GPIb-IX-V interactions. Rolling of platelets leads to adhesion by the collagen receptors integrin $\alpha_2\beta_1$ and GPVI. Subsequent GPVI-induced signalling results in the release of ADP and thromboxane A_2 (TxA_2), acting as autocrine agents. Under shear conditions, both VWF and fibrinogen establish interactions with other platelets, after which formed fibrin fibres stabilise the thrombus and promote clot retraction.

Both mediators act via G-protein coupled receptors, *i.e.* the $P2Y_1$ and $P2Y_{12}$ receptors for ADP, and the TP receptors for thromboxane A_2 ⁵¹. Specifically the $P2Y_{12}$ receptor inhibits the enzyme adenylyl cyclase, which acts by antagonising the inhibitory effect of prostaglandin I_2 ^{52,53}. Not further detailed here, but important for the establishment of platelet aggregates is the signalling-dependent activation of integrin $\alpha IIb\beta_3$ to mediate fibrinogen binding⁵⁴. As a bivalent ligand, fibrinogen enables the interaction between two platelets, and hence the aggregate formation⁵⁵.

Platelets and coagulation

Platelet activation is linked to the coagulation process, the latter forming thrombin via the intrinsic and extrinsic coagulation pathways (Figure 3). In the intrinsic pathway, the coagulation factors XII and XI become activated, *e.g.* via collagen or platelet-produced polyphosphates. In the presence of traces of thrombin, the ensuing activation of factors VIII and IX is greatly enhanced by phosphatidylserine (PS), exposed at the surface of highly activated platelets⁵⁶. The PS exposure allows assembly of the tenase complex, consisting of factor IXa and cofactor VIIIa, producing factor Xa⁵⁷.

The extrinsic pathway starts with exposed tissue factor⁵⁸. Tissue factor interacts with factor VII(a)⁵⁹ to produce factor Xa. Traditionally, both the extrinsic and intrinsic pathways are considered to join into a common pathway, where higher amounts of factor X are activated into factor Xa, which in turn generates thrombin via the prothrombinase complex of factor Xa and cofactor Va⁶⁰. Thrombin accomplishes the cleavage of fibrinogen into fibrin monomers, which assemble into in the fibrin mesh⁶¹. Thrombin in addition triggers positive feedback routes by activating factor XI⁶², factor VIII⁶³ and factor V⁶⁴ (Figure 3). Additional anticoagulant and fibrinolysis pathways are less relevant for this thesis.

An important other role of thrombin is to cleave and activate the protease-activated receptors PAR-1 and PAR-4 on human platelets^{65,66}. As G-protein coupled receptors these induce strong platelet activation. The platelet activation induced by thrombin thus forms another cross-talk mechanism of the coagulation process and platelets.

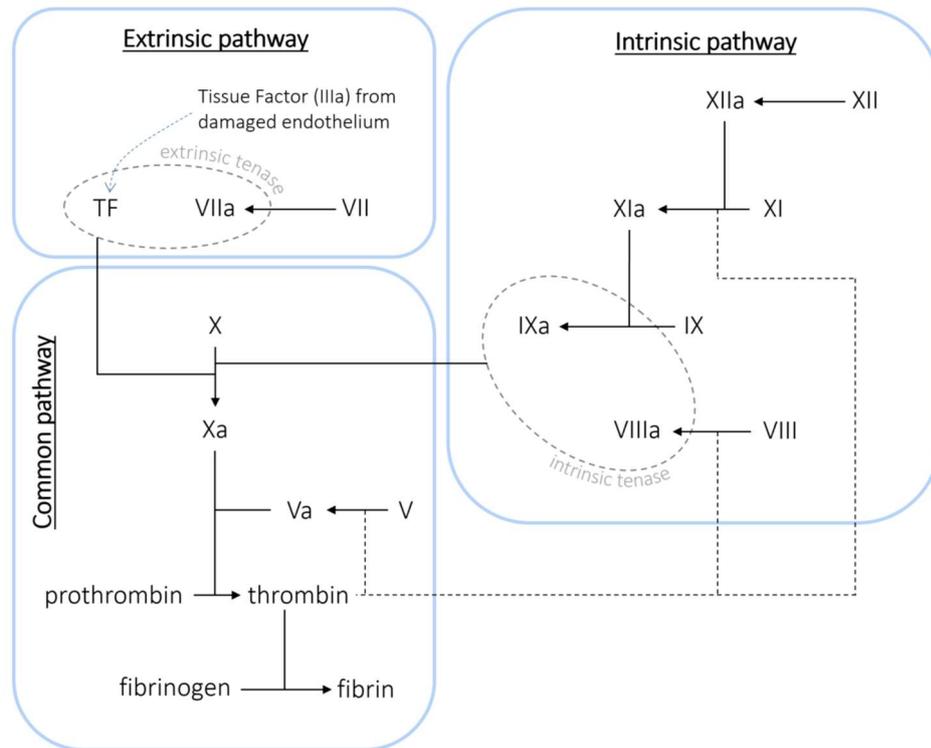


Figure 3. Simplified scheme of the coagulation cascade. The coagulation cascade is classically divided into three parts: the extrinsic, intrinsic and common pathways. The extrinsic pathway is triggered by the complex of tissue factor and factor VII(a), which cleaves traces of factor X into factor Xa. The intrinsic pathway is started by factor XII activation, subsequently involving factors XIa, IXa and VIIIa. On a procoagulant PS phospholipid surface, factors IXa and VIIIa form the tenase complex, which produces larger amounts of factor Xa in the common pathway. Also in a PS-dependent way, factor Xa with cofactor Va massively cleaves prothrombin into thrombin. The formed thrombin acts as a positive feedback on activation of factors V, VIII and XI, and furthermore cleaves fibrinogen to fibrin monomers.

GPVI structure and membrane location

The platelet GPVI receptor is a transmembrane protein with an estimated molecular weight of 62 kDa⁶⁷. Its extracellular region consists of two C2-type immunoglobulin domains, D1 and D2, and a mucin-rich-glycosylated stalk. Also present are a transmembrane region and a short intracellular tail (Figure 4A). In the membrane, GPVI forms a complex with the Fc receptor γ -chain (FcR γ), which contains an immunoreceptor tyrosine-based activation motif

(ITAM), required for intracellular signalling⁴⁷.

Antibody studies have indicated that GPVI on platelets is present more in a monomeric than a dimeric form. Flow cytometric quantitation indicated that the GPVI dimers account for 2–29% in resting platelets, while after stimulation with the GPVI ligand collagen-related peptide (CRP) this number increases to 39%⁶⁸. Also the presence of multiple GPVI binding sites in collagen fibres results in the formation of GPVI dimers, as shown by super-resolution microscopy⁶⁹.

Structural aspects appear to favour the dimerisation of GPVI. Thus, the homodimer presence of FcR γ allows for binding of two GPVI chains^{70,71}. In addition, the extracellular D1 and D2 domains, when dimerised, form a back-to-back conformation, thus creating a binding groove for interaction with multimeric ligands like collagen (Figure 4B)^{72,73}.

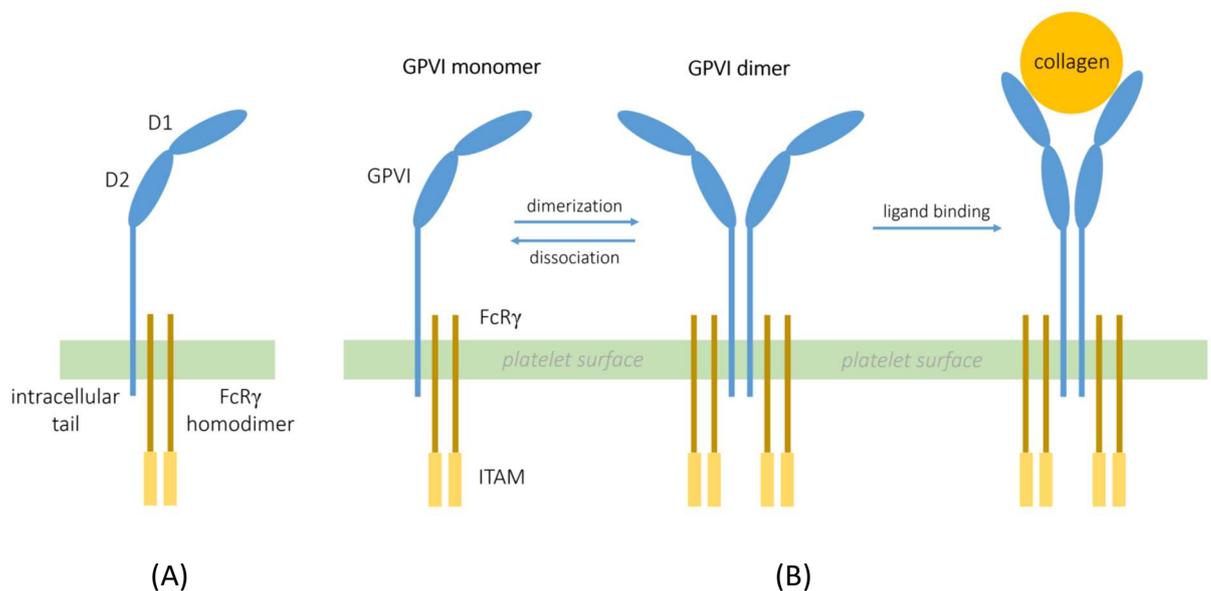


Figure 4. Structure of GPVI receptor complex. Figure adapted from Ref⁷¹. (A) The extracellular region of GPVI consists of D1 and D2 immunoglobulin-like domains and a mucin-rich stalk. The transmembrane sequence and intracellular tail are linked to the dimeric FcR γ -chain, which carries the ITAM signalling sequence. (B) In response to multimeric ligands, such as collagen, GPVI assembles into dimers increasing the ligand affinity. Panel adapted from Ref⁷².

A direct relation between GPVI dimerisation and ligand binding has been measured. Surface plasmon resonance analysis indicated that a dimeric recombinant GPVI binds to immobilised collagen with a K_D of 5.76×10^{-7} M, while the binding of a corresponding monomeric GPVI was too weak for quantification⁷⁴. In other words, the affinity of GPVI for collagen appeared to be higher in the dimeric than the monomeric form.

Recent evidence showed that GPVI can also interact with other ligands than collagens, for instance with laminins⁷⁵, factor XIIIa⁷⁶, fibrinogen and fibrin^{77,78}. For these ligands, it is still unclear to which extent GPVI dimerisation occurs. Whether dimerising GPVI ligands induce platelet activation or *vice versa* platelet activation leads to GPVI dimerisation is still a matter of debate.

A related topic/question in this context is whether specific domains in the plasma membrane can contribute to GPVI dimerisation. Interest herein is the potential role of lipid rafts, which are small and transient nanodomains, enriched in cholesterol and glycolipids, known to regulate membrane fluidity and protein trafficking^{79,80}. In general, is it considered that lipid rafts are important for the transmission of signals. For several receptors, including Fc receptors and integrins, signal transduction in a raft-dependent manner has been observed⁸¹.

Signalling routes induced by GPVI

In platelets, ligands of GPVI induce a complex pathway of protein tyrosine kinase-dependent signalling events, such as reviewed by others^{82,83}. Below, I focus on parts that are relevant for this thesis.

Ligand-binding to GPVI initially results in phosphorylation via Src-family kinases (SFKs) of the tyrosine residues of the ITAM domain of FcR γ , thereby allowing the kinase Syk to dock to this domain⁸⁴. Since in other cell types SFKs concentrate in lipid rafts⁸⁵, also for platelets a

role of rafts in this phosphorylation can be assumed. Subsequently Syk becomes sequentially phosphorylated at two tyrosine residues, namely Y³²³ and Y⁵²⁵ (Figure 5)^{86,87}. There is evidence for a negative feedback, in that the phospho-Y³²³ site of Syk allows binding of the ubiquitin ligase c-Cbl^{88,89,90}, which attracts the phosphatase TULA-2 (T-cell ubiquitin ligand-2) to establish the dephosphorylation of the Y⁵²⁵ residue⁹¹.

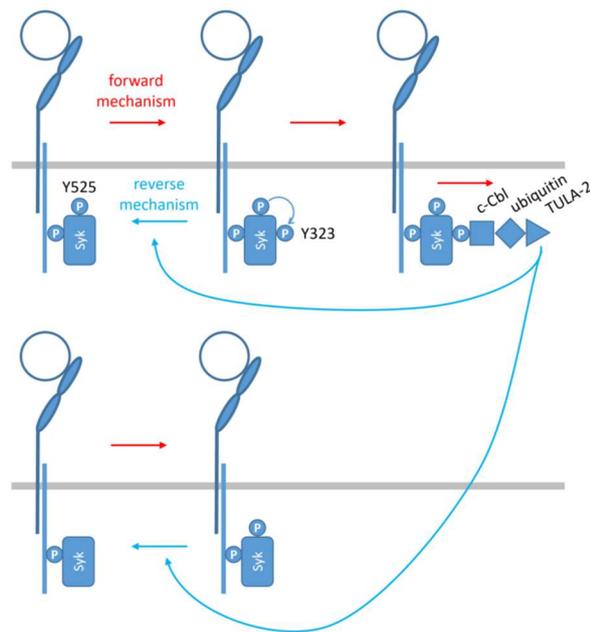


Figure 5. Pathway of transient Syk tyrosine phosphorylation. Redrawn from Ref⁹². Proposed pathway of a stepwise phosphorylation of residues Y⁵²⁵ and Y³²³ of Syk to allow complex formation of c-Cbl, ubiquitin and TULA-2. The latter protein phosphatase mediates ubiquitylation-dependent Syk dephosphorylation.

Once phosphorylated, Syk propagates the cascade via tyrosine phosphorylation of several other signalling proteins, including Btk (Bruton's tyrosine kinase), the adaptor protein LAT, and the effector protein phospholipase C γ 2 (PLC γ 2) (Figure 6)⁹³. In spite of the recognition that the tyrosine-phosphorylated Btk is involved in the GPVI signalling pathway, the precise molecular mechanism of its action is not well understood.

Phosphoinositides, *i.e.* inositol-containing phospholipids, play important roles in agonist-induced platelet signalling events. Within the plasma membrane, the LAT-associated phosphoinositide 3-kinase (PI3K) converts phosphatidylinositol 4,5-bisphosphate (PIP₂) into phosphatidylinositol 3,4,5-trisphosphate (PIP₃)⁹⁴. The formed PIP₃ provides an anchor for the binding of several PH-domain containing signalling proteins. In addition, the activated PLC γ 2 cleaves PIP₂ into the second messengers inositol 1,4,5-triphosphate (IP₃) and 1,2-diacylglycerol (DAG)⁹⁵. The DAG in the membrane activates protein kinase C (PKC) isoforms⁹⁶, whereas the soluble IP₃ binds to IP₃ receptors on the DTS, promoting Ca²⁺ release from the DTS⁷⁷ (Figure 6). While all of this points to a central role of PIP₂ in the GPVI-induced signalling scheme, how PI3K and PLC γ 2 compete for the phosphoinositide still needs to be resolved.

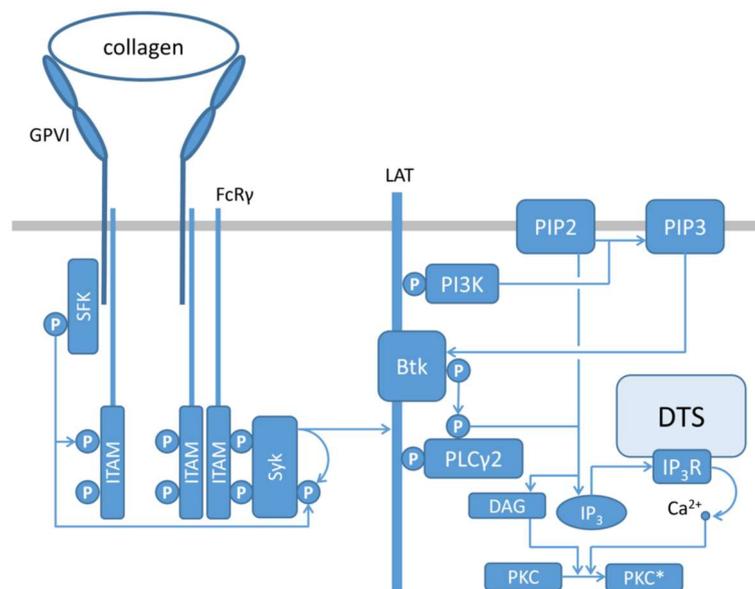


Figure 6. Schematic diagram of GPVI signalling to phosphoinositide turnover. GPVI ligands induce the SFK-mediated phosphorylation and docking of Syk to the ITAM domain of FcR γ . Once phosphorylated, Syk in turn phosphorylates and activates the adaptor protein LAT, and also Btk, PI3K and PLC γ 2. In the membrane, the phosphoinositide PIP₂ is phosphorylated by PI3K into PIP₃, whereas it is cleaved by PLC γ 2 into DAG and IP₃. The IP₃ mediates Ca²⁺ mobilisation. Both DAG and elevated Ca²⁺ contribute to activation of PKC isoforms. For abbreviations and references, see text.

Of note, the platelet Ca^{2+} responses are regulated not only by PLC-induced Ca^{2+} release from the DTS, but also by other mechanisms (Figure 7). These include back-pumping of elevated cytosolic Ca^{2+} into the DTS via SERCA pumps⁷⁷ and into the extracellular medium by PMCA pumps in the plasma membrane. Furthermore, extracellular Ca^{2+} can enter the cytosol via ORAI1 channels, which are linked to STIM1 proteins in the DTS membrane⁹⁷. Of note, IP_3 -induced Ca^{2+} release is also triggered by the PLC β -activating agonists thrombin,⁹⁸ thromboxane A_2 and ADP⁹⁹. ATP-triggered Ca^{2+} entry also occurs via the P2X_1 ion channels¹⁰⁰. So far, no systematic studies are available comparing the relative contributions of all these channels and pumps to the cytosolic Ca^{2+} rises induced via GPVI or other platelet receptors.

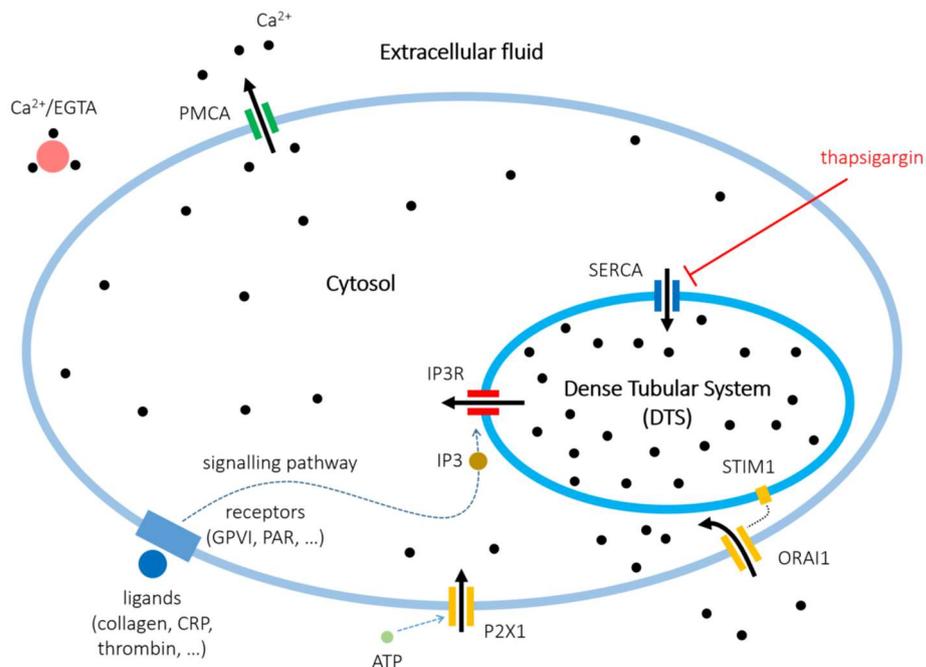


Figure 7. Regulation of agonist-induced cytosolic Ca^{2+} rises in platelets. Receptors signalling via $\text{PLC}\gamma 2$ or $\text{PLC}\beta$ isoforms generate the second messenger IP_3 , which opens IP_3 -receptor (IP_3R) Ca^{2+} channels in the Ca^{2+} -loaded DTS, causing a rise in cytosolic Ca^{2+} . A fast and quickly desensitised manner of Ca^{2+} entry is mediated by ATP, activating the P2X_1 Ca^{2+} ion channel¹⁰⁰. The Ca^{2+} -store depletion, sensed by STIM1 in the DTS membrane, evokes Ca^{2+} entry from the extracellular medium via ORAI1 Ca^{2+} channels in the plasma membrane. Back-pumping of cytosolic Ca^{2+} occurs by SERCA isoforms (inhibited by thapsigargin), whereas Ca^{2+}

externalisation occurs by plasma membrane Ca²⁺-ATPases (PMCA). In isolated platelets, the presence of extracellular EGTA acts as a Ca²⁺ chelator prevents Ca²⁺ entry. For references, see text.

Platelet disorders and antiplatelet medication

Both congenital and acquired bleeding disorders are linked to platelet dysfunction. The most prevalent hereditary bleeding disorder is von Willebrand disease (VWD), caused by a qualitative or quantitative defect in VWF¹⁰¹, which indirectly leads to a reduced platelet adhesion. Other bleeding diseases can be due to low platelet counts or to receptor or signalling defects of platelets.

Whereas venous thrombosis is commonly treated with anticoagulants, arterial thrombosis is treated with platelet inhibitors, in agreement with the specific role of platelets in thrombus formation within the arterial part of the circulation. Arterial thrombosis or atherothrombosis is mostly the result of plaque rupture or erosion. It can manifest as heart infarction (occluded coronary artery), angina (semi-occlusion), or ischemic stroke (occluded carotid artery)¹⁰².

A most prescribed drug for the secondary prevention of atherothrombosis is aspirin, blocking (platelet) cyclooxygenase and thromboxane A₂ release¹⁰³. Other common antiplatelet drugs are P2Y₁₂ receptor inhibitors, such as clopidogrel and prasugrel, which prevent a substantial part of the platelet activation process mediated by autocrine ADP. The antiplatelet drugs, while not completely protective, however come at the expense of (pre)clinical bleeding in a non-negligible number of patients^{104,105,106}. This problem has led to the search for better antiplatelet drugs.

One of the promising new drug targets is GPVI, since its inhibition *in vivo* effectively reduces experimental arterial thrombosis and thrombo-inflammation in mice, while leaving

the haemostatic response unaffected^{107–109}. Its unique expression on platelets makes GPVI to a good therapeutic target with expected limited side effects^{108,110}. At present, two GPVI-related drugs are investigated in clinical trials, *i.e.* the heterodimer recombinant GPVI construct Revacept and the antibody-based drug Glenzocimab¹¹¹. In the last few years also anti-GPVI nanobodies have been developed with a therapeutic potential as well¹¹².

Modelling ways of receptor-dependent processes

Several modelling approaches are available to analyse the molecular interactions and pathways of biological systems. The choice of approach being influenced by the complexity of the model, the time and spatial components, the data available for calibration and validation and the questions requiring an answer¹¹³. The mathematical models that have been used to better understand the processes of thrombus formation and blood coagulation^{114,115} use a range of techniques from models comprising ordinary differential equations (ODE) that describe the processes underlying coagulation and subcellular interactions to spatial models in the form of partial differential equations and agent based models, more details of which are described below.

Modelling of the GPVI structure, interactions and movements

Spatiotemporal modelling techniques can be used to investigate the mobility and dimerisation of proteins such as the GPVI receptors. The mathematical equation for molecular diffusion is based on the partial differential equation (PDE) for a two-dimensional diffusion process, assuming a continuous distribution of molecules¹¹⁶:

$$\frac{\partial \phi}{\partial t} = D \nabla^2 \phi$$

where ϕ is the molecular concentration, D is the diffusivity and ∇ is the del or nabla operator. The outcome is continuous in both space and time with no discrimination between individual receptor molecules.

An alternative molecular dynamics (MD) approach is possible, which describes the dynamics of a single protein or peptide by calculating the trajectory of each atom or group of atoms. The MD simulation of target proteins has been helpful in drug discovery¹¹⁷. Limitations of the MD technique are the restrictions in spatial and time scales. Even for small sections of a given protein, high computational resources are required for calculating a configuration, and then still result in short trajectory times, in the order of nanoseconds.

A related technique to simulate movements of individual (receptor) molecules is provided by Langevin dynamics¹¹⁸, which describe the Brownian motion of a particle by combining the standard Newton's law equation of motion with random noise:

$$m \frac{d\vec{v}}{dt} = -\nabla U(\vec{r}) - \lambda \vec{v} + \vec{\eta}(t)$$

where m is the particle's mass, \vec{r} is its positional vector, t is the time variable, λ stand for the viscosity; $\vec{\eta}$ is the noise term representing random collisions with other molecules in a fluid; and U is a function for interactions between molecules. The equation describes movements of individual molecules, for instance, receptors in a membrane. In constructed sets of ODEs, it is possible to simulate the movement and interactions per complete molecule. However, deterministically solving the equation of motion for all molecules present again requires high computational power.

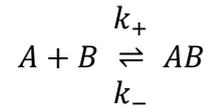
Yet another approach is that of agent-based modelling (ABM). This treats molecules (receptors) as moving agents in an area (membrane). Instead of calculating the exact path of movement for each agent, ABM stochastically simulates the changes in time; this comes with

lower complexity and less computational costs^{119,120}. In general terms, the ABM method is an approach to simulate a real-world system by transforming objects into interacting agents with fixed properties in the environment led by definable mathematical rules. ABM thereby provides a micro-scale simulation that predicts for changes in a larger-scale system.

When applied to receptor interactions, ABM can introduce different kinds of agents (receptors and other membrane proteins), each with a defined size, mass and movement pattern (stochastic or deterministic), and with a specified interaction spectrum (attractive or repulsive). Authors, introducing the term protein diffusivity, have applied ABM to simulate the interactions between membrane proteins based on certain hypotheses^{121,122}. Because of the availability of prior data, ABM simulation is an attractive approach for modelling the movements of GPVI molecules on a platelet membrane. A limitation though is that simplification of the agent's (molecular) properties is needed. In conclusion, ABM, as a preferable method for receptor clustering simulation due to its spatiotemporal property, stochasticity, simplicity, computational cost, and discretisation, is used in Chapter 2.

Mathematically modelling of GPVI signalling

The GPVI-induced signalling pathway involves several components and steps with parallel positive and negative feedback loops. A type of mathematical modelling that can capture this is creating a set of ordinary differential equations (ODEs) for each of these steps. The underlying assumption is that the components are evenly distributed in space, and that the reactions follow mass action kinetics. The set of ODEs then comes with time-dependent predictions for each protein in the signalling system¹²³. A given biochemical reaction in the form of:



can be translated in an ODE for the component AB as:

$$\frac{d[AB]}{dt} = k_+[A][B] - k_-[AB]$$

$$\frac{d[A]}{dt} = \frac{d[B]}{dt} = -k_+[A][B] + k_-[AB]$$

For given values of k_+ and k_- (binding and dissociation constants), and known initial levels of [A], [B] and [AB] at time $t = 0$, the equation can be solved numerically. The result is a time-dependent simulation for each variable. This is a useful approach for biochemical reactions where otherwise spatial-dependent information is absent, such as for the distribution of proteins across a cell membrane. In prior work, the GPVI-induced signalling events of Syk (de)phosphorylation have been modelled in this way⁹². The system of ordinary differential equations being inferred from and validated against experimental data. This Bayesian approach (called Approximate Bayesian Computation (ABC)¹²⁴) has been utilised extensively to compare the outputs from systems of ODEs to experimental data^{125,126}. It allows the systematic updating of beliefs in the structure of the model in light of experimental data the fundamental theorem behind this being Bayes that states the probability of an event A given some evidence B is equal to the probability of B given A multiplied by the probability of A, divided by the probability of B. In mathematical terms, it can be expressed as:

$$P(A|B) = P(B|A) \frac{P(A)}{P(B)}$$

where $P(A|B)$ is the probability of A given B, which is the value we want to calculate, $P(B|A)$ is the probability of B given A, which is the likelihood of the evidence given the prior knowledge, $P(A)$ is the prior probability of A, which is our prior belief or experience about the probability of A before considering the evidence, and $P(B)$ is the probability of the evidence B, which

serves as a normalising constant to ensure that the result is a valid probability distribution.

This approach allows us to determine the best set of parameters that enables the model to simulate the experimental data. The advantage of ABC approaches is that it is not necessary to analytically describe the likelihood of parameter dependence on data. Rather, approximates are calculated by running many simulations (we use $n = 50000$ in Chapter 3 and 4) with parameters drawn randomly from biologically feasible ranges (priors). This provides a systematic way to assess if a model's structure is capable of generating the given experimental data and in this thesis we supplement this by utilising model comparison techniques to¹²⁴. compare many different structural models. The criteria used for determining the best model for the system, in this work, is Akaike Information Criteria with a correction of small sample size (AICc)¹²⁷ that is used to pick the best model that balances goodness of fit and model complexity by the sum of squares error (SSE) via:

$$AICc = n \ln \left(\frac{SSE}{n} \right) + 2k + \frac{2k^2 + 2k}{n - k - 1}$$

where k is the number of parameters and n is the number of datapoints used in the fitting process. The numerical value of AICc reflect the 'distance' from the true model, more AICc mean the model is far from 'true model'.

There is a tension between the need to incorporate biological detail into models such as those of sub-cellular signalling pathways and the necessity of developing robust, predictive and well-constrained models. ABC approaches offer a way to balance this, allowing assessment of the inferred parameter values in light of data and in light of this provide a description of the variability of predictions. We use this approach in Chapters 3 and 4. Of course ABC can be utilised to assess the ability of agent based models to reproduce experimental data¹²⁸ and would be useful in validating outputs from the agent based model developed in Chapter 2

Once more data comes available.

Neural network modelling of platelet Ca²⁺ responses

Modelling of the GPVI-evoked platelet Ca²⁺ response is a challenge, because of the involvement of multiple signalling proteins, Ca²⁺ channels and pumps, and secondary mediators, which are present in or are acting on distinct intracellular regions. Due to synergistic effects between these actors, a conventional ODE modelling approach is unsuitable. Solution here is to use a machine learning technique, known as neural network generation^{129,130}. This method acts as a black box, since the data processing relies on creation of layers with hidden parameters. On the other hand, with appropriate training, a neural network can generate useful output for a given input with valuable predictive power.

As illustrated in Figure 8, a standard neural network consists of three partitions, *i.e.* an input layer, one or more hidden layers, and an output layer. The input layer is fed with experimental datasets, which are processed by adding different weights, bias effects and summations. Several activation functions can be called upon, depending on the fitting purpose, such as training acceleration, accuracy and linearity^{131,132}. Finally, the training output is compared to known results, after which the weights and sums are tuned to minimise the difference between the experimental data and the predicted output.

A recurrent neural network is a more complex version that employs the output of a processing node, and then feeds it back into the network. Each node records the previous operations, and uses the information for an improved data processing. This type of neural network is useful for time series predictions. Other researchers have used it to predict specific agonist-induced rises in cytosolic Ca²⁺ upon platelet activation^{133,134}. In that work, the recurrent neural network was trained by providing input of the Ca²⁺ responses of platelets from several donors.

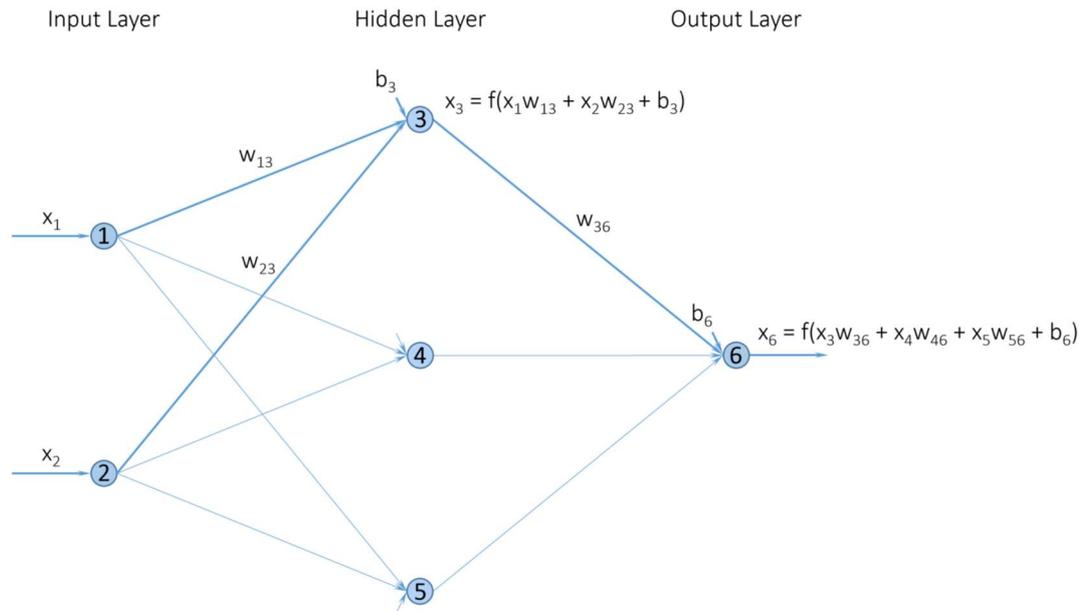


Figure 8. Example organisation of a neural network. The illustrated neural network consists of three partitions, namely an input layer, one or more hidden layers, and the output layer. The number of inputs is set at 2, with hidden layer consisting of 3 nodes, and 1 output layer. Shown is how the network feeds input values x_1 , x_2 into nodes 1 and 2 of the input layer, after which combined values pass through each node in the hidden layer (3-5) with different weights (w_{ij} connected nodes i and j). Further processing works by summing up the node values with additional bias (b_i for node i), and passing these through an activation function f . The function output feeds the output layer with node 6. Compared of the output with the target value, results in an error estimate. Training of the neural network is to tune the weights and biases to minimise this error.

Partial least square (PLS) regression of platelet Ca^{2+} responses

Also useful, but less predictive than a neural network, is the method of partial least square (PLS) regression analysis^{135,136}. It provides a data-driven modelling strategy with multiple predictor and response variables. A PLS regression analysis, searching for multiple correlative relationships, can for instance be used to make quantitative predictions of the effects of new signalling interventions. The method is frequently used in computational biology, in high-dimensional genomics and in cellular signalling¹³⁷⁻¹³⁹. In the PLS regression

analysis, a linear model is used that finds coordinates for multiple independent variables in order to minimise the difference between input and output variables (Figure 9). The result is expressed as components C_i , with C_1 providing the highest contribution to the output variable Y . The maximum number of C_i equals the number of input variables X_i . Commonly, the first 2-3 PLS components cover most of the variance of Y . Although the constructed model is linear and may not have the same predictive power as a neural network, it is easier to interpret the effect of X_i on Y .

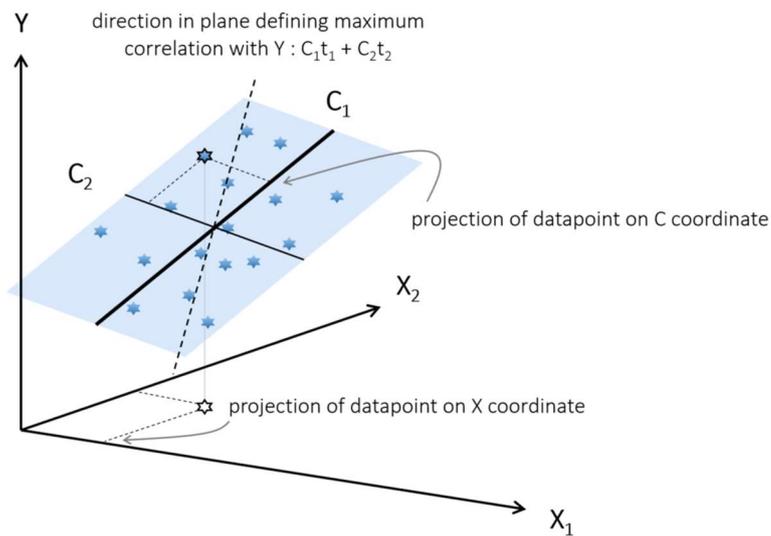


Figure 9. Geometrical interpretation of a partial least square regression analysis. Figure adapted from Ref¹³⁵. The X coordinates represent the input values, while the C coordinates define the transformed set by PLS regression. From an N -dimensional X input, PLS regression converts values from each data point by constructing a K ($\leq N$) dimensional hyperplane, which maximises the variance between the new coordinate and the target Y . Projection of the data to this new coordinate results in a C -coordinate ($C_1 = a_1X_1 + a_2X_2 + \dots$). The maximal correlation hyperplane is written as a linear combination of C -components ($C_1t_1 + C_2t_2 + \dots$, where the score t_i refers to coordinate C_i). The new coordinates are sorted such that the C_1 axis has the highest contribution to the explained variance of outcome Y , followed by C_2 and so on. The contribution of X_i to Y is expressed as a coefficient for each x_i component in C_1 , etc.

Aims and outline of this thesis

The overarching aim of this thesis is to understand the overall process of glycoprotein VI (GPVI)-induced platelet activation by applying different mathematical/computational modelling and validation approaches. Basic assumption is that the obtained insight can help to identify and precise the mode of actions of GPVI-dependent signalling processes that have a therapeutic potential. The introduction of **Chapter 1** gives relevant background information on platelets and on platelet activation processes. Subsequent sections describe the structure, the localisation and the signalling mechanisms of GPVI. The chapter furthermore describes different ways of (mathematical) modelling approaches that are relevant for and that are being used in this thesis (see Figure 10).

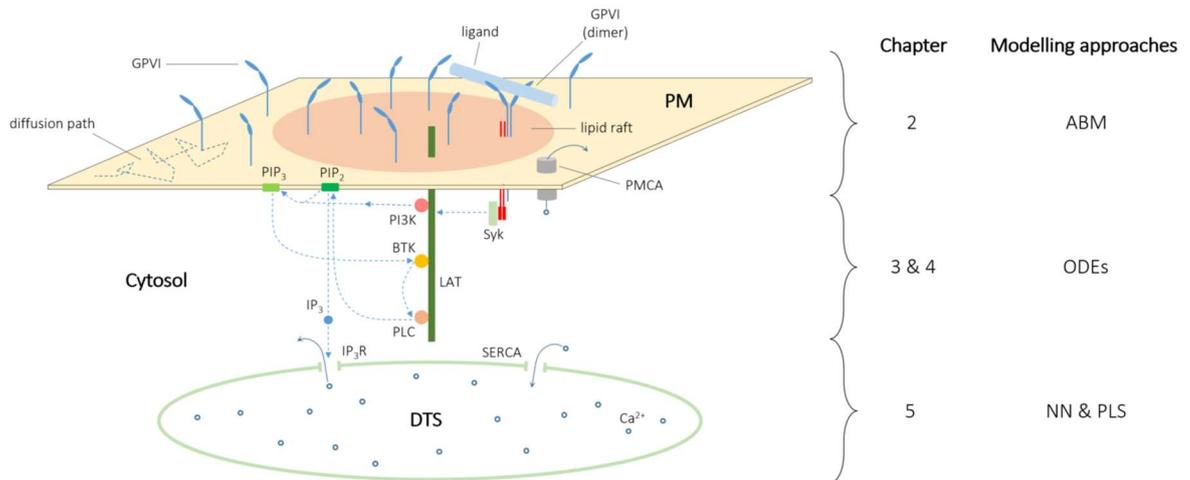


Figure 10. Overview of mathematical modelling approaches in this thesis. Chapter 2 uses ABM to model the dimer formation of GPVI in the plasma membrane taking into account the impact of confined domains and other molecular interactors. Chapter 3 uses modelling by ODEs of the time-dependent GPVI-induced tyrosine (de)phosphorylation events up to activation of Btk and PLCg2 and of IP₃ formation. Chapter 4 continues on this by developing an ODE-based model of the enzymatic regulation of phosphoinositide turnover. Chapter 5 uses a recurrent NN and PLS analysis to model the GPVI-induced and other Ca²⁺ responses in activated platelets, to prelude on the systematic Ca²⁺ response measurements of Chapter 6. Abbreviations: ABM,

agent-based modelling; ODEs, ordinary differential equations; NN, neural network; PLS, partial least squares.

The application of ABM to study factors influencing GPVI dimerisation on the plasma membrane is the main topic of **Chapter 2**. Based on prior information on the heterogeneity of GPVI movement patterns, a confined domain (representing lipid rafts) is introduced with restricted receptor diffusivity, acting as a GPVI-dimer forming entity. **Chapter 3** uses non-linear ODE-based modelling approaches to describe and predict the changes over time of GPVI-induced and tyrosine kinase-dependent protein activation steps. These include the changes in (phosphorylated) LAT, Btk and PLC γ 2, leading to PIP $_2$ -dependent IP $_3$ generation. The modelling work links to the question of whether the phosphoinositide-modulating kinases are suitable targets for suppression of GPVI activation. The subsequent **Chapter 4** presents an ODE-based model for the generation of time-dependent profiles of phosphoinositide changes in GPVI-activated platelets, including levels of PIP $_2$ isomers and PIP $_3$. The ODE model is validated by comparing its predictions with new experimental data on pharmacological inhibitors that are known to affect the phosphoinositide turnover.

Chapter 5 addresses the time-dependent Ca $^{2+}$ responses of Fura-2-loaded platelets stimulated with GPVI or PAR agonists. Both a recurrent neural network and a PLS regression model are developed to simulate and predict the nanomolar changes in cytosolic Ca $^{2+}$. Questions to be answered are if non-equation based modelling approaches can capture the complex regulation of cytosolic Ca $^{2+}$ rises, involving many triggers, channels, pumps and modulators. In other words, can a model be developed that simplifies the system's complexity, but still has sufficient predictive power and an understandable interpretability.

The general discussion in **Chapter 6** compares and integrates the various modelling and

experimental approaches in order to better understand the GPVI signalling mechanism, such in comparison to the current knowledge. Opportunities and limitations of the models are discussed, and follow-up steps are proposed for advancing the research field.

References

1. Holinstat, M. Normal platelet function. *Cancer Metastasis Rev.* **36**, 195–198 (2017).
2. Periyah MH, Halim AS & Saad AZM. Mechanism action of platelets and crucial blood coagulation pathways in Hemostasis. *Int. J. Hematol. Stem Cell Res.* **11**, 319–327 (2017).
3. Paciaroni, M., Ince, B., Hu, B., Jeng, J. S., Kutluk, K., Liu, L., *et al.* Benefits and Risks of Clopidogrel vs. Aspirin Monotherapy after Recent Ischemic Stroke: A Systematic Review and Meta-Analysis. *Cardiovasc. Ther.* **2019**, 1–12 (2019).
4. Bender, M., May, F., Lorenz, V., Thielmann, I., Hagedorn, I., Finney, B. A., *et al.* Combined in vivo depletion of glycoprotein VI and C-type lectin-like receptor 2 severely compromises hemostasis and abrogates arterial thrombosis in mice. *Arterioscler. Thromb. Vasc. Biol.* **33**, 926–934 (2013).
5. Frojmovic, M. M. & Milton, J. G. Human platelet size, shape, and related functions in health and disease. *Physiol. Rev.* **62**, 185–261 (1982).
6. Hanson, S. & Slichter, S. Platelet kinetics in patients with bone marrow hypoplasia: evidence for a fixed platelet requirement. *Blood* **66**, 1105–1109 (1985).
7. Edward Quach, M., Chen, W. & Li, R. Mechanisms of platelet clearance and translation to improve platelet storage. *Blood* **131**, 1512–1521 (2018).
8. Italiano, J. E., Bergmeier, W., Tiwari, S., Falet, H., Hartwig, J. H., Hoffmeister, K. M., *et al.* Mechanisms and implications of platelet discoid shape. *Blood* **101**, 4789–4796 (2003).
9. Shin, E. K., Park, H., Noh, J. Y., Lim, K. M. & Chung, J. H. Platelet shape changes and cytoskeleton dynamics as novel therapeutic targets for anti-thrombotic drugs. *Biomol. Ther.* **25**, 223–230 (2017).
10. Kunicki, T. J. Platelet membrane glycoproteins and their function: An overview. *Blut* **59**, 30–34 (1989).
11. Kovacsovics, T. J. & Hartwig, J. H. Thrombin-induced GPIb-IX centralization on the platelet surface requires actin assembly and myosin II activation. *Blood* **87**, 618–629 (1996).
12. Bearer, E. L., Prakash, J. M. & Li, Z. Actin dynamics in platelets. *Int. Rev. Cytol.* **217**, 137–182 (2002).

13. Melchinger, H., Jain, K., Tyagi, T. & Hwa, J. Role of platelet mitochondria: Life in a nucleus-free zone. *Front. Cardiovasc. Med.* **6**, 1–11 (2019).
14. Yadav, S. & Storrie, B. The cellular basis of platelet secretion: Emerging structure/function relationships. *Platelets* **28**, 108–118 (2017).
15. Brass, L. F. The effect of Na⁺ on Ca²⁺ homeostasis in unstimulated platelets. *J. Biol. Chem.* **259**, 12571–12575 (1984).
16. Brass, L. F. & Joseph, S. K. A role for inositol triphosphate in intracellular Ca²⁺ mobilization and granule secretion in platelets. *J. Biol. Chem.* **260**, 15172–15179 (1985).
17. Horiguchi, M., Kimura, M., Lytton, J., Skurnick, J., Nash, F., Awad, G., *et al.* Ca²⁺ in the dense tubules: A model of platelet Ca²⁺ load. *Hypertension* **31**, 595–602 (1998).
18. Crescente, M., Pluthero, F. G., Li, L., Lo, R. W., Walsh, T. G., Schenk, M. P., *et al.* Intracellular trafficking, localization, and mobilization of platelet-borne thiol isomerases. *Arterioscler. Thromb. Vasc. Biol.* **36**, 1164–1173 (2016).
19. Patel, S. R., Hartwig, J. H. & Italiano, J. E. The biogenesis of platelets from megakaryocyte proplatelets Find the latest version : Review series The biogenesis of platelets from megakaryocyte proplatelets. *J. Clin. Invest.* **115**, 3348–3354 (2005).
20. Lefrançois, E., Ortiz-Muñoz, G., Caudrillier, A., Mallavia, B., Liu, F., Sayah, D. M., *et al.* The lung is a site of platelet biogenesis and a reservoir for haematopoietic progenitors. *Nature* **544**, 105–109 (2017).
21. Kaufman, R. M., Airo, R., Pollack, S. & Crosby, W. H. Circulating megakaryocytes and platelet release in the lung. *Blood* **26**, 720–731 (1965).
22. Harker, L. A. & Finch, C. A. Thrombokinetics in man. *J. Clin. Invest.* **48**, 963–974 (1969).
23. Trowbridge, E. A., Martin, J. F., Slate, D. N., Kishk, Y. T., Warren, C. W., P.J.Harley, *et al.* The origin of platelet count and volume. *Clin. Phys. Physiol. Meas.* **5**, 145–170 (1984).
24. Jackson, C. W. Megakaryocyte endomitosis: A review. *Int. J. Cell Cloning* **8**, 224–226 (1990).
25. Becker, R. P. & De Bruyn, P. P. H. The transmural passage of blood cells into myeloid sinusoids and the entry of platelets into the sinusoidal circulation; a scanning electron microscopic investigation. *Am. J. Anat.* **145**, 183–205 (1976).
26. Bender, M., Thon, J. N., Ehrlicher, A. J., Wu, S., Mazutis, L., Deschmann, E., *et al.* Microtubule sliding drives proplatelet elongation and is dependent on cytoplasmic dynein. *Blood* **125**, 860–868 (2015).
27. Dunois-Lardé, C., Capron, C., Fichelson, S., Bauer, T., Cramer-Bordé, E. & Baruch, D. Exposure of human megakaryocytes to high shear rates accelerates platelet production. *Blood* **114**, 1875–1883 (2009).

28. Thon, J. N., Montalvo, A., Patel-Hett, S., Devine, M. T., Richardson, J. L., Ehrlicher, A., *et al.* Cytoskeletal mechanics of proplatelet maturation and platelet release. *J. Cell Biol.* **191**, 861–874 (2010).
29. Chesnutt, J. K. W. & Han, H.-C. Effect of red blood cells on platelet activation and thrombus formation in tortuous arterioles. *Front. Bioeng. Biotechnol.* **1**, 1–12 (2013).
30. Zhao, R., Kameneva, M. V. & Antaki, J. F. Investigation of platelet margination phenomena at elevated shear stress. *Biorheology* **44**, 161–177 (2007).
31. Van Hinsbergh, V. W. M. Endothelium - Role in regulation of coagulation and inflammation. *Semin. Immunopathol.* **34**, 93–106 (2012).
32. Chesterman, C., Owe-Young, R., Macpherson, J. & Krilis, S. Substrate for endothelial prostacyclin production in the presence of platelets exposed to collagen is derived from the platelets rather than the endothelium. *Blood* **67**, 1744–1750 (1986).
33. Davis, G. E. & Senger, D. R. Endothelial extracellular matrix: Biosynthesis, remodeling, and functions during vascular morphogenesis and neovessel stabilization. *Circ. Res.* **97**, 1093–1107 (2005).
34. van der Meijden, P. E. J. & Heemskerk, J. W. M. Platelet biology and functions: new concepts and clinical perspectives. *Nat. Rev. Cardiol.* **16**, 166–179 (2019).
35. De Groot, P. G. & Sixma, J. J. Platelet-vessel wall interaction. *Vasc. Med. Rev.* **4**, 145–155 (1993).
36. Ruggeri, Z. M. & Mendolicchio, G. L. Adhesion mechanisms in platelet function. *Circ. Res.* **100**, 1673–1685 (2007).
37. Golebiewska, E. M. & Poole, A. W. Platelet secretion: From haemostasis to wound healing and beyond. *Blood Rev.* **29**, 153–162 (2015).
38. Hawiger, J. Formation and regulation of platelet and fibrin hemostatic plug. *Hum. Pathol.* **18**, 111–122 (1987).
39. Mitchell, J. A., Ali, F., Bailey, L., Moreno, L. & Harrington, L. S. Role of nitric oxide and prostacyclin as vasoactive hormones released by the endothelium. *Exp. Physiol.* **93**, 141–147 (2008).
40. Weksler, B. B., Marcus, A. J. & Jaffe, E. A. Synthesis of prostaglandin I₂ (prostacyclin) by cultured human and bovine endothelial cells. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 3922–3926 (1977).
41. Cheung, P. Y., Salas, E., Schulz, R. & Radomski, M. W. Nitric oxide and platelet function: Implications for neonatology. *Semin. Perinatol.* **21**, 409–417 (1997).
42. Weiss, H. J. & Turitto, V. T. Prostacyclin (prostaglandin I₂, PGI₂) inhibits platelet adhesion and

- thrombus formation on subendothelium. *Blood* **53**, 244–250 (1979).
43. Chen, J. & López, J. A. Interactions of platelets with subendothelium and endothelium. *Microcirculation* **12**, 235–246 (2005).
 44. Denorme, F., Vanhoorelbeke, K. & De Meyer, S. F. von Willebrand factor and platelet glycoprotein Ib: A thromboinflammatory axis in stroke. *Front. Immunol.* **10**, 1–8 (2019).
 45. Schick, P. K., Walker, J., Profeta, B., Denisova, L. & Bennett, V. Synthesis and secretion of von Willebrand factor and fibronectin in megakaryocytes at different phases of maturation. *Arterioscler. Thromb. Vasc. Biol.* **17**, 797–801 (1997).
 46. Watson, S. Platelet activation by extracellular matrix proteins in haemostasis and thrombosis. *Curr. Pharm. Des.* **15**, 1358–1372 (2009).
 47. Gibbins, J. M., Okuma, M., Farndale, R., Barnes, M. & Watson, S. P. Glycoprotein VI is the collagen receptor in platelets which underlies tyrosine phosphorylation of the Fc receptor γ -chain. *FEBS Lett.* **413**, 255–259 (1997).
 48. Watson, S. P., Reep, B., Mcconnell, R. T. & Lapetina, E. G. Collagen stimulates [3H]inositol trisphosphate formation in indomethacin-treated human platelets. *Biochem. J.* **226**, 831–837 (1985).
 49. Hamberg, M., Svensson, J. & Samuelsson, B. Thromboxanes: a new group of biologically active compounds derived from prostaglandin endoperoxides. *Proc. Natl. Acad. Sci. U. S. A.* **72**, 2994–2998 (1975).
 50. Paul, B. Z. S., Jianguo, J. & Kunapuli, S. P. Molecular mechanism of thromboxane A₂-induced platelet aggregation. *J. Biol. Chem.* **274**, 29108–29114 (1999).
 51. Hechler, B. & Gachet, C. P₂ receptors and platelet function. *Purinergic Signal.* **7**, 293–303 (2011).
 52. Shepherd, G., Lewis, P., Blair, I., de Mey, C. & MacDermot, J. Epoprostenol (prostacyclin, PGI₂) binding and activation of adenylate cyclase in platelets of diabetic and control subjects. *Br. J. Clin. Pharmacol.* **15**, 77–81 (1983).
 53. Colli, S., Lombroso, M., Maderna, P., Tremoli, E. & Nicosia, S. Effects of PGI₂ on platelet aggregation and adenylate cyclase activity in human type IIa hypercholesterolemia. *Biochem. Pharmacol.* **32**, 1989–1993 (1983).
 54. Takagi, J., Petre, B. M., Walz, T. & Springer, T. A. Global conformational arrangements in integrin extracellular domains in outside-in and inside-out signaling. *Cell* **110**, 599–611 (2002).
 55. Weisel, J. W. & Litvinov, R. I. Fibrin formation, structure and properties. *Subcell Biochem* **82**, 405–456 (2017).
 56. Heemskerk, J. W. M., Siljander, P. R. M., Bevers, E. M., Farndale, R. W. & Lindhout, T.

- Receptors and signalling mechanisms in the procoagulant response of platelets. *Platelets* **11**, 301–306 (2000).
57. Reddy, E. C. & Rand, M. L. Procoagulant phosphatidylserine-exposing platelets in vitro and in vivo. *Front. Cardiovasc. Med.* **7**, 1–11 (2020).
 58. Szotowski, B., Antoniak, S., Poller, W., Schultheiss, H. P. & Rauch, U. Procoagulant soluble tissue factor is released from endothelial cells in response to inflammatory cytokines. *Circ. Res.* **96**, 1233–1239 (2005).
 59. Kovalenko, T. A., Panteleev, M. A. & Sveshnikova, A. N. The mechanisms and kinetics of initiation of blood coagulation by the extrinsic tenase complex. *Biophys. (Russian Fed.)* **62**, 291–300 (2017).
 60. Schreuder, M., Reitsma, P. H. & Bos, M. H. A. Blood coagulation factor Va's key interactive residues and regions for prothrombinase assembly and prothrombin binding. *J. Thromb. Haemost.* **17**, 1229–1239 (2019).
 61. Jarvis, G. E., Atkinson, B. T., Frampton, J. & Watson, S. P. Thrombin-induced conversion of fibrinogen to fibrin results in rapid platelet trapping which is not dependent on platelet activation or GPIb. *Br. J. Pharmacol.* **138**, 574–583 (2003).
 62. Gailani, D. & Bronze, G. J. Factor XI activation in a revised model of blood coagulation. *Science (80-.)*. 909–912 (1991).
 63. Jesty, J. & Beltrami, E. Positive feedbacks of coagulation: Their role in threshold regulation. *Arterioscler. Thromb. Vasc. Biol.* **25**, 2463–2469 (2005).
 64. Balandina, A. N., Shibeko, A. M., Kireev, D. A., Novikova, A. A., Shmirev, I. I., Panteleev, M. A., *et al.* Positive feedback loops for factor v and factor vii activation supply sensitivity to local surface tissue factor density during blood coagulation. *Biophys. J.* **101**, 1816–1824 (2011).
 65. Öberg, K. V. The role of platelet thrombin receptors PAR1 and PAR4 in platelet activation. *Linköping University Medical Dissertations* (2009).
 66. Heuberger, D. M. & Schuepbach, R. A. Protease-activated receptors (PARs): Mechanisms of action and potential therapeutic modulators in PAR-driven inflammatory diseases. *Thromb. J.* **17**, 1–24 (2019).
 67. Jung, S. M., Tsuji, K. & Moroi, M. Glycoprotein (GP) VI dimer as a major collagen-binding site of native platelets: Direct evidence obtained with dimeric GPVI-specific Fabs. *J. Thromb. Haemost.* **7**, 1347–1355 (2009).
 68. Jung, S. M., Moroi, M., Soejima, K., Nakagaki, T., Miura, Y., Berndt, M. C., *et al.* Constitutive dimerization of glycoprotein VI (GPVI) in resting platelets is essential for binding to collagen and activation in flowing blood. *J. Biol. Chem.* **287**, 30000–30013 (2012).

69. Poulter, N. S., Pollitt, A. Y., Owen, D. M., Gardiner, E. E., Andrews, R. K., Shimizu, H., *et al.* Clustering of glycoprotein VI (GPVI) dimers upon adhesion to collagen as a mechanism to regulate GPVI signaling in platelets. *J. Thromb. Haemost.* **15**, 549–564 (2017).
70. Moroi, M. & Jung, S. M. Platelet glycoprotein VI: Its structure and function. *Thromb. Res.* **114**, 221–233 (2004).
71. Slater, A., Perrella, G., Onselaer, M., Martin, E. M., Gauer, J. S., Xu, R., *et al.* Does fibrin(ogen) bind to monomeric or dimeric GPVI, or not at all? *Platelets* **30**, 281–289 (2019).
72. Loyau, S., Dumont, B., Ollivier, V., Boulaftali, Y., Feldman, L., Ajzenberg, N., *et al.* Platelet glycoprotein VI dimerization, an active process inducing receptor competence, is an indicator of platelet reactivity. *Arterioscler. Thromb. Vasc. Biol.* **32**, 778–785 (2012).
73. Horii, K., Kahn, M. L. & Herr, A. B. Structural basis for platelet collagen responses by the immune-type receptor glycoprotein VI. *Blood* **108**, 936–942 (2006).
74. Miura, Y., Takahashi, T., Jung, S. M. & Moroi, M. Analysis of the interaction of platelet collagen receptor glycoprotein VI (GPVI) with collagen: A dimeric form of GPVI, but not the monomeric form, shows affinity to fibrous collagen. *J. Biol. Chem.* **277**, 46197–46204 (2002).
75. Inoue, O., Suzuki-Inoue, K., McCarty, O. J. T., Moroi, M., Ruggeri, Z. M., Kunicki, T. J., *et al.* Laminin stimulates spreading of platelets through integrin $\alpha 6\beta 1$ -dependent activation of GPVI. *Blood* **107**, 1405–1412 (2006).
76. De Simone, I., Baaten, C. C. F. M. J., Jandrot-Perrus, M., Gibbins, J. M., Ten Cate, H., Heemskerk, J. W. M., *et al.* Coagulation factor XIIIa and activated protein C activate platelets via GPVI and PAR1. *Int. J. Mol. Sci.* **23**, 1–14 (2022).
77. Mammadova-Bach, E., Ollivier, V., Loyau, S., Schaff, M., Dumont, B., Favier, R., *et al.* Platelet glycoprotein VI binds to polymerized fibrin and promotes thrombin generation. *Blood* **126**, 683–691 (2015).
78. Onselaer, M.-B., Hardy, A. T., Wilson, C., Sanchez, X., Babar, A. K., Miller, J. L. C., *et al.* Fibrin and D-dimer bind to monomeric GPVI. *Blood Adv.* **1**, 1495–1504 (2017).
79. Munro, S. Lipid rafts: Elusive or illusive? *Cell* **115**, 377–388 (2003).
80. Brown, D. A. Membrane organization | Lipid rafts. *Encycl. Biol. Chem. Third Ed.* **2**, 817–820 (2021).
81. Simons, K. & Toomre, D. Lipid rafts and signal transduction. *Nat. Rev. Mol. Cell Biol.* **1**, 31–41 (2000).
82. Bye, A. P., Unsworth, A. J. & Gibbins, J. M. Platelet signaling: A complex interplay between inhibitory and activatory networks. *J. Thromb. Haemost.* **14**, 918–930 (2016).
83. Versteeg, H. H., Heemskerk, J. W. M., Levi, M. & Reitsma, P. H. New Fundamentals in

- hemostasis. *Physiol. Rev.* **93**, 327–358 (2013).
84. Asselin, J., Gibbins, J. M., Achison, M., Lee, Y. H., Morton, L. F., Farndale, R. W., *et al.* A collagen-like peptide stimulates tyrosine phosphorylation of syk and phospholipase Cy2 in platelets independent of the integrin $\alpha 2\beta 1$. *Blood* **89**, 1235–1242 (1997).
 85. Hitosugi, T., Sato, M., Sasaki, K. & Umezawa, Y. Lipid raft-specific knockdown of Src family kinase activity inhibits cell adhesion and cell cycle progression of breast cancer cells. *Cancer Res.* **67**, 8139–8148 (2007).
 86. Tsuji, M., Ezumi, Y., Arai, M. & Takayama, H. A novel association of Fc receptor γ -chain with glycoprotein VI and their co-expression as a collagen receptor in human platelets. *J. Biol. Chem.* **272**, 23528–23531 (1997).
 87. Bergmeier, W. & Stefanini, L. Platelet ITAM signaling. *Curr. Opin. Hematol.* **20**, 445–450 (2013).
 88. Auger, J. M., Best, D., Snell, D. C., Wilde, J. I. & Watson, S. P. c-Cbl negatively regulates platelet activation by glycoprotein VI. *J. Thromb. Haemost.* **1**, 2419–2426 (2003).
 89. Dangelmaier, C. A., Quinter, P. G., Jin, J., Tsygankov, A. Y., Kunapuli, S. P. & Daniel, J. L. Rapid ubiquitination of Syk following GPVI activation in platelets. *Blood* **105**, 3918–3924 (2005).
 90. Paolini, R., Molfetta, R., Beitz, L. O., Zhang, J., Scharenberg, A. M., Piccoli, M., *et al.* Activation of Syk tyrosine kinase is required for c-Cbl-mediated ubiquitination of Fc ϵ RI and Syk in RBL cells. *J. Biol. Chem.* **277**, 36940–36947 (2002).
 91. Chen, X., Ren, L., Kim, S., Carpino, N., Daniel, J. L., Kunapuli, S. P., *et al.* Determination of the substrate specificity of protein-tyrosine phosphatase TULA-2 and identification of Syk as a TULA-2 substrate. *J. Biol. Chem.* **285**, 31268–31276 (2010).
 92. Dunster, J. L., Mazet, F., Fry, M. J., Gibbins, J. M. & Tindall, M. J. Regulation of early steps of GPVI signal transduction by phosphatases: A systems biology approach. *PLoS Comput. Biol.* **11**, 1–26 (2015).
 93. Humphries, L. A., Dangelmaier, C., Sommer, K., Kipp, K., Kato, R. M., Griffith, N., *et al.* Tec kinases mediate sustained calcium influx via site-specific tyrosine phosphorylation of the phospholipase Cy Src homology 2-Src homology 3 linker. *J. Biol. Chem.* **279**, 37651–37661 (2004).
 94. Watson, S. P., Auger, J. M., McCarty, O. J. T. & Pearce, A. C. GPVI and integrin $\alpha 11\beta 3$ signaling in platelets. *J. Thromb. Haemost.* **3**, 1752–1762 (2005).
 95. Jackson, J. T., Mulazzani, E., Nutt, S. L. & Masters, S. L. The role of PLC γ 2 in immunological disorders, cancer, and neurodegeneration. *J. Biol. Chem.* **297**, 1–15 (2021).
 96. Vermeer, J. E. M., Van Wijk, R., Goedhart, J., Geldner, N., Chory, J., Gadella, T. W. J., *et al.* In

- vivo imaging of diacylglycerol at the cytoplasmic leaflet of plant membranes. *Plant Cell Physiol.* **58**, 1196–1207 (2017).
97. Martin-Romero, F. J., Pascual-Caro, C., Lopez-Guerrero, A., Espinosa-Bermejo, N. & Pozo-Guisado, E. Regulation of calcium signaling by STIM1 and ORAI1. *Calcium Signal Transduct.* 3–22 (2018) doi:10.5772/intechopen.78587.
 98. Watson, S. P., McConnell, R. T. & Lapetina, E. G. The rapid formation of inositol phosphates in human platelets by thrombin is inhibited by prostacyclin. *J. Biol. Chem.* **259**, 13199–13203 (1984).
 99. Daniel, J. L., Dangelmaier, C. A., Selak, M. & Smith, J. B. ADP stimulates IP3 formation in human platelets. *FEBS Lett.* **206**, 299–303 (1986).
 100. Oury, C., Toth-Zsamboki, E., Thys, C., Tytgat, J., Vermylen, J. & Hoylaerts, M. F. The ATP-gated P2X1 ion channel acts as a positive regulator of platelet responses to collagen. *Thromb. Haemost.* **86**, 1264–1271 (2001).
 101. Budde, U., Metzner, H. J. & Müller, H. G. Comparative analysis and classification of von Willebrand factor/factor VIII concentrates: Impact on treatment of patients with von Willebrand disease. *Semin. Thromb. Hemost.* **32**, 626–635 (2006).
 102. Tutwiler, V., Peshkova, A. D., Andrianova, I. A., Khasanova, D. R., Weisel, J. W. & Litvinov, R. I. Blood clot contraction (Retraction) is impaired in acute ischemic stroke. *Blood* vol. 128 271–279 (2016).
 103. Ornelas, A., Zacharias-Millward, N., Menter, D. G., Davis, J. S., Lichtenberger, L., Hawke, D., et al. Beyond COX-1: the effects of aspirin on platelet biology and potential mechanisms of chemoprevention. *Cancer Metastasis Rev.* **36**, 289–303 (2017).
 104. Di Minno, A., Spadarella, G., Prisco, D., Scalera, A., Ricciardi, E. & Di Minno, G. Antithrombotic drugs, patient characteristics, and gastrointestinal bleeding: Clinical translation and areas of research. *Blood Rev.* **29**, 335–343 (2015).
 105. CAPRIE Steering Committee. A randomised, blinded, trial of clopidogrel versus aspirin in patients at risk of ischaemic events (CAPRIE). *Lancet* **348**, 1329–1339 (1996).
 106. De Oliveira, E. I. & Bhatt, D. L. Clinical evaluation of clopidogrel across the whole spectrum of indications: Primary and secondary prevention of coronary artery disease. *Eur. Hear. Journal, Suppl.* **8**, 10–14 (2006).
 107. Gardiner, E., Arthur, J. & Andrews, R. Targeting GPVI as a novel antithrombotic strategy. *J. Blood Med.* **5**, 59–68 (2014).
 108. Borst, O. & Gawaz, M. Glycoprotein VI - novel target in antiplatelet medication. *Pharmacol. Ther.* **217**, 1–8 (2021).

109. Nieswandt, B., Schulte, V., Bergmeier, W., Mokhtari-Nejad, R., Rackebrandt, K., Cazenave, J. P., *et al.* Long-term antithrombotic protection by in vivo depletion of platelet glycoprotein VI in mice. *J. Exp. Med.* **193**, 459–469 (2001).
110. Harbi, M. H., Smith, C. W., Nicolson, P. L. R., Watson, S. P. & Thomas, M. R. Novel antiplatelet strategies targeting GPVI, CLEC-2 and tyrosine kinases. *Platelets* **32**, 29–41 (2021).
111. Crunkhorn, S. Novel antibody blocks thrombus formation. *Nat. Rev. Drug Discov.* **7**, 476 (2008).
112. Slater, A., Di, Y., Clark, J. C., Jooss, N. J., Martin, E. M., Alenazy, F., *et al.* Structural characterization of a novel GPVI-nanobody complex reveals a biologically active domain-swapped GPVI dimer. *Blood* **137**, 3443–3453 (2021).
113. Fischer, H. P. Mathematical modeling of complex biological systems: From parts lists to understanding systems behavior. *Alcohol Res. Heal.* **31**, 49–59 (2008).
114. Leiderman, K. & Fogelson, A. An overview of mathematical modeling of thrombus formation under flow. *Thromb. Res.* **133**, S12–S14 (2014).
115. Perdomo, J. L. & Pillis, L. G. De. Mathematical modeling of blood coagulation. (Harvey Mudd College, 2016).
116. Philibert, J. One and a half century of diffusion: Fick, Einstein, before and beyond. *J. Basic Princ. Diffus. Theory, Exp. Appl.* (2006).
117. Durrant, J. D. & McCammon, J. A. Molecular dynamics simulations in drug discovery. *Encycl. Bioinforma. Comput. Biol. ABC Bioinforma.* **1–3**, 652–665 (2018).
118. Klipp, E., Liebermeister, W., Wierling, C. & Kowald, A. *Systems biology: A textbook.* Wiley-VCH (2016).
119. Niazi, M. & Hussain, A. Agent-based computing from multi-agent systems to agent-based models: A visual survey. *Scientometrics* **89**, 479–499 (2011).
120. Soheilypour, M. & Mofrad, M. R. K. Agent-based modeling in molecular systems biology. *BioEssays* **40**, 1–8 (2018).
121. Das, A. A., Darsana, T. A. & Jacob, E. Agent-based re-engineering of ErbB signaling: A modeling pipeline for integrative systems biology. *Bioinformatics* **33**, 726–732 (2017).
122. Juliette, G., Peters, R. & Owen, D. M. An agent-based model of molecular aggregation at the cell membrane. *PLoS One* **15**, 1–17 (2020).
123. Städter, P., Schälte, Y., Schmiester, L., Hasenauer, J. & Stapor, P. L. Benchmarking of numerical integration methods for ODE models of biological systems. *Sci. Rep.* **11**, 1–11 (2021).
124. Beaumont, M. A. Approximate Bayesian Computation in Evolution and Ecology. *Annu. Rev. Ecol. Evol. Syst.* **41**, 379–406 (2010).

125. Toni, T., Welch, D., Strelkowa, N., Ipsen, A. & Stumpf, M. P. H. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *J. R. Soc. Interface* **6**, 187–202 (2009).
126. Alahmadi, A. A., Flegg, J. A., Cochrane, D. G., Drovandi, C. C. & Keith, J. M. A comparison of approximate versus exact techniques for Bayesian parameter inference in nonlinear ordinary differential equation models. *R. Soc. Open Sci.* **7**, (2020).
127. Burnham, K. P. & Anderson, D. R. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. (2002).
128. van der Vaart, E., Johnston, A. S. A. & Sibly, R. M. Predicting how many animals will be where: How to build, calibrate and evaluate individual-based models. *Ecol. Modell.* **326**, 113–123 (2016).
129. Gurney, K. *An introduction to neural networks*. (UCL Press, 1997).
130. Leshno, M., Lin, V. Y., Pinkus, A. & Schocken, S. Multilayer feedforward networks with non-polynomial activation functions can approximate any function. *Neural Networks* **6**, 861–867 (1993).
131. Feng, J. & Lu, S. Performance analysis of various activation functions in artificial neural networks. *J. Phys. Conf. Ser.* **1237**, (2019).
132. Sharma, S., Sharma, S. & Athaiya, A. Activation functions in neural networks. *Int. J. Eng. Appl. Sci. Technol.* **04**, 310–316 (2020).
133. Chatterjee, M. S., Purvis, J. E., Brass, L. F. & Diamond, S. L. Pairwise agonist scanning predicts cellular signaling responses to combinatorial stimuli. *Nat. Biotechnol.* **28**, 727–732 (2010).
134. Lee, M. Y. & Diamond, S. L. A human platelet calcium calculator trained by pairwise agonist scanning. *PLoS Comput. Biol.* **11**, 1–24 (2015).
135. Wold, S., Sjöström, M. & Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **58**, 109–130 (2001).
136. Abdi, H. Partial least squares regression and projection on latent structure regression (PLS Regression). *Wiley Interdiscip. Rev. Comput. Stat.* **2**, 97–106 (2010).
137. Boulesteix, A. L. & Strimmer, K. Partial least squares: A versatile tool for the analysis of high-dimensional genomic data. *Brief. Bioinform.* **8**, 32–44 (2007).
138. Kreeger, P. K. Using partial least squares regression to analyze cellular response data. *Sci. Signal.* **6**, 1–6 (2013).
139. Bourgeois, D. L. & Kreeger, P. K. *Partial least squares regression models for the analysis of kinase signaling*. *Kinase Signaling Networks* (2017).

Chapter 2 (published)

An agent-based approach for modelling and simulation of glycoprotein VI receptor diffusion, localisation and dimerisation in platelet lipid rafts

Receptor clustering and dimerisation are important for receptor activation because they facilitate the formation of signalling complexes and enhances the efficiency of downstream signalling. Clustering brings receptors in close proximity, allowing them to interact and initiate intracellular signalling cascades, leading to cellular responses such as gene expression, enzyme activation, or membrane trafficking, while dimerisation is a prerequisite for some receptor activation and ligand binding. This chapter presents an agent-based model (ABM) approach to investigate the clustering and dimerisation of the platelet- and megakaryocyte-specific receptor for collagen glycoprotein VI (GPVI) and explores the role of glycolipid-enriched raft-like domains in regulating receptor diffusion.

Sci Rep 13, 3906 (2023).

<https://doi.org/10.1038/s41598-023-30884-6>

An agent-based approach for modelling and simulation of glycoprotein VI receptor diffusion, localisation and dimerisation in platelet lipid rafts

Chukiat Tantiwong^{1,5}, Joanne L. Dunster¹, Rachel Cavill², Michael G. Tomlinson³, Christoph Wierling⁴, Johan W. M. Heemskerk^{5,6}, Jonathan M. Gibbins^{1,*}

¹ School of Biological Sciences, University of Reading, Reading, UK

² Department of Data Science and Knowledge Engineering, Maastricht University, Maastricht, The Netherlands

³ School of Biosciences, University of Birmingham, Birmingham, UK

⁴ Alacris Theranostics GmbH, Berlin, Germany

⁵ Department of Biochemistry, CARIM, Maastricht University, Maastricht, The Netherlands

⁶ Synapse Research Institute, Maastricht, Netherlands

* Corresponding author: j.m.gibbins@reading.ac.uk

Abstract

Receptor diffusion plays an essential role in cellular signalling via the plasma membrane microenvironment and receptor interactions, but the regulation is not well understood. To aid in understanding of the key determinants of receptor diffusion and signalling, we developed agent-based models (ABMs) to explore the extent of dimerisation of the platelet- and megakaryocyte-specific receptor for collagen glycoprotein VI (GPVI). This approach assessed the importance of glycolipid enriched raft-like domains within the plasma membrane that lower receptor diffusivity. Our model simulations demonstrated that GPVI dimers preferentially concentrate in confined domains and, if diffusivity within domains is

decreased relative to outside of domains, dimerisation rates are increased. While an increased amount of confined domains resulted in further dimerisation, merging of domains, which may occur upon membrane rearrangements, was without effect. Modelling of the proportion of the cell membrane which constitutes lipid rafts indicated that dimerisation levels could not be explained by these alone. Crowding of receptors by other membrane proteins was also an important determinant of GPVI dimerisation. Together, these results demonstrate the value of ABM approaches in exploring the interactions on a cell surface, guiding the experimentation for new therapeutic avenues.

Keywords: Agent-based modelling, confined domain, glycoprotein VI, receptor diffusion, receptor dimerisation, transmembrane receptors, lipid raft

Introduction

The plasma membrane of eukaryotic cells provides a physical and biochemical interface¹ that allows the precise control of cell functions, facilitates shape change and movement², attachment to the extracellular matrix or other cells, the controlled transfer of solutes outside-in and inside-out³, and the onset of the signalling mechanisms that regulate a cell⁴. Through the basic structure of its phospholipid bilayer, the plasma membrane provides a specialised environment in which cell surface receptors engage with extracellular ligands to trigger the transduction of signals in the cytosol. These signals are then propagated and amplified through enzyme cascades culminating in a controlled change in cell behaviour, for instance in gene expression, migration, secretion, proliferation, survival and apoptosis⁵⁻⁷.

Transmembrane receptors may move laterally within the phospholipid plane of the plasma membrane, although there are movement restraints due to the presence of and

linkage to other surface proteins, as well as due to the presence of intracellular proteins, such as the membrane actin-myosin and tubular cytoskeletons^{8,9}. The receptors may also be restricted in their movements due to the position of ligands, for instance, in the extracellular matrix¹⁰, or due to ligand-induced dimerisation or clustering, as in cases of the insulin and antibody receptors^{11,12}. Interactions of plasma membrane receptors with other proteins inside the cell are furthermore controlled via biochemical processes such as post-translational modifications of proteins (phosphorylation, acetylation, ubiquitination, sumoylation, glycosylation, lipidation), ultimately leading to precisely regulated temporal and spatial control of cell signalling mechanisms¹³⁻¹⁶.

The ability of receptors to initiate cell signalling is influenced by the membrane phospholipid composition and distribution⁷. Small and transient nanodomains of the membrane enriched in cholesterol and glycolipids, known as lipid rafts, present unique physicochemical properties, enabling a highly localised enrichment of cholesterol and other lipid molecular species to influence membrane fluidity and the ability of proteins to move within¹⁷. The concept of intra-membrane heterogeneities and lipid rafts has thereby facilitated our understanding of the spatiotemporal orchestration of receptor signalling mechanisms.

Limited efforts have been made so far to develop theoretical models that combine the effects of intra-membrane constraints and ligand-induced actions for understanding of the critical elements of receptor localisation and movement. One approach that can be used to study the dynamics of a particle on a membrane is agent-based modelling (ABM). Previous work has utilised this approach to investigate the formation of generalised molecular clusters¹⁸, finding that protein diffusion is influenced by its neighbourhood, or to investigate more specific questions about particular receptor classes (such as integrins¹⁹) without

recourse to data. Das and coworkers¹⁸ have developed an in-house code to link data and agent-based models to answer specific questions centring on the activation of trafficking of EGFR-HER2 receptors.

For this study we constructed a simple and effective model, based on experimental evidence, for predicting receptor movements on the anucleate platelets using the ABM approach. Our chosen target was the receptor for collagen, glycoprotein (GP) VI, which is uniquely expressed on blood platelets and megakaryocytes^{20,21}. The binding of collagen to this receptor leads to GPVI dimerisation and clustering, and to a signalling response that culminates in rapid thrombus formation, which contributes to haemostasis²². The monomeric GPVI receptor has a weak affinity for collagen and is non-covalently associated with the Fc receptor γ -chain, through which it transmits signals²³. Receptor dimerisation results in the formation of a complex with a higher affinity for collagen (Figure 1A), thus facilitating ligand binding and signalling responses²⁴⁻²⁶.

Developing an ABM with distinct regions of membrane lipid composition - here referred to as confined domains that are proxy entities for lipid rafts^{28,29} - we studied how GPVI receptors on the platelet plasma membrane can switch between monomer and dimeric entities. Our modelling studies support the preferential enrichment for GPVI in lipid rafts, in agreement with experimental observations³⁰. Through simulation of multiple facets of the plasma membrane and membrane proteins, we thus provide a basis for understanding how receptor complexes form and function, and can impact altered receptor signalling processes in disease.

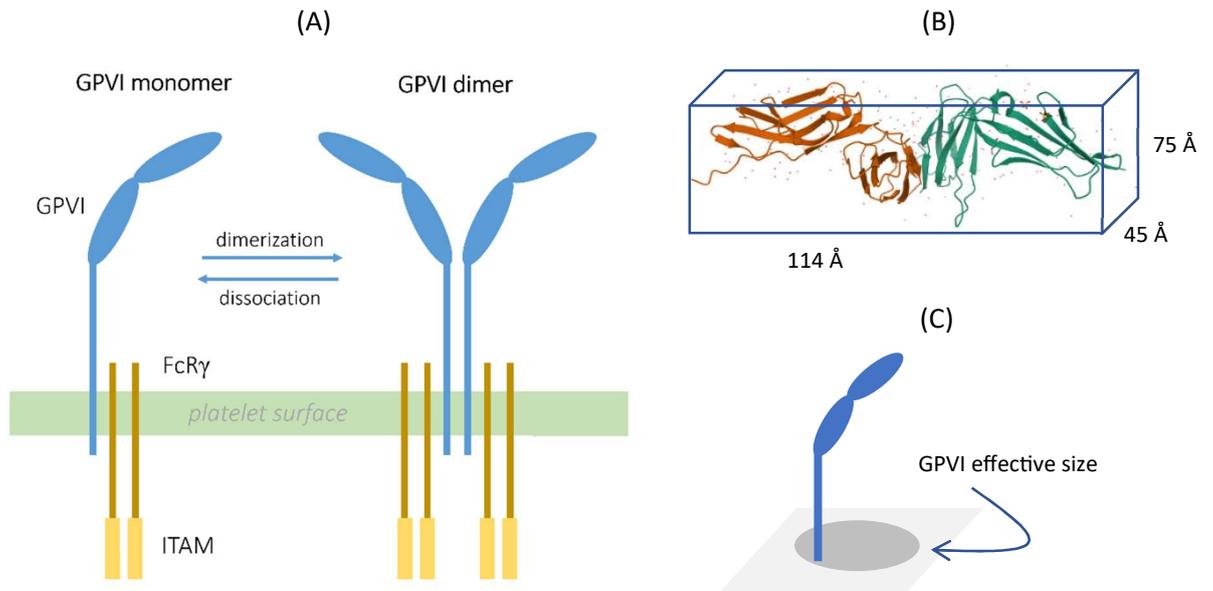


Figure 1. Structure and dimerisation of platelet GPVI receptors. (A) The extracellular domain of monomeric GPVI on platelets comprises of two IgG domains and a connection to the transmembrane domain (blue). The GPVI protein is stably connected to two chains of the FcR γ -chain, forming ITAM-containing signalling domains. Monomers of GPVI can dimerise with other monomers (dimerisation), a process that is reversible (dissociation). Adapted from Induruwa et al. (2016)²⁷. (B) Crystal structure of human platelet GPVI. Image taken from RCSB PDB (rcsb.org), annotation PDB ID 2GI7²³. (C) Projected illustration of GPVI as a transmembrane protein with an assigned effective area in two dimensions.

Methodology

1. Application of agent-based modelling (ABM)

An ABM approach was used to simulate agents (receptors and lipids) on the cell surface³¹. This approach has been used in different fields of physical science, biological science, social science, and finances³². For example, several recently published works used ABM to study the spreading of the COVID-19 pandemic^{33–35}. There are several ways to implement ABM, either by coding the model from scratch or using existing software. A commonly used ABM software package is NetLogo, which is multi-purpose, computationally efficient and easy to use, offering the advantage of being easily implemented and modified by

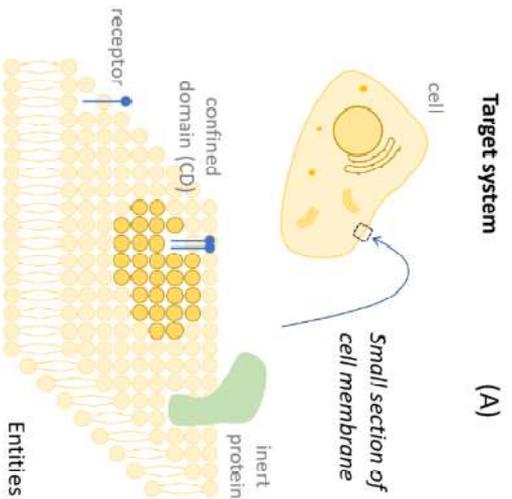
non-theoretical experimentalists³⁶. Using NetLogo, we simulated the diffusion of receptors in a two-dimensional plasma membrane. The implementation of this is demonstrated in Figure 2A-D, and a flowchart is provided in the Supplement. The generated models can be easily modified to model different kinds of receptors and transmembrane proteins, by adjusting properties such as size, mass and diffusivity. To ease this modification, the code to run simulations is made available, and details on how to install and implement it are given in the Supplement. In our ABM approach, receptors are able to move with an assigned behaviour, which is either deterministic or stochastic as modelled. Certain areas of the plasma membrane were considered as confined areas with reduced diffusivity. By default, components in the system were studied in a two-dimensional box with periodic boundary conditions to imitate an infinite membrane³⁷.

2. Brownian motion

Agents (receptor molecules and other membrane proteins) were considered to move freely in the two-dimensional surface in random directions. By applying a mean square distance (MSD) of Brownian motion on a two-dimensional surface as time (t) dependent³⁸:

$$MSD = 4Dt,$$

a given step size ($dS \sim MSD^{1/2}$) was taken, depending on the agent's diffusivity (D) as $dS \sim D^{1/2}$. Herein, the constant of variation was a function of the applied scaling. Agents in the simulation were modelled as circular discs, which never overlapped. It was assumed that the area occupied by one receptor is conserved during dimerisation, and that the space occupied by two monomers is equal to that occupied by one dimer, $\pi R_{dimer}^2 = 2\pi R_{monomer}^2$. The sizes (radii) of dimer R_{dimer} and monomer were then related as $R_{dimer} = \sqrt{2} R_{monomer}$ (Figure 2C). The relationship of diffusivity and particle size was retrieved from the Stokes-Einstein relation³⁹:



(A)

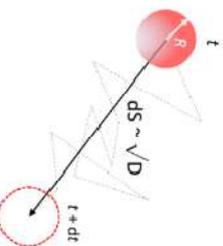
- From Literatures**
- What is the size of the cell?
 - What is # receptor per cell?
 - What is the size of receptor?
 - What is % of CD?
 - What is the size of CD?

- Parameterisation**
- Determine #box per cell
 - Determine box size
 - Determine # receptor per box

- Scaling**
- Length-scaling: box $\rightarrow N^2$ grids
 - $x \rightarrow (x/\text{box-size}) * N$

(C)

Area conserved in dimerisation



Particle step-size
 $ds^2 \sim \text{Mean Square Distance (MSD)}$
 $= 4Ddt \sim D$



Stokes-Einstein's relation
 $D = k_B T / 6\pi\eta R \sim 1/R$
 $D_{\text{dimer}} = D_{\text{monomer}} / \sqrt{2}$

Simulation scaling

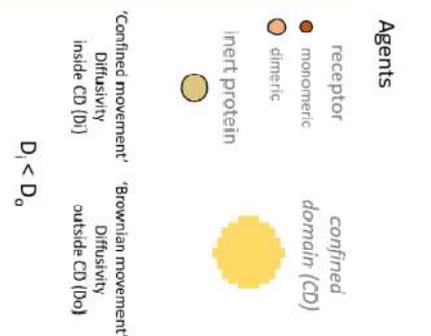
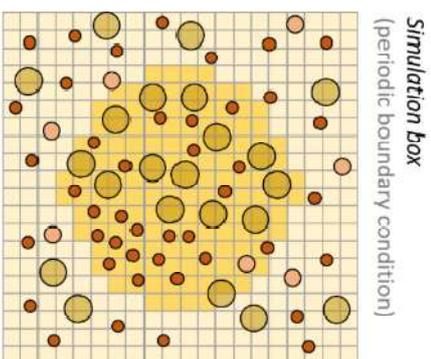
- Diffusivity: $D = 1$
- Time-step: $dt = 1$
- Step-size: $ds \sim D^{1/2}$

Monomer & Dimer

- $R_{\text{dimer}} = \sqrt{2} R_{\text{monomer}}$
- $D_{\text{dimer}} = D_{\text{monomer}} / \sqrt{2}$
- $ds_{\text{dimer}} = ds_{\text{monomer}} / \sqrt{2}$

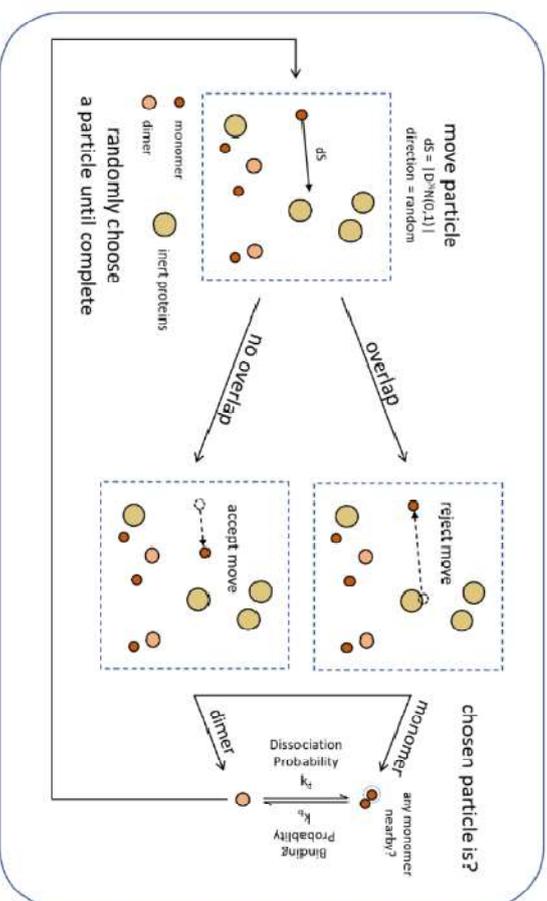
Agent based model

(B)



in each time-step

(D)



◀ **Figure 2. Overview of ABM simulation procedure.** (A) The target system, i.e. the platelet membrane. The simplified version of a membrane consists of two areas, i.e. parts where molecules are confined in movements (confined domains), and the remaining part where they move freely (Brownian motion). In addition to inert proteins, the receptors of interest are indicated as transmembrane proteins. (B) Application of ABM to target receptor dimerisation. The membrane in the simulation box consists of agents (receptor molecules) in monomeric or dimeric forms and inert proteins. The confined domains are considered to represent lipid rafts. All agents are treated as independent, of which mathematical rules determine their properties and interactions. (C) Assignment of agent parameters. The simulation parameters included diffusivity, particle size and step size. (D) Rules for agent movements. Each simulation step consists of a randomly placed agent with random walk (rejected in case of overlapping), dimerisation and dissociation. Steps are repeated until all agents are selected, after which movements follow. The role of the confined domain is to capture differences in diffusivity between receptors on and off lipid rafts. The dimensionless/scaled parameter values are discussed later in Methodology section 4.

$$R = k_B T / 6\pi\eta D \sim 1/D,$$

where k_B , T , and η are Boltzmann constant, temperature, and viscosity, respectively. Although this formula is modelled in a 3-dimensional case, we presumed that the inversely proportional relationship between R and D was retained in 2-dimensions, the coefficient being absorbed in the scaling process. Combining these assumptions, the relationship between step size of monomer and dimer was:

$$dS_{\text{monomer}} = 2^{3/4} dS_{\text{dimer}}$$

Note that the step size of an agent (receptor) in each time step may not be equal. In the calculation above, the maximum step size was set, but the actual step size in each movement could be generated according to a Wiener's process, $dS_{\text{actual}} \sim |N(0,1)| dS_{\text{maximum}}$. Herein $N(0,1)$ forms a random variable with a standard normal distribution (Figure 2D).

Experimentally, using single-particle tracking, it has been seen that the diffusivity of GPVI molecules on mouse platelets decreased by approximately ten times, when present in regions with confined membrane properties⁴⁰, with the receptor's mode of motion changing from Brownian movement to restricted movement. In our ABM implementation, the mode of motion of the receptor inside and outside the confined domain remained the same; the only difference being the diffusivity. While the presence of this domain confined the movement of the receptor, we assumed that the receptor was effectively moved slower, with a smaller diffusivity within the domain.

3. Receptor dimerisation

The effects of dimerisation and dissociation of receptors were captured by the probabilities k_b and k_d , respectively. Herein, dimerisation was defined as the conversion from two monomers to one dimer. The proximity that two monomers have to fall within in order to form a dimer was arbitrarily set at 10% of the monomer's diameter. For calculating the conversion, a random number $R_{[0,1]} \in [0,1]$ was generated. Dimerisation occurred if this number met the condition of $R_{[0,1]} < k_b$. Conversely, dissociation was imputed as the change from one dimer to two monomers. For dimer movements, also a random number $R_{[0,1]}$ was generated, and dissociation occurred when $R_{[0,1]} < k_d$ (Figure 2D).

4. Parameterisation and scaling analysis

The following section explains how values were assigned to parameters. Note that when precise values for parameters were not available, order of magnitude estimates needed to be made, applicable to the platelet surface and the collagen receptor GPVI. The simulation conversion parameters estimated in the following section are summarised in Table 1.

Parameter	Real-world scale	Simulation scale
Size of simulation box	$3.0 \times 10^{-7} \text{ m}$	30
GPVI effective size	114 Å	3.8% of 30 \sim 1.14
GPVI diffusivity	$0.091 \times 10^{-12} \text{ m}^2/\text{s}$	1
Expected step size of GPVI in a time-step	12.5 nm	$(\pi/2)^{1/2} \sim 1.25$
Time-step	0.43 ms	1

Table 1. List of real-world and simulation parameters. See estimates in Methodology, section 4.

4.1. Platelet surface area.

The platelet volume based on previous work⁴¹ was taken to be $V_p \approx 7.4 \times 10^{-18} \text{ m}^3$, allowing us to determine (by assuming that platelets are perfect spheres) the radius R and the surface area A :

$$R \approx 1.2 \times 10^{-6} \text{ m},$$

$$A \approx 1.8 \times 10^{-11} \text{ m}^2.$$

Some assumptions needed to be made in considering the shape and volume of platelets, since their activation results in changes in morphology and membrane organisations. We reasoned that with the open canicular system exposed, following activation, the morphology of a platelet is closer to a sphere than a discoid. If an average discoid platelet is considered to have an average diameter of $\sim 3 \mu\text{m}$, the thickness of the cell can be determined as $\sim 1 \mu\text{m}$. Thus, the surface area of a platelet would be $\sim 2.4 \times 10^{-11} \text{ m}^2$ ($\sim 33\%$ more than a spherical shape). If we account for the contribution of the open canicular system (estimated to be $\sim 25\%$ of the plasma membrane surface)⁴², the total surface area will increase to $3.2 \times 10^{-11} \text{ m}^2$. However, since the open canicular system is continuous with the plasma membrane, we assumed that the volume of a platelet remains constant during shape

change. The consequences of a different receptor surface density is addressed in section 9 of the results. For the remaining simulations, we maintained a platelet surface area of $1.8 \times 10^{-11} \text{ m}^2$, consistent with spherical shape with a diameter of $2.4 \text{ }\mu\text{m}$.

4.2. Simulation box size.

The (transient) confined domain diameter for a lipid raft of $d \approx 100 - 300 \text{ nm}$ was obtained from an earlier publication²⁸. For convenience, we used a raft size of 200 nm . Note that the size did not affect the model outcomes (see results, section 4). A model limitation is the assumption of the confined domain as a single circular area in the centre of a periodic box, implying that a too-small box can result in simulation artefacts. In other words, if a raft size is smaller than 30 nm , less than one receptor molecule will be present inside a box. Too-small number of receptors per box could also lead to high fluctuations in the simulation results. We further assumed lipid rafts occupy about 35% of the plasma membrane surface area⁴³. The total area occupied by lipid rafts was then calculated as:

$$A_{\text{raft}} \approx 35\% \times A \approx 6.4 \times 10^{-12} \text{ m}^2.$$

Considering this as the total area of confined domains, with $d \approx 200 \text{ nm}$, the count of domains was:

$$\text{box}_{\text{per platelet}} = A_{\text{raft}} / (\pi(d/2)^2) \approx 205.$$

In the simulation, the box area (consisting of one confined domain per box) was:

$$A_{\text{box}} = A / \text{box}_{\text{per platelet}} \approx 9.0 \times 10^{-14} \text{ m}^2,$$

with a box length of:

$$L = (A_{\text{box}})^{1/2} \approx 3.0 \times 10^{-7} \text{ m}.$$

4.3. Receptor count per box.

The number of GPVI molecules in a single platelet was estimated as ≈ 9600 copies⁴⁴.

This gave as a number of GPVI monomers per simulation box:

$$\text{GPVI}_{\text{per box}} = \text{GPVI}_{\text{per platelet}} / \text{box}_{\text{per platelet}} \approx 9600/117 \approx 47$$

4.4. GPVI receptor molecule dimensions.

The molecular dimensions of GPVI were taken from its crystal structure²³: $114 \text{ \AA} \times 45 \text{ \AA} \times 75 \text{ \AA}$. Considering the extremum case that its longest side is the projected diameter of the GPVI on the platelet surface (Figure 1B,C), we choose a $d_{\text{GPVI}} \approx 114 \text{ \AA}$. The size of a GPVI monomer scaled to the box size then was:

$$d_{\text{GPVI}}^{\text{scaled}} = d_{\text{GPVI}}/L \approx (114 \times 10^{-10})/(3.0 \times 10^{-7}) \approx 3.8\%.$$

4.5. Step size and time scale of modelling.

The diffusivity of a single GPVI molecule in the membrane (off lipid raft) has been measured before⁴⁰, $D_{\text{exp}} \approx 0.091 \times 10^{-12} \text{ m}^2\text{s}^{-1}$. From the mean square distance of particle on a two-dimensional surface moving in Brownian motion, the step size can be scaled as:

$$dS^2 \approx 4Ddt.$$

According to this equation, we could either pick dS and determine the scale of dt from dS or vice versa. To simplify the simulation, we scaled the step size to order $O(1)$ by setting $D \sim 1$, $dt = 1$, and $dS = D^{1/2} |N(0,1)|$. Note that the constant 4 was absorbed in the D scaling and that random Brownian motion was assumed to have a random standard normal distribution, $N(0,1)$. With these definitions, the expected step size was calculated at:

$$dS_{\text{expected}} \approx (\pi/2)^{1/2} \times (L/30) \approx 12.5 \times 10^{-9} \text{ m},$$

where $r_{\text{expected}} = (\pi/2)^{1/2}$ is the expected distance determined by a standard normal distribution

function, and 30 comes from the defined scaled box size. Hence, the time scale was set as:

$$dt \approx dS^2/4D \approx (12.5 \times 10^{-9})^2 / (4 \times 0.091 \times 10^{-12}) \approx 0.43 \text{ ms.}$$

This time scale was small enough to capture the confined behaviour of particles, which occurs in seconds²⁸.

4.6. Inert proteins.

The modelling further included an unknown number of transmembrane proteins that have no interaction with the receptor of interest. The effect of a collision between proteins was already incorporated in the diffusion simulation via Brownian motion. The motion direction and step size changed randomly due to random encounters, implying that the presence of inert proteins was included by default. Additional parameters such as additional inert proteins (in arbitrary numbers) were used to check for effects on receptor dimerisation, while they essentially block receptors from moving into occupied spaces.

5. Standard setup of ABM simulations

Simulations were performed in NetLogo 6.2.2 (Supplementary Figure S1), using the algorithm illustrated in Figure 2 (for details see Supplementary Figure S2). A list of simulation parameters per research question is provided in Table 2. The default start setting was 47 receptor monomers that were uniformly distributed in a box representing the plasma membrane. Of note, this default did not take into account the heterogeneities caused by membrane cytoskeletal connections and receptor complexes, although the model may reach a non-uniform equilibrium after the simulation. Based on the calculations above, the diameter of GPVI was approximated as 3.8% of the length of the simulation box (scaled as 30 x 30 pixels). The movement speed of monomers was set to $D^{1/2}$. Each simulation was run for $\geq 200,000$ steps

to ensure equilibrium. An average of the last 50,000 steps was used for the analysis.

Except where indicated otherwise, binding and dissociating probabilities were arbitrarily defined as $k_b = 0.05 \text{ molecule}^{-1}$ per unit time and $k_d = 0.01$ per unit time (or per timestep, $dt \sim 0.43 \text{ ms}$ as calculated above). These numbers were chosen to ensure balancing of the time scale of dimerisation and dissociation, *i.e.*, to prevent an equilibrium without dimers or monomers. This also ensured that the number of GPVI molecules in dimeric form in the simulations were broadly consistent with the dimeric levels measured experimentally²⁶. The impact of variation of these parameters is shown in section 10 of the results. All simulations were repeated three times. The code for this model, together with the setup for each simulation, is available in the Supplement.

Simulation type	Binding rate (k_b) (per molecule per unit time)	Dissociation rate (k_d) (per unit time)	Diffusivity outside CD (D_o) (unit length ² /per unit time)	Diffusivity inside CD (D_i) (unit length ² /per unit time)	CD occupied area (%)	Number of added inert proteins (molecule)	Number of inert protein packs (dimensionless)	Fold number of CD merging (dimensionless)	GPVI number per platelet (% of 9600)
1. Receptors in CD area vs D_i	0	0	1	$1, 2^{-1}, \dots, 2^{-9}, 2^{-10}$	35	0	0	1	100
2. Dimerisation (with CD) vs D_i	0.05	0.01	1	$1, 2^{-1}, \dots, 2^{-9}, 2^{-10}$	35	0	0	1	100
3. Dimerisation vs %CD & D_i	0.05	0.01	1	$1, 2^{-1}, \dots, 2^{-4}, 2^{-5}$	0,5...,75,80	0	0	1	100
4. Dimerisation vs CD merging	0.05	0.01	1	0.1	35	0	0	0.5,1...,7.5,8	100
5. Added inert proteins	0.05	0.01	1	1	0	0,25...,175,200	0,25...,175,200	1	100
6. Disintegrated inert proteins	0.05	0.01	1	1	0	100	$1, 2^1, \dots, 2^7, 2^8$	1	100
7. Dimerisation (w/o CD) vs D	0.05	0.01	$2^{-5}, 2^{-4}, \dots, 2^4, 2^5$	1	0	0	0	1	100
8. Receptors in CD area vs %CD	0, 0.01, 0.005, ..., 0.000625	$k_b/5$	1	0.1	15,16...,22	0	0	1	100
9. Dimerisation vs receptor number	0.05	0.01	1	0.1	35	0	0	1	25, 50, ..., 150, 175, 200
10. Dimerisation vs k_b & k_d	60, 70, ..., 130, 140 % of 0.05	60, 70, ..., 130, 140 % of 0.01	1	0.1	35	0	0	1	100

Table 2. List of parameters used in each simulation. Abbreviation: CD, confined domain. Please note in this context, k_d and k_b are implemented as a rate in unit time as described in the Methodology, section 3.

Results

In the present study, we aimed to understand how a fluid-mosaic plasma membrane influences receptor diffusion, interaction or dimerisation, and the initiation of cell signalling. According to the mosaic model, the phospholipids and proteins are not uniformly distributed. Lipid patches (rafts) are considered to concentrate signalling proteins, including receptors, thereby permitting or enhancing cell signalling processes^{17,30}. Precisely how this occurs has not yet been resolved. The agent-based modelling (ABM) approach allowed us to explore the impact of confined lipid domains within the plasma membrane on the enrichment and clustering of the collagen receptor GPVI. The model can easily be applied to other receptors and cell types of interest, with adapted parameters as in the methods section.

To address ABM simulations, we designed a receptor-containing simulation box, representing a defined square part of the plasma membrane with mobile GPVI molecules and initially a single confined domain ("raft"). With the chosen parameters, we assumed that GPVI monomers have no inherent tendency to form dimers or clusters. Note that the confined domain was not included in the simulation from sections 5 to 7.

1. Simulated receptors preferentially localise to confined domain areas

Differential diffusivity in the lipid domains of a membrane may result in an uneven distribution of transmembrane proteins. In the present ABM, we assumed that the confined domains contain a higher level of proteins that are free to move inside or outside^{45,46}. In a series of simulations, we tested this idea.

The proportion occupied by the confined domain, as assumed in rafts, was estimated as 35% by Prior *et al.*³⁹. The size of the confined domain was fixed as a circle, which represented a domain of lower protein diffusion. The diffusivity ratio outside and inside the

confined domain was varied to simulate effects on receptor diffusion.

If the receptor localisation is not affected by diffusivity, the relative numbers of receptors located inside or outside the confined domain will be similar for all diffusivity ratios. In Figure 3A, the ratio of diffusivity of receptors between the outside and inside of a confined domain, expressed as $D_{out}:D_{in}$, was taken as an independent variable and then changed from 2^0 , 2^1 , 2^2 , ... to 2^{10} (i.e., 1024). The actual ratio can be estimated to be ~ 10 , according to single particle tracking results of GPVI molecules in mouse platelets⁴⁰. In our studies we varied this ratio from 1 to 1024 to explore the extreme relationships between diffusivity and location preference of GPVI. The number of receptors located inside the domain, as a dependent variable, was found to asymptotically reach 100%, with 50% at a $D_{out}:D_{in}$ of in the range of 8 to 16 (Figure 3B). Note that if a different diffusivity ratio does not affect the receptor localisation, this number should not deviate from the starting value of 35%. Based on the obtained changes at default model settings, we concluded that GPVI receptors will preferentially localise to the confined domains, *i.e.*, the areas with lower diffusivity.

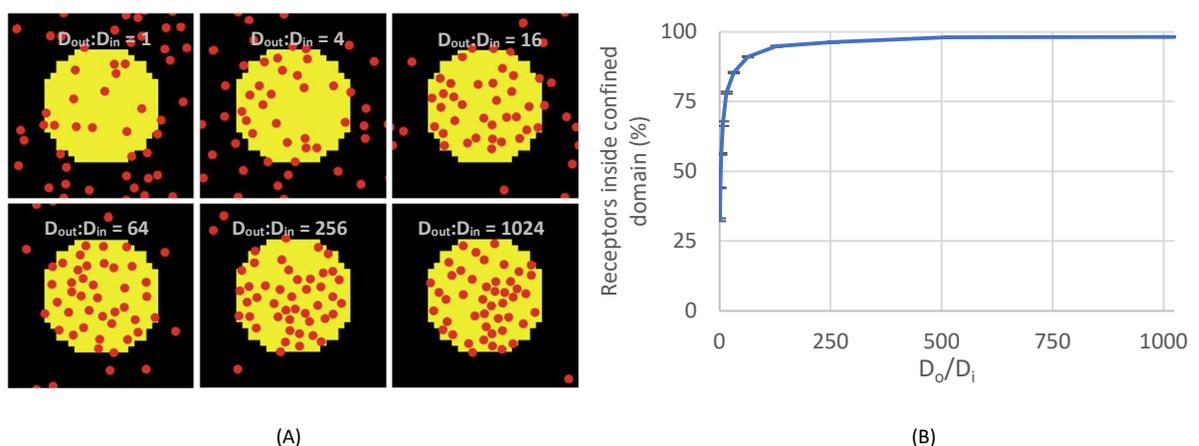


Figure 3. Preferential localisation of single receptors in the confined domain. (A) Snapshots of 11 simulations of 47 receptors (red dots) moving on the simulated membrane with confined domain (yellow circle). Note the sub-micrometer size of the simulation box of $0.3 \times 0.3 \mu\text{m}$, and the initial random distribution of GPVI receptors. Simulations were run for $\geq 200,000$ steps, with

an average of the last 50,000 steps shown. The diffusivity ratio between outside and inside confined domains, $D_{out}:D_{in}$, was changed from 1 (2^0) to 1024 (2^{10}). The snapshots shown are for $D_{out}:D_{in} = 1, 4, 16, 64, 256$ and 1024. (B) Effect of an altered ratio $D_{out}:D_{in}$ on number of receptors inside the confined domain. Each simulation was repeated three times, means \pm SD.

Previous studies have demonstrated that GPVI is present in cholesterol-rich lipid rafts. GPVI recruitment occurs upon platelet adhesion to collagen⁴⁷, a process which can lead to GPVI clustering⁴⁸. While these membrane structures concentrate specific signalling proteins within, recent studies reveal that lipid rafts also cage or restrict protein and receptors diffusion^{49,50}, which may be a prerequisite for GPVI clustering. Indeed, a heterotypic interaction of GPVI with PECAM1 is increased in lipid rafts⁵¹. Considering that lipid rafts can orchestrate the GPVI signalling⁵², we hypothesized that lowered diffusivity in rafts compared to non-raft domains results in an increased GPVI dimerisation within.

2. Decreasing diffusivity in the confined domain increases receptor dimerisation

We then explored how the confined domain affected the likeliness of receptor dimerisation, a process that is known to enhance GPVI ligand-binding properties^{24–26}. For simplicity in the ABM simulation, we assumed that dimerisation is not modulated by other proteins in the plasma membrane or actin cytoskeleton. We thus assumed that the fraction of receptors in dimeric form remains the same for all $D_{out}:D_{in}$ ratios.

As illustrated in simulation snapshots (Figure 4A), we found that an increase in the diffusivity ratio (i.e., lower diffusivity in the confined domain with $D_{out}:D_{in}$ set from 2^0 to 2^{10}) yielded a higher number of receptor dimers. Herein, the ratio of diffusivity of receptors outside or inside the confined domains was taken as an independent variable. The dimeric receptors increased non-linearly with the diffusivity ratio to reach a saturation level of 80%

(Figure 4B). The simulation thus pointed to a main effect of intra-membrane differences in receptor diffusivity for promoting receptor dimerisation.

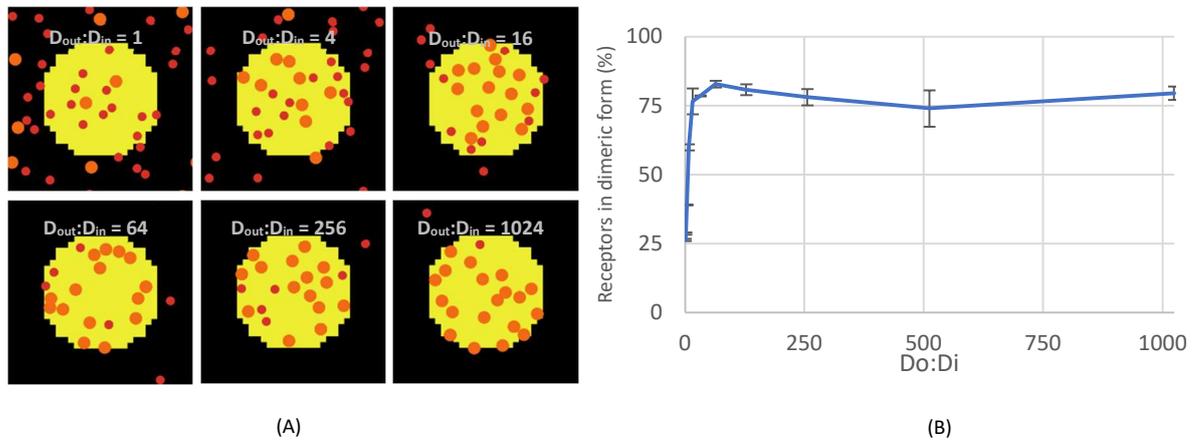


Figure 4. Higher diffusivity ratio enhances receptor confinement and dimerisation. (A) Snapshots of simulation of receptors in monomeric (red) or dimeric (orange) forms in the presence of a confined domain (yellow circle). Initially, 47 monomeric receptors were randomly distributed without dimeric form. Simulations were run for $\geq 200,000$ steps, with an average of the last 50,000 steps shown. The diffusivity ratio between outside and inside confined domains, $D_{out}:D_{in}$, was varied from 1 (2^0) to 1024 (2^{10}). Snapshots are shown for $D_{out}:D_{in} = 1, 4, 16, 64, 256$ and 1024. **(B)** Effect of altering the ratio of $D_{out}:D_{in}$ on the number of receptors in dimeric form. Simulations were repeated three times, means \pm SD.

3. Total area of the confined domain influences receptor dimerisation

To explore whether the relative size of a confined domain affected dimerisation, this domain was again set as a circular area, of which the relative radius was altered to make up an increasing part of the membrane box size (Figure 5A). The area occupied by the confined domain was then modelled from 0-75%, *i.e.*, up to twice the estimated area of lipid rafts, while the diffusivity ratio $D_{out}:D_{in}$ was varied from 1 to 32. We found that both the area occupied by confined domains and the diffusivity ratio greatly affected the average number of dimers. Interestingly, the number of receptor dimers increased substantially from 10% to plateau to 40%, when the $D_{out}:D_{in}$ increased (Figure 5B). The highest dimer levels were reached at the

two highest $D_{out}:D_{in}$ ratios of 16 and 32. In addition, a larger area occupied by the confined domains was needed to plateau at lower $D_{out}:D_{in}$ ratios. In other words, the level of dimerisation increased with the diffusivity ratio, with curves reaching a saturation point at the lower domain area in case of a higher diffusivity ratio. Translated to receptor biology, this suggested that both the attraction strength and the size of raft-like structures can determine the extent of receptor dimerisation.

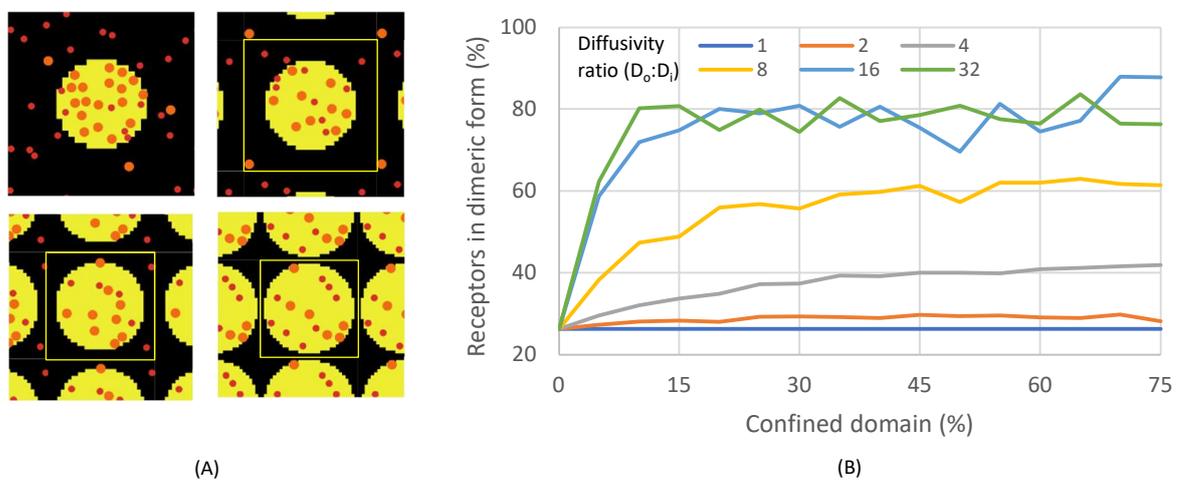


Figure 5. Increasing confined domain area induces more receptor dimerisation. Effect of increasing the confined domain area at different diffusivity ratios. (A) Snapshots of the occupied area of the confined domain (yellow circle) from upper left at 20%, 40%, 60% and 75%. Red and orange dots represent monomeric and dimeric receptors, respectively. The size of the confined domain was kept constant while the confined domain increased, resulting in an increasing number of confined domains. Note that at higher area percentages, the number of receptors per box reduces, when the number of boxes per cell increases. (B) Results of simulation for receptor fractions in dimeric form. Simulations were run for $\geq 200,000$ steps, with an average of the last 50,000 steps shown.

4. Merging of confined domains does not influence receptor dimerisation

Since membrane rafts are temporary structures that can reversibly merge⁵³, we hypothesised that the merging could affect receptor dimerisation. To assess this, we varied the number of confined domains while fixing the total area occupied, and then simulated the receptor organisation. Herein, we set the ratio of outside/inside diffusivity of receptors $D_{out}:D_{in}$ to 10, knowing that about half of the GPVI receptors on mouse platelets have a diffusivity approximately ten times lower than the other half of receptors with Brownian motion⁴⁰. The simulation is visualized by snapshots in Figure 6A. When extending this domain number to higher fold merging, we observed no change in dimer formation (Figure 6B). Translating to real life, for platelets this suggests that the mere merging of membrane rafts does not impact receptor dimerisation.

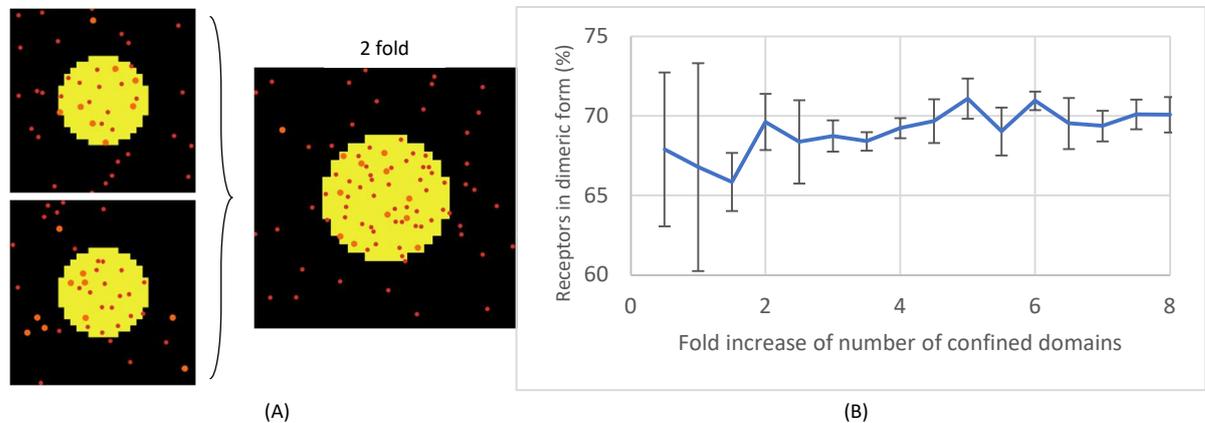


Figure 6. Merging of confined domains has no effect on receptor dimerisation. Simulated was the effect of merging two confined domains while fixing the total occupied area size. **(A)** Snapshots of two confined domains merged into one (yellow circles). The red and orange dots represent monomeric and dimeric receptors, respectively. **(B)** Simulation for determining dimeric receptors as a function of the fold merging of confined domains. Simulations were repeated three times, mean \pm SD; Pearson correlation of 0.40 indicates a weak positive correlation between confined domain folds and dimerisation.

5. Inert protein crowding in the membrane increases receptor dimerisation

As the platelet membrane contains other moving transmembrane proteins without interaction with the GPVI receptor, we also added free-moving membrane proteins to the ABM, acting as obstacles to receptor diffusion. In our simulation, the number of inert proteins per box varied from 0, 25, 50, ... to 200 (Figure 7A). The size of inert proteins was arbitrarily set at 0.05 of the box size, and their speed was set at $0.5D^{1/2}$. The average number of receptors in dimeric form, as an outcome variable, almost linearly increased from 25% to 45%, while the number of inert proteins increased from 0 to 200 (Figure 7B). This is explained by the space occupied by the inert proteins, thus tightening the diffusion room of monomeric receptors, which then leads to a higher encounter rate between receptors.

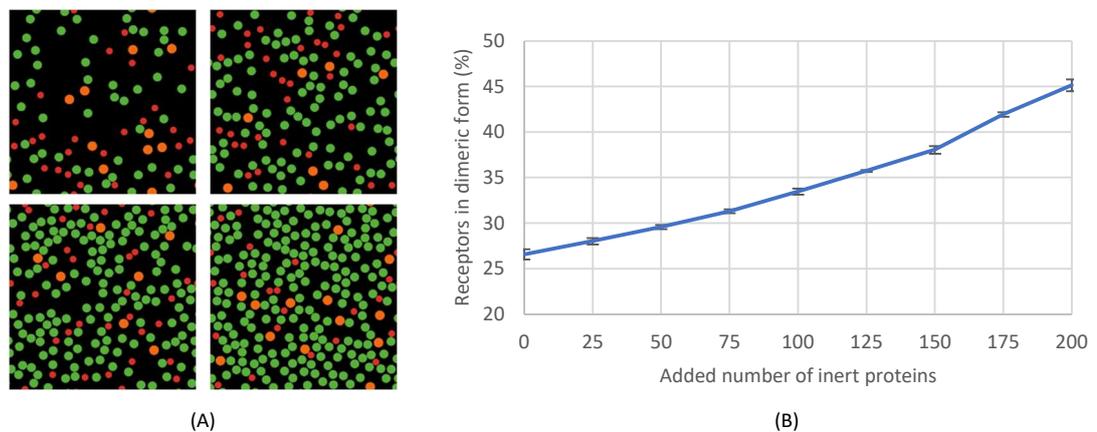


Figure 7. Increasing inert protein crowding induces more receptor dimerisation. Simulation of added inert proteins on the receptor dimerisation. Red and orange dot represents monomeric and dimeric receptors, respectively; green dots represent inert proteins. Simulations were run for $\geq 100,000$ steps, with an average of the last 50,000 steps shown. (A) Snapshots from the left top with inert proteins of 50, 100, 150 and 200. (B) Plot of dimer counts versus number of inert proteins. Note the more abundant dimeric receptors, when protein crowdedness increases. Each simulation was repeated three times, mean \pm SD.

6. Disintegration of inert proteins has a minor impact on receptor dimerisation

We then considered that inert proteins could differ upon platelet activation, *i.e.*, the proteins can become aggregated or disintegrated⁵⁴. This was simulated by splitting the space size into smaller components while not changing the total space occupied by inert proteins. Inert proteins were placed randomly in the simulation box, and the inert protein size was initially set as one large circle with a diameter half of the box size. Then the protein number was increased from 1 to 256 (2^0 to 2^8), while the size was proportionally decreased with a total conserved area (Figure 8A). According to the Stokes-Einstein relation³⁹, diffusivity may be expected to increase since smaller particles move faster. Yet, our ABM simulations showed a minor increase from 30% to 36% of dimeric receptors, when the inert protein disintegrated from 1 to 256 pieces (Figure 8B). To verify that this was not statistical noise, we determined a Pearson correlation coefficient of +0.94. Accordingly, it appears that the disintegration of inert proteins exhibits only a minimal effect on receptor dimerisation.

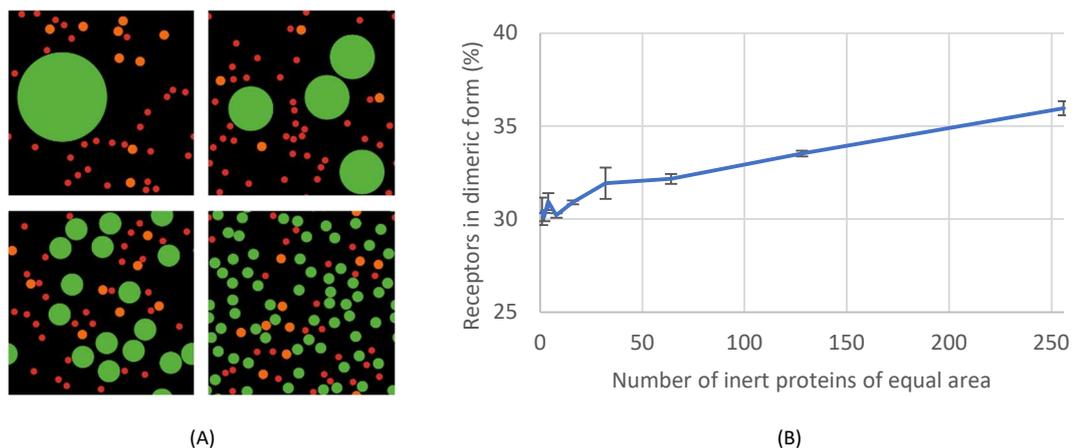


Figure 8. Disintegration of inert protein slightly affects receptor dimerisation. Simulation of the effect of size of inert proteins on receptor dimerisation. The total area occupied by inert proteins was kept constant, while subareas of smaller size were created. See further Figure 5. (A) Snapshots for 1, 4, 16 and 64 splits of inert proteins. (B) Plot of receptor dimer counts versus

the number of disintegrated inert proteins. Each simulation was repeated three times, mean \pm SD (Pearson correlation +0.96).

7. Decreasing receptor diffusivity increases the level of receptor dimerisation

According to work by Haining et al.⁴⁰, the activation of GPVI decreased in Tspan9 knock-out mice, while also the overall diffusivity of GPVI decreased. This suggested that a reduced diffusivity per se can lead to reduced dimer formation. To test this, we simulated the variation of receptor diffusivities from 2^{-5} , 2^{-6} , ... to 2^5 ; and then measured dimerisation, taken as a proxy for receptor activation. It appeared that the number of dimeric receptors, as a dependent variable decreased substantially from 85% to 25%, when the diffusivity increased (Figure 9). In other words, a slower-moving agent has a higher chance of encountering other agents. This suggests that the phenotype of reduced GPVI signalling of Tspan9-knock-out platelet is unlikely to be explained by changes other than membrane diffusion alone.

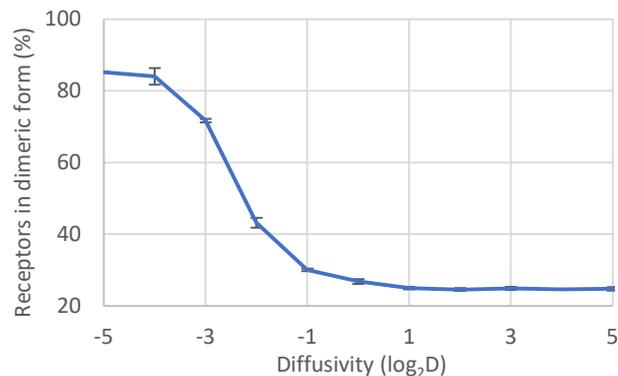


Figure 9. Decreasing receptor diffusivity increases dimerisation. Simulated testing of altered receptor diffusivity to assess receptor dimerisation, with otherwise fixed parameters. For convenience, receptor velocity was taken as $v = D^{1/2}$. Diffusivity varied from 2^{-5} , 2^{-3} , ..., to 2^5 . The simulation shows a decrease in dimer number at an increased diffusivity. Data are shown in a semi-log₂ scale; mean \pm SD.

8. Estimation of the plasma membrane area of confined domains

Haining *et al.*⁴⁰ deduced the proportion of the plasma membrane that constitutes confined domains using a single particle tracking microscopy, noting that GPVI exhibited distinctive Brownian and confined movement without and within confined domains, respectively. The number of GPVI molecules in Brownian and restricted movement mode was approximately equal⁴⁰. A temporal change in the proportion of the membrane confined domains may also impact the localisation of a receptor. Using electron microscopy and spatial point pattern analysis, previously lipid rafts were estimated to comprise approximately 35% of the total membrane surface⁴³.

We used ABM to ask what proportion range of the plasma membrane should comprise a confined domain, such in accordance with the 50% of GPVI receptors with restricted movements⁴⁰. To answer this, we fixed the diffusivity ratio to $D_{\text{out}}:D_{\text{in}} = 10:1$ and the size of the domain to 200 nm, and then varied the percentage of the plasma membrane occupied by a confined domain. The number 10:1 was obtained from the diffusion coefficients of two pools of GPVI⁴⁰. A first run of the simulation gave 20-21% of confined domains, which is below the estimation of lipid rafts of 35%⁴³. Subsequently, the effects of enforced dimerisation were added (Figure 10). The adding of dimerisation somewhat decreased the corresponding domain area (with $\text{GPVI}_{\text{inside}} \sim 50\%$), meaning that the area occupied by the confined domain would not exceed 21%, based on the model prediction. We therefore concluded that, while confined domains govern the receptor dimerisation rate, the physicochemical properties of these do not alone control receptor function. Other constraining features such as more complex receptor interactions, including the actin-based membrane skeleton within lipid rafts⁵⁵ and receptor crowding, are also important.

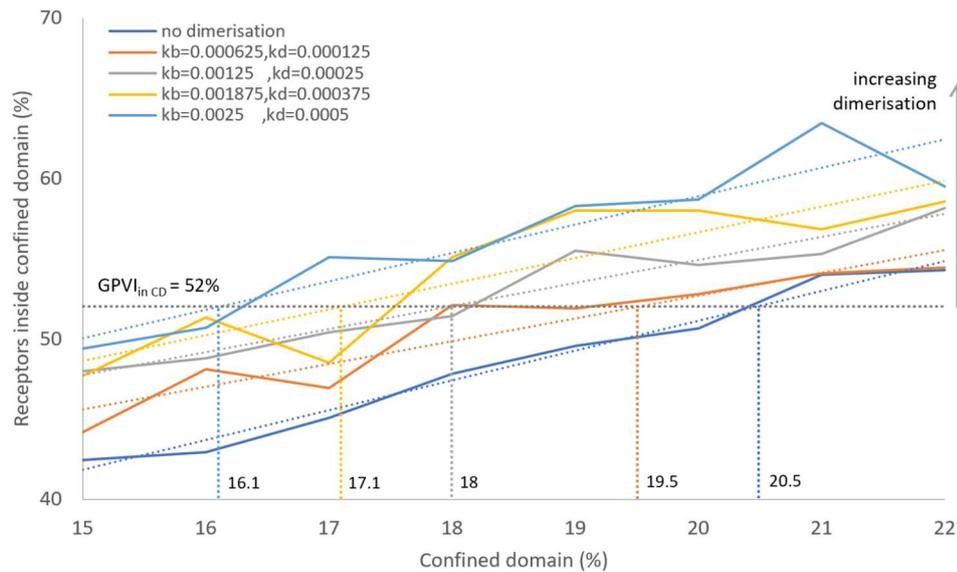


Figure 10. Enforced GPVI dimerisation reduces the confined domain area for a given GPVI localisation ratio. Plot of simulation of GPVI localisation in the presence of a confined domain with variable occupied area. Note the near linear increase of receptors inside the confined domain when this area increases. The reported value of GPVI with restricted movements⁴⁰ is about 50%, pointing to a confined domain size of 20–21%. In the presence of dimerisation, this area slightly decreases to 19–20%, with a $k_b = 0.000625$ and $k_d = 0.000125$ (least square regression analysis).

9. Increased receptor surface density results in higher predicted dimerisation levels

Several estimates may affect the number of receptors on the cell surface used in the current model. The first variable is the number of GPVI receptor copies. We set this number at 9,600 per platelet, following Burkhardt's work⁴⁴, which was obtained by quantitative mass spectrometry. Other studies using flow cytometry reported different figures ranging from 3,000 to 9,000^{56,57}, while also different *GP6* alleles lead to altered membrane-expressed GPVI levels⁵⁸. Furthermore, even within a given subject, platelet sub-populations exist with >10 fold differences in GPVI level, related to ageing cells⁵⁹, differential cell size, receptor internalisation and shedding⁶⁰.

Parameter estimation in this model assumed the platelet to be a perfect sphere; in

reality, the disc-like shape of platelet leads to a higher surface area given the same volume. Also human and mouse platelets differ in this respect. For mice, the GPVI density can be estimated as 575 molecules per μm^2 (mouse platelet volume of $\sim 4.7 \text{ fl}^{61}$, with GPVI ~ 7800 molecules per platelet⁶²). Considering that the dimerisation rate depends on receptor density, inter-species differences can also be captured by the current simulation.

While setting for human the GPVI density as 9,600 per platelet surface area $\approx 1.8 \times 10^{-11} \text{ m}^2$ ($\sigma_0 = 9600 / 1.8 \times 10^{-11} \text{ m}^2 = 533 \text{ molecules per } \mu\text{m}^2$) as a reference, we varied this density from -75% to +100% from σ_0 and measured the percentage of receptors in dimeric form (Figure 11). The simulation predicted that an increased surface density of GPVI elevates the dimeric GPVI from 50% to 75% (over a range of -75% to +100% of reference levels). This implies that in the model dimerisation does not increase proportionately with the ratio of density.

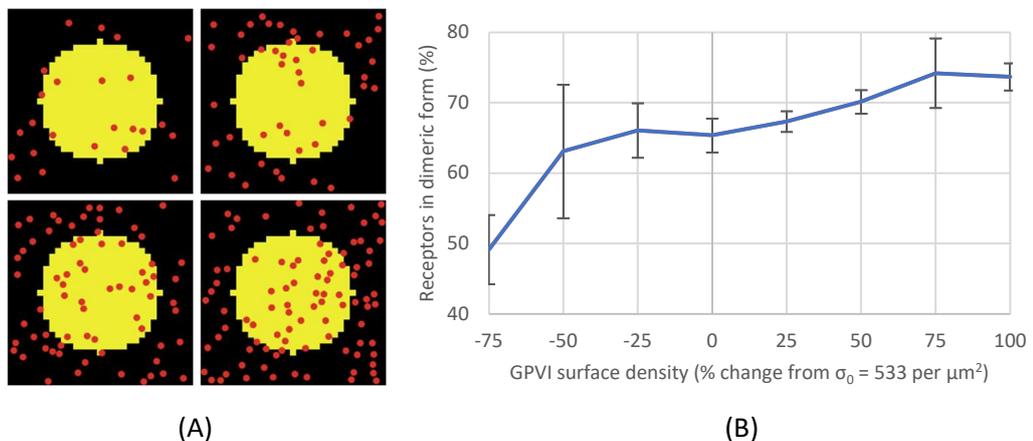


Figure 11. Increasing GPVI surface density increases dimerisation. (A) Simulation setups with various surface densities of GPVI: from left to right, top to bottom 50%, 100%, 150%, and 200% of σ_0 . (B) The plot shows that receptor dimerisation (as proportional to total GPVI, in %) increases with the GPVI surface density (as proportional to σ_0 in %). Reference density σ_0 was set as 533 molecules per μm^2 (9600 receptors divided by platelet surface area $\approx 1.8 \times 10^{-11} \text{ m}^2$; spherical assumption).

10. Simulation of ligand binding increases GPVI dimerisation

A way to simulate the effect of ligand binding is to increase the GPVI binding rate and/or dissociation rate. In real life, we expect GPVI in dimeric form to increase and to remain dimeric on collagen-adhered platelets²⁶. To simulate this, we varied the k_b and k_d from the initial values ($k_b = 0.05 \text{ molecule}^{-1} \text{ unit time}^{-1}$ and $k_d = 0.01 \text{ unit time}^{-1}$) by $\pm 10\%$, $\pm 20\%$, $\pm 30\%$, and $\pm 40\%$. The percentage of receptors in dimeric form was then simulated, as displayed in Figure 12. In this case, the dimerisation rate to increased when the binding rate was increased and/or dissociation rate decreased – both may illustrate the effect of ligand binding. A decrease in dissociation rate means that a formed dimer is more stable (*e.g.* stabilised by a multimeric ligand), while an increase in binding rate allows monomers to form into dimers with greater probability (induced by receptor-associated proteins).



Figure 12. Increase in binding rate and decrease in dissociation rate increase GPVI dimerisation. Simulation varying the binding and dissociation rate value from -40%, -30%, ..., +30%, +40%, deviating from the initial values of $k_b = 0.05 \text{ molecule}^{-1} \text{ unit time}^{-1}$ and $k_d = 0.01 \text{ unit time}^{-1}$. Values represent percentages of GPVI in dimeric form, as a proportion of total GPVI. Red colour represents higher dimerisation, blue lower dimerisation.

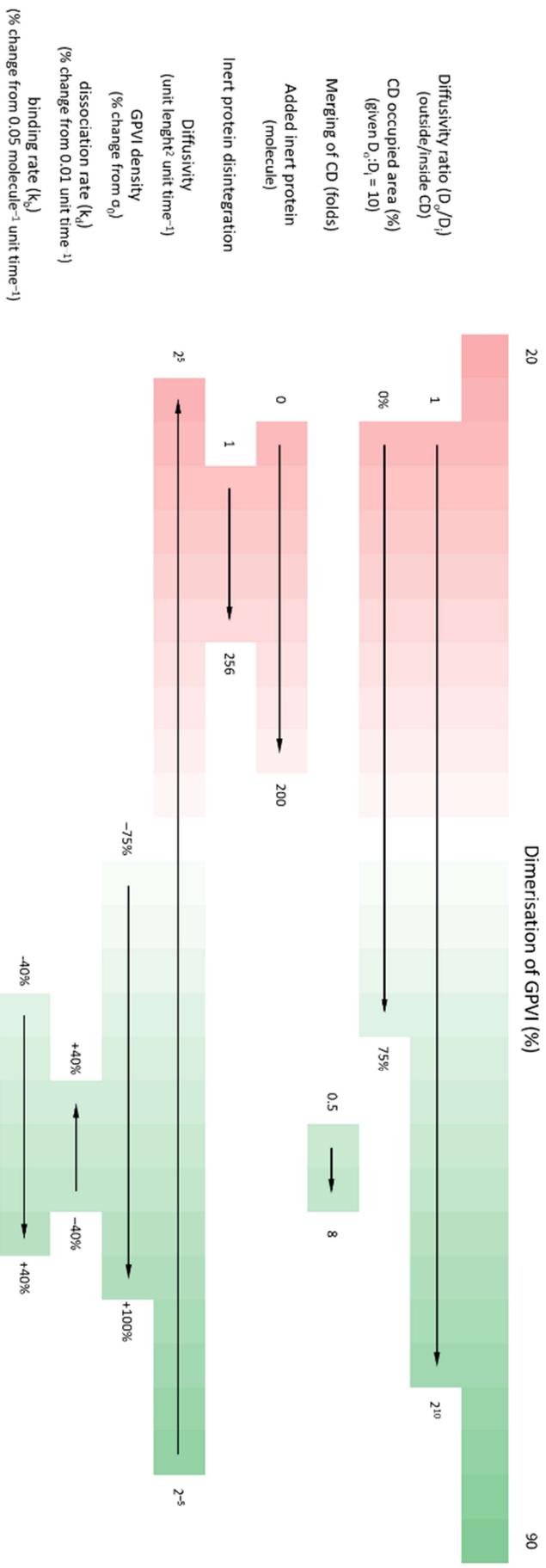


Figure 13 Relative contributions of various modelled factors to GPVI dimer formation. Arrows show the direction of increasing the indicated factor. Other conditions were assumed to be fixed, when varying the parameter of interest. The level of dimerisation (in %) displayed in colour scale: from 20% (red) to 90% (right). Numbers indicate parameter ranges tested with units or dimension indicated in parentheses.

Concluding remarks

In this study, we have demonstrated the abilities of a simple ABM technique to understand the constraints of receptor localisation and movement in the plasma membrane. Receptor dimerisation and subsequent clustering upon ligation are key initiators of signal transduction by many receptors that regulate cell function, including cell adhesion, migration and activation, for instance in the context of haemostasis and immunity. The ABM illustrates the presence of different lipid domains with distinctive properties (as confined domains), the space that these occupy on the cell surface, and the importance of the plethora of additional proteins on the cell surface, that form crowds and influence a given receptor's ability to interact with partner proteins. The relative contributions of the functionally relevant parameters tested in the ABM to GPVI dimerisation levels are summarised in Figure 13.

Due to its simplicity, computational efficiency and ease of use, ABM has the potential to be developed and generalised also to other cell types and more complex systems of receptor/protein or cell membranes. It can be applied to various studies by adapting the properties of agents (e.g., mass, size, environment), and how these affect agent diffusivity and interaction rules. Moreover, the present still simple ABM can be further developed into a more complex system with more agents and conditions. Useful additions such as receptor interactions with the cytoskeleton can be added in by utilising a computational cluster^{63,64}.

Taken together, this study forms an initial step to model and define membrane properties and their influences on receptor function. This will highlight specific processes that may be targeted therapeutically to increase or decrease receptor function and may be used for teaching, enabling the impact of modulation of various model components to be tested or demonstrated *in silico*.

Authors contributions

CT performed the modelling and coding part, with mathematical advice from RC, JLD and CW and biological advice and supervision from JMG and JWMH. All co-authors participated in drafting the manuscript.

Acknowledgements

CT received funding from the European Union's Horizon 2020 research and innovation Marie Skłodowska-Curie program, grant agreement No. 766118. CT registered in a joined PhD program at the University of Reading and Maastricht University. JLD is supported by the British Heart Foundation (RG/20/7/34866) and the MRC (MR/W015293/1).

Conflict of interest

The authors have no conflicts of interest to disclose.

Data Availability Statement

All data generated or analysed during this study available from the corresponding author on reasonable request.

References

1. Lombard, J. Once upon a time the cell membranes: 175 years of cell boundary research. *Biol. Direct.* **9**, 1–35 (2014).
2. Radhakrishnan, K., Halász, Á., Vlachos, D. & Edwards, J. S. Quantitative understanding of cell signaling: The importance of membrane organization. *Curr. Opin. Biotechnol.* **21**, 677–682 (2010).
3. Stein, W. D. *The Movement of Molecules across Cell Membranes*. Academic Press, New York (1967).

4. Klammt, C. & Lillemeier, B. F. How membrane structures control T cell signaling. *Front. Immunol.* **3**, 1–9 (2012).
5. Juliano, R. L. & Haskill, S. Signal transduction from the extracellular matrix. *J. Cell Biol.* **120**, 577–585 (1993).
6. Goñi, F. M. The basic structure and dynamics of cell membranes: An update of the Singer-Nicolson model. *Biochim. Biophys. Acta.* **1838**, 1467–1476 (2014).
7. Sunshine, H. & Iruela-Arispe, M. L. Membrane lipids and cell signaling. *Curr. Opin. Lipidol.* **28**, 408–413 (2017).
8. Owen, D. M., Williamson, D., Rentero, C. & Gaus, K. Quantitative microscopy: protein dynamics and membrane organisation. *Traffic* **10**, 962–971 (2009).
9. Trimble, W. S. & Grinstein, S. Barriers to the free diffusion of proteins and lipids in the plasma membrane. *J. Cell Biol.* **208**, 259–271 (2015).
10. Dityatev, A., Seidenbecher, C. I. & Schachner, M. Compartmentalization from the outside: the extracellular matrix and functional microdomains in the brain. *Trends Neurosci.* **33**, 503–512 (2010).
11. Ward, C. *et al.* Structural insights into ligand-induced activation of the insulin receptor. *Acta Physiol.* **192**, 3–9 (2008).
12. Lemmon, M. A. Ligand-induced ErbB receptor dimerization. *Exp. Cell Res.* **315**, 638–648 (2009).
13. Ardito, F., Giuliani, M., Perrone, D., Troiano, G. & Muzio, L. Lo. The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy. *Int. J. Mol. Med.* **40**, 271–280 (2017).
14. Foot, N., Henshall, T. & Kumar, S. Ubiquitination and the regulation of membrane proteins. *Physiol. Rev.* **97**, 253–281 (2017).
15. Torres, M. *et al.* The implications for cells of the lipid switches driven by protein–membrane interactions and the development of membrane lipid therapy. *Int. J. Mol. Sci.* **21**, 2377 (2020).
16. Okada, A. K. *et al.* Lysine acetylation regulates the interaction between proteins and membranes. *Nat. Commun.* **12**, 1–12 (2021).
17. Sezgin, E., Levental, I., Mayor, S. & Eggeling, C. The mystery of membrane organization: Composition, regulation and roles of lipid rafts. *Nat. Rev. Mol. Cell Biol.* **18**, 361–374 (2017).
18. Das, A. A., Darsana, T. A. & Jacob, E. Agent-based re-engineering of ErbB signaling: A modeling pipeline for integrative systems biology. *Bioinformatics* **33**, 726–732 (2017).
19. Juliette, G., Peters, R. & Owen, D. M. An agent-based model of molecular aggregation at the cell membrane. *PLoS One* **15**, 1–17 (2020).

20. Watson, S. Platelet activation by extracellular matrix proteins in haemostasis and thrombosis. *Curr. Pharm. Des.* **15**, 1358–1372 (2009).
21. Mammadova-Bach, E., Nagy, M., Heemskerk, J. W., Nieswandt, B. & Braun, A. Store-operated calcium entry in thrombosis and thrombo-inflammation. *Cell Calcium* **77**, 39–48 (2019).
22. Rayes, J., Watson, S. P. & Nieswandt, B. Functional significance of the platelet immune receptors GPVI and CLEC-2. *J. Clin. Invest.* **129**, 12–23 (2019).
23. Horii, K., Kahn, M. L. & Herr, A. B. Structural basis for platelet collagen responses by the immune-type receptor glycoprotein VI. *Blood* **108**, 936–942 (2006).
24. Miura, Y., Takahashi, T., Jung, S. M. & Moroi, M. Analysis of the interaction of platelet collagen receptor glycoprotein VI (GPVI) with collagen: A dimeric form of GPVI, but not the monomeric form, shows affinity to fibrous collagen. *J. Biol. Chem.* **277**, 46197–46204 (2002).
25. Jung, S. M., Tsuji, K. & Moroi, M. Glycoprotein (GP) VI dimer as a major collagen-binding site of native platelets: direct evidence obtained with dimeric GPVI-specific Fabs. *J. Thromb. Haemost.* **7**, 1347–1355 (2009).
26. Jung, S. M. *et al.* Constitutive dimerization of glycoprotein VI (GPVI) in resting platelets is essential for binding to collagen and activation in flowing blood. *J. Biol. Chem.* **287**, 30000–30013 (2012).
27. Induruwa, I., Jung, S. M. & Warburton, E. A. Beyond antiplatelets: The role of glycoprotein VI in ischemic stroke. *Int. J. Stroke* **11**, 618–625 (2016).
28. Dietrich, C., Yang, B., Fujiwara, T., Kusumi, A. & Jacobson, K. Relationship of lipid rafts to transient confinement zones detected by single particle tracking. *Biophys. J.* **82**, 274–284 (2002).
29. Triantafilou, M., Morath, S., Mackie, A., Hartung, T. & Triantafilou, K. Lateral diffusion of Toll-like receptors reveals that they are transiently confined within lipid rafts on the plasma membrane. *J. Cell Sci.* **117**, 4007–4014 (2004).
30. Bodin, S., Tronchère, H. & Payrastre, B. Lipid rafts are critical membrane domains in blood platelet activation processes. *Biochim. Biophys. Acta.* **1610**, 247–257 (2003).
31. Bonabeau, E. Agent-based modeling: Methods and techniques for simulating human systems. *Proc. Natl. Acad. Sci. U S A.* **99**, 7280–7287 (2002).
32. Niazi, M. & Hussain, A. Agent-based computing from multi-agent systems to agent-based models: A visual survey. *Scientometrics* **89**, 479–499 (2011).
33. Cuevas, E. An agent-based model to evaluate the COVID-19 transmission risks in facilities. *Comput. Biol. Med.* **121**, 103827 (2020).

34. Truszkowska A., Behring B., Hasanyan J., Zino L., Butail S., Caroppo E., Jiang Z.P., Rizzo A., P. M. High-resolution agent-based modeling of COVID-19 spreading in a small town. (2021). ArXiv q-bio. doi:10.1002/adts.202000277
35. Shamil, M. S., Farheen, F., Ibtehaz, N., Khan, I. M. & Rahman, M. S. An agent-based modeling of COVID-19: validation, analysis, and recommendations. *Cognit. Comput.* (2021). doi:10.1007/s12559-020-09801-w
36. Wilensky, U. NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL. (1999). Available at: <http://ccl.northwestern.edu/netlogo/>.
37. Pratt, L. R., Haan, S. W., Pratt, L. R. & Haan, S. W. Effects of periodic boundary conditions on equilibrium properties of computer simulated fluids. *J. Chem. Phys.* **1864**, (1981).
38. Michalet, X. Brownian motion in an isotropic medium. *Phys. Rev. E. Interdiscip. Top.* 1–13 (2010). doi:10.1103/PhysRevE.82.041914
39. Einstein, A. On the motion of small particles suspended in liquids at rest required by the molecular-kinetic theory of heat. *Annu. Phys.* (1905).
40. Haining, E. J. *et al.* Tetraspanin Tspan9 regulates platelet collagen receptor GPVI lateral diffusion and activation. *Platelets* **28**, 629–642 (2017).
41. Dunster, J. L., Mazet, F., Fry, M. J., Gibbins, J. M. & Tindall, M. J. Regulation of early steps of GPVI signal transduction by phosphatases: a systems biology approach. *PLoS Comput. Biol.* **11**, 1–26 (2015).
42. Frojmovic, M. M. & Milton, J. G. Human platelet size, shape, and related functions in health and disease. *Physiol. Rev.* **62**, 185–261 (1982).
43. Prior, I. A., Muncke, C., Parton, R. G. & Hancock, J. F. Direct visualization of Ras proteins in spatially distinct cell surface microdomains. *J. Cell Biol.* **160**, 165–170 (2003).
44. Burkhart, J. M. *et al.* The first comprehensive and quantitative analysis of human platelet protein composition allows the comparative analysis of structural and functional pathways. *Blood* **120**, 73–82 (2012).
45. Honigsmann, A. & Pralle, A. Compartmentalization of the cell membrane. *J. Mol. Biol.* **428**, 4739–4748 (2016).
46. Krapf, D. Compartmentalization of the plasma membrane. *Curr. Opin. Cell Biol.* **53**, 15–21 (2018).
47. Komatsuya, K., Kaneko, K. & Kasahara, K. Function of platelet glycosphingolipid microdomains /lipid rafts. *Int. J. Mol. Sci.* **21**, 1–18 (2020).
48. Jooss, N. J. *et al.* Anti- GPVI nanobody blocks collagen- and atherosclerotic plaque- induced GPVI clustering, signaling, and thrombus formation. *J. Thromb. Haemost.* **20**, 2617–2631 (2022).

49. Goiko, M., De Bruyn, J. R. & Heit, B. Short-lived cages restrict protein diffusion in the plasma membrane. *Sci. Rep.* **6**, 1–13 (2016).
50. Mosqueira, A., Camino, P. A. & Barrantes, F. J. Cholesterol modulates acetylcholine receptor diffusion by tuning confinement sojourns and nanocluster stability. *Sci. Rep.* **8**, 1–11 (2018).
51. Lee, F. A. *et al.* Lipid rafts facilitate the interaction of PECAM-1 with the glycoprotein VI-FcR γ -chain complex in human platelets. *J. Biol. Chem.* **281**, 39330–39338 (2006).
52. Locke, D., Chen, H., Liu, Y., Liu, C. & Kahn, M. L. Lipid rafts orchestrate signaling by the platelet receptor glycoprotein VI. *J. Biol. Chem.* **277**, 18801–18809 (2002).
53. Saitov, A. *et al.* Determinants of lipid domain size. *Int. J. Mol. Sci.* 1–14, 3502 (2022).
54. Rönnlund, D., Yang, Y., Blom, H., Auer, G. & Widengren, J. Fluorescence nanoscopy of platelets resolves platelet-state specific storage, release and uptake of proteins, opening up future diagnostic applications. *Adv. Healthc. Mater.* **1**, 707–713 (2012).
55. Fujiwara, T., Ritchie, K., Murakoshi, H., Jacobson, K. & Kusumi, A. Phospholipids undergo hop diffusion in compartmentalized cell membrane. *J. Cell Biol.* **157**, 1071–1081 (2002).
56. Best, D. *et al.* GPVI levels in platelets: relationship to platelet function at high shear. *Blood* **102**, 2811–2818 (2003).
57. Barrachina, M. N. *et al.* GPVI surface expression and signalling pathway activation are increased in platelets from obese patients: Elucidating potential anti-atherothrombotic targets in obesity. *Atherosclerosis* **281**, 62–70 (2019).
58. Joutsu-Korhonen, L. *et al.* The low-frequency allele of the platelet collagen signaling receptor glycoprotein VI is associated with reduced functional responses and expression. *Blood* **101**, 4372–4379 (2003).
59. Veninga, A. *et al.* GPVI expression is linked to platelet size, age, and reactivity. *Blood Adv.* **6**, 4162–4173 (2022).
60. Selvadurai, M. V. & Hamilton, J. R. Structure and function of the open canalicular system—the platelet’s specialized internal membrane network. *Platelets* **29**, 319–325 (2018).
61. Emeis, J. J. *et al.* A guide to murine coagulation factor structure, function, assays, and genetic alterations. *J. Thromb. Haemost.* **5**, 670–679 (2007).
62. Huang, J. *et al.* Assessment of a complete and classified platelet proteome from genome-wide transcripts of human platelets and megakaryocytes covering platelet functions. *Sci. Rep.* **11**, 1–18 (2021).
63. Fachada, N., Lopes, V. V., Martins, R. C. & Rosa, A. C. Parallelization strategies for spatial agent-based models. *Int. J. Parallel Program.* **45**, 449–481 (2017).

64. Kabiri Chimeh, M., Heywood, P., Pennisi, M., Pappalardo, F. & Richmond, P. Parallelisation strategies for agent based simulation of immune systems. *BMC Bioinformatics* **20**, 1–14 (2019).

Supplemental

The program implemented in this work is uploaded to the following reference.

Tantiwong, Chukiat (2022):

ABM_for_transmembrane_receptor_diffusion_and_dimerisation.nlogo. figshare. Software.

<https://doi.org/10.6084/m9.figshare.20444046.v2>

In order to use the program, you need to install NetLogo. Please download NetLogo via

<https://ccl.northwestern.edu/netlogo/download.shtml>

Open the downloaded program with NetLogo; you should be able to see an interface, as shown in Figure S1. You can change the value of parameters, click on setup and run to visualise the movement of agents. The tick's sliding tab can be used to speed up the simulation. The plot on the right-hand side can be exported with right-click. The info tab shows the description of how the model work, the definition of each parameter, things to notice, and things to try. The code tab contains raw code to run the model, with descriptive comments.

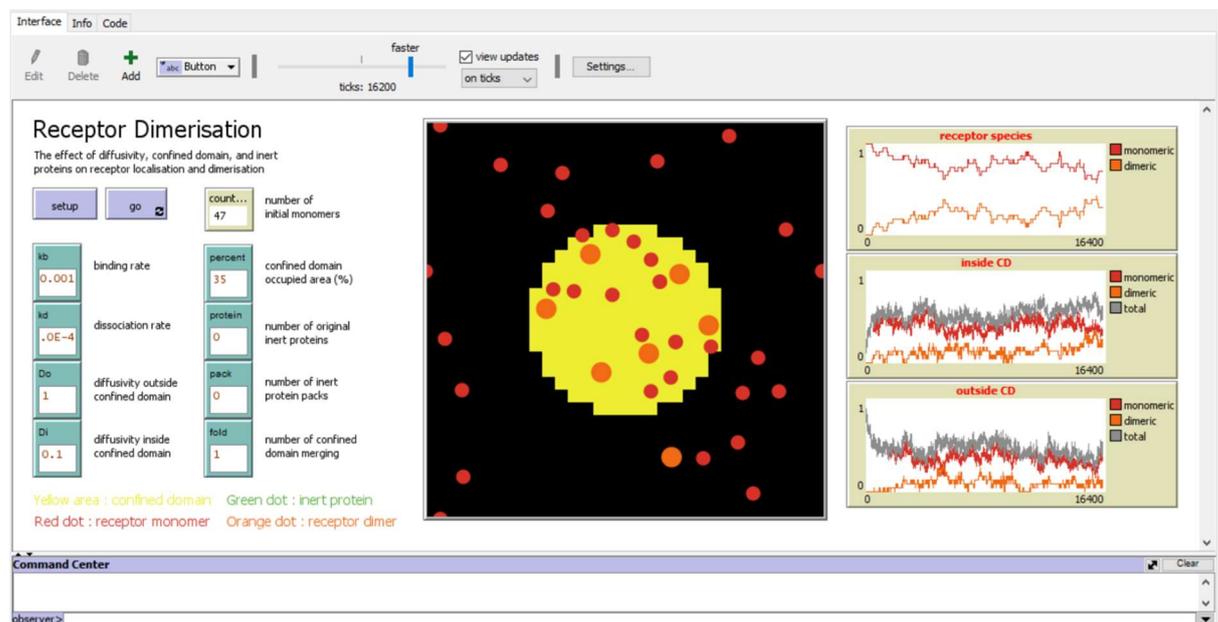


Figure S1. NetLogo screenshot of applied ABM. A screenshot was taken from the NetLogo program, and contains parts of interfaces, information, and codes. The shown interface part visualises the simulation box, the time-plots of receptor monomers and dimers, and the adjustable parameters. The information part describes the model and parameters. The coding part contains relevant details.

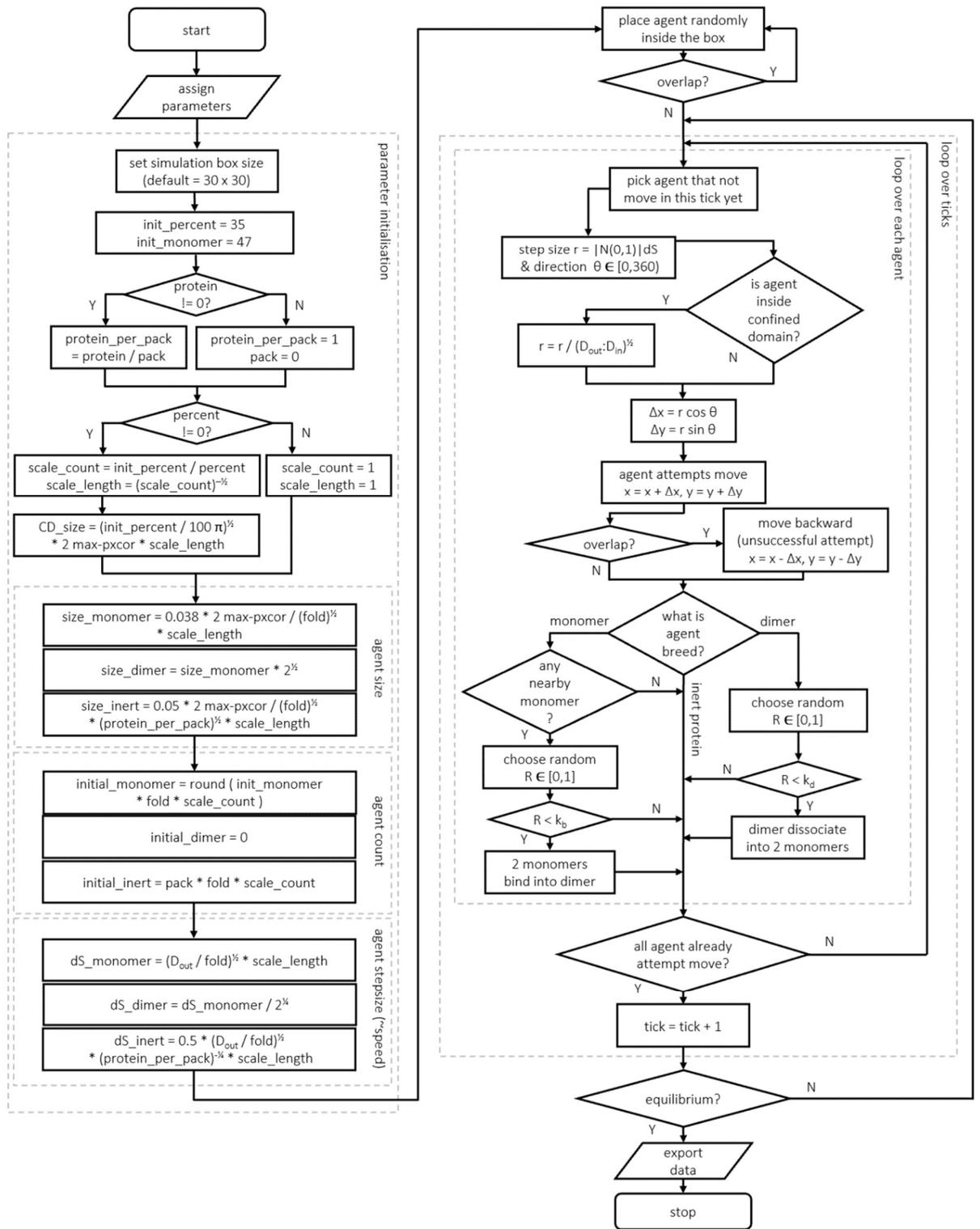


Figure S2. Flowchart of algorithm used in the GPVI-receptor ABM. The algorithm includes various steps in the simulation from randomly assigned initial positions up to stochastically movements per agent.

Chapter 3 (to be submitted)

The importance of regulation on controlling GPVI signal transduction: a systems biology approach

Following GPVI activation triggered by dimerisation and collagen binding, signalling from GPVI/ligand complex to Syk-phosphorylation has been modelled in our previous work. Building upon previous work, we aimed to construct a mechanistic description of the complex pathway leading from Syk phosphorylation to the release of inositol trisphosphate (IP3) into the platelet cytosol through the LAT signalosome. The aim was a more comprehensive model, leading to the initiation of inositol phospholipid metabolism and Ca^{2+} mobilisation. In this chapter, we present a mathematical model that combines new experimental data to capture time-dependent changes in key components of the pathway and unravel the underlying regulatory mechanisms.

The Importance of Regulation on Controlling GPVI Signal Transduction:

A Systems Biology Approach

Joanne L. Dunster¹, Chukiat Tantiwong^{1,2}, Rachel Cavill³, Johan W. M. Heemskerk^{2,4}, Christoph Wierling⁵, Amanda J Unsworth⁶, Alexander P Bye⁷, Jonathan M Gibbins^{1,*}

¹ School of Biological Science, University of Reading, Reading, UK

² Department of Biochemistry, CARIM, Maastricht University, Maastricht, Netherlands

³ Department of Data Science and Knowledge Engineering, Maastricht University,
Maastricht, Netherlands

⁴ Synapse Research Institute, Maastricht, Netherlands

⁵ Evotec, Göttingen, Germany

⁶ Manchester Metropolitan University, Manchester, UK

⁷ Reading School of Pharmacy, University of Reading, Reading, UK

* Corresponding author: j.m.gibbins@reading.ac.uk

Abstract

The platelet collagen receptor GPVI is a key trigger for the activation of platelets that provide the core component of the blood clots that stop bleeding at sites of injury but can, when formed inappropriately, lead to heart attacks and strokes. Although the subcellular components downstream of the GPVI receptor are known, how they interact to control platelet activation has not yet been fully unravelled, hindering the ability to predict therapeutic interventions. Experimentally deciphering interactions that change quickly over time is challenging, and so here, we combine new experimental data that describes time-dependent changes in six key components within the signalling pathway with mathematical methods to

help elucidate regulatory mechanisms. We extend previous work describing the early events downstream of the platelet GPVI receptor, attempting to construct a mechanistic description of the complex pathway up to the release of InsP3 to the platelet cytosol. Bayesian techniques are used to compare competing hypotheses regarding how the components interact with experimental data. The modelling sheds light on inositol InsP3 activity, a key step that is hard to describe experimentally and is able to predict the effect therapeutic targetting of Btk. The sets of interaction parameters in each model were estimated, and the best candidate signalling model was selected by Akaike information criteria. Model simulations reveal how nonlinear regulation of Btk phosphorylation and subsequent stabilisation of membrane lipids are key regulatory components of the pathway.

Introduction

The changes that platelets undergo in response to extracellular signals are essential for health; they activate, change shape and become sticky, allowing them to aggregate into clots that plug damage and prevent excessive bleeding at sites of injury, but when triggered inappropriately, lead to the formation of clots that can trigger heart attacks and strokes¹. The development of new drugs to suppress platelet function and thereby prevent thrombosis has been shown to be an effective strategy. The collagen receptor glycoprotein VI (GPVI) drives thrombo-inflammatory platelet responses². It triggers a subcellular pathway of events that is particularly long and complex, centring around the proteins Spleen tyrosine kinase (Syk) and Bruton's tyrosine kinase (BTK) that are both targetted therapeutically³.

A cell's response to its environment involves a complex dance between subcellular molecules, proteins, lipids, and ions, culminating in changes to the cell's behaviour. While the key players of these pathways are now often well known, the choreography of the events is

often obscured from view, inaccessible to experimental methods that only provide snapshots of events in space and time. Mathematical models can join the data together to describe the time-dependent interactions between components, providing predictions that are testable, information on those aspects of a system that are not directly accessible to experimental observation and insight into the effect of therapeutic interventions⁴. The ability of these models to be useful in a laboratory or clinical setting relies on their ability to make valid predictions. But most models are not supported by sufficient quantitative biological data. They are often informed by data describing events in other cells or experimental setups.

We previously developed a mathematical model of the early events downstream of GPVI⁷. Here, we extend this work in an attempt to decipher and order interactions that occur later in this pathway and culminate in the release of inositol InsP3 (IP3) to the cytosol. The aim is to provide a framework that is easily adapted and extended to test and direct future experimental data. There is uncertainty in the structure of the signalling pathway, and therefore, we compiled a number of hypotheses about how the proteins and lipids may interact, combining them into eight competing model structures. We use a data-driven approach to test the different model structures. The data that we used was generated in-house specifically to describe the events downstream of GPVI stimulation; it is unusually dense, describing the changes in six components over eleven timepoints. The model structures were translated into mathematical models and tested and calibrated against the data using Bayesian techniques. The use of this framework avoids point estimates for parameters, instead characterising the uncertainty in their inference from the model and available data, pointing to the most informative experimental directions⁵. Simulations from the competing models were compared to glean insight into the importance of individual events on the ability of models to fit the experimental data, uncovering the importance of nonlinear regulation of Btk

and that experimental discrimination between the competing models could be achieved by the collection of data that differentiates between cytosolic and membrane bound proteins.

Methods and Materials

Experimental data

Experimental data traditionally used for calibration, inference and validation of mathematical models that capture subcellular changes are often lacking in density, describing the changes to just a few components (or indeed just one) over a limited number of timepoints. Higher-density data is known to increase parameter identifiability and the ability of the model to make valid predictions⁶, and our previous work demonstrated how data dense in time-dependent changes can aid in the discovery of the often neglected and less understood negative feedbacks within signalling pathways that can play a role in regulating signal transduction⁷. Therefore here, experimental data was collected that was dense in both the timepoints over which it was assessed (11 spread over 3 minutes) and the number of components (4 proteins (Syk, LAT, PLC γ 2, Btk and a PKC substrate) and 2 lipids (PIP2 and PIP3)), all of which are known to play a role downstream of the GPVI receptor.

Data characterising the time-dependent changes to Syk, LAT, PLC γ 2, Btk and a PKC substrate were performed on blood samples from healthy donors that had given consent, using procedures approved by University of Reading Research Ethics Committee. Blood was collected by venepuncture into 4% sodium citrate and washed platelets prepared by differential centrifugation, as described previously. In short platelet rich plasma (PRP) was separated from whole blood via centrifugation at 100 xg for 20 minutes at room temperature. Acid citrate dextrose was added to the PRP (1:8 v/v) prior to an additional centrifugation step at 350 xg for 20 minutes to pellet the platelets. The platelet pellet was then resuspended in

Modified-Tyrodes HEPEs buffer (134mM NaCl, 0.34mM Na₂HPO₄, 2.9mM KCl, 12mM NaHCO₃, 20mM N-2-hydroxyethylpiperazine-N-2-ethanesulfonic acid, 5mM glucose and 1mM MgCl₂, pH 7.3), containing 100μM MRS2179, 1μM cangrelor, 1mM EGTA and 10 μM Indomethacin; to suppress secondary signalling and secretion and used immediately.

Samples were stimulated with CRP-XL (Cambcol, Cambridge, UK) at a final concentration of 1 μg/ml, then lysed in Laemmli sample buffer, denatured by boiling and loaded onto 4-12% NuPAGE Bis-Tris gradient gels (BioRad) and separated by SDS-PAGE electrophoresis. Immunoblotting was performed using standard techniques. In short, proteins separated by electrophoresis were then transferred via semi-dry (Bio-Rad) western blotting to Immobilon-FL membranes (Millipore). Membranes were blocked with 5% (w/v) BSA and probed with phosphospecific antibodies that recognise the phosphorylated state of proteins of interest ((p)LAT-Y200, (p)Syk Y525/526, (p)Btk Y223 , (p)Ser-PKC substrates, (p)PLC – Y759). Following primary incubation immunoblots were treated with either a fluorescent Cy5 dye-labelled goat anti-Rabbit or a Alexa-Fluor 647 dye-labelled donkey anti-mouse (Life Technologies) antibodies. Proteins were detected using fluorophore conjugated secondary antibodies and visualised using a Typhoon Fluorimager and Image Quant software (GE Healthcare). Band intensities were quantified and levels of total protein were used to normalize the phosphorylation data using Image Quant software and ImageJ.

The phosphoinositides (PtdIns(4,5)P₂ and PtdIns(3,4,5)P₃) were measured via a newly developed method that utilises a quantitative targeted ion chromatography-mass spectrometry-based workflow that separates phosphoinositide isomers and increases the quantitative accuracy of measured phosphoinositides, described in⁸. In summary washed platelets from healthy donors were pretreated with apyrase and indomethacin, and subsequently stimulated with CRP. High concentrations of CRP (30μg/mL) being used to

improve the detection of these low abundance phosphoinositides. Stimulation was stopped at specified time points (0,30,60,90,120,180 and 600s), with ice cold 1M HCl⁹ and the samples were analysed and quantified using ion chromatography tandem mass spectrometry system.

Model development

The aim is to develop a model that captures the key components that play a role in conducting signals downstream of a platelet's GPVI receptor. We extend our previous model that captures the early steps in the GPVI pathway⁷, and culminated in the activation of the protein Syk, to incorporate eight new components: the trans-membrane protein Linker for Activation of T cells (LAT), a key member of a signalosome assembled following Syk phosphorylation, the protein Phosphoinositide 3-kinase (PI3K), the phosphoinositides PtdIns(4,5)P₂ (hereafter called PIP₂) and PtdIns(3,4,5)P₃ (PIP₃), a member of the Tec family of tyrosine kinases Bruton tyrosine kinase (Btk), PLC γ 2, the second messengers diacylglycerol (DAG) and protein kinase C (PKC) and, lastly, the inositol 1,4,5-trisphosphate (IP₃) that when released to the platelets cytosol triggers calcium release from internal stores that leads to platelet activation.

While the key components of the GPVI pathway are well known, there is some uncertainty in how they interact to control IP₃ release. To reflect this uncertainty, we compiled a core set of reactions that reflect current biological knowledge on the interactions (that we call our base model) and three modifications that vary methods of interactions and regulation between the components.

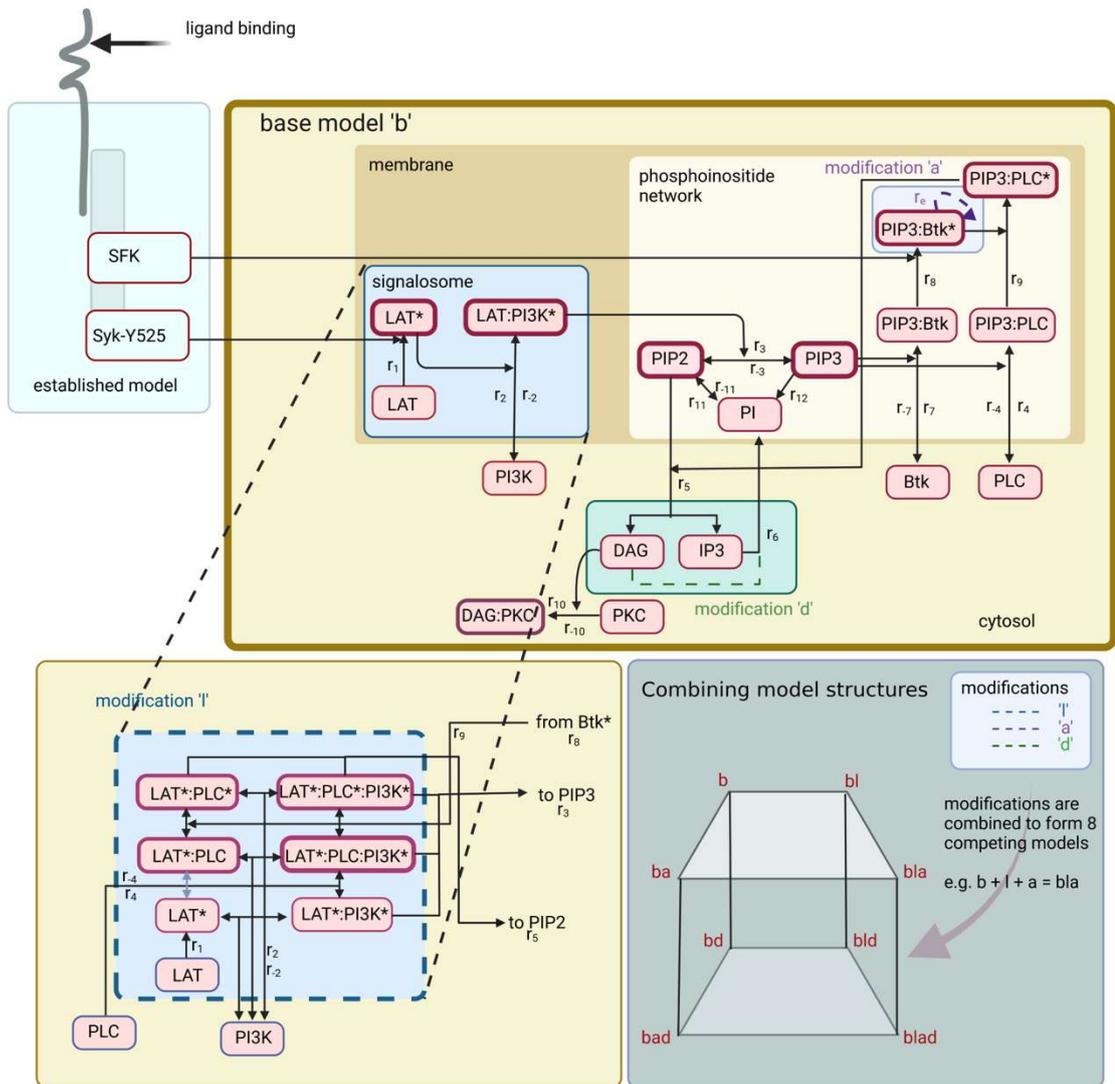


Figure 1 Schematic depicting the interactions and reactions of the GPVI signalling pathway that are captured in the mathematical models. The reactions captured in the base model structure (denoted *b*) are depicted in the top panel. There are 3 modifications to this, denoted with an additional 'a', 'd' and 'l'. The inset in black broken lines demonstrates the modification denoted *l* and requires the removal of PLC binding to PIP3 (right-hand side). The modification 'd' captures the removal of recycling of DAG to the membrane, and 'a' the addition of nonlinear phosphorylation of Btk. The modifications are combined to form eight competing model structures. Variables are represented by boxes, and the parameter associated with each process is placed next to the relevant arrow. Green boxes are outputs from the early model, red indicate that they are described experimentally, and blue are those without data. The equations for all models are described in the supplemental

The base model

These reactions are depicted in Figure 1 (top panel). In summary, following activation of GPVI, Syk-dependent events close to the platelet membrane (and captured in our earlier work) lead to formation of a signalosome, a key component of which is LAT. The cytosolic protein PI3K can be recruited to this cluster of proteins at the cell membrane. Its primary substrate is PIP2, a lipid of the phosphoinositide pathway, which through the activity of PI3K, is converted to PIP3. The phosphoinositide pathway is a large complex network of reactions with many other components, but we focus here on only those known to influence the GPVI signalling pathway and include a pool (denoted PI) to represent the other members of this network. Btk and PLC γ 2, cytosolic proteins, both have PH domains that can bind to PIP3, facilitating their recruitment to the cell membrane. Btk itself is able to mediate phosphorylation and activation of PLC γ 2. Once activated PLC γ 2 feeds back to influence the phosphoinositide pathway by mediating cleavage of PIP2, releasing it in the form of DAG and IP3 to the cytosol. The formation of DAG leads to the binding and activation of PKC, a regulator of platelet granule secretion, and IP3 is responsible for the release of calcium from internal stores into the platelet cytosol.

The following describes three modifications to the base model:

Modification I: changes the mechanism of recruitment of PLC γ 2 to the platelet membrane.

In the base model, recruitment of PLC γ 2 to the platelet membrane is through a PH domain that binds PLC γ 2 to the phosphoinositide PIP3. But, LAT has been proposed to recruit PLC γ 2 to the membrane through its C-terminal SH2 domain and this modification, denoted by an 'I', reflects this¹⁰.

Modification d: neglects the recycling of DAG from the cytosol to the cell membrane.

It is generally believed that cytosolic DAG and IP3 are recycled slowly up to the membrane, combining to join the pool of phosphoinositides, and this was captured in the base model. However, IP3 and DAG are known to be recycled via different pathways that likely involve differing rates¹¹. Therefore, assuming a linear combination of IP3 and DAG may be misleading, and all models with a 'd' neglect DAG recycling.

Modification a: includes nonlinear regulation of Btk phosphorylation.

The recruitment of Btk to the platelet membrane, via its PH domain, and its subsequent phosphorylation is known to be an important step downstream of GPVI, being crucial for the activation of phospholipase C- γ 2 (PLC γ 2). In the base model, phosphorylation of Btk (on Tyr 551) is assumed to be linear, being mediated by SFKs present upstream close to the GPVI receptor. But, there is evidence that this process is nonlinear, with membrane recruitment triggering trans-autophosphorylation¹²⁻¹⁴. Therefore all models denoted with an 'a' include an additional nonlinear process of Btk phosphorylation.

Combining model structures

The four different scenarios denoted 'b', 'l', 'a' and 'd', were combined to form eight networks of interactions (denoted as b, ba, bd, bl, bad, bla, bld and blad). For example

model b + modification l + modification d = model bld,

that incorporates the base reactions, has PLC γ 2 being recruited to the LAT signalosome, and DAG recycling is neglected. Each network of reactions (Figure 1) was translated into a system of ordinary differential equations. In addition to the equations and parameters in the original model, models 'b', 'ba', and 'bad' comprise sixteen equations and all models with an 'l' ('bl',

'bla', 'bld', 'blad') have eighteen equations, models with an 'a' having nineteen parameters and those without an 'a' having eighteen (Table 1 provides a description of the models parameters while full details of the equations are given in Supplemental). These eight mathematical models capture our ideas on the dominant mechanisms by which the key proteins downstream of a platelet's GPVI receptor interact to release IP3 to the cytosol, leading the calcium flux and platelet activation.

Parameter	Description	Unit of measure
r_1	Syk phosphorylation of LAT signalosome	molecules ⁻¹ s ⁻¹
r_2	Rate of binding for PI3K	m ³ moles ⁻¹ s ⁻¹
r_{-2}	Dissociation rate for PI3K	s ⁻¹
r_3	Rate of conversion of PIP2 to PIP3	molecules ⁻¹ s ⁻¹
r_{-3}	Rate of conversion of PIP3 to PIP2	s ⁻¹
r_4	Rate of binding for PLC	m ³ moles ⁻¹ s ⁻¹
r_{-4}	Dissociation rate for PLC	s ⁻¹
r_5	Hydrolysis of PIP2 to IP3 and DAG	molecules ⁻¹ s ⁻¹
r_6	Recycling of IP3 to the membrane (PI)	molecules ⁻¹ s ⁻²
r_7	Rate of binding for Btk	m ³ moles ⁻¹ s ⁻¹
r_{-7}	Dissociation rate for Btk	s ⁻¹
r_8	Rate for Btk phosphorylation	molecules ⁻¹ s ⁻¹
r_9	Rate for PLC phosphorylation	molecules ⁻¹ s ⁻¹
r_{10}	Rate of binding for DAG (to PKC)	molecules ⁻¹ s ⁻¹
r_{-10}	Dissociation rate for DAG	s ⁻¹
r_{11}	PI conversion to PIP2	s ⁻¹
r_{-11}	PIP2 conversion to PI	s ⁻¹
r_{12}	PIP3 conversion to PI	s ⁻¹
r_e	Btk non-linear phosphorylation & molecules	molecules ⁻¹ s ⁻¹

Table 1 A summary of new models parameters. Their definitions and units of measure, prior distributions are uniform in log-scale in range 10^{-10} to 10^{+3} .

Model inference and validation

To assess which, if any, of the eight networks of interactions was supported by the experimental data, the mathematical models were calibrated and compared to the data using a form of Approximate Bayesian Computation. The steps in this process are depicted in Figure 2.

We estimated all unknown parameter values for our models as previously⁷. Parameter values were drawn randomly from wide ranges via a Latin Hypercube. The parameter set was passed to a constrained local optimisation routine (MATLAB's `fmincon`) that varies all unknown parameters to minimise the differences between the model structures and experimental data via a cost function (SSE)

$$\text{Sum Squared due to Error (SSE)} = \sum_{i,j} (1/N_j) * (y_i(\theta)^j - \text{Data}_i^j)^2$$

where $y_i(\theta)^j$ is the model's prediction for the model variable j (that were used in fitting, i.e. LAT*, PLC*) at time point i (which depends on the parameters θ), Data_i^j represents the experimental observations, and N_j is the number of data points corresponding to the variable j . This cost function allows us to define each models ability to fit the data where we mitigate against a more complex model, with more parameters, being able to better fit the data by using Akaike Information Criteria (AICc)¹⁵.

$$AICc = AIC + \frac{2k^2 + 2k}{n - k - 1}$$

where

$$AIC = n \ln \left(\frac{SSE}{n} \right) + 2k$$

Here, k be the number of estimated parameters in the model, and n be the number of data (fitting points). This modified criterion taking into account the experimental sample size by increasing the relative penalty for model complexity with small datasets. The value of AICc has

no meaning in isolation, its relevance only becoming apparent when it is used to compare (and rank) models fitted to the same experimental data¹⁵. Parameter values that enable the model to 'best' describe the experimental data were saved. The process was repeated 50,000 times for each model. It yields sets of parameter values that allow a particular model to achieve its best description of the experimental data. The distance (captured in this cost function) for simulations produced from the set of 'best' parameter values for each candidate model can then be compared to assess which model structure is best able to fit the available experimental data. The numerical simulations generated from these parameter values, which cover the time-dependent changes in all of the proteins and lipids included in the model, can then be compared to assess if they produce biologically consistent and feasible predictions. The ranges of the best parameter values are then assessed, providing insight into the mechanisms that allow a particular model to describe the data better than an alternative model.

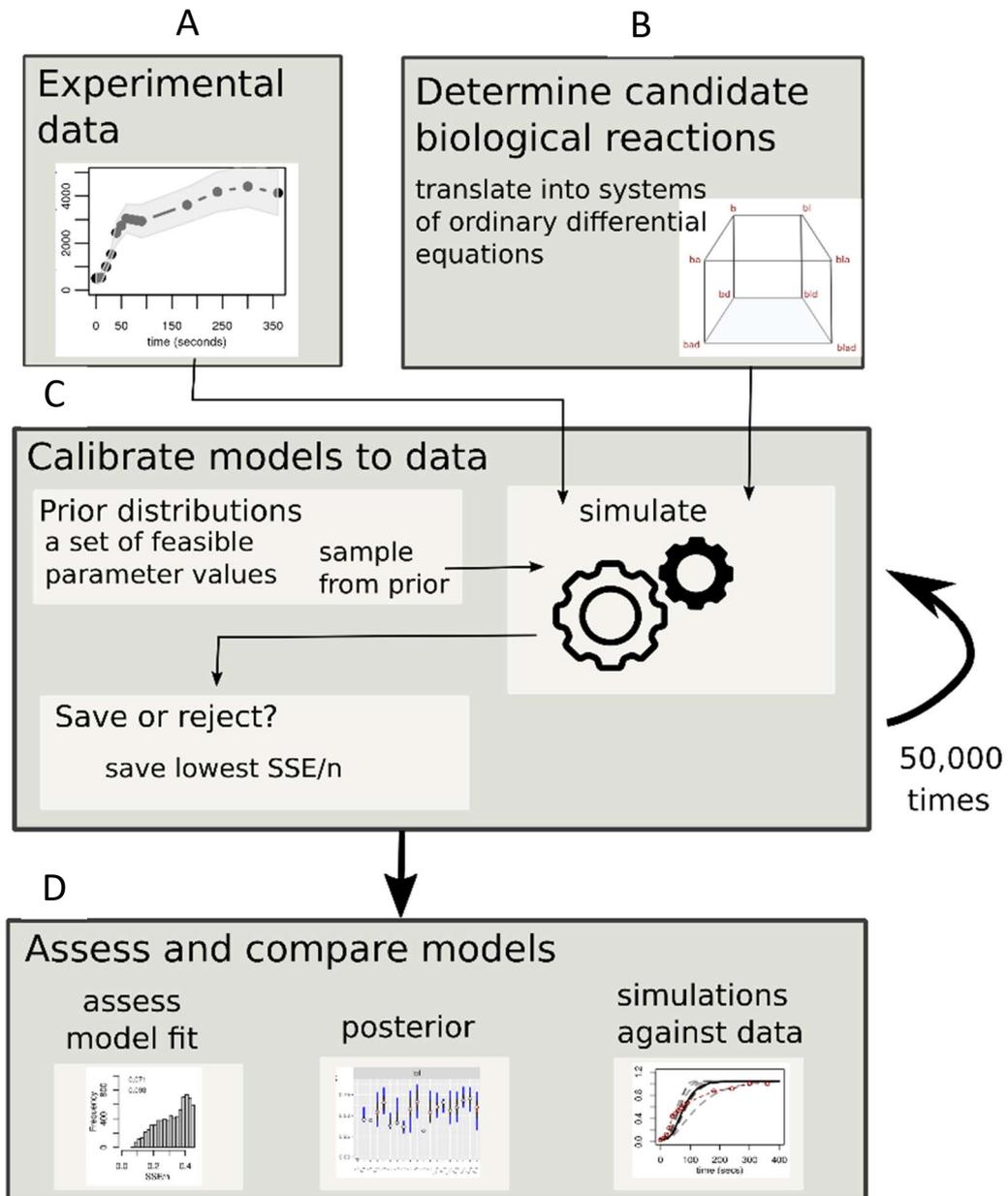


Figure 2 The steps involved in model calibration and comparison to experimental data. A) Time-resolved data are collected to describe 6 key steps in the pathway. B) The four alternative hypotheses regarding the interactions and reactions between the signalling proteins and plasma membrane lipids are combined and translated into systems of ordinary differential equations. C) Bayesian inference is used to infer parameters that have a high or appreciable probability of having generated the observed data. A parameter set is sampled from prior distributions and used to numerically simulate a model structure. The cost function (SSE) is used to decide if the parameters allow the model to generate predictions close to the data. D) The distance between each model's predictions and the data provides information on which

model structure is best able to replicate the data. Posterior distributions provide information on the uncertainty in parameter estimates that are learned from observed data and also on the range of realistic parameter values that can produce observed data. Model simulations provide predictions for unobservable components and additional information on the gap between observable components and the data.

Results

Experimental data

The experimental data (Figure 3) describes the changes that occur in platelets following a ligand binding to the GPVI receptor. Standard laboratory methods to describe post-translational modifications usually comprise a limited number of timepoints. Here our data comprises 11 timepoints describing changes over 3 minutes, capturing the full dynamics of post-translational changes. The proteins (LAT, Btk, PLCy2 and PKC) and the lipids (PIP2 and PIP3) were selected based on their known importance in GPVI signal transduction. Phosphorylation of Btk is quantified, while the data describing changes in LAT, PLCy2 and PKC is normalised to maximal levels.

Phosphorylation of LAT and PLCy2 monotonically increases before achieving a steady state approximately 2 minutes after stimulation. The data indicate that phosphorylated Btk follows a similar pattern, though hints at an early peak (and regulation) at 1 minute before phosphorylation increases at later timepoints. The quantification of phosphorylated Btk indicates that it comprises only a fraction (<40%) of the total pool, as occurs in other cell types¹⁶, a further indication of regulation of Btk phosphorylation. In agreement with studies in alternative cell types, the abundance of PIP2 is shown to be three orders of magnitude higher than for PIP3.

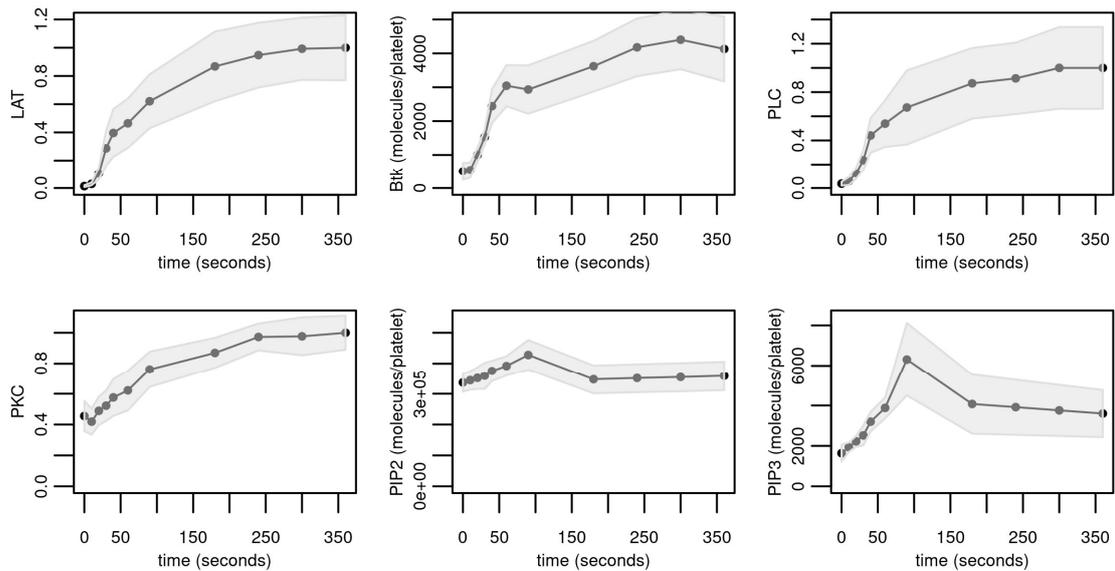


Figure 3 Experimental data describing CRP-induced changes to two phosphoinositides and four proteins that play a role in the GPVI signalling pathway. Results for Btk, PIP2 and PIP3 are expressed as molecules/platelet, while results for LAT, PLC and PKC are normalised to the maximum. Results are mean \pm SD from multiple donors (LAT, 6 donors; Btk, 8; PLC, 4; PKC, 6; PIP2 and PIP3, 4).

The importance of non-linear regulation of Btk in differentiating the reactions that dominate the GPVI signalling pathway

We are initially interested in which of our eight competing mathematical models are best supported by the experimental data. Table 2 provides information (output from the cost function) on the ability of each model to fit the experimental data. This summarises the distance between the experimental data and the models simulations, a lower number reflecting a closer fit to the data. The results indicate that a single structural modification that incorporates PLC γ 2 binding to the signalosome or neglects DAG recycling does not greatly improve the fit to the data (minimum AICc for model 'b'~'bl'~'bd'~ -185), the same being true for a combination of these ('bld'~ -185). The inclusion of nonlinear phosphorylation of Btk improves the fit of any of these combinations, with a model that combines all modifications

(blad) being able to produce simulations closest to the data ('blad', -244), it is worth noting that removal of nonlinear regulation of Btk phosphorylation reduces to a model that is least able to simulate the data (bld, -184).

Model	b	ba	bd	bad	bl	bla	bld	blad
Median (10) $n \log(\text{SSE}/n)$	-186.71	-238.5	-188.23	-240.63	-185.82	-239.18	-184.17	-244.37
Median (100) $n \log(\text{SSE}/n)$	-182.21	-234.56	-183.73	-240.63	-181.77	-236.75	-176.8	-241.08
Median (10) AICc	-136.16	-183.97	-137.68	-186.1	-135.27	-184.66	-126.24	-189.85
K	18	19	18	19	18	19	18	19

Table 2 Results of fitting each model to the data. Results of fitting each model to the data. SSE = distance of model simulations from experimental observations, n is the number of experimental observations (11 timepoints \times 6 datasets = 66) and K = the number of the model's parameters that are inferred from the data. Median (10) indicates the median of the lowest 10 results, Median (100) the lowest 100. Metrics in bold denote the models with the lowest $n \log(\text{SSE}/n)$ or AICc.

The posteriors (Figure 4) describe the parameter values inferred from the data, they are shown normalised against the prior distributions (statistics are provided in Supplemental). These sets hold information on which parts of the signalling pathway can be best inferred from the data as well as a parameters influence on the general ability of the model to fit the data, two concepts that are intimately linked. If a posterior distribution is broad and, therefore, not very different from the prior, then the parameter is not inferable from the data and the fitting process is insensitive to its variation. In contrast, parameters with a narrower posterior distribution (lower values, denoted by asterisks indicate the best constrained) they have had

information returned from the experimental data, and the predictions for the data are sensitive to its variation.

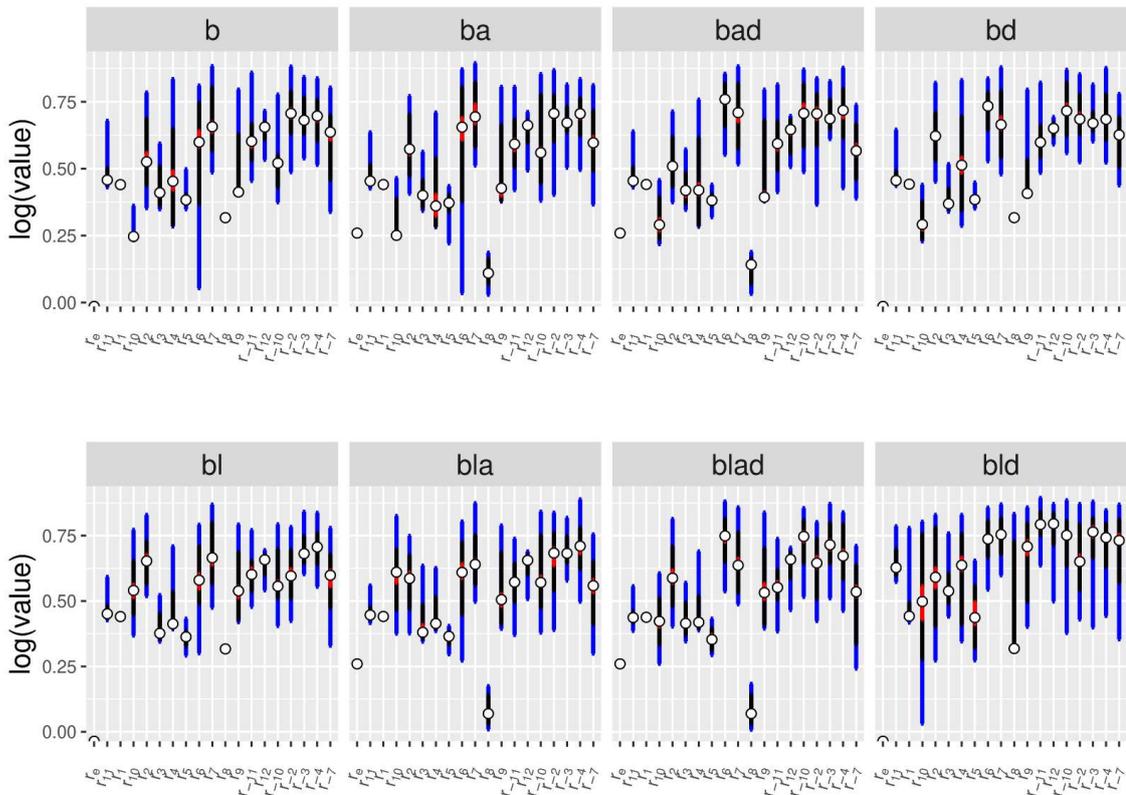


Figure 4: An illustration of the approximated posterior parameter ranges. These posteriors show parameter ranges for lowest SSE normalised to the prior distribution. Vertical lines represent the 90% credible interval, black the interquartile, red the inner 10%. Circles represent medians. To avoid crowding 90% intervals are described numerically in Table 5 of Supplemental.

Generally parameter posterior distributions were well dispersed, making point estimates unreliable but, despite this unidentifiability that is present in all of the models, some models are better at constraining the data than others. Models 'bd' and 'blad' are best able to constrain the parameters (the 90% percentiles being restricted to 40% of their priors for 16 parameters), while 'bld', the model that was noted above to have difficulty fitting the data is also least able to constrain the parameters (the 90% percentiles being restricted to 40% of

their priors for 8 parameters, only 1 of these being restricted to a range of 25% of their priors). However, we cannot infer that the better model always yields a narrower range of parameters posterior. The goodness of fit still needs to be quantified and evaluated by the cost function. In general the models are better at inferring forward reactions than their reverse. Models are best at inferring parameters r_1 (LAT phosphorylation) and the parameters controlling Btk phosphorylation (r_8 and its addition r_e , which no doubt reflects the quantity of surrounding data (for r_1) or its quality, the Btk data being quantified. Other parameters that have been consistently restricted by most models are those of the phosphoinositide pathway (r_{11} , r_{12} , r_3 and r_{-3}).

Comparing the 10 best simulations for each model against the experimental data reveals that all models are able to capture the general trend displayed in the experimental data describing phosphorylation (an example of this is shown in Figure 5, A), but some models are less well able to describe the dynamics of the phosphoinositides PIP2 and PIP3 (Figure 5, B). Models 'b', 'bd', 'bl' and 'bld' predict sharp early transient peaks in both phosphoinositide species while the inclusion of nonlinear rates of Btk phosphorylation (models with an 'a') are able to produce simulations without these sharp peaks.

In summary model 'blad' is best supported by the experimental data and removing non-linear regulation from this model results in a model ('bld') that is least able to generate the data. The selection of model 'blad' as representing the most likely set of dominant reactions that lead to the data is supported by the notable difference between the posteriors of the two models with model 'blad', being one of two models best able to restrain the parameter values, and model 'bld' the worst. Simulations from models with non-linear regulation of Btk phosphorylation predict that Btk binding to the phosphoinositide PIP3

stabilises the phosphoinositide pathway, which is reflected in the lack of early transient peaks in simulations of PIP2 and PIP3.

Binding kinetics of proteins can differentiate model predictions

Comparing model predictions reveals other differentiating predictions. Simulations for all three species of Btk, when it is in the cytosol, bound to the membrane and phosphorylated (Figure 5, B) are different from models that include non-linear regulation of Btk phosphorylation to those without. Models without an 'a' predict that cytosolic Btk slowly decays, but once Btk recruited to the membrane it is quickly becomes phosphorylated. In contrast, models with an 'a' predict that Btk quickly transfers to the membrane where it is slowly phosphorylated. In the former, it is the slow availability of Btk at the membrane that controls the rate of phosphorylation, while in the latter, with nonlinear rates of phosphorylation, this is not required.

In a similar manner there are differing predictions for the rate that PLC is recruited to the membrane (Figure 5, panel C). Models where PLC is recruited to the signalosome (denoted with 'l') predict that cytosolic PLC is slowly recruited to the membrane where it is quickly phosphorylated whereas models where PLC binds to PIP3 (without an 'l') predict that cytosolic PLC is quickly recruited to PIP3 at the membrane, where it is slowly activated (Figure 5).

Model simulations generally predict that the phosphoinositide pool (PI) is not depleted, remaining at high levels. The exception is in simulations from models that neglect recycling of DAG back to the membrane and have linear Btk phosphorylation, namely models 'bd' and 'bld'. The 10 best simulations from these models can be categorised into two groups, those that show little depletion of the phosphoinositide pool that are more in line with data describing positive levels of these alternative phosphoinositides 10 minutes after stimulation

with CRP^{8,17,18} and those that do, the latter being accompanied by subsequent large rises in IP3 and DAG (Figure 5, D).

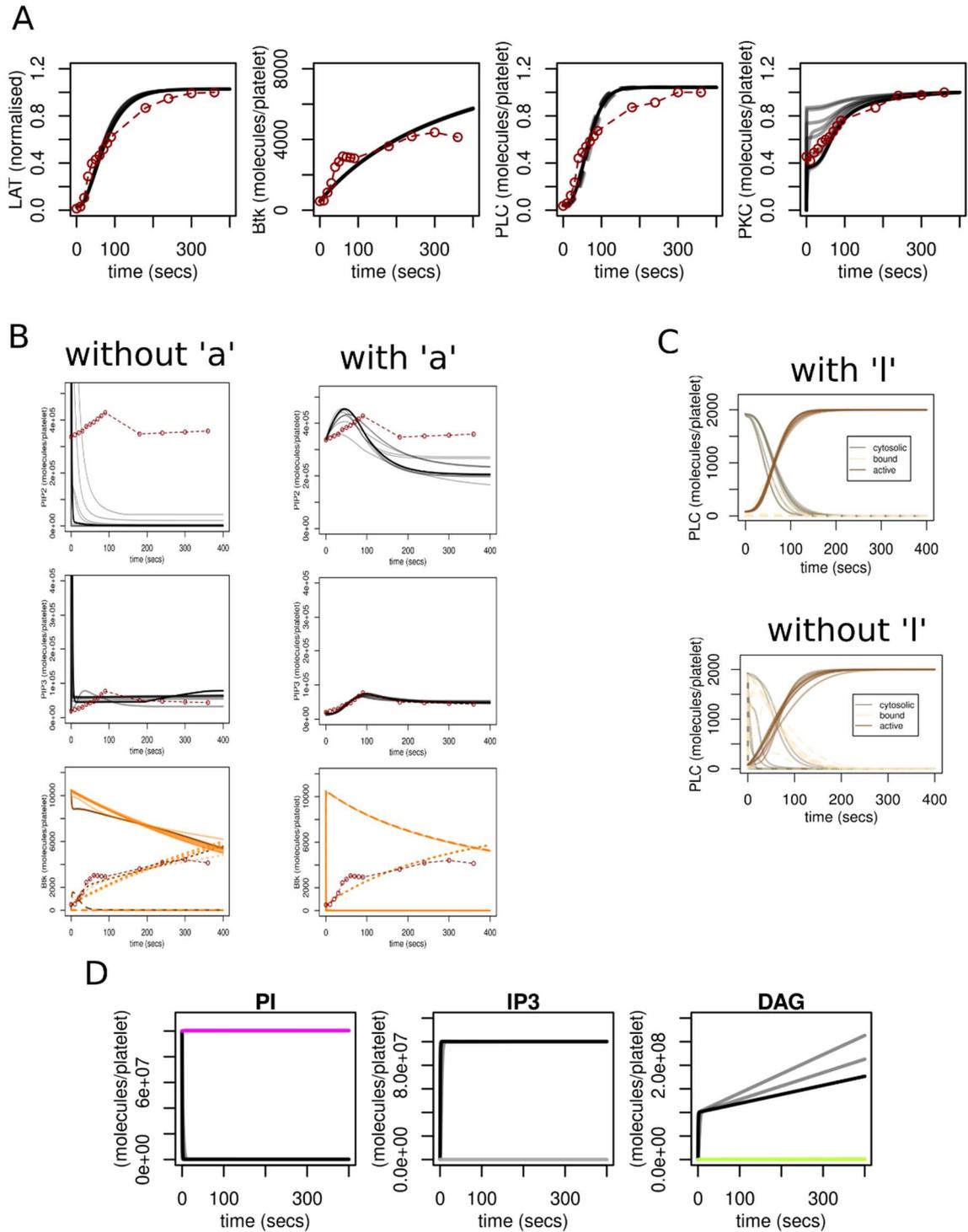


Figure 5 Comparison of model simulations. A) An example of simulations of LAT, Btk, PLC, and PKC activity (black lines) compared to data (red lines). B) Simulations from a model that does not incorporate nonlinear phosphorylation (left) compared to simulations from a model that

does (right). C) Simulations of PLC activity comparing the outputs from a model where PLC binds to the signalosome compared to a model where PLC binds to PIP3. D) Simulations demonstrating divergent predictions for PI, IP3 and DAG. All simulations show the 10 best fits. The different colours highlight the different predictions for subsets of simulations.

In summary, while non-linear regulation of Btk is supported by the current experimental dataset differential predictions of cytosolic and membrane bound proteins points the way to a further experimentation that could be informative in discriminating between competing hypothesis regarding the reactions that dominate in the GPVI signalling pathway.

The impact of collecting experimental measurements of IP3

In simulations from all models, predictions of the IP3 inositol trisphosphate are highly variable. While measurement of IP3 is difficult, due to its rapid hydrolysis, we hypothesised that experimental data describing this key molecule may increase the model identifiability (and therefore predictive ability). Data describing changes in IP3 in platelets downstream of an alternative receptor are available in literature¹⁹. The cost measure associate with IP3 has been added to SSE calculation. This data (Figure 6, right panels (in red)) re-inferred models 'bld' (the model least able to replicate the original datasets) and 'blad' (the best).

The inclusion of the additional data did not improve the agreement between model simulations and data for both models (Supplemental Figures 10 and 11). Parameter ranges, compared to priors, were also similar with only minor reductions in the uncertainty in their values. Both models failed to fit the IP3 data. But, model 'blad' shows consistency in its predictions, displaying a slow-prolonged release of IP3 to the cytosol that is in contradiction to the experimental data from literature which shows thrombin stimulated IP3 release with a sharp transient before falling to a lower positive steady state by 100 seconds after stimulation.

Calcium flux from intracellular stores is known to closely follow the release of IP3 and thrombin stimulated calcium flux shows a similar pattern to the data used for model fitting and CRP stimulated calcium flux reminiscent of the predictions from model 'blad'²⁰.

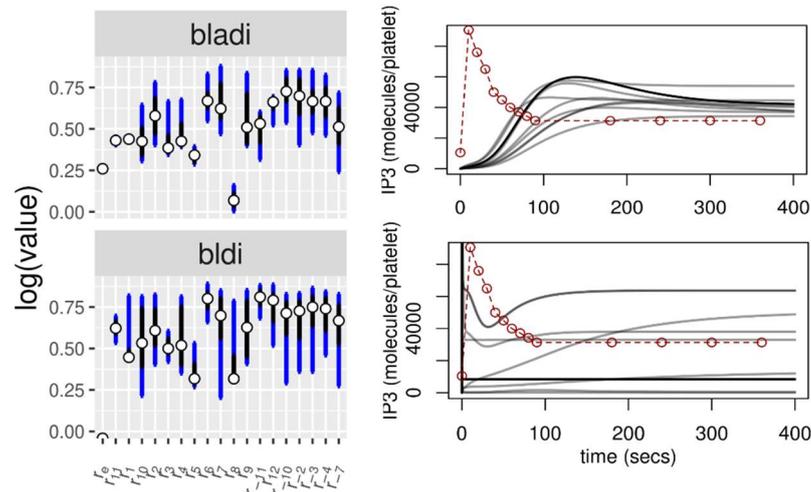


Figure 6 The impact of fitting the models 'blad' and 'bld' to datasets that include changes in IP3, the 'i' indicating the model was inferred to IP3 data. Left panels show the parameter ranges of the 100 lowest SSE normalised to the prior distribution. Vertical lines represent the 95% credible interval, black the inter-quartile, red the inner 10%. Circles represent medians. Right panels show model simulations for IP3 (black lines) against IP3 data (red) (top, from model blad, bottom from bld).

Sensitivity analysis

The sensitivity analysis has been performed on model blad by varying both parameters (from -90% and +90%) and initial conditions (-50% and +50%) to see their effect on steady state level (SS), maximum level (max), time reach maximum (time), and slope ratio (shape) of each specie in the model, defined in Figure 7C. It reveals that the levels of the phosphoinositides are unsurprisingly controlled by the internal rates of conversion between phosphoinositides (r_3 , r_{-3} , r_{11} , r_{-11} and r_{12}). Rising levels of PI, or its rates of conversion to other species (r_{11} , r_{-11} and r_{12}),

results in the raising of PIP₂, PIP₃ and its downstream products IP₃ and DAG. Levels of PI3K and the rate it converts PIP₂ to PIP₃ (r_3 , r_{-3}) have a differential effect, with rising levels increasing PIP₃ and decreasing PIP₂ and its downstream products. Similar divergent effects are seen from changing levels of PLC or its action on PIP₂ (r_5), which increase PIP₂ and PIP₃ but lower IP₃, DAG and PKC activity.

Another important note is the decrease/increase in the rate that IP₃ is recycled back to the membrane (r_6) and only affects IP₃ levels (both steady state and its maximum), it being returned to the pool of phosphoinositides that are greatly in excess of IP₃ – the difference in order of magnitude. Additionally, the nonlinear phosphorylation term (k_e) of the model dominates the behaviour of Btk. The activity of the signalosome, which in this model comprises of LAT, PI3K and PLC, has effects on most downstream events. The detailed results are presented in Figure 7A.

Figure 7B shows the effect of initial conditions variation on the same set of observations (SS, max, shape, and time). The first thing we could spot is the variation of ligand concentration (L) largely affects the response time throughout the signalling cascade (PI3L, LAT, PLC, and PKC). Steady state (SS) of some components is not only sensitive to upstream components but also downstream; for example, $PI3K_{SS}$ is sensitive to $[PI3K]_0$ (itself), $[LAT]_0$ (upstream), and $[PLC]_0$ (downstream). This could be due to the reverse reaction; thus lower PLC could decrease the rate of spanning the PI3K* species (which can be $LAT^*:PI3K^*$, $LAT^*:PLC:PI3K^*$, and $LAT^*:PLC^*:PI3K^*$, see figure 1). While LAT signalosome interactions are complicated, BTK_{max} is only controlled by $[BTK]_0$ itself, as it comparatively independent. However, we could see that the variation of other variables / initial conditions, either positively or negatively, affect the BTK_{shape} such that the 'shape' is getting larger, meaning that the BTK curve can quickly become a plateau.

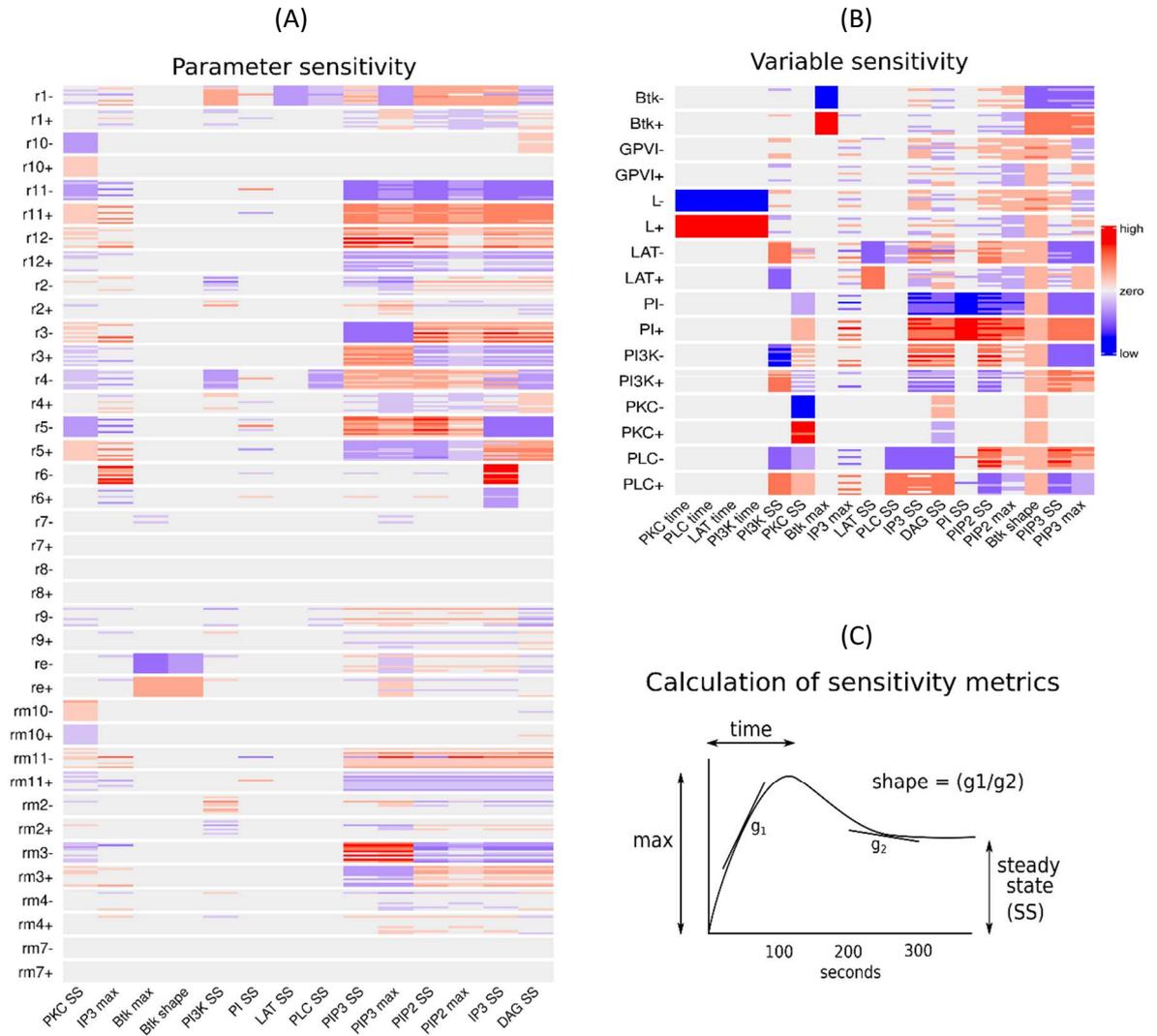


Figure 7 Heatmaps resulting from the sensitivity analysis. A) shows the effect of varying the models parameters (-90% and +90%). B) the effect of varying the initial values of the variables (-50% and +50%). Each column represents a model's output, and each row a parameter or variable (where a negative sign indicates a decrease and a positive increase). The summary statistic defined as $value = ((manipulated - original)/original) \times 100$. In the heatmap the data are binned into 5 groups with cut offs $\pm 10, 50, 100, 500$. Red and blue colours reflect the increasing and decreasing of value, respectively. C) the illustration of the definition of observe sensitivities: SS, max, time, and shape. SS is the average concentration during a steady state (from $t = 320$ s to 360 s). max is the maximum concentration of the time series. time is the time when the concentration hits its maximum value. shape is the ratio of the slope of tangential lines before and after the curve reaches its maximum point.

Discussion

In this study, we described a mathematical modelling approach to investigate the signaling pathway activated by CRP in platelets. The modelling approach provides insights into the role of key molecules and processes in the signaling pathway and highlights the limitations of the current understanding of the pathway. We aimed to test current knowledge of the structure of the GPVI signalling pathway using mathematical models. Our goal was not to create a final and complete model, but to form a framework that could be modified and extended as more data and ideas were generated. The complexity of the pathway and the data made intuition-based analyses difficult, so we turned to mathematical models as a way to test hypotheses and direct resources to experiments that were likely to provide the most information.

We encountered a number of challenges while creating these models. Firstly, we found that it was relatively easy to generate mathematical models and predictions, but the problem was making predictions that could be relied on in a research or even clinical setting. We knew that for most models, many different parameter sets could generate results that fit data, but some may not be valid, and some may be excluded when tested against alternative data or produce divergent predictions under perturbations. To address this issue, we used the Bayesian method, which provides a framework for assessing uncertainty in models in both the parameters and the structure of the model²¹. In a Bayesian framework, the prior of the parameters play a crucial role in the estimation of parameters, and the prior was selected based on the available literature and expert knowledge.

Additionally, as the models describing subcellular pathways became more complex, it was easy for them to be over-parameterised with respect to the experimental data used for

parameter estimation, resulting in uncertainty in the individual parameter estimates as well as in the predictions made from the model. To address this, we sought to assess whether the amount and quality of the experimental data constrained the parameter estimates, which is known as practical identifiability²²⁻²⁴. Even for our model that shows the best ability to reflect the available experimental data, only 3 parameters, i.e. r_1 , r_8 , and r_e , were appreciably narrowed from their priors, indicating that there is always an inherent tension often neglected when constructing mathematical descriptions of biological systems between the desire to include every possible interaction into a model, increasing uncertainty, and the availability of supporting data that allow such a model to make valid predictions. The parameter's posterior could also depend on the proximity of the corresponding interactions to the data being measured; for example, r_1 is closest to measured LAT* without other parameters' interference. The uncertainty in our parameter estimates raises the possibility that a structurally simpler model may fit the empirical data equally well. However, we do not take this approach, preferring to retain biological components that are the focus of research, instead using the unidentifiability to point to areas of further research.

The model was calibrated and tested against an unusually comprehensive set of dense data that describes the time-dependent changes in six key components obtained from the use of multi-omics techniques, such as phospho-proteomics, which enables characterisation of protein copy numbers and phosphorylation of signalling components. These models were designed to test our knowledge of the pathway against our data and represented a first step in understanding the complex interactions within the pathway. Of the competing models, a model that includes nonlinear regulation of Btk was found to be the best able to describe the experimental data, uncovering the importance of nonlinear regulation of Btk in stabilizing the pathway. Additionally, our attempts to extend the model highlighted the importance of the

membrane in the regulation of the pathway, both in terms of stabilization and speed to activation. We discovered that while PIP2 and PIP3 represent less than 1% of membrane phospholipids, their ability to act as a tether controls the regulation of Btk with feedback through the stabilization of PIP3 that, in turn, regulates the phosphatidylinositol pathway. The importance of membrane tethering in the regulation of the pathway points the way to the inclusion of spatial effects that, through lack of spatially resolved data, was necessarily neglected here. Furthermore, Calcium data is readily available, but given the uncertainty of the intermediate steps, it was also not incorporated. However, with well-controlled experimental set up, it is a future priority to incorporate calcium time series into the model for purposes of calibration. This development could also extend the model's validation capability by experiment on ligand concentration as has been done in²⁵. In conclusion, this study paves the way for future research to enhance the model by incorporating calcium time series data, enabling a more comprehensive understanding of the CRP/collagen-activated signalling pathway in platelets.

Authors contributions

CT and JLD performed the modelling and coding part, with mathematical advice from RC, and CW, and biological advice and supervision from JMG and JWMH. APB and AJU collected all experimental data used in this work. CT, JLD, JMG, and JWMH participated in drafting the manuscript.

Acknowledgements

CT received funding from the European Union's Horizon 2020 research and innovation Marie Skłodowska-Curie program, grant agreement No. 766118. CT registered in a joined PhD

program at the University of Reading and Maastricht University. JLD is supported by the British Heart Foundation (RG/20/7/34866) and the MRC (MR/W015293/1).

Conflict of interest

The authors have no conflicts of interest to disclose.

References

1. Holinstat, M. Normal platelet function. *Cancer Metastasis Rev.* **36**, 195–198 (2017).
2. Perrella, G., Nagy, M., Watson, S. P. & Heemskerk, J. W. M. Platelet GPVI (Glycoprotein VI) and Thrombotic Complications in the Venous System. *Arterioscler. Thromb. Vasc. Biol.* **41**, 2681–2692 (2021).
3. Zheng, T. J., Lofurno, E. R., Melrose, A. R., Lakshmanan, H. H. S., Pang, J., Phillips, K. G., *et al.* Assessment of the effects of Syk and BTK inhibitors on GPVI-mediated platelet signaling and function. *Am. J. Physiol. - Cell Physiol.* **320**, C902–C915 (2021).
4. Joanna L. Dunster, Mikhail A. Panteleev, Jonathan M. Gibbins, A. N. S. Mathematical Techniques for Understanding Platelet Regulation and the Development of New Pharmacological Approaches. *Platelets and Megakaryocytes* **1812**, 255–279 (2018).
5. Liepe, J., Filippi, S., Komorowski, M. & Stumpf, M. P. H. Maximizing the Information Content of Experiments in Systems Biology. *PLoS Comput. Biol.* **9**, (2013).
6. Babbie, A. C. & Stumpf, M. P. H. How to deal with parameters for whole-cell modelling. *J. R. Soc. Interface* **14**, (2017).
7. Dunster, J. L., Mazet, F., Fry, M. J., Gibbins, J. M. & Tindall, M. J. Regulation of early steps of GPVI signal transduction by phosphatases: A systems biology approach. *PLoS Comput. Biol.* **11**, 1–26 (2015).
8. Cheung, H. Y. F., Coman, C., Westhoff, P., Manke, M., Sickmann, A., Borst, O., *et al.* Targeted Phosphoinositides Analysis Using High-Performance Ion Chromatography-Coupled Selected Reaction Monitoring Mass Spectrometry. *J. Proteome Res.* **20**, 3114–3123 (2021).
9. Kielkowska, A., Niewczas, I., Anderson, K. E., Durrant, T. N., Clark, J., Stephens, L. R., *et al.* A new approach to measuring phosphoinositides in cells by mass spectrometry. *Adv. Biol. Regul.* **54**, 131–141 (2014).
10. Pasquet, J.-M., Gross, B., Quek, L., Asazuma, N., Zhang, W., Sommers, C. L., *et al.* LAT Is Required

- for Tyrosine Phosphorylation of Phospholipase C α 2 and Platelet Activation by the Collagen Receptor GPVI. *Mol. Cell. Biol.* **19**, 8326–8334 (1999).
11. Majerus, P. W. Inositol phosphate biochemistry. *Annu. Rev. Biochem.* **61**, 225–250 (1992).
 12. Rawlings, D. J., Scharenberg, A. M., Park, H., Wahl, M., Lin, S., Kato, R. M., *et al.* Activation of Btk by autophosphorylation mechanism initiated by SRC Family Kinases. *Science (80-.)*. **271**, 822–825 (1996).
 13. Wang, Q., Vogan, E. M., Nocka, L. M., Rosen, C. E., Zorn, J. A., Harrison, S. C., *et al.* Autoinhibition of Bruton’s tyrosine kinase (Btk) and activation by soluble inositol hexakisphosphate. *Elife* **2015**, 1–31 (2015).
 14. Chung, J. K., Nocka, L. M., Decker, A., Wang, Q., Kadlecsek, T. A., Weiss, A., *et al.* Switch-like activation of Bruton’s tyrosine kinase by membrane-mediated dimerization. *Proc. Natl. Acad. Sci. U. S. A.* **166**, 10798–10803 (2019).
 15. Anderson, D. R., Burnham, K. P. & White, G. C. Comparison of Akaike information criterion and consistent Akaike information criterion for model selection and statistical inference from capture-recapture studies. *J. Appl. Stat.* (1998).
 16. Wahl, M. I., Fluckiger, A. C., Kato, R. M., Park, H., Witte, O. N. & Rawlings, D. J. Phosphorylation of two regulatory tyrosine residues in the activation of Bruton’s tyrosine kinase via alternative receptors. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 11526–11533 (1997).
 17. Mujalli, A., Chicanne, G., Bertrand-Michel, J., Viars, F., Stephens, L., Hawkins, P., *et al.* Profiling of phosphoinositide molecular species in human and mouse platelets identifies new species increasing following stimulation. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1863**, 1121–1131 (2018).
 18. Valet, C., Chicanne, G., Severac, C., Chaussade, C., Whitehead, M. A., Cabou, C., *et al.* Essential role of class II PI3K-C2 α in platelet membrane morphology. *Blood* **126**, 1128–1137 (2015).
 19. Rittenhouse, S. E. & Sasson, J. P. Mass changes in myoinositol triphosphate in human platelets stimulated by thrombin. Inhibitory effects of phorbol ester. *J. Biol. Chem.* **260**, 8657–8660 (1985).
 20. Fernández, D. I., Provenzale, I., Cheung, H. Y. F., van Groningen, J., Tullemans, B. M. E., Veninga, A., *et al.* Ultra-high-throughput Ca $^{2+}$ assay in platelets to distinguish ITAM-linked and G-protein-coupled receptor activation. *iScience* **25**, (2022).
 21. Lintusaari, J., Gutmann, M. U., Dutta, R., Kaski, S. & Corander, J. Fundamentals and recent developments in approximate Bayesian computation. *Syst. Biol.* **66**, e66–e82 (2017).
 22. Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., *et al.* Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile

- likelihood. *Bioinformatics* **25**, 1923–1929 (2009).
23. Shotwell, M. S. & Gray, R. A. Estimability Analysis and Optimal Design in Dynamic Multi-scale Models of Cardiac Electrophysiology. *J. Agric. Biol. Environ. Stat.* (2016) doi:10.1007/s13253-016-0244-7.
 24. Sher, A., Niederer, S. A., Mirams, G. R., Kirpichnikova, A., Allen, R., Pathmanathan, P., *et al.* A Quantitative Systems Pharmacology Perspective on the Importance of Parameter Identifiability. *Bull. Math. Biol.* **84**, 1–15 (2022).
 25. Jooss, N. J., Simone, I. De, Provenzale, I., Fernández, D. I., Brouns, S. L. N., Farndale, R. W., *et al.* Role of platelet glycoprotein VI and Tyrosine Kinase Syk in thrombus formation on collagen-like surfaces. *Int. J. Mol. Sci.* **20**, (2019).

Supplemental

Here we present details of the mathematical models and provide the results of inferring the models to the experimental data.

1. Mathematical model structures

1.1 Previous model

We extend our previous work describing the early events downstream of the GPVI receptor to capture current biological knowledge of the events up to, and including, the production of IP3. The details of the model of the early reactions, taken from¹, are included for completeness. The variables, initial conditions, parameters and values are summarised in Table 1 and the equations are:

$$\frac{d[l]}{dt} = -\frac{k_1}{V_e A_v} [l][g] + \frac{k_{-1}}{V_e A_v} [G] \quad (1a)$$

$$\frac{d[g]}{dt} = -k_1 [l][g] + k_{-1} [G] \quad (1b)$$

$$\frac{d[G]}{dt} = k_1 [l][g] - k_{-1} [G] - k_2 [G] \quad (1c)$$

$$\frac{d[G^p]}{dt} = k_2 [G] - \frac{k_3}{V_p A_v} [G^p][s] \quad (1d)$$

$$\frac{d[G_{0,0}^b]}{dt} = \frac{k_3}{V_p A_v} [G^p][s] - p_2 [G_{0,0}^b] + p_{-2} [G_{1,0}^b] - p_1 [G_{0,0}^b] + \gamma_2 ([G_{1,1}^u] + [G_{1,0}^r]) [G_{0,1}^b] \quad (1e)$$

$$\begin{aligned} \frac{d[G_{1,0}^b]}{dt} = & p_2 [G_{0,0}^b] - p_{-2} [G_{1,0}^b] - \frac{p_3}{V_p A_v} [G_{1,0}^b][c] + p_{-3} [G_{1,0}^c] - p_1 [G_{1,0}^b] \\ & + \gamma_2 ([G_{1,1}^u] + [G_{1,0}^r]) [G_{1,1}^b] \end{aligned} \quad (1f)$$

$$\begin{aligned} \frac{d[G_{1,0}^c]}{dt} = & \frac{p_3}{V_p A_v} [G_{1,0}^b][c] - p_{-3} [G_{1,0}^c] - p_1 [G_{1,0}^c] - p_4 [G_{1,0}^c] + p_{-4} [G_{1,0}^u] \\ & + \gamma_2 ([G_{1,1}^u] + [G_{1,0}^r]) [G_{1,1}^c] \end{aligned} \quad (1g)$$

$$\begin{aligned} \frac{d[G_{1,0}^u]}{dt} &= p_4[G_{1,0}^c] - p_{-4}[G_{1,0}^u] - \frac{p_5}{V_p A_v} [G_{1,0}^u][r] - p_1[G_{1,0}^u] + p_{-5}[G_{1,0}^r] \\ &\quad + \gamma_2([G_{1,1}^u] + [G_{1,0}^r])[G_{1,1}^u] \end{aligned} \quad (1h)$$

$$\frac{d[G_{1,0}^r]}{dt} = \frac{p_5}{V_p A_v} [G_{1,0}^u][r] - p_1[G_{1,0}^r] + \gamma_1[G_{1,1}^u] - p_{-5}[G_{1,0}^r] \quad (1i)$$

$$\frac{d[G_{0,1}^b]}{dt} = p_1[G_{0,0}^u] - q_2 p_2 [G_{0,1}^b] + p_{-2}[G_{1,1}^b] - \gamma_2([G_{1,1}^r] + [G_{1,0}^r])[G_{0,1}^b] \quad (1j)$$

$$\begin{aligned} \frac{d[G_{1,1}^b]}{dt} &= p_1[G_{1,0}^b] + q_2 p_2 [G_{0,1}^b] - p_{-2}[G_{1,1}^b] - \frac{p_3}{V_p A_v} [G_{1,1}^b][c] + p_{-3}[G_{1,1}^c] \\ &\quad - \gamma_2([G_{1,1}^r] + [G_{1,0}^r])[G_{1,1}^b] \end{aligned} \quad (1k)$$

$$\begin{aligned} \frac{d[G_{1,1}^c]}{dt} &= p_1[G_{1,0}^c] + \frac{p_3}{V_p A_v} [G_{1,1}^b][c] - p_{-3}[G_{1,1}^c] - p_4[G_{1,1}^c] + p_{-4}[G_{1,1}^u] \\ &\quad - \gamma_2([G_{1,1}^r] + [G_{1,0}^r])[G_{1,1}^c] \end{aligned} \quad (1l)$$

$$\begin{aligned} \frac{d[G_{1,1}^u]}{dt} &= p_1[G_{1,0}^u] + p_4[G_{1,1}^c] - p_{-4}[G_{1,1}^u] - \frac{p_5}{V_p A_v} [G_{1,1}^u][r] + p_{-5}[G_{1,1}^r] \\ &\quad - \gamma_2([G_{1,1}^r] + [G_{1,0}^r])[G_{1,1}^u] \end{aligned} \quad (1m)$$

$$\frac{d[G_{1,1}^r]}{dt} = p_1[G_{1,0}^r] + \frac{p_5}{V_p A_v} [G_{1,1}^u][r] - \gamma_1[G_{1,1}^r] - p_{-5}[G_{1,1}^r] \quad (1n)$$

$$\frac{d[s]}{dt} = -\frac{k_3}{V_p A_v} [G^p][s] \quad (1o)$$

$$\frac{d[c]}{dt} = -\frac{p_3}{V_p A_v} [G_{1,0}^b][c] + p_{-3}[G_{1,0}^c] - \frac{p_3}{V_p A_v} [G_{1,1}^b][c] + p_{-3}[G_{1,1}^c] \quad (1p)$$

$$\frac{d[r]}{dt} = -\frac{p_5}{V_p A_v} [G_{1,0}^u][r] - \frac{p_5}{V_p A_v} [G_{1,1}^u][r] + p_{-5}[G_{1,1}^u] + p_{-5}[G_{1,0}^r] \quad (1q)$$

where, in the equations below, $[Syk - Y525^*] = [G_{0,0}^b] + [G_{0,1}^b] + [G_{1,1}^c] + [G_{1,1}^u] + [G_{1,1}^r]$

and $[SFK] = [g] + [G] + [G_p] + [G_{0,0}^b] + [G_{1,0}^b] + [G_{1,0}^c] + [G_{1,0}^u] + [G_{1,0}^r] + [G_{0,1}^b] + [G_{1,1}^b]$

$+ [G_{1,1}^c] + [G_{1,1}^u] + [G_{1,1}^r]$.

Variable	Initial condition	Parameter	Value
l	3×10^{-4}	k_1	8
g	5000	k_{-1}	0.0302
s	2763	k_2	0.211
G	0	k_3	614.8
G^p	0	V_e	3.3×10^{-9}
$G^{k_{i,j}}$	0	V_p	7.4×10^{-18}
r	7800	p_1	67.954
c	2581	p_2	22.526
		p_{-2}	47.008
		p_3	20812
		p_{-3}	27.954
		p_4	21.689
		p_{-4}	54.011
		p_5	6.4052
		p_{-5}	0.23309
		q2	17.02
		γ_1	99.846
		γ_2	2.5969

Table 1 Variables, initial conditions (left) and parameters and values (right) used for simulations of equations (1), our previous model of the early events in signalling through the GPVI receptor. $G^{k_{i,j}}$ represents eight variables where i and j indicate phosphorylation on Y323 and Y525 respectively (0, unphosphorylated; 1, phosphorylated) and k denotes the sequential processes of (b), binding of Syk; (c), binding of c-Cbl; (u), ubiquitination and (r), binding of TULA-2.

1.2 Baseline model (b)

The extension to this model is initiated by the phosphorylation of LAT and the assembly of the LAT signalosome that includes the recruitment of PI3K. The equations describing these events are given by:

$$\frac{d[LAT]}{dt} = -r_1[Syk - Y525^*][LAT] \quad (2a)$$

$$\frac{d[LAT^*]}{dt} = r_1[Syk - Y525^*][LAT] - \frac{r_2}{V_p A_v}[LAT^*][PI3K] + r_{-2}[LAT^*:PI3K] \quad (2b)$$

$$\frac{d[PI3K]}{dt} = -\frac{r_2}{V_p A_v}[LAT^*][PI3K] + r_{-2}[LAT^*:PI3K] \quad (2c)$$

$$\frac{d[LAT^*:PI3K]}{dt} = \frac{r_2}{V_p A_v}[LAT^*][PI3K] - r_{-2}[LAT^*:PI3K] \quad (2d)$$

these being conserved such that $[LAT]_T - [LAT] - [LAT^*] = \frac{1}{V_p A_v}[LAT^*:PI3K]$ and $[PI3K]_T - [PI3K] = \frac{1}{V_p A_v}[LAT^*:PI3K]$. The output $[LAT^*:PI3K]$ alters the balance of the

phosphoinositide network, for which the equations for the key components are described by

$$\frac{d[PI]}{dt} = r_{-11}[PIP2] - r_{11}[PI] + r_6[IP3][DAG] + r_{-12}[PIP3] \quad (3a)$$

$$\begin{aligned} \frac{d[PIP2]}{dt} = & -r_3[LAT^*:PI3K][PIP2] + r_{-3}[PIP3] - r_{-11}[PIP2] + r_{11}[PI] \\ & -r_5[PIP3:PLC^*][PIP2] \end{aligned} \quad (3b)$$

$$\begin{aligned} \frac{d[PIP3]}{dt} = & r_3[LAT^*:PI3K][PIP2] - r_{-3}[PIP3] - \frac{r_4}{V_p A_v}[PIP3][PLC] + r_{-4}[PIP3:PLC] \\ & - \frac{r_7}{V_p A_v}[PIP3][Btk] + r_{-7}[PIP3:Btk] - r_{-12}[PIP3] \end{aligned} \quad (3c)$$

The equations for Btk and PLC are given by

$$\frac{d[Btk]}{dt} = -\frac{r_7}{V_p A_v}[PIP3][Btk] + r_{-7}[PIP3:Btk] \quad (3d)$$

$$\frac{d[PIP3:Btk]}{dt} = \frac{r_7}{V_p A_v}[PIP3][Btk] - r_{-7}[PIP3:Btk] - r_8[PIP3:Btk][SFK] \quad (3e)$$

$$\frac{d[PIP3:Btk^*]}{dt} = r_8[PIP3:Btk][SFK] \quad (3f)$$

$$\frac{d[PLC]}{dt} = -\frac{r_4}{V_p A_v} [PIP3][PLC] + r_{-4}[PIP3:PLC] \quad (3g)$$

$$\frac{d[PIP3:PLC]}{dt} = \frac{r_4}{V_p A_v} [PIP3][PLC] - r_{-4}[PIP3:PLC] - r_9[PIP3:Btk^*][PIP3:PLC] \quad (3h)$$

$$\frac{d[PIP3:PLC^*]}{dt} = r_9[PIP3:Btk^*][PIP3:PLC] \quad (3i)$$

and those for IP3, DAG and PKC

$$\frac{d[IP3]}{dt} = r_5[PIP3:PLC^*][PIP2] - r_6[IP3][DAG] \quad (4a)$$

$$\frac{d[PKC]}{dt} = -r_{10}[DAG][PKC] + r_{-10}[DAG:PKC^*] \quad (4b)$$

$$\frac{d[DAG]}{dt} = r_5[PIP3:PLC^*][PIP2] - r_6[IP3][DAG] - r_{10}[DAG][PKC] + r_{-1} [DAG:PKC^*] \quad (4c)$$

$$\frac{d[DAG:PKC^*]}{dt} = r_{10}[DAG][PKC] - r_{-10}[DAG:PKC^*] \quad (4d)$$

Conservation within these equations (3),(4) is described by

$$\begin{aligned} [PI]_T &= [PI] + [PIP2] + [PIP3] + [PIP3:PLC] + [PIP3:PLC^*] + [PIP3:Btk] \\ &\quad + [PIP3:Btk^*] + [IP3] \end{aligned} \quad (5a)$$

$$[Btk]_T - [Btk] = \frac{1}{V_p A_v} ([PIP3:Btk] + [PIP3:Btk^*]) \quad (5b)$$

$$[PLC]_T - [PLC] = \frac{1}{V_p A_v} ([PIP3:PLC] + [PIP3:PLC^*]) \quad (5c)$$

These seventeen equations comprise model "b".

1.3 Autophosphorylation model (a)

Models denoted with an additional "a" capture autophosphorylation of Btk and as such have an additional term added to equations (3e) and (3f) such that

$$\begin{aligned} \frac{d[PIP3:Btk]}{dt} &= \frac{r_7}{V_p A_v} [PIP3][Btk] - r_{-7}[PIP3:Btk] - r_8[PIP3:Btk][SFK] \\ &\quad - r_e [PIP3:Btk]^2 \end{aligned} \quad (6a)$$

$$\frac{d[PIP3:Btk^*]}{dt} = r_8[PIP3:Btk][SFK] + r_e [PIP3:Btk]^2 \quad (6b)$$

1.4 LAT signalosome model (I)

Models denoted with an additional "I" assume that PLC is recruited to the membrane via LAT signalosome, in place of recruitment to PIP3. Equations (2) therefore read

$$\begin{aligned} \frac{d[LAT^*]}{dt} = & r_1[Syk - Y525^*][LAT] - \frac{r_2}{V_p A_v}[LAT^*][PI3K] + r_{-2}[LAT^*:PI3K] \\ & - \frac{r_4}{V_p A_v}[LAT^*][PLC] + r_{-4}[LAT^*:PLC] \end{aligned} \quad (7a)$$

$$\begin{aligned} \frac{d[PI3K]}{dt} = & - \frac{r_2}{V_p A_v}[LAT^*][PI3K] + r_{-2}[LAT^*:PI3K] - \frac{r_2}{V_p A_v}[LAT^*:PLC][PI3K] \\ & + r_{-2}[LAT^*:PI3K:PLC] - \frac{r_2}{V_p A_v}[LAT^*:PLC^*][PI3K] + r_{-2}[LAT^*:PI3K:PLC^*] \end{aligned} \quad (7b)$$

$$\begin{aligned} \frac{d[LAT^*:PI3K]}{dt} = & \frac{r_2}{V_p A_v}[LAT^*][PI3K] - r_{-2}[LAT^*:PI3K] - \frac{r_4}{V_p A_v}[LAT^*:PI3K][PLC] \\ & + r_{-4}[LAT^*:PI3K:PLC] \end{aligned} \quad (7c)$$

$$\begin{aligned} \frac{d[PLC]}{dt} = & - \frac{r_4}{V_p A_v}[LAT^*][PLC] + r_{-4}[LAT^*:PLC] - \frac{r_4}{V_p A_v}[LAT^*:PI3K][PLC] \\ & + r_{-4}[LAT^*:PI3K:PLC] \end{aligned} \quad (7d)$$

$$\begin{aligned} \frac{d[LAT^*:PLC]}{dt} = & \frac{r_4}{V_p A_v}[LAT^*][PLC] - r_{-4}[LAT^*:PLC] - \frac{r_2}{V_p A_v}[LAT^*:PLC][PI3K] \\ & + r_{-2}[LAT^*:PI3K:PLC] - r_9[PIP3:Btk^*][LAT^*:PLC] \end{aligned} \quad (7e)$$

$$\begin{aligned} \frac{d[LAT^*:PI3K:PLC]}{dt} = & \frac{r_4}{V_p A_v}[LAT^*:PI3K][PLC] - r_{-4}[LAT^*:PI3K:PLC] \\ & + \frac{r_2}{V_p A_v}[LAT^*:PLC][PI3K] - r_{-2}[LAT^*:PI3K:PLC] \\ & - r_9[LAT^*:PI3K:PLC][PIP3:Btk^*] \end{aligned} \quad (7f)$$

$$\begin{aligned} \frac{d[LAT^*:PLC^*]}{dt} = & r_9[PIP3:Btk^*][LAT^*:PLC] + \frac{r_2}{V_p A_v}[LAT^*:PLC^*][PI3K] \\ & - r_{-2}[LAT^*:PI3K:PLC^*] \end{aligned} \quad (7g)$$

$$\begin{aligned} \frac{d[LAT^*:PI3K:PLC^*]}{dt} = & \frac{r_2}{V_p A_v}[LAT^*:PLC^*][PI3K] - r_{-2}[LAT^*:PI3K:PLC^*] \\ & + r_9[LAT^*:PI3K:PLC][PIP3:Btk^*] \end{aligned} \quad (7h)$$

and the equations (3h) and (3i) are dropped and equations (3b) and (3c) are modified such that

$$\begin{aligned} \frac{d[PIP2]}{dt} = & -r_3([LAT^*:PI3K] + [LAT^*:PI3K:PLC] + [LAT^*:PI3K:PLC^*])[PIP2] \\ & +r_{-3}[PIP3] - r_{-1}[PIP2] + r_{11}[PI] \\ & -r_5([LAT^*:PLC^*] + [LAT^*:PI3K:PLC^*])[PIP2] \end{aligned} \quad (7i)$$

$$\begin{aligned} \frac{d[PIP3]}{dt} = & r_3([LAT^*:PI3K] + [LAT^*:PI3K:PLC] + [LAT^*:PI3K:PLC^*])[PIP2] \\ & -r_{-3}[PIP3] - \frac{r_7}{V_p A_v}[PIP3][Btk] + r_{-7}[PIP3:Btk] - r_{12}[PIP3] \end{aligned} \quad (7j)$$

1.5 DAG recycling model (d)

Models with a "d" neglect DAG recycling. This requires the alteration to equations (3a)

and (4) such that

$$\frac{d[PI]}{dt} = r_{-11}[PIP2] - r_{11}[PI] + r_6[IP3] + r_{12}[PIP3] \quad (8a)$$

$$\frac{d[DAG]}{dt} = r_5[PIP3:PLC^*][PIP2] - r_{10}[DAG][PKC] + r_{-10}[DAG:PKC^*] \quad (8b)$$

$$\frac{d[IP3]}{dt} = r_5[PIP3:PLC^*][PIP2] - r_6[IP3] \quad (8c)$$

Or the models with both "l" and "d", (8b) and (8c) must be

$$\begin{aligned} \frac{d[DAG]}{dt} = & r_5([LAT^*:PLC^*] + [LAT^*:PI3K:PLC^*])[PIP2] - r_{10}[DAG][PKC] \\ & +r_{-10}[DAG:PKC^*] \end{aligned} \quad (9a)$$

$$\frac{d[IP]}{dt} = r_5([LAT^*:PLC^*] + [LAT^*:PI3K:PLC^*])[PIP2] - r_6[IP3] \quad (9b)$$

In the above A_v stands for Avogadro's number, V_e and V_p are the volume of extracellular medium per cell under the experimental condition and cytosolic volume per platelet respectively. The variables and parameters for the models of this new section of the GPVI pathway are summarised in Tables 2 and Table 1 in main text, respectively.

Variable	Initial conditions	Source
LAT	4800	2
PI3K	3400	2
Btk	11100	2
PLC	2000	2
PI	10^8	3-5
PIP2	336666	in-house
PIP3	20000	in-house
IP3	0	6
PKC	17600	2
DAG	59000	2

Table 2 Initial conditions for the variables of the new models.

All units are molecules per platelet.

2. Additional plots and parameters

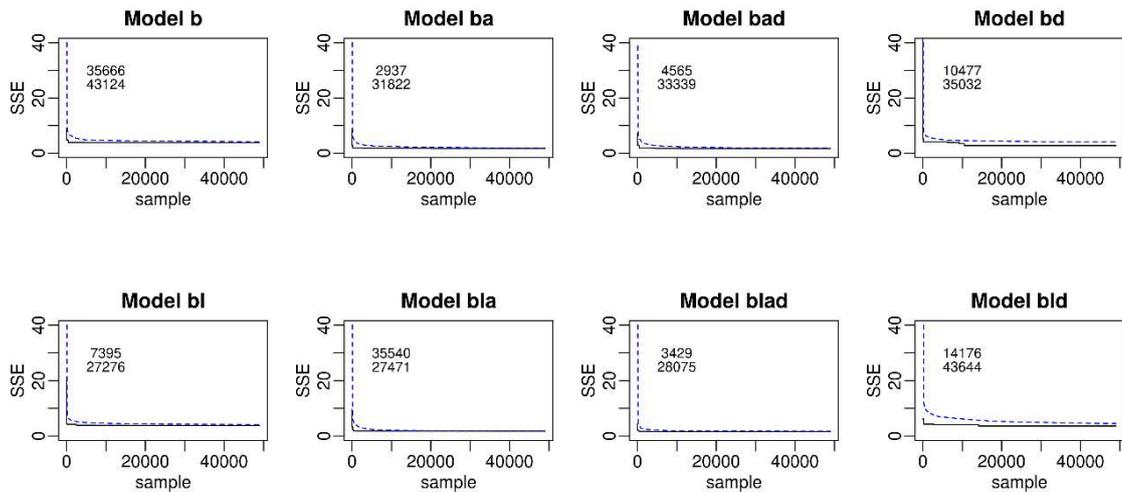


Figure 1 Samples until cost function convergence for the different models. Solid line represents the minimum 'best' fit (lowest cost function) with upper number in plot indicating the corresponding sample number to achieve this. Broken line represents the mean of the lowest 100 'best' fits (lowest cost function) with lower number indicating the corresponding sample number to achieve this

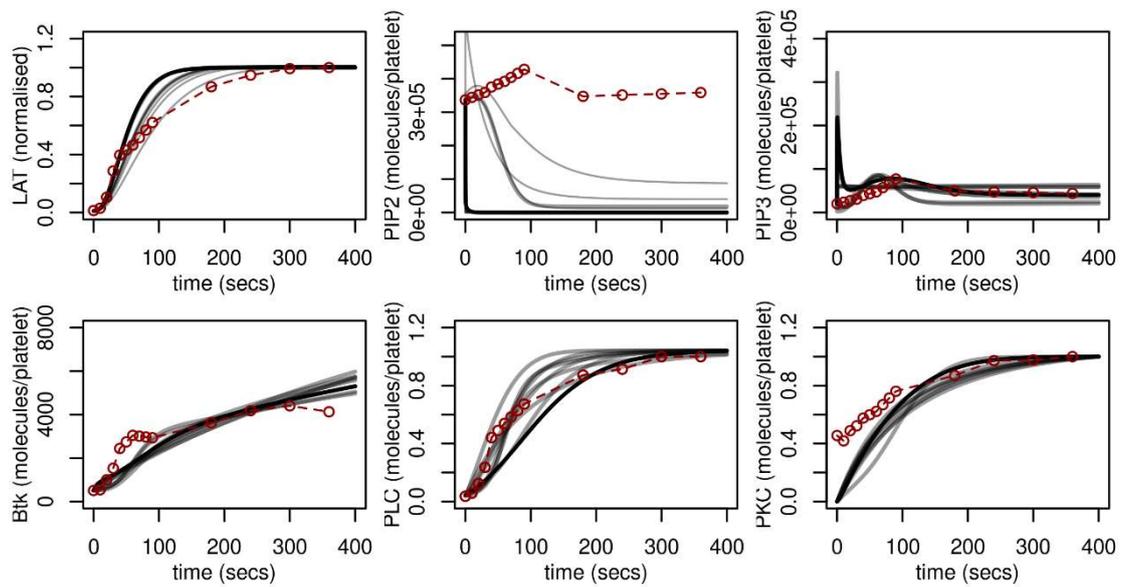


Figure 2 Model b simulations (solid lines) compared to experimental results (broken lines).
 Simulations are based on the parameter values from the 10 'best' fits.

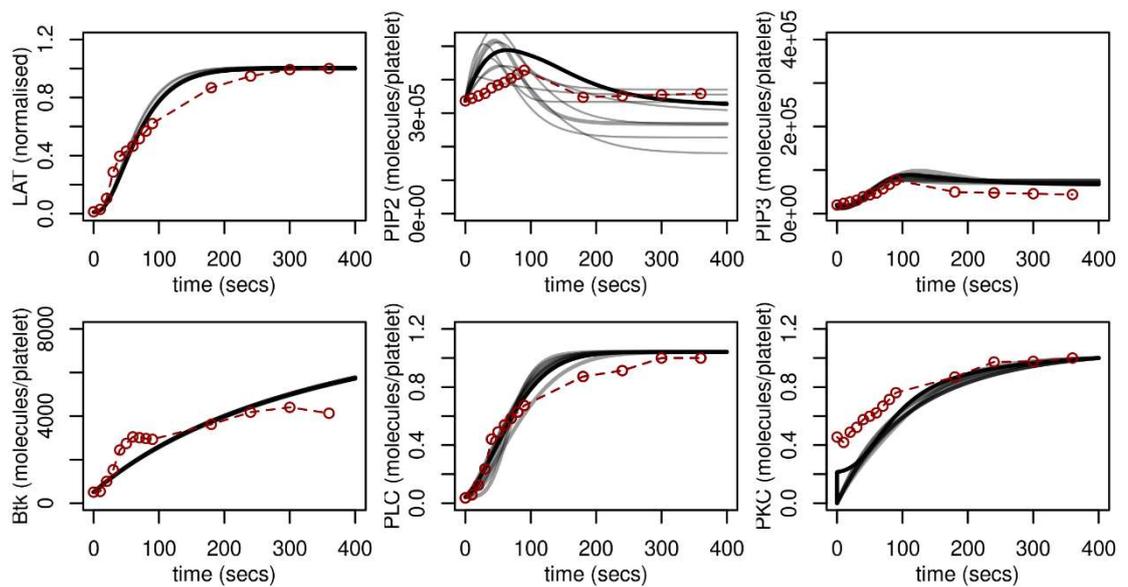


Figure 3 Model ba simulations (solid lines) compared to experimental results (broken lines).
 Simulations are based on the parameter values from the 10 'best' fits.

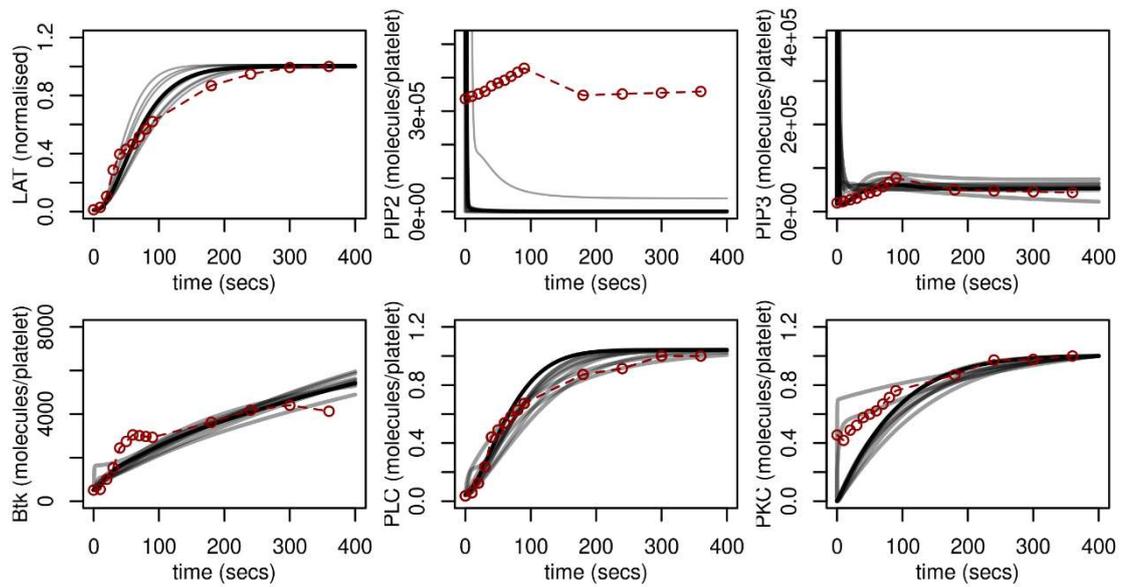


Figure 4 Model bd simulations (solid lines) compared to experimental results (broken lines).
 Simulations are based on the parameter values from the 10 'best' fits.

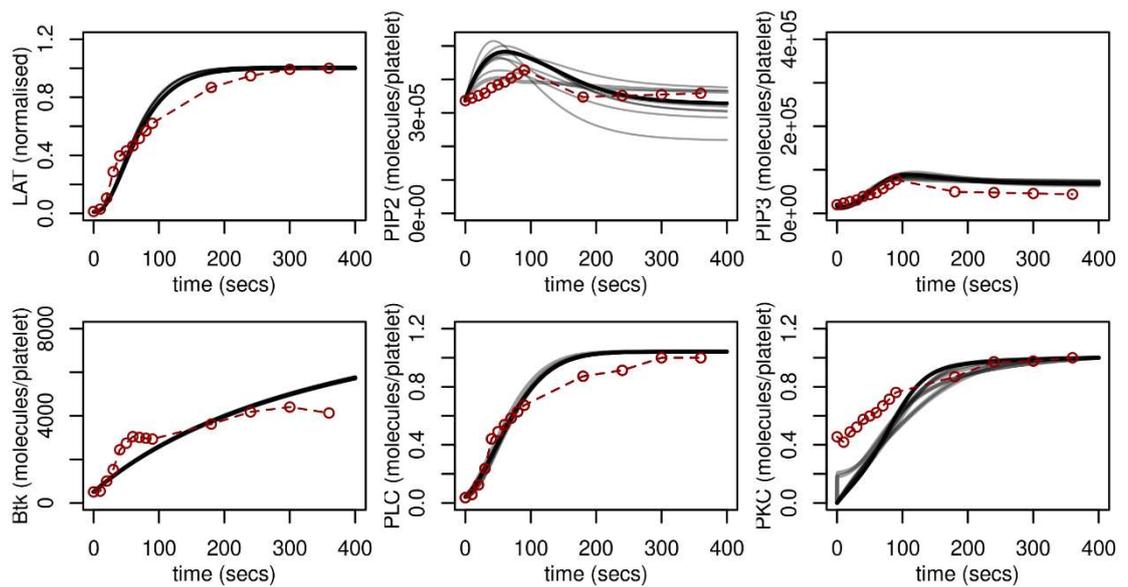


Figure 5 Model bad simulations (solid lines) compared to experimental results (broken lines).
 Simulations are based on the parameter values from the 10 'best' fits.

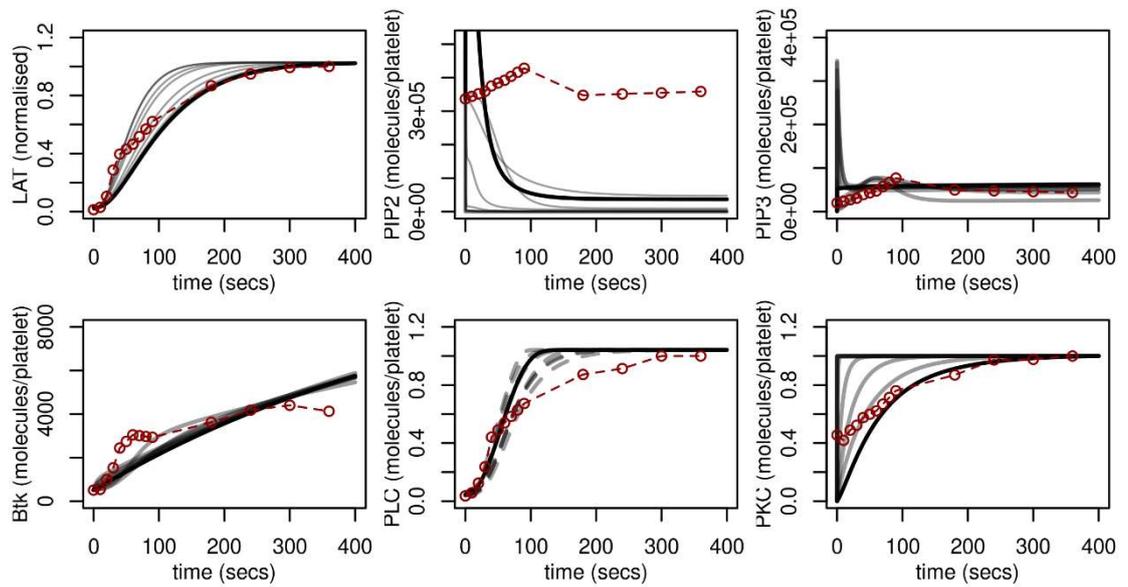


Figure 6 Model bl simulations (solid lines) compared to experimental results (broken lines).
 Simulations are based on the parameter values from the 10 'best' fits.

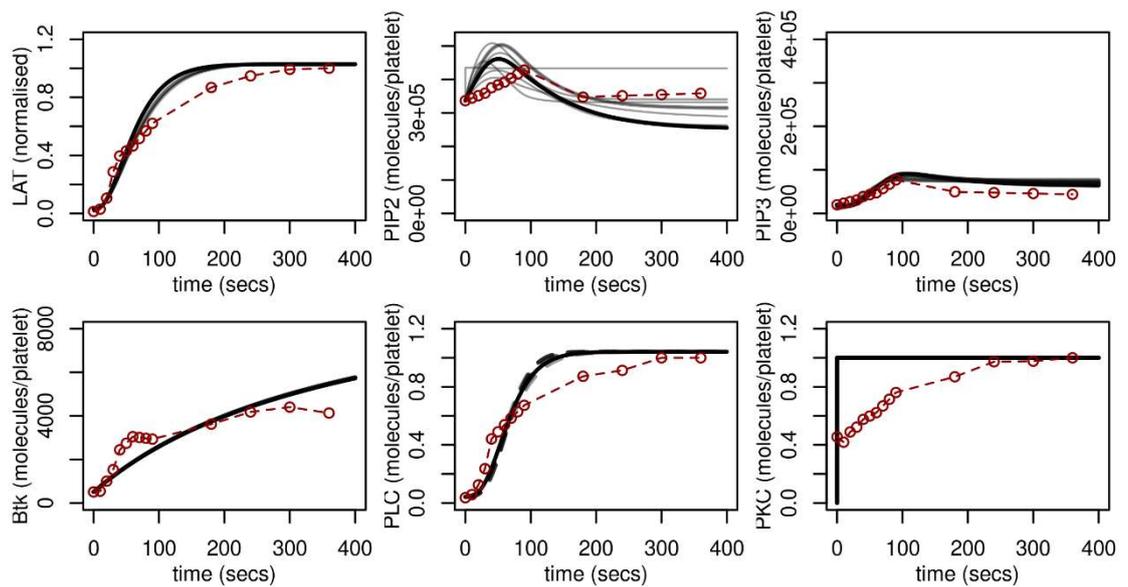


Figure 7 Model bla simulations (solid lines) compared to experimental results (broken lines).
 Simulations are based on the parameter values from the 10 'best' fits.

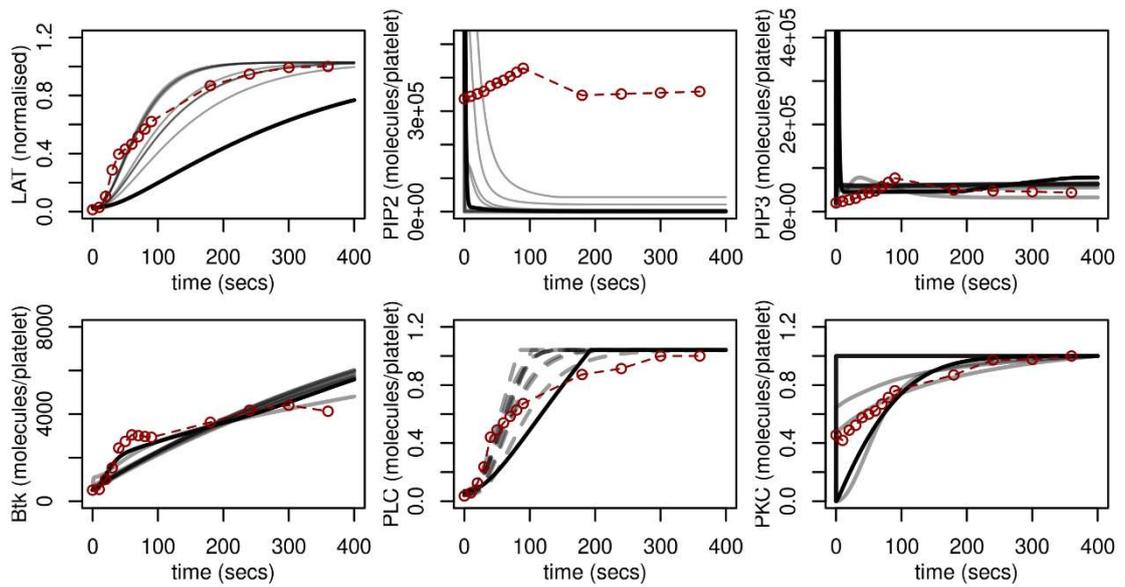


Figure 8 Model bld simulations (solid lines) compared to experimental results (broken lines). Simulations are based on the parameter values from the 10 'best' fits.

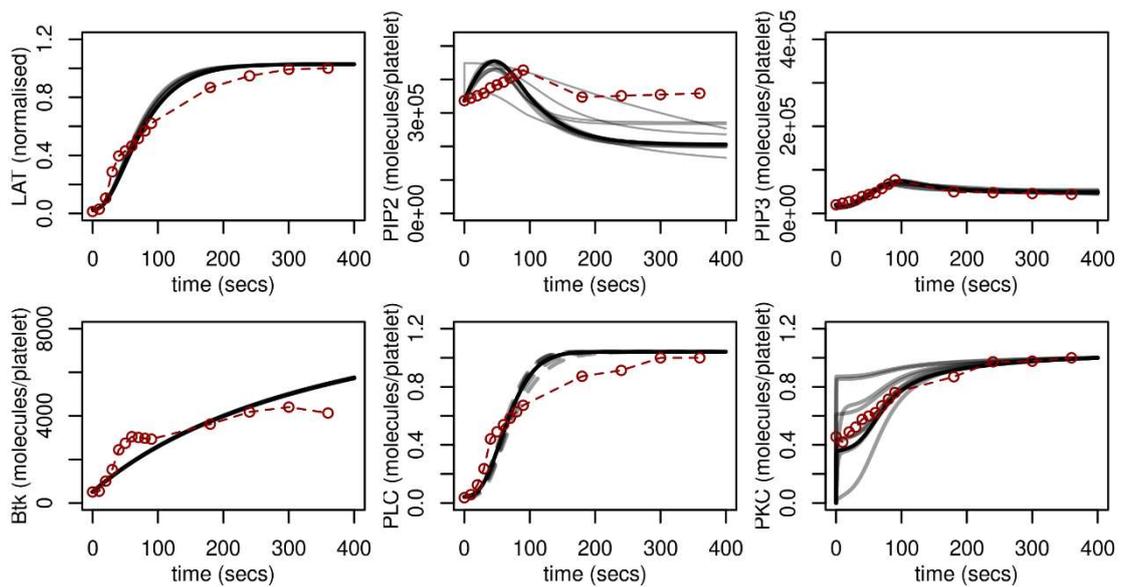


Figure 9 Model blad simulations (solid lines) compared to experimental results (broken lines). Simulations are based on the parameter values from the 10 'best' fits.

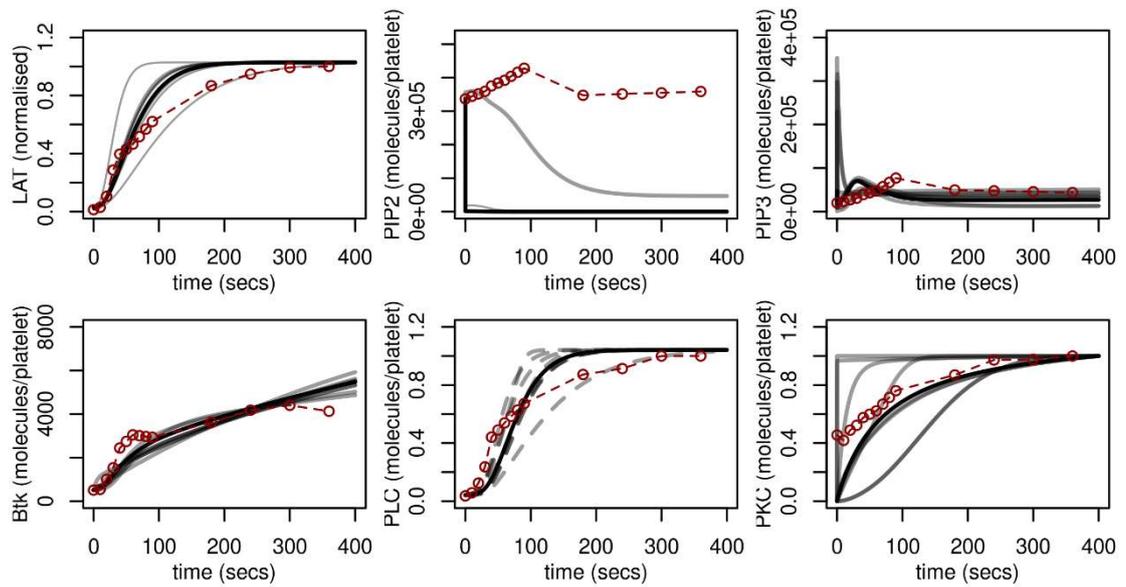


Figure 10 Model bladi simulations (solid lines) compared to experimental results (broken lines). Simulations are based on the parameter values from the 10 'best' fits.

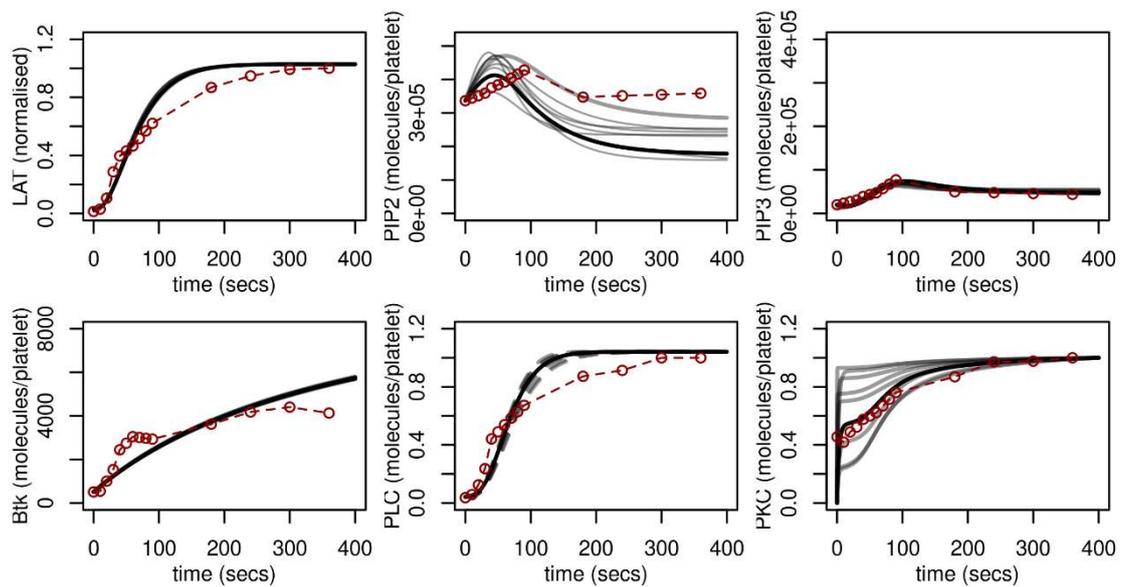


Figure 11 Model bladi simulations (solid lines) compared to experimental results (broken lines). Simulations are based on the parameter values from the 10 'best' fits.

Parameter	b	ba	bd	bad	bl	bla	bld	blad
r ₁	0.0125***	0.015***	0.013***	0.016***	0.009***	0.012***	0.362*	0.012***
r ₂	0.432	0.365*	0.370*	0.340*	0.310*	0.374*	0.557	0.411
r ₋₂	0.246**	0.220**	0.177**	0.226**	0.179**	0.291*	0.317*	0.224**
r ₃	0.551	0.429	0.545	0.476	0.317*	0.241**	0.484	0.301*
r ₋₃	0.150**	0.214**	0.097**	0.121**	0.141**	0.109**	0.381*	0.139**
r ₄	0.757	0.836	0.310*	0.303*	0.491	0.531	0.314*	0.344*
r ₋₄	0.397*	0.382*	0.400	0.366*	0.394*	0.377*	0.271*	0.371*
r ₅	0.001***	0.158**	0.001***	0.158**	0.001***	0.164**	0.516	0.175**
r ₆	0.395*	0.429	0.340*	0.418	0.373*	0.400	0.457	0.444
r ₇	0.129**	0.225**	0.211**	0.240**	0.404	0.450	0.772	0.345*
r ₋₇	0.248**	0.210**	0.210**	0.210**	0.166**	0.143**	0.213**	0.171**
r ₈	0.183**	0.218**	0.103**	0.191**	0.153**	0.181**	0.373*	0.239**
r ₉	0.395*	0.466	0.329*	0.473	0.359*	0.448	0.441	0.378*
r ₁₀	0.303*	0.311*	0.213**	0.216**	0.238**	0.240**	0.482	0.361*
r ₋₁₀	0.324*	0.339*	0.353*	0.449	0.283*	0.391*	0.404	0.377*
r ₁₁	0.465	0.447	0.342*	0.350*	0.450	0.456	0.515	0.470
r ₋₁₁	0.403	0.472	0.311*	0.383*	0.389*	0.463	0.508	0.340*
r ₁₂	0.404	0.386*	0.337*	0.402	0.293*	0.367*	0.258*	0.354*
r _e	-	0.003***	-	0.002***	-	0.001***	-	0.001***

Table 3 90% range of posterior (Q90). Scaled between 0 and 1 of prior. ***, ** and * indicates restriction in range $Q90 < 0.02$, $Q90 < 0.25$ and $Q90 < 0.40$ respectively.

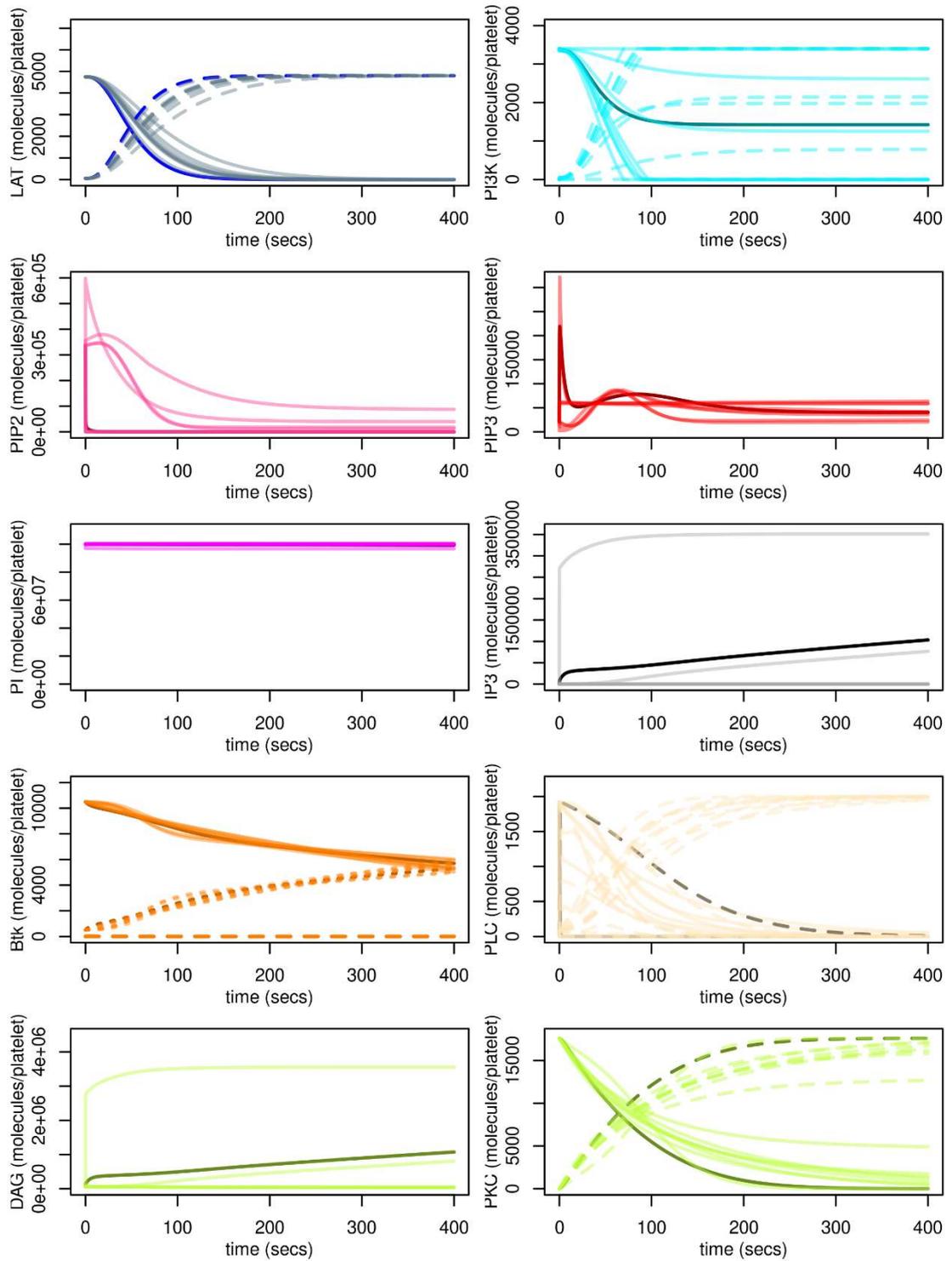


Figure 12 Model b simulations for all model variables (not just those that can be compared to experimental data). Simulations are based on the parameter values from the 10 'best' fits. Dark line denotes simulation with lowest cost function. Top left panel, [LAT] solid line, [LAT*] broken line. Top right, [PI3K] solid line, [LAT*:PI3K] broken line. Second from bottom (left), [Btk] solid line, [PIP3:Btk] broken, [PIP3:Btk*] small broken. Second from bottom (right), [PLC] solid line, [PIP3:PLC] broken, [PIP3:PLC*] small broken. Bottom right, [PKC] solid line, [DAG:PKC*] broken line.

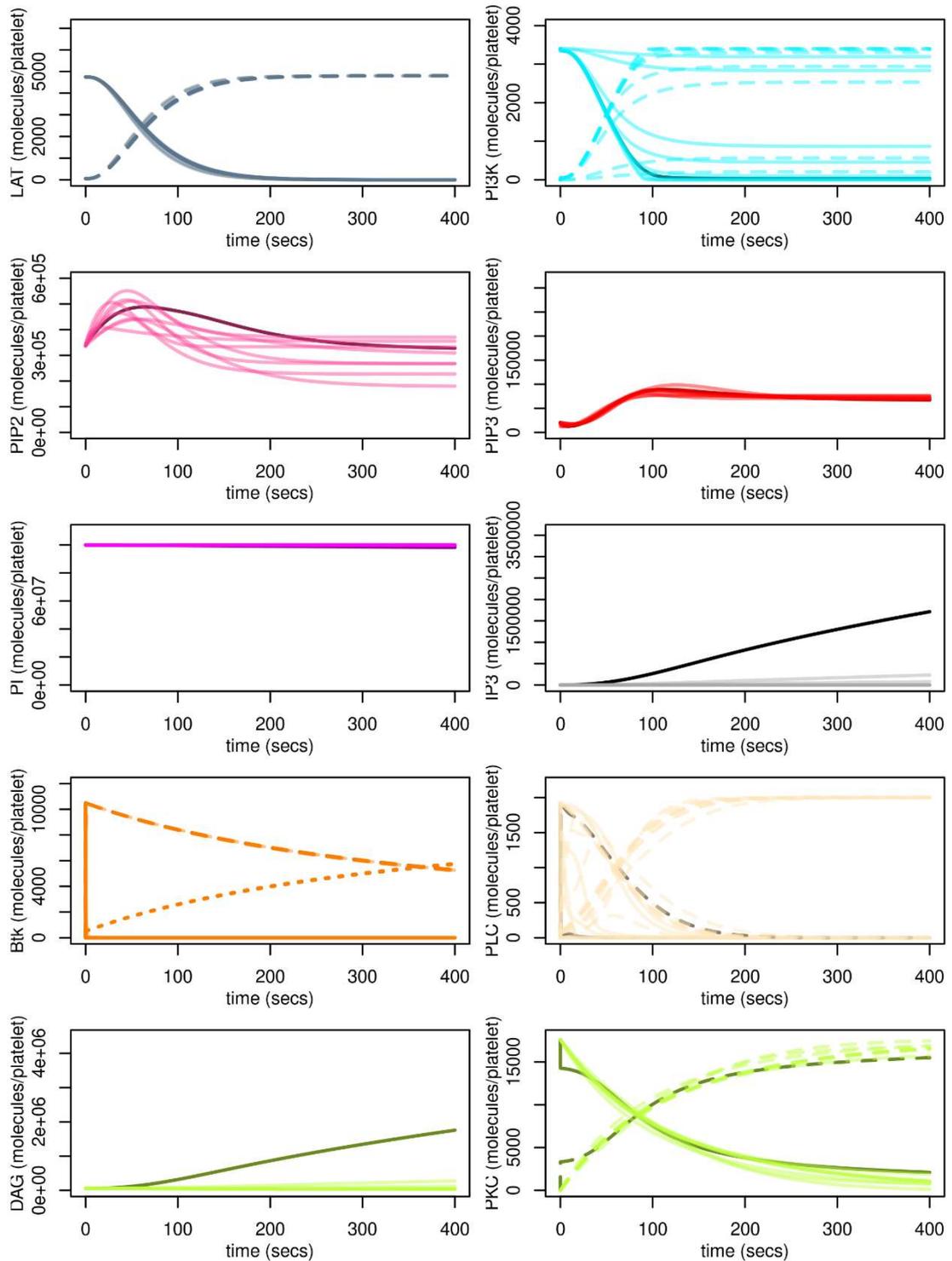


Figure 13 Model ba simulations for all model variables (not just those that can be compared to experimental data). Simulations are based on the parameter values from the 10 'best' fits. Dark line denotes simulation with lowest cost function. Top left panel, [LAT] solid line, [LAT*] broken line. Top right, [PI3K] solid line, [LAT*:PI3K] broken line. Second from bottom (left), [Btk] solid line, [PIP3:Btk] broken, [PIP3:Btk*] small broken. Second from bottom (right), [PLC] solid line, [PIP3:PLC] broken, [PIP3:PLC*] small broken. Bottom right, [PKC] solid line, [DAG:PKC*] broken line.

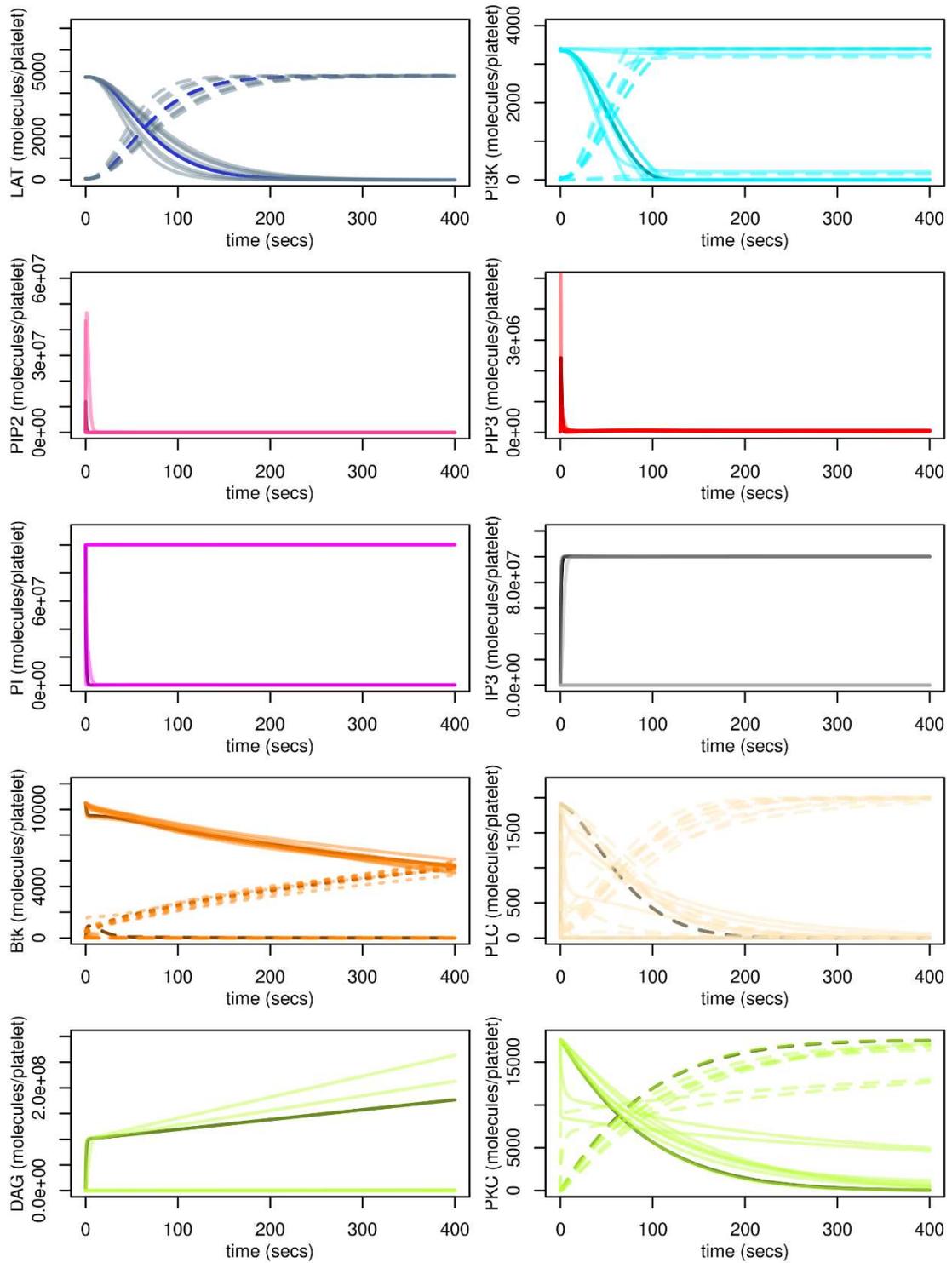


Figure 14 Model bd simulations for all model variables (not just those that can be compared to experimental data). Simulations are based on the parameter values from the 10 'best' fits. Dark line denotes simulation with lowest cost function. Top left panel, [LAT] solid line, [LAT*] broken line. Top right, [PI3K] solid line, [LAT*:PI3K] broken line. Second from bottom (left), [Btk] solid line, [PIP3:Btk] broken, [PIP3:Btk*] small broken. Second from bottom (right), [PLC] solid line, [PIP3:PLC] broken, [PIP3:PLC*] small broken. Bottom right, [PKC] solid line, [DAG:PKC*] broken line.

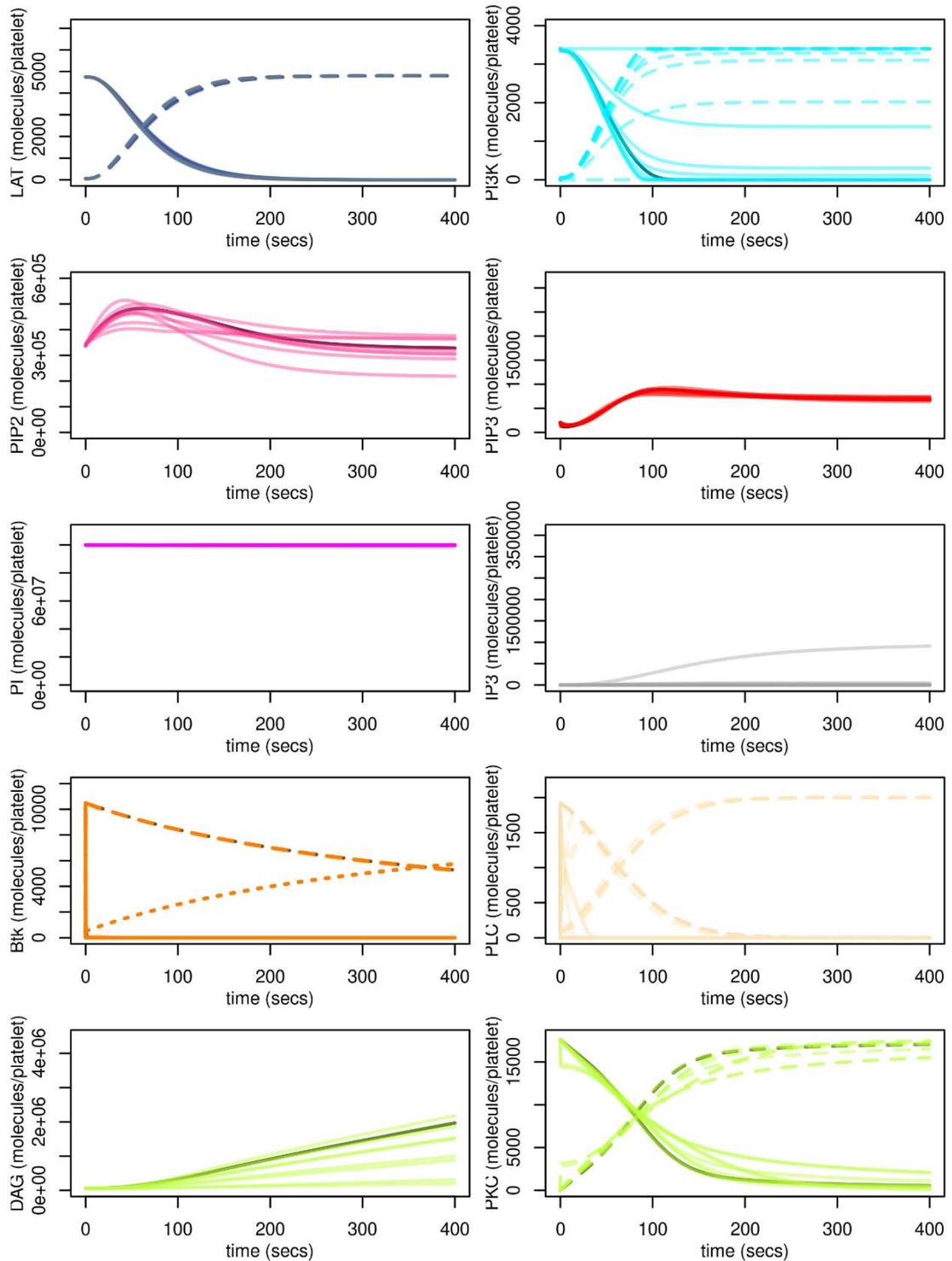


Figure 15 Model bad simulations for all model variables (not just those that can be compared to experimental data). Simulations are based on the parameter values from the 10 'best' fits. Dark line denotes simulation with lowest cost function. Top left panel, [LAT] solid line, [LAT*] broken line. Top right, [PI3K] solid line, [LAT*:PI3K] broken line. Second from bottom (left), [Btk] solid line, [PIP3:Btk] broken, [PIP3:Btk*] small broken. Second from bottom (right), [PLC] solid line, [PIP3:PLC] broken, [PIP3:PLC*] small broken. Bottom right, [PKC] solid line, [DAG:PKC*] broken line.

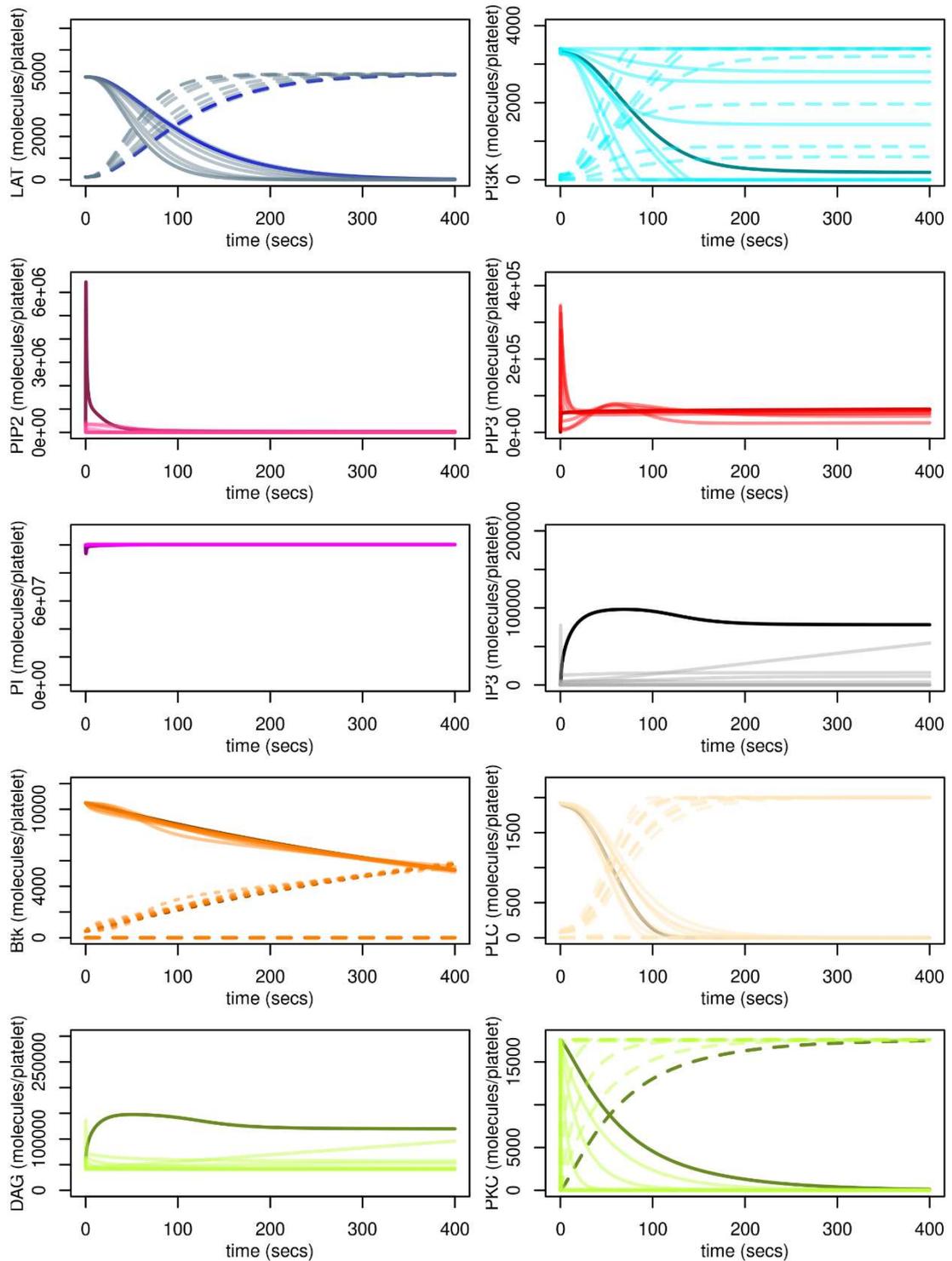


Figure 16 Model bl simulations for all model variables (not just those that can be compared to experimental data). Simulations are based on the parameter values from the 10 'best' fits. Dark line denotes simulation with lowest cost function. Top left panel, [LAT] solid line, [LAT*] broken line. Top right, [PI3K] solid line, [LAT*:PI3K] broken line. Second from bottom (left), [Btk] solid line, [PIP3:Btk] broken, [PIP3:Btk*] small broken. Second from bottom (right), [PLC] solid line, [PIP3:PLC] broken, [PIP3:PLC*] small broken. Bottom right, [PKC] solid line, [DAG:PKC*] broken line.

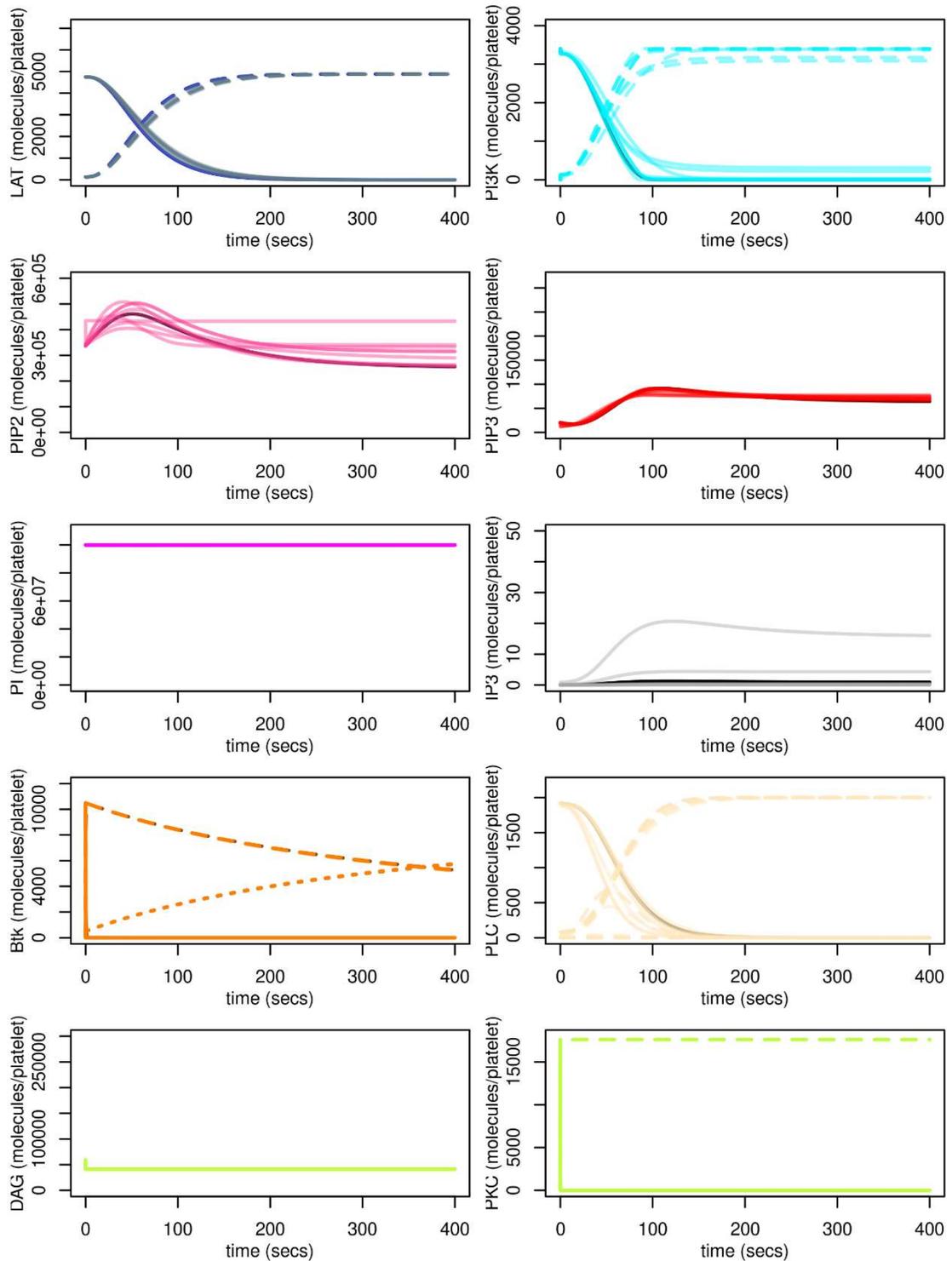


Figure 17 Model bla simulations for all model variables (not just those that can be compared to experimental data). Simulations are based on the parameter values from the 10 'best' fits. Dark line denotes simulation with lowest cost function. Top left panel, [LAT] solid line, [LAT*] broken line. Top right, [PI3K] solid line, [LAT*:PI3K] broken line. Second from bottom (left), [Btk] solid line, [PIP3:Btk] broken, [PIP3:Btk*] small broken. Second from bottom (right), [PLC] solid line, [PIP3:PLC] broken, [PIP3:PLC*] small broken. Bottom right, [PKC] solid line, [DAG:PKC*] broken line.

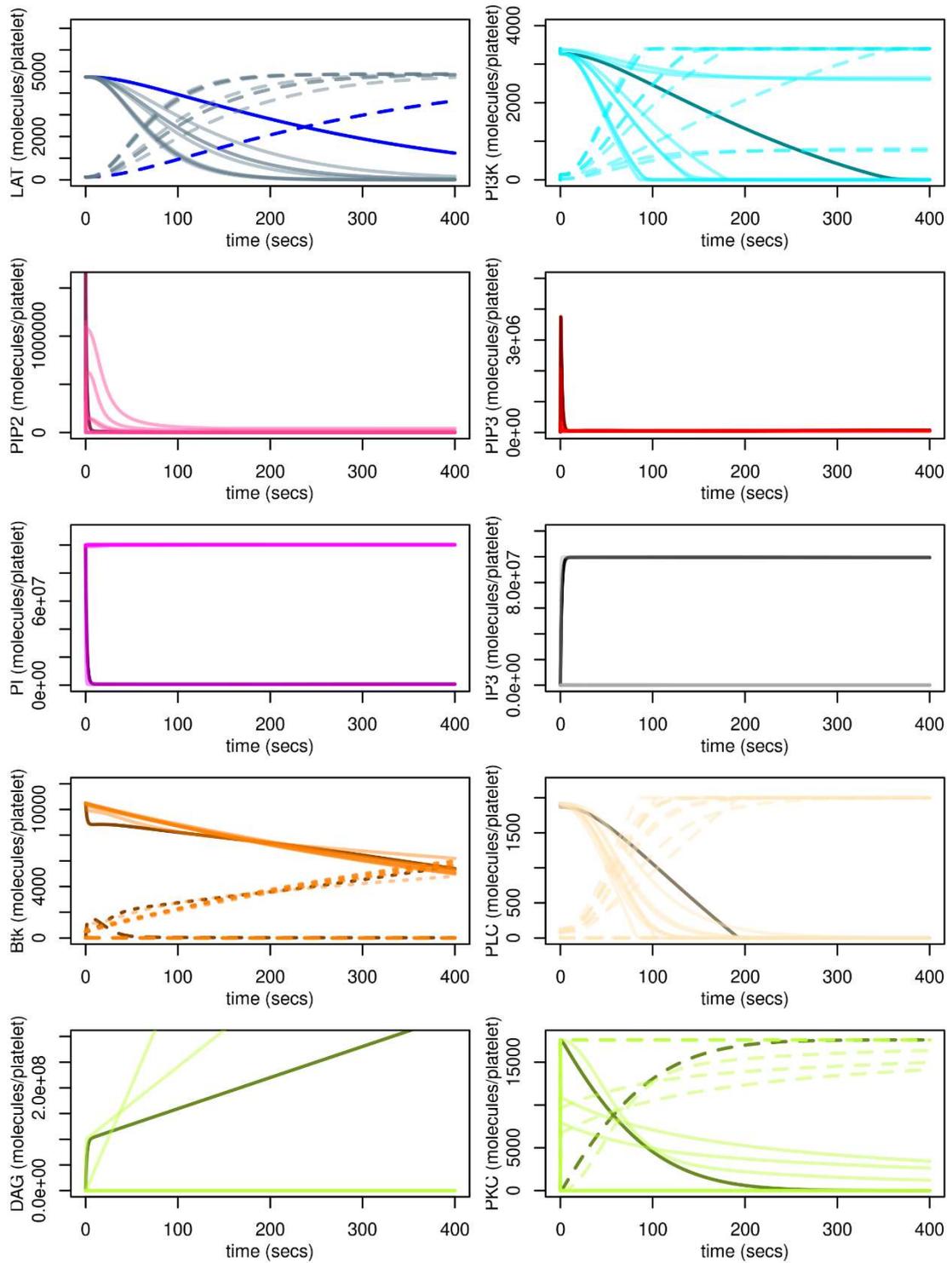


Figure 18 Model bld simulations for all model variables (not just those that can be compared to experimental data). Simulations are based on the parameter values from the 10 'best' fits. Dark line denotes simulation with lowest cost function. Top left panel, [LAT] solid line, [LAT*] broken line. Top right, [PI3K] solid line, [LAT*:PI3K] broken line. Second from bottom (left), [Btk] solid line, [PIP3:Btk] broken, [PIP3:Btk*] small broken. Second from bottom (right), [PLC] solid line, [PIP3:PLC] broken, [PIP3:PLC*] small broken. Bottom right, [PKC] solid line, [DAG:PKC*] broken line.

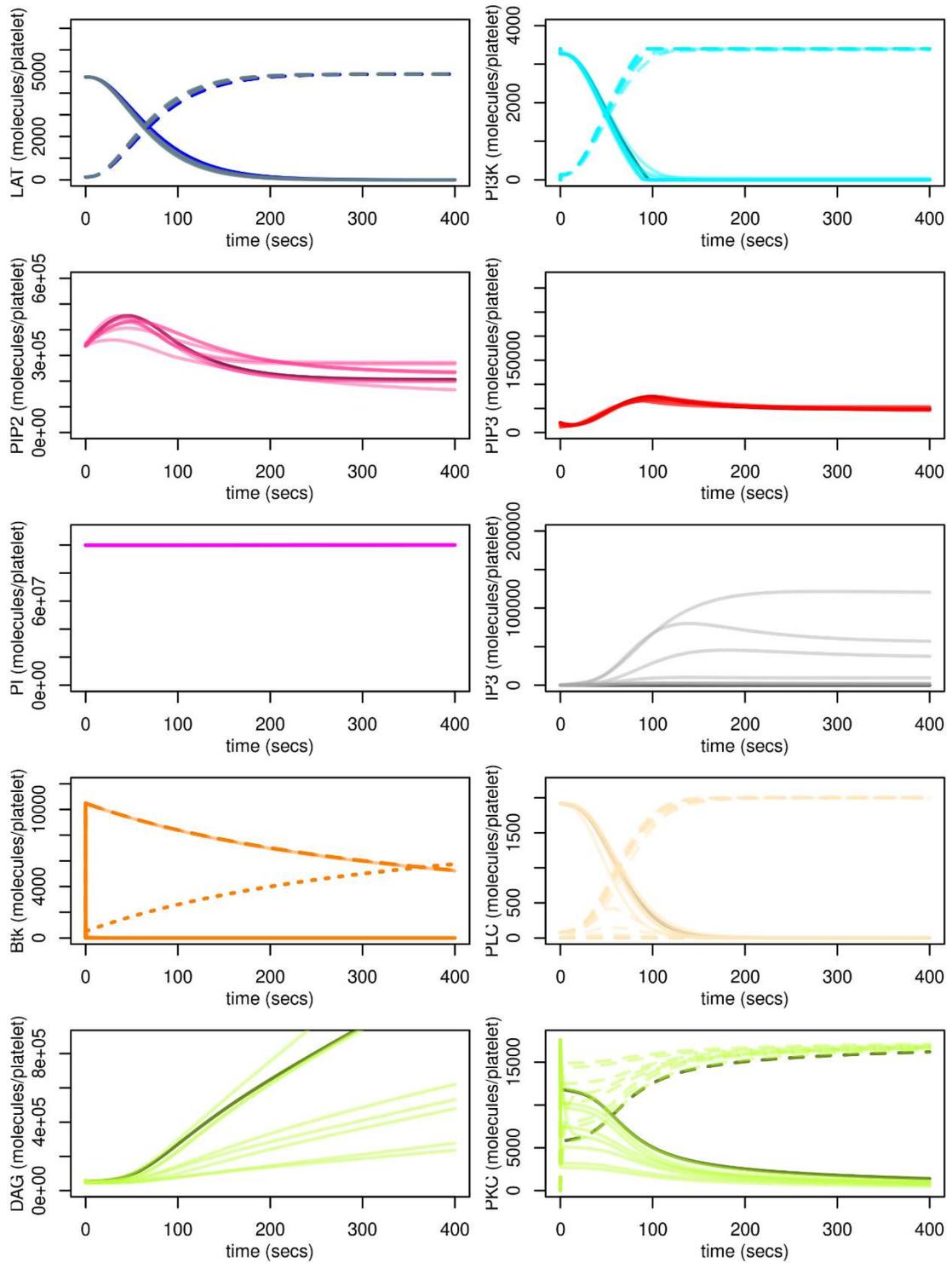


Figure 19 Model blad simulations for all model variables (not just those that can be compared to experimental data). Simulations are based on the parameter values from the 10 'best' fits. Dark line denotes simulation with lowest cost function. Top left panel, [LAT] solid line, [LAT*] broken line. Top right, [PI3K] solid line, [LAT*:PI3K] broken line. Second from bottom (left), [Btk] solid line, [PIP3:Btk] broken, [PIP3:Btk*] small broken. Second from bottom (right), [PLC] solid line, [PIP3:PLC] broken, [PIP3:PLC*] small broken. Bottom right, [PKC] solid line, [DAG:PKC*] broken line.

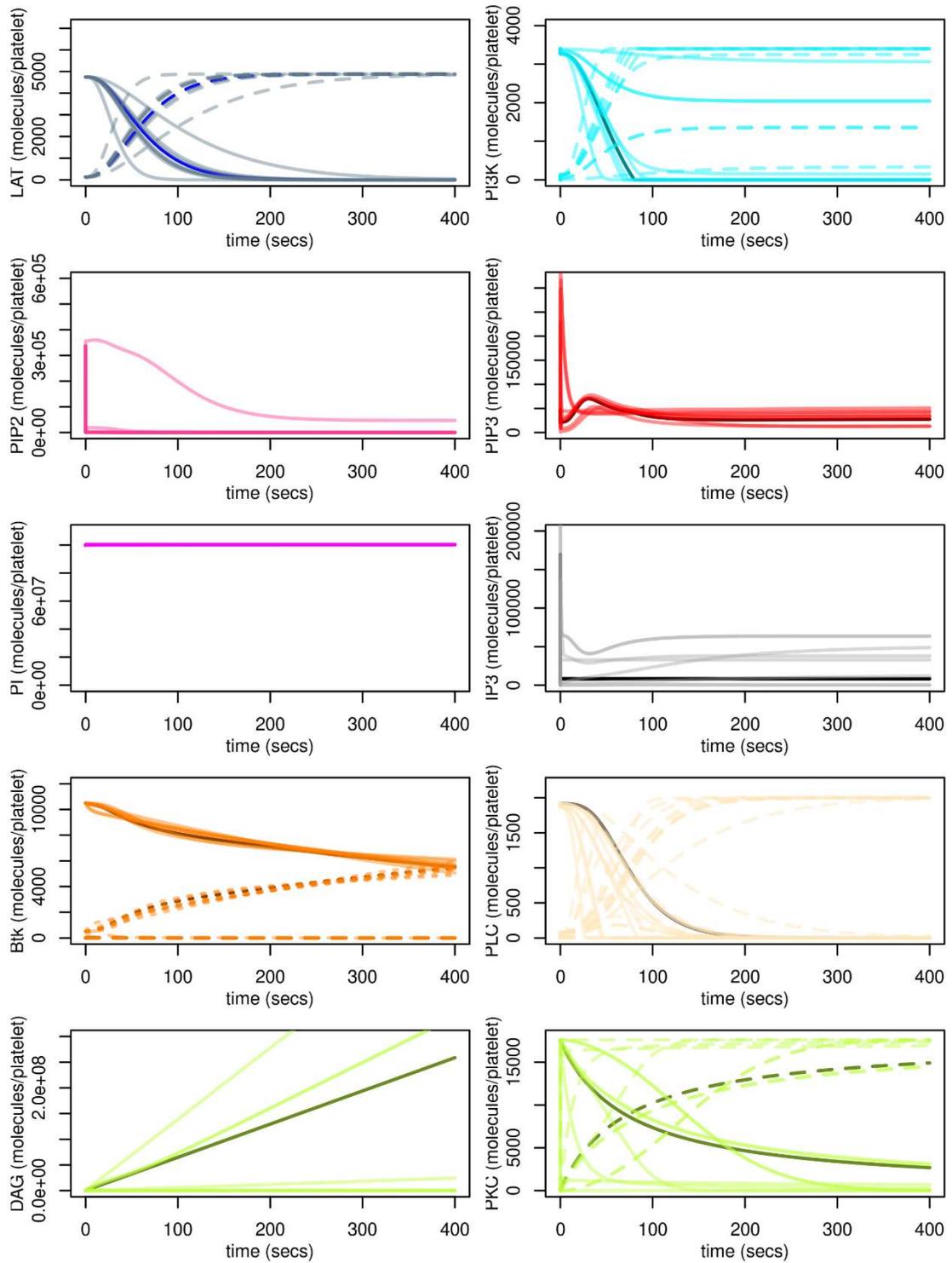


Figure 20 Model bldi simulations for all model variables (not just those that can be compared to experimental data). Simulations are based on the parameter values from the 10 'best' fits. Dark line denotes simulation with lowest cost function. Top left panel, [LAT] solid line, [LAT*] broken line. Top right, [PI3K] solid line, [LAT*:PI3K] broken line. Second from bottom (left), [Btk] solid line, [PIP3:Btk] broken, [PIP3:Btk*] small broken. Second from bottom (right), [PLC] solid line, [PIP3:PLC] broken, [PIP3:PLC*] small broken. Bottom right, [PKC] solid line, [DAG:PKC*] broken line.

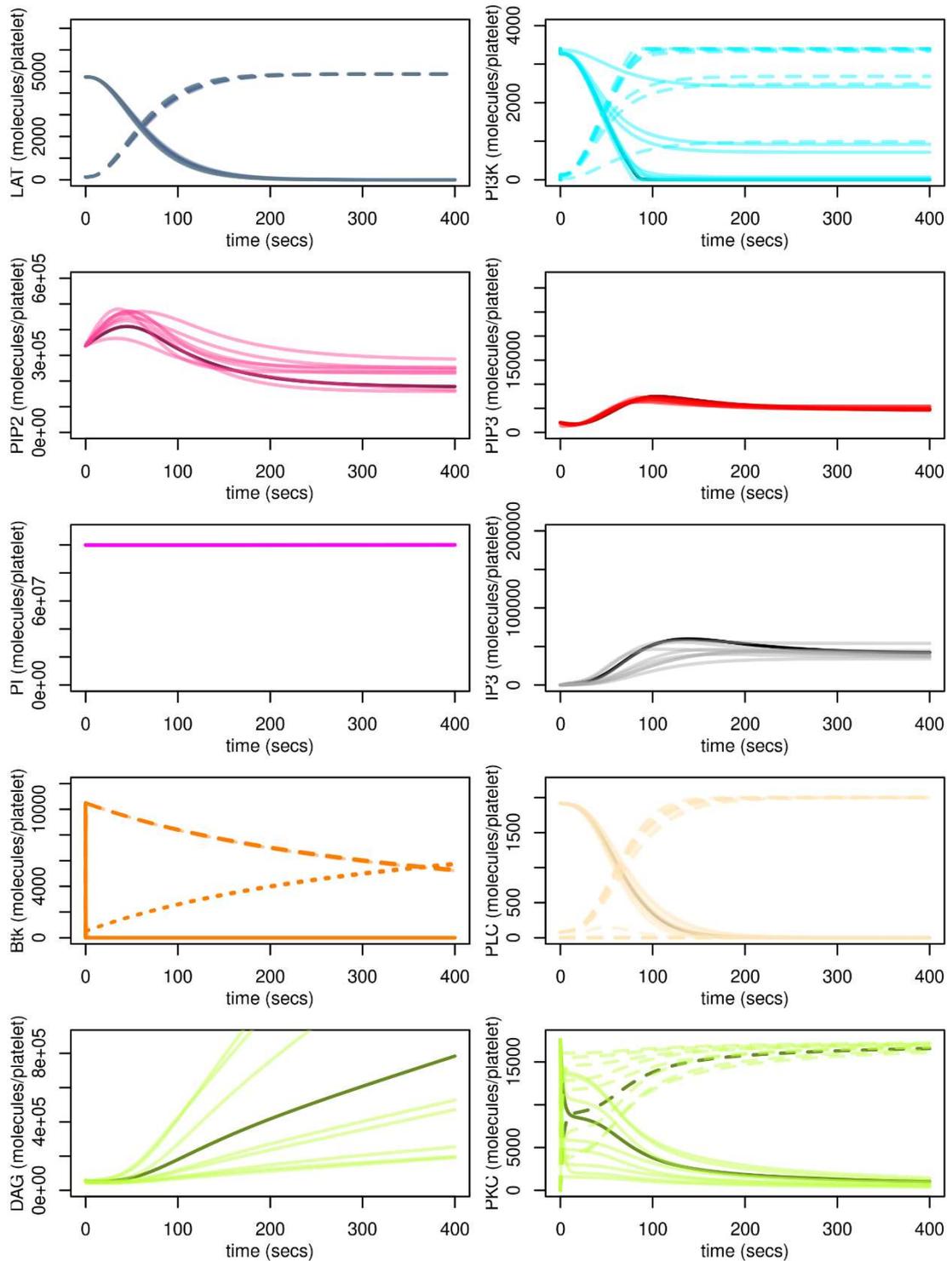


Figure 21 Model bladi simulations for all model variables (not just those that can be compared to experimental data). Simulations are based on the parameter values from the 10 'best' fits. Dark line denotes simulation with lowest cost function. Top left panel, [LAT] solid line, [LAT*] broken line. Top right, [PI3K] solid line, [LAT*:PI3K] broken line. Second from bottom (left), [Btk] solid line, [PIP3:Btk] broken, [PIP3:Btk*] small broken. Second from bottom (right), [PLC] solid line, [PIP3:PLC] broken, [PIP3:PLC*] small broken. Bottom right, [PKC] solid line, [DAG:PKC*] broken line.

Parameter	bldi	bladi
r ₁	0.395*	0.009***
r ₂	0.423	0.379*
r ₋₂	0.187**	0.324*
r ₃	0.466	0.288*
r ₋₃	0.267*	0.110**
r ₄	0.232**	0.286*
r ₋₄	0.648	0.412
r ₅	0.471	0.159**
r ₆	0.448	0.441
r ₇	0.602	0.340*
r ₋₇	0.165**	0.054**
r ₈	0.368*	0.176**
r ₉	0.476	0.455
r ₁₀	0.507	0.380*
r ₋₁₀	0.386*	0.366*
r ₁₁	0.556	0.481
r ₋₁₁	0.54	0.319*
r ₁₂	0.202**	0.292*
r _e	-	0.001***

Table 4 Variation of posterior.

References

1. Dunster, J. L., Mazet, F., Fry, M. J., Gibbins, J. M. & Tindall, M. J. Regulation of early steps of GPVI signal transduction by phosphatases: A systems biology approach. *PLoS Comput. Biol.* **11**, 1–26 (2015).
2. Burkhardt, J. M., Vaudel, M., Gambaryan, S., Radau, S., Walter, U., Martens, L., *et al.* The first comprehensive and quantitative analysis of human platelet protein composition allows the comparative analysis of structural and functional pathways. *Blood* **120**, 73–82 (2012).
3. Nasuhoglu, C., Feng, S., Mao, J., Yamamoto, M., Yin, H. L., Earnest, S., *et al.* Nonradioactive

- analysis of phosphatidylinositides and other anionic phospholipids by anion-exchange high-performance liquid chromatography with suppressed conductivity detection. *Anal. Biochem.* **301**, 243–254 (2002).
4. Wenk, M. R., Lucast, L., Di Paolo, G., Romanelli, A. J., Suchy, S. F., Nussbaum, R. L., *et al.* Phosphoinositide profiling in complex lipid mixtures using electrospray ionization mass spectrometry. *Nat. Biotechnol.* **21**, 813–817 (2003).
 5. Balla, T., Baukal, A. J., Guillemette, G. & Catt, K. J. Multiple pathways of inositol polyphosphate metabolism in angiotensin-stimulated adrenal glomerulosa cells. *J. Biol. Chem.* **263**, 4083–4091 (1988).
 6. Rittenhouse, S. E. & Sasson, J. P. Mass changes in myoinositol triphosphate in human platelets stimulated by thrombin. Inhibitory effects of phorbol ester. *J. Biol. Chem.* **260**, 8657–8660 (1985).

Chapter 4 (to be submitted)

Development of a mathematical model of platelet phosphoinositide metabolism

The model introduced in the previous chapter simplified the complexity of the phosphoinositide (PI) cycle due to the lack of availability of PI metabolism data. In this chapter, the potential interplay between GPVI downstream signalling, PI and their associated kinase/phosphatases was investigated. Here, we present a mathematical model of PI metabolism in human platelets in response to GPVI activation. The model was constructed and calibrated against experimental data capturing transient time-course changes in tyrosine kinase phosphorylation, phosphoinositide positional isomers, IP3 levels, and calcium mobilisation.

Development of a mathematical model of platelet

phosphoinositide metabolism

Hilaire Y. F. Cheung^{1,2,3}, Chukiat Tantiwong^{1,4}, Albert Sickmann², Steve P. Watson³, Johan W. M. Heemskerk^{1,5}, Jonathan M. Gibbins⁴, Joanne L. Dunster⁴

¹ Department of Biochemistry, Maastricht University, Maastricht, Netherlands

² Leibniz-Institut für Analytische Wissenschaften-ISAS-e.V, Dortmund, Germany

³ Institute of Cardiovascular Sciences, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK

⁴ School of Biological Sciences, University of Reading, Reading, UK

⁵ Synapse Research Institute Maastricht, Maastricht, Netherlands

Abstract

Current antiplatelet therapies come with a risk of bleeding, so treatments targeting new pathways that preserve haemostasis are needed. Two potential inter-related candidates being investigated in this thesis are collagen and fibrin(ogen) receptor GPVI, which has a major role in thrombosis but a minor one in haemostasis, and phosphoinositides and their associated kinase/phosphatases, which are involved in Ca^{2+} mobilisation and regulation of pleckstrin homology (PH) domains-containing proteins. The key to efficient GPVI and phosphoinositides targeting is better understanding of the interconnections between the involved enzymes and lipids. Here, we describe a mathematical model of the phosphoinositide metabolism in human platelets in response to activation of glycoprotein VI (GPVI). The model was constructed and calibrated against experimental data, covering transient time-course changes in tyrosine kinase phosphorylation, phosphoinositide positional isomers, inositol trisphosphate (InsP_3)

level and Ca^{2+} mobilisation. The developed model was able to simulate the phosphoinositide metabolism upon GPVI activation, and predict the effect of phosphatidylinositol 4-kinase A (PI4KA) inhibitor GSK-A1 on inositol triphosphate (InsP_3) levels and Ca^{2+} mobilisation, demonstrating its function in the $\text{PtdIns}(4,5)\text{P}_2$ resynthesis to sustain downstream signalling.

Introduction

The phosphoinositide metabolism is deeply integrated with major signalling pathways such as G protein-coupled receptors (GPCRs) or tyrosine kinase receptors in all cell types, including the glycoprotein VI (GPVI) signalling cascade and Ca^{2+} mobilisation. Given the importance of phosphoinositides, several mathematical models have tried to incorporate these lipids in signalling pathways, though often with a limited outcome. Some mathematical models captured the involvement of a lamellipodium and cell polarisation, which include the molecules $\text{PtdIns}(4,5)\text{P}_2$, $\text{PtdIns}(3,4,5)\text{P}_3$, phosphoinositide 3-kinase (PI3K) and the protein phosphatase and tensin homolog (PTEN), and then explored the role in cellular shape change and chemotaxis.^{1,2} Olivença *et al.* developed a detailed mathematical model that captured the complete phosphoinositide pathway, accounting for all species of phospho-inositides and the interconverting enzymes.³ The authors validated their model by predicting the effects of small interfering ribonucleic acid (siRNA) species for the knockdown of phosphoinositide kinases and phosphatases on the steady-state $\text{PtdIns}(4,5)\text{P}_2$ level in human alveolar epithelial cells. This time-dependent model comprised a system of 10 ordinary differential equations (ODEs); herein 19 kinetic parameters were calibrated to steady-state data, but neglecting the transient nature of the pathway.

Other recent studies attempt to describe mathematical elements of phosphoinositide metabolism in platelets. Diamond *et al.* developed a homogeneous mathematical model that

captures GPCR signalling and the phosphatidylinositol cycle (PI cycle), which can predict the platelet dose-dependent response of Ca^{2+} mobilisation and inositol trisphosphate (InsP_3) production.⁴ However, the author utilised copy numbers of 11 enzymes and parameter values for model simulations that were taken from data from cells and tissues that are unrelated to platelets, such as rat liver, human cerebrospinal fluid, plant cells and yeast cells. The approximations covered enzymes participating in the phosphoinositide metabolism, including phosphatidylinositol-4-phosphate kinase, inositol monophosphatase, inositol-1,4-bisphosphate 1-phosphatase, diacyl-glycerol kinase and cytidine diphosphate-diacylglycerol synthase. In addition, the model included a reduced representation of the phosphoinositide pathway, but neglected other phosphoinositide species such as $\text{PtdIns}(3,4,5)\text{P}_3$ and $\text{PtdIns}(3,4)\text{P}_2$.

Regarding platelets, Mazet *et al.* developed an ODE-based mathematical model that captured the phosphoinositide cycle (PI cycle).⁵ The authors took parameter values from the literature and adjusted the simulations to match the experimental data, which led to the idea that lipid- and protein-binding proteins help to regulate the levels of $\text{PtdIns}(4,5)\text{P}_2$ and InsP_3 in GPCR signalling. The model predicted changes of 10^8 molecules of inositol or PtdIns in less than 10 s, which seems to be too drastic to be physiologically relevant, as the highest enzyme turnover reported is purified catalase which can convert 2.8×10^6 molecules of hydrogen peroxide per second.^{6,7}

The protein phosphorylation cascade and phosphoinositide metabolism downstream of GPVI signalling in platelets is relatively well understood. Many of the earlier studies, however, focused on single routes and did not combine several pathways or involve kinetic analyses. This paper aims to develop a new dynamic mathematical description of the phosphoinositide metabolism in platelets, which is based on and validated against

experimentally consistent high-density data. The model is calibrated against data describing the time course of five phosphoinositide species, *i.e.* PtdIns, PtdIns4P, PtdIns(4,5)P₂, PtdIns(3,4)P₂ and PtdIns(3,4,5)P₃, and the tyrosine phosphorylations of Linker for activation of T cells (LAT) at Y²⁰⁰ and phospholipase C γ 2 (PLC γ 2) at Y¹²¹⁷, which represent the triggers of collagen-related peptide (CRP) on the phosphoinositide metabolism. The model is subsequently used to predict the effect of the phosphatidylinositol 4-kinase A (PI4KA) inhibitor GSK-A1 and the inositol polyphosphate-5-phosphatase (OCRL) phosphatase inhibitor YU142670 on phosphoinositide metabolism and platelet activation. The model is further cross-validated with inositol monophosphate (InsP₁) and Ca²⁺ measurements.

Results

Development of mathematical model on phosphoinositides metabolism downstream of GPVI signalling

Details of the model network, reactions and parameters. The mathematical model describing phosphoinositides was developed using ODEs; as such it is based on two assumptions: *(i)* all species that are located in the same compartment are evenly distributed, and *(ii)* all reactions in the model follow the law of mass action which states the rate of reaction is proportional to the concentration of the reactant and the rate constant. In the model, the reactions that comprise the phospho-inositide pathway are determined based on the copy number of the relevant enzyme in platelets,^{8,9} as shown in Figure 1. This is a simplified interpretation of the phosphoinositide pathway that neglects regulation via kinases and phosphatases and any diffusion or trafficking present in the cytosol and inner membranes at the same rate constant.

In Model 1, we integrated the phosphoinositides metabolism with CRP-induced

activities of PI3K and PLC γ 2, which respectively converts PtdIns(4,5)P₂ into PtdIns(3,4,5)P₃, and hydrolyses PtdIns(4,5)P₂ into InsP₃ in the plasma membrane (PM) compartment. We herein assume that all phosphoinositide species are in the PM. This assumption is supported by a staining study showing that PtdIns4P and PtdIns(4,5)P₂ are localised in the PM of resting and activated platelets.¹⁰ We also incorporate the phosphoinositide cycle (PI cycle) to recycle the hydrolysed PtdIns(4,5)P₂ with no dead-end species. To increase the ability of the model to be inferred from the data, the rest of the PI cycle is reduced to three components InsP₃, InsP₁ and an inositol pool (I_{pool}), which comprises inositol and other inositol phosphates, residing in the cytosol (Cyt) compartment. Similarly, the variable phosphatidylinositol pool in the PM (P_{pool}) comprises the PM contribution of PtdIns3P, PtdIns5P, and PtdIns(3,5)P₂, which molecules can not be measured by the IC-MS method due to a low abundance and overlap with the more prominent species PtdIns4P and PtdIns(4,5)P₂.

This model was modified to Model 2, in order to consider the transportation of phosphatidylinositol from Cyt to PM compartment by phosphatidyl-inositol transfer protein type α (PITP α), which was shown to contribute to thrombin-induced InsP₃ production in platelets.¹¹ The goal of this adjustment is to slow down the flux of PtdIns conversion to PtdIns4P in an effort to improve the simulations.

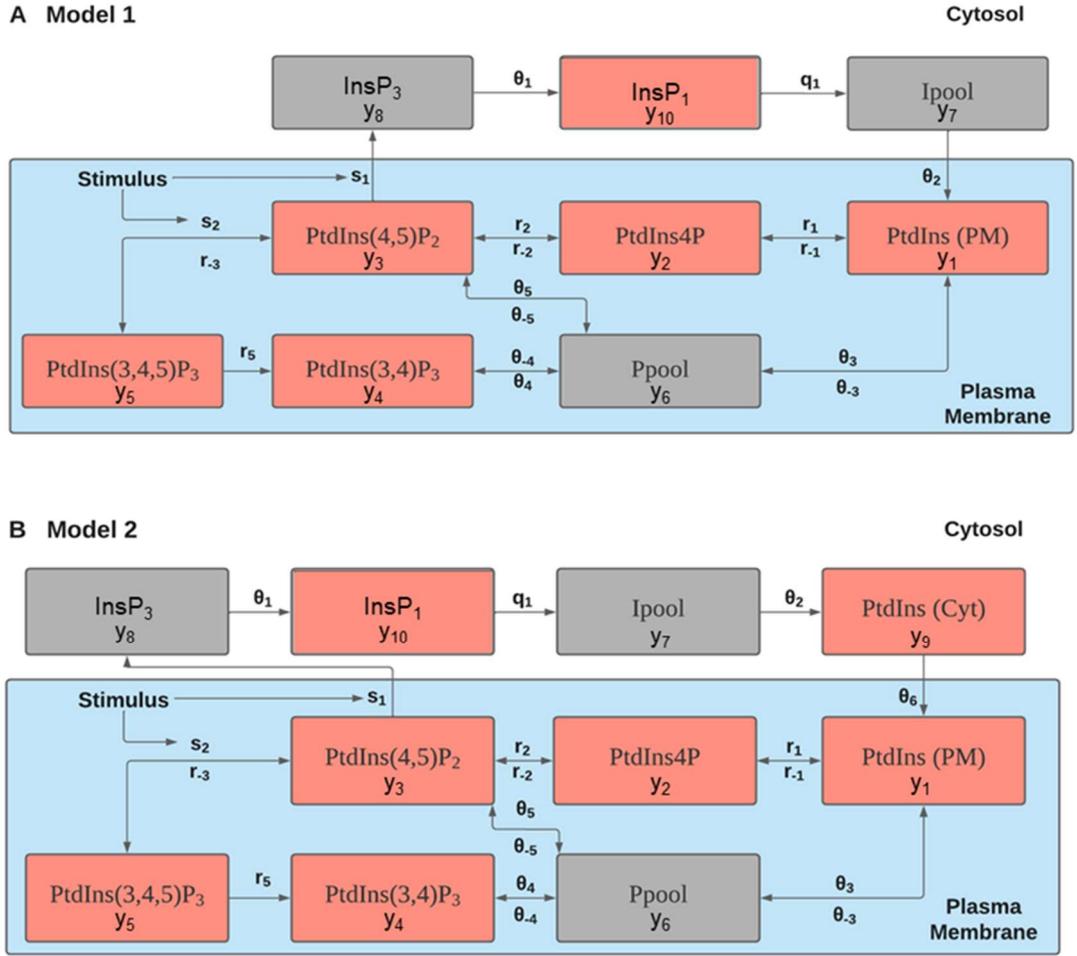


Figure 1. Network diagram of the phosphoinositides metabolism model. Network diagram of Model 1 (A) and Model 2 (B). Variables are represented by square boxes and the parameter associated with each process is placed next to the relevant arrow. See Tables 1 and 2 for a description of the variables and parameters. In Model 2, an additional species *PtdIns (Cyt)* and 1 new parameter θ_6 are added to incorporate the transportation of phosphatidylinositol from the Cyt to PM compartment.

Utilising mass-action kinetics, the network diagram in Figure 1A is translated into the following system of nine ODEs for Model 1:

$$\frac{dy_1}{dt} = \theta_2 y_7 - r_1 y_1 + r_{-1} y_2 - \theta_3 y_1 + \theta_{-3} y_6, \quad (1.1)$$

$$\frac{dy_2}{dt} = r_1 y_1 - r_{-1} y_2 - r_2 y_2 + r_{-2} y_3, \quad (1.2)$$

$$\frac{dy_3}{dt} = r_2y_2 - r_{-2}y_3 - s_1stim(t)y_3 - s_2stim(t)y_3 + r_{-3}y_5 - \theta_5y_3 + \theta_{-5}y_6, \quad (1.3)$$

$$\frac{dy_4}{dt} = r_5y_5 - \theta_4y_4 + \theta_{-4}y_6, \quad (1.4)$$

$$\frac{dy_5}{dt} = s_2stim(t)y_3 - r_{-3}y_5 - r_5y_5, \quad (1.5)$$

$$\frac{dy_6}{dt} = \theta_3y_1 - \theta_{-3}y_6 + \theta_5y_3 - \theta_{-5}y_6 + \theta_4y_4 - \theta_{-4}y_6, \quad (1.6)$$

$$\frac{dy_7}{dt} = q_1y_{10} - \theta_2y_7, \quad (1.7)$$

$$\frac{dy_8}{dt} = s_1stim(t)y_3 - \theta_1y_8, \quad (1.8)$$

$$\frac{dy_{10}}{dt} = \theta_1y_8 - q_1y_{10}. \quad (1.9)$$

The function $stim(t)$ that appears in equations 1.3, 1.4 and 1.8 captures the effect of CRP stimulus on the phosphoinositide metabolism, *i.e.* the activation of PI3K and PLC γ 2. This assumes that any changes in phosphoinositides metabolism do not affect tyrosine phosphorylation. Similarly, the network diagram in Figure 1B can be translated into the following system of ten ODEs for Model 2:

$$\frac{dy_1}{dt} = \theta_6y_9 - r_1y_1 + r_{-1}y_2 - \theta_3y_1 + \theta_{-3}y_6, \quad (2.1)$$

$$\frac{dy_2}{dt} = r_1y_1 - r_{-1}y_2 - r_2y_2 + r_{-2}y_3, \quad (2.2)$$

$$\frac{dy_3}{dt} = r_2y_2 - r_{-2}y_3 - s_1stim(t)y_3 - s_2stim(t)y_3 + r_{-3}y_5 - \theta_5y_3 + \theta_{-5}y_6, \quad (2.3)$$

$$\frac{dy_4}{dt} = r_5y_5 - \theta_4y_4 + \theta_{-4}y_6, \quad (2.4)$$

$$\frac{dy_5}{dt} = s_2stim(t)y_3 - r_{-3}y_5 - r_5y_5, \quad (2.5)$$

$$\frac{dy_6}{dt} = \theta_3y_1 - \theta_{-3}y_6 + \theta_5y_3 - \theta_{-5}y_6 + \theta_4y_4 - \theta_{-4}y_6, \quad (2.6)$$

$$\frac{dy_7}{dt} = q_1 y_{10} - \theta_2 y_7, \quad (2.7)$$

$$\frac{dy_8}{dt} = s_1 \text{stim}(t) y_3 - \theta_1 y_8, \quad (2.8)$$

$$\frac{dy_9}{dt} = \theta_2 y_7 - \theta_6 y_9, \quad (2.9)$$

$$\frac{dy_{10}}{dt} = \theta_1 y_8 - q_1 y_{10}. \quad (2.10)$$

The model's variables, their units of measure, and initial conditions are summarised in Table 1. The model parameters and the units are summarised in Table 2.

Variable	Description	Initial amount	Reference
y ₁	PtdIns (PM)	1,350,000	This study, Ref ¹²
y ₂	PtdIns4P	640,000	This study
y ₃	PtdIns(4,5)P ₂	310,000	This study
y ₄	PtdIns(3,4)P ₂	5,200	This study
y ₅	PtdIns(3,4,5)P ₃	1,900	This study
y ₆	Ppool	25000	Ref ^{13,14}
y ₇	lpool	100,000,000	Ref ¹⁵
y ₈	InsP ₃	0	This study
y ₉	PtdIns (Cyt)	1,350,000	This study, Ref ¹²
y ₁₀	InsP ₁	2500	This study

Table 1. Variables and initial conditions for the mathematical Models 1 and 2. Variable y₉, PtdIns(Cyt) is only present in Model 2. All inositol compounds, except for InsP₃, were measured in (and have units of measure in the model as) molecules per platelet. The InsP₃ data were normalised, as are the model simulations.

Parameter	Description	Units	Model 1	Model 2	Fitting constraint
r_1	Rate of conversion of PtdIns to PtdIns4P	s^{-1}	✓	✓	$10^{-4} - 10^2$
r_{-1}	Rate of conversion of PtdIns4P to PtdIns	s^{-1}	✓	✓	$10^{-4} - 10^2$
r_2	Rate of conversion of PtdIns4P to PtdIns(4,5)P ₂	s^{-1}	✓	✓	$10^{-4} - 10^2$
r_{-2}	Rate of conversion of PtdIns(4,5)P ₂ to PtdIns4P	s^{-1}	✓	✓	$10^{-4} - 10^2$
r_{-3}	Rate of conversion of PtdIns(3,4,5)P ₃ to PtdIns(4,5)P ₂	s^{-1}	✓	✓	$10^{-4} - 10^2$
r_5	Rate of conversion of PtdIns(3,4,5)P ₃ to PtdIns(3,4)P ₂	s^{-1}	✓	✓	$10^{-4} - 10^2$
θ_1	Rate of conversion of InsP ₃ to InsP ₁	s^{-1}	✓	✓	$10^{-2} - 10^{-1}$
θ_2	Rate of conversion of Ipool to PtdIns (PM)	s^{-1}	✓	N/A	$10^{-4} - 10^2$
θ_2	Rate of conversion of Ipool to PtdIns (Cyt)	s^{-1}	N/A	✓	$10^{-4} - 10^2$
θ_3	Rate of conversion of PtdIns to Ppool	s^{-1}	✓	✓	$10^{-4} - 10^2$
θ_{-3}	Rate of conversion of Ppool to PtdIns	s^{-1}	✓	✓	$10^{-4} - 10^2$
θ_4	Rate of conversion of PtdIns(3,4)P ₂ to Ppool	s^{-1}	✓	✓	$10^{-4} - 10^2$
θ_{-4}	Rate of conversion of Ppool to PtdIns(3,4)P ₂	s^{-1}	✓	✓	$10^{-4} - 10^2$
θ_5	Rate of conversion of PtdIns(4,5)P ₂ to Ppool	s^{-1}	✓	✓	$10^{-4} - 10^2$
θ_{-5}	Rate of conversion of Ppool to PtdIns(4,5)P ₂	s^{-1}	✓	✓	$10^{-4} - 10^2$
θ_6	Rate of conversion of PtdIns (Cyt) to PtdIns (PM)	s^{-1}	N/A	✓	$10^{-8} - 10^2$
s_1	Rate of hydrolysis of PtdIns(4,5)P ₂ into InsP ₃	s^{-1}	✓	✓	$10^{-6} - 10^{-2}$
s_2	Rate of conversion of PtdIns(4,5)P ₂ into PtdIns(3,4,5)P ₃	s^{-1}	✓	✓	$10^{-4} - 10^2$
q_1	Rate of conversion of InsP ₁ to Ipool	s^{-1}	✓	✓	$10^{-4} - 10^{-3}$

Table 2. Model parameters and fitting constraints in the phosphoinositide metabolism model. Abbreviation: N/A, not applicable in the specified model. For parameter values of the 10 best fits of the model, see the supplemental.

Time course profiling of protein phosphorylation and Ca²⁺ mobilisation in CRP-stimulated platelets

Experiments were conducted to determine the time-course phosphorylation of LAT at Y²⁰⁰ and PLC γ 2 at Y¹²¹⁷ downstream of GPVI activation for the determination of the function $stim(t)$, that represents stimulus of phosphoinositide metabolism in response to platelet activation with CRP. The phosphorylation of LAT was chosen as a proxy for PI3K, as this enzyme recruits and activates PI3K.¹⁶ Washed platelets at 4×10^8 cells were simulated with 10 μ g/mL CRP in the presence of 9 μ M eptifibatide, and were lysed at the stated time after CRP. The cell lysates were probed against phospho-specific antibodies to determine the extent of tyrosine

phosphorylation. The addition of eptifibatide prevented the interference of integrin $\alpha\text{IIb}\beta\text{3}$ outside-in signalling, which acts through the kinases Src and Syk,¹⁷ and hence prevented platelet aggregation under the stirring conditions.

As shown in Figure 2B, platelet stimulation with CRP induced a rapid increase in phosphorylation for LAT at Y²⁰⁰, and PLC γ 2 at Y¹²¹⁷, both of which plateaued at 45 s and sustained for up to 50 min. The *stim(t)* function was determined by fitting to the phosphorylation data:

$$stim(t) = (0.001te^{-0.0002t^2}) + \tanh(0.02t),$$

with *t* representing the time following CRP stimulation. The obtained *stim* function (dotted line) closely matched the LAT Y²⁰⁰ and PLC γ 2 Y¹²¹⁷ phosphorylation profiles.

Experiments were also conducted to measure Ca²⁺ mobilisation in CRP-stimulated platelets. Herein, Fura-2-loaded washed platelets were pre-treated for 10 min with the ADP scavenger apyrase and the cyclooxygenase-1 inhibitor indomethacin to prevent the interference of secondary mediators in signalling. The platelets were again stimulated with in the presence of CaCl₂, to provide physiological extracellular Ca²⁺, which allowed extracellular Ca²⁺ entry. As shown in Figure 2C, similar to the phosphorylation of LAT at Y²⁰⁰ and PLC γ 2 at Y¹²¹⁷, the platelet Ca²⁺ mobilisation plateaued at 120 s (rising from 35 ± 13 to 233 ± 68 nM) and slightly decreased to 283 ± 70 nM over 10 min.

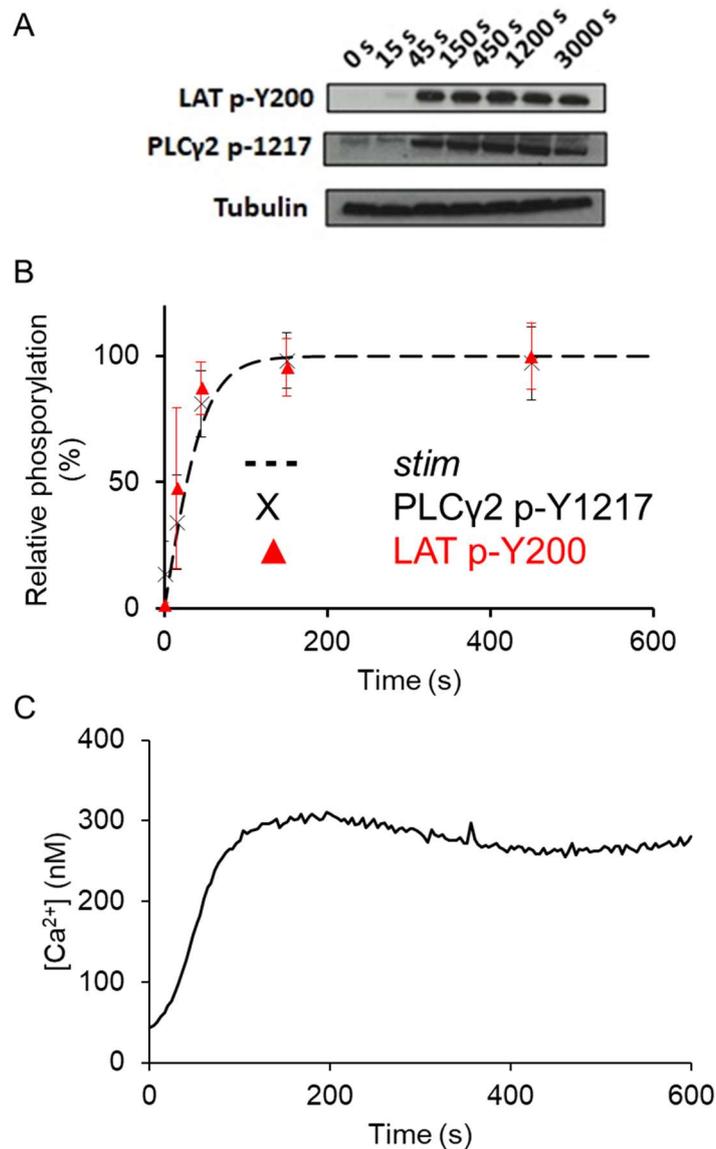


Figure 2. CRP-induced sustained tyrosine phosphorylation and Ca²⁺ mobilisation and the determination of stim function. (A) Washed platelets were stimulated with 10 μg/mL CRP in the presence of 9 μM eptifibatide. Platelets were lysed with 5x reducing lysis buffer at the stated time after addition of CRP. Representative phosphorylation blots from whole-cell lysates after probing with the stated antibodies. (B) Plot of relative phosphorylation for LAT at Y²⁰⁰, PLCγ2 at Y¹²¹⁷, and the extrapolated stim function (mean ± SD, n = 3). (C) Fura 2-loaded platelets at 2 x 10⁸ cells/mL were pre-treated for 10 min with apyrase (2.5 U/mL) and indomethacin (20 μM). Platelets were then stimulated with 10 μg/mL CRP in the presence of 1 mM CaCl₂. Representative traces of changes in [Ca²⁺] over the 600 s were recorded.

Time course profiling of phosphoinositides in CRP-stimulated platelets

Experiments were then conducted to profile the time-courses of phosphorylation of tyrosine kinases downstream of GPVI activation. Washed platelets again were pretreated with apyrase and indomethacin, and subsequently stimulated with CRP; at certain time points the stimulation was stopped with icecold HCl.¹⁸ In this experiment higher concentrations of both platelets and CRP were used (Figure 3), when compared to the other assays to improve the detection of low abundance phosphoinositides, especially PtdIns(3,4,5)P₃.

As shown in Figure 3, PtdIns showed a slight initial increase (rising from 2.3 ± 0.5 to $2.9 \pm 0.6 \times 10^5$ molecules/platelet over the first 60 s) before declining to below base level ($1.5 \pm 0.2 \times 10^5$ molecules/platelet). In contrast, the amount of PtdIns(4,5)P₂ gradually increased 1.6-fold over the first 120 s (from 2.9 ± 0.6 to $4.8 \pm 0.4 \times 10^6$ molecules/platelet), and remained elevated over basal for up to 30 min. PtdIns(3,4)P₂ increased by 6-fold (from 0.5 ± 0.2 to $3.0 \pm 0.8 \times 10^4$ molecules/platelet) over the first 180 s and this elevated level was sustained for 10 min before dropping to $1.7 \pm 0.1 \times 10^4$ molecules/platelet after 30 min. For PtdIns(3,4,5)P₃, a 2.6-fold increase was observed over the first 180 s, from 2.2 ± 0.8 to $5.8 \pm 2.8 \times 10^4$ molecules/platelet. The large error bar can be attributed to donor variability and/or the low abundance of PtdIns(3,4,5)P₃ which increase the impact of background noise and lower the accuracy of measurement. No significant change was observed for PtdIns4P, and its level remained near the basal level for 30 min.

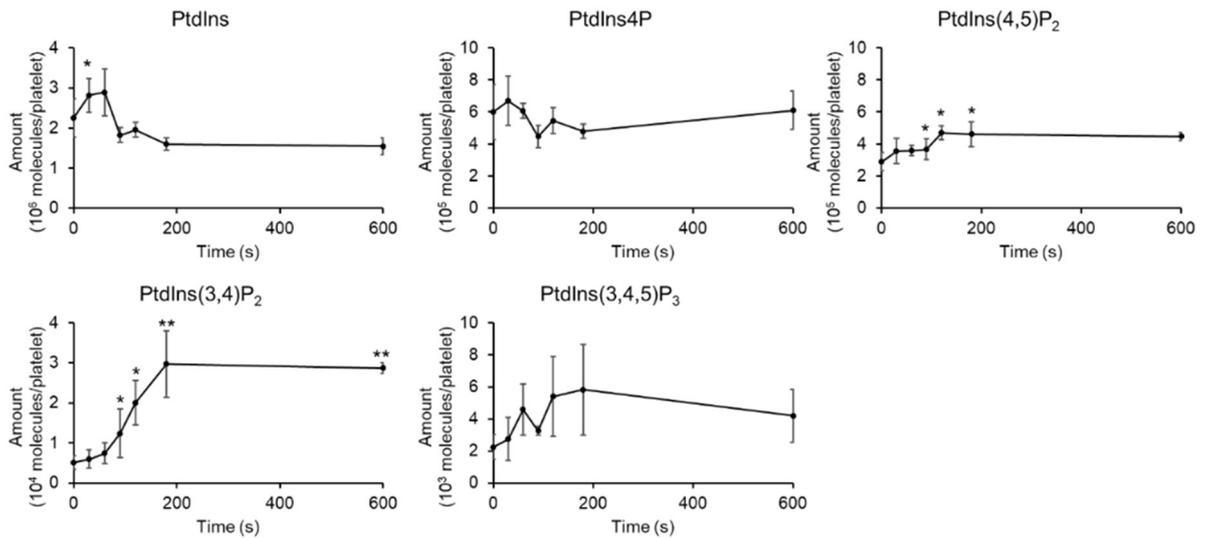


Figure 3. CRP-induced changes in phosphoinositides profile in human platelet. Washed platelets at 1.2×10^9 cells were stimulated with $30 \mu\text{g}/\text{mL}$ CRP in the presence of apyrase and indomethacin. Stimulation was stopped at the stated time with icecold 1 M HCl. Phosphoinositides in the samples were analysed and quantified using IC-MS. Results are expressed in molecules/platelet and are means \pm SD from 3 experiments. $*P < 0.05$ and $**P < 0.01$, one-way ANOVA followed by Tukey's test, compared to $t = 0 \text{ s}$.

Comparison of model simulations with experimental profile

Figure 4A shows that both models are able to fit the data equally well, with similar minimum (both at 0.01) and median SSE/n (0.014 for Model 1 and 0.012 for Model 2). In Figure 4B, the posterior distributions of the parameter values show that this dataset cannot constrain the parameter values and that the approximated parameters in Model 2 are more widespread compared to Model 1. The 10 best simulations (*i.e.* parameter values with the lowest SSE) are compared to the experimental data as shown in Figure 5. In both models, we observed that the top 10 simulations (black lines) fit well with the experimental profile (red lines). In addition, most best-fit lines had similar trajectories and were able to converge to their steady state over 600 s, except those of I_{pool} and P_{pool} , which converge to different levels, and InsP1 which is increasing function.

In Model 1, the best-fitting simulations predicted a rapid depletion of I_{pool} from 10^8 molecules to 10^4 molecules in the first 200 s. These molecules are subsequently converted to P_{pool} , which rapidly increases from 2×10^4 molecules to 10^8 molecules in the first 200 s through PtdIns as an intermediate. In contrast, Model 2 predicted a much less drastic change in I_{pool} and P_{pool} . The number of molecules in I_{pool} stayed constant over 600 s of simulations, while the amount of P_{pool} slightly increased to around 10^5 molecules. The observed variations in the model simulations are due to the lack of experimental data available for calibration. Nevertheless, the simulations of Model 1 seem too drastic to be physiologically possible within 200 s, while the predictions of Model 2 fit better to our understanding of phosphoinositide metabolism in platelets. The prediction in Model 2 of the P_{pool} increase is also comparable to a study by Valet *et al.*, which showed a 3-fold increase in PtdIns3P after 3 min of CRP stimulation,¹⁹ assuming that PtdIns5P and PtdIns(3,5)P₂ remain unchanged.

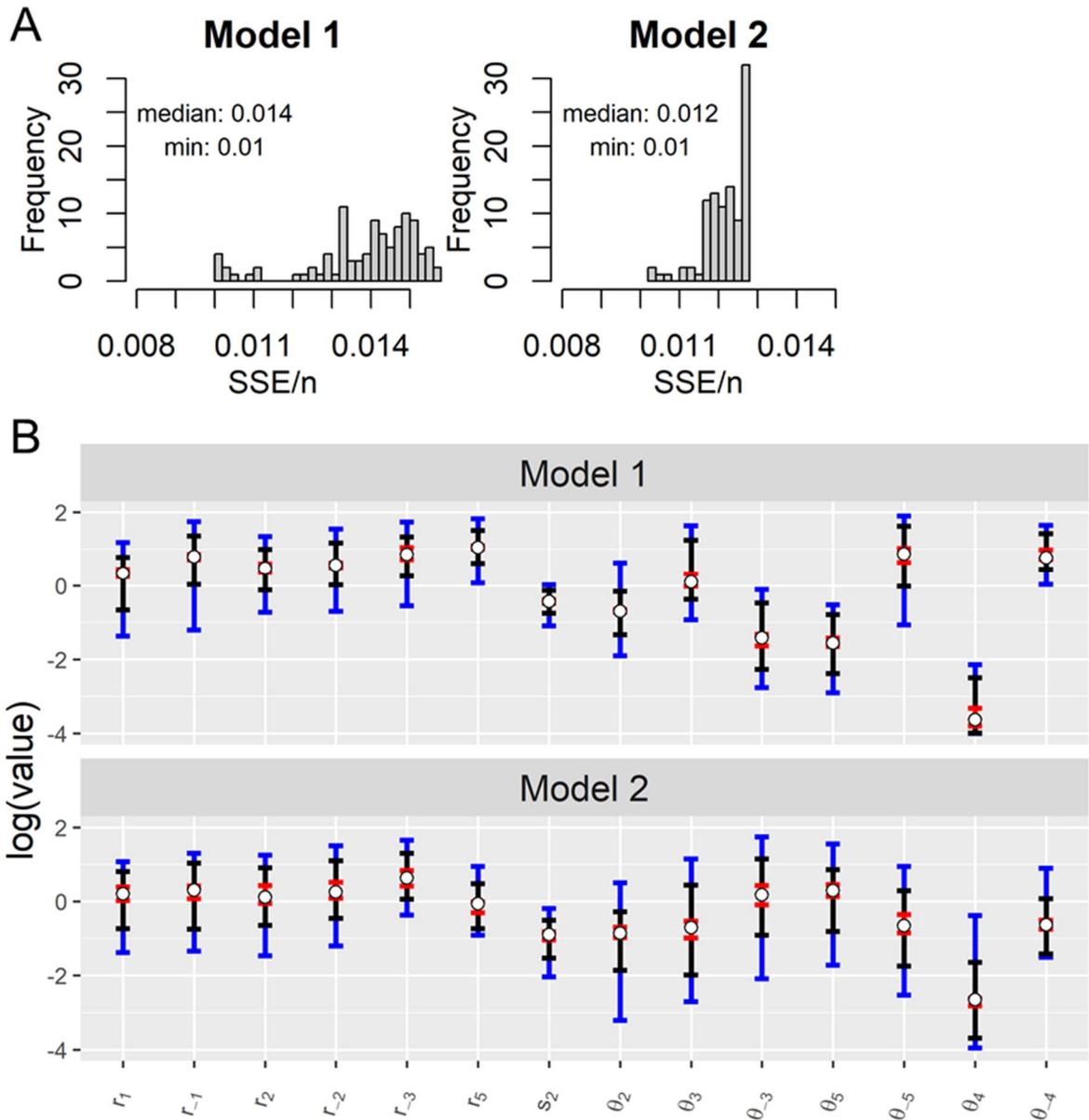


Figure 4. Assessing parameter fitting with cost function and range of parameters approximations. (A) Histograms showing the frequency of cost function SSE/n for the best 100 simulations for Models 1 (left) and 2 (right). Median and minimum SSE/n are also listed. (B) Posterior parameter ranges are shown for Models 1 (top) and 2 (bottom). The blue, black, and red lines and hollow circles show the full range, interquartile range, the inner 10% range and median of each estimated parameter. Posteriors for s_1 , θ_1 , θ_6 and q_1 are constrained to other ranges and are shown in the supplemental.

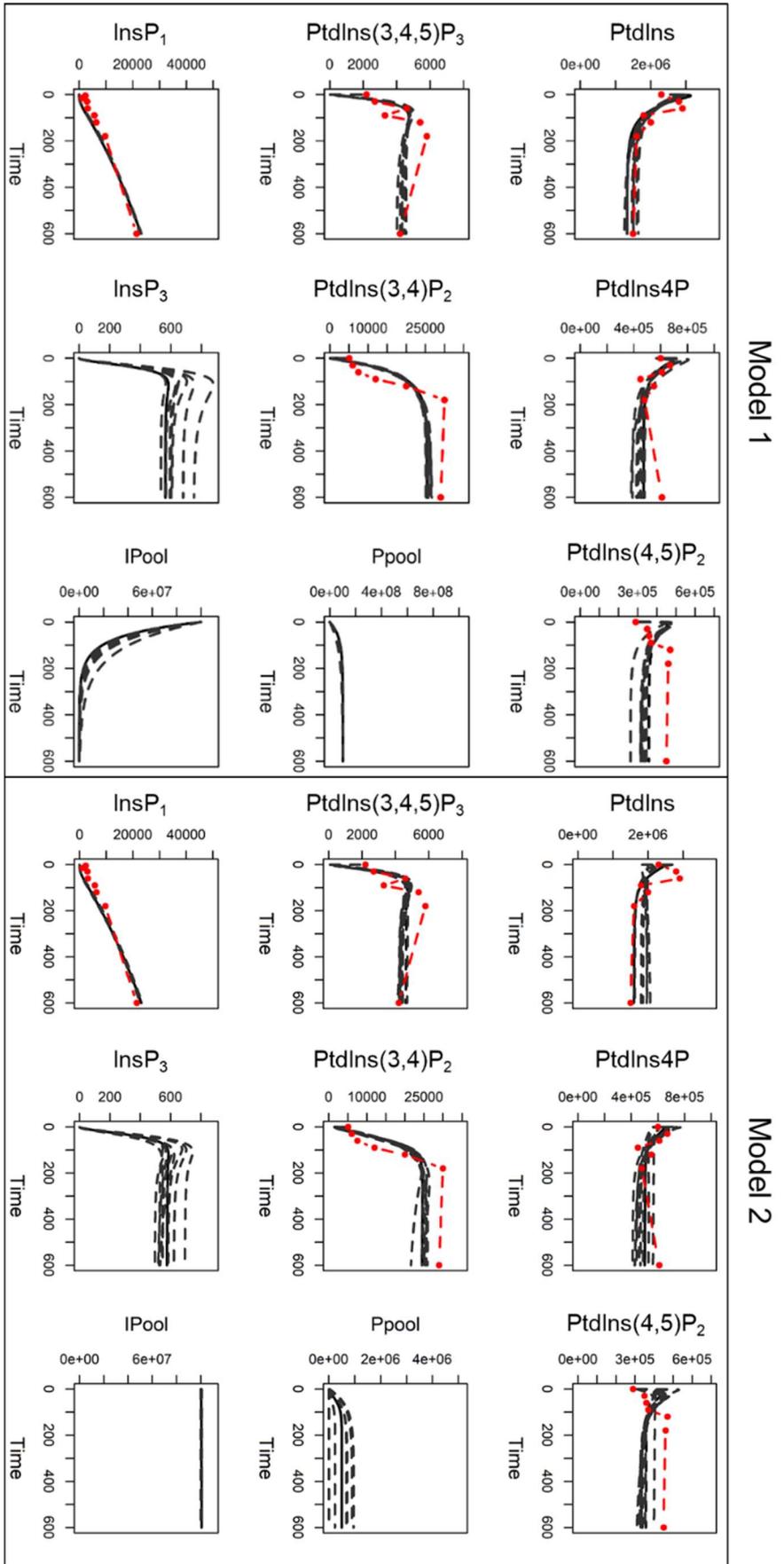


Figure 5. Comparison of model simulation profile with experimental data. The best-fitting model profiles of Model 1 (left) and Model 2 (right). The 10 best simulations (black and grey dotted lines) were compared to experimental observations in Figure 3 (red dotted lines).

Model prediction of the effect of phosphoinositides kinase and phosphatase inhibitors

The mathematical models capture the relationship between phosphoinositides in agonist-stimulated platelets. To test the models under alternative conditions, functional studies were undertaken using the PI4KA inhibitor GSK-A1²⁰ and the OCRL inhibitor YU142670²¹ on platelet activation. The results were compared with the model predictions. The use of these inhibitors was based on the hypothesis that they affect the resynthesis of PtdIns(4,5)P₂ or the dephosphorylation of PtdIns(4,5)P₂, respectively, thereby lowering the CRP-induced InsP₃ production and Ca²⁺ mobilisation. To predict the effects of the phosphatase inhibitor YU122670, the model was simulated using the original parameters listed in Table 2 except for lowering the rate of conversion of PtdIns(4,5)P₂ to PtdIns4P r_{-2} to 10% of the original value, and lowering the rate of conversion of PtdIns to PtdIns4P r_1 to 10% of the original value for GSK-A1.

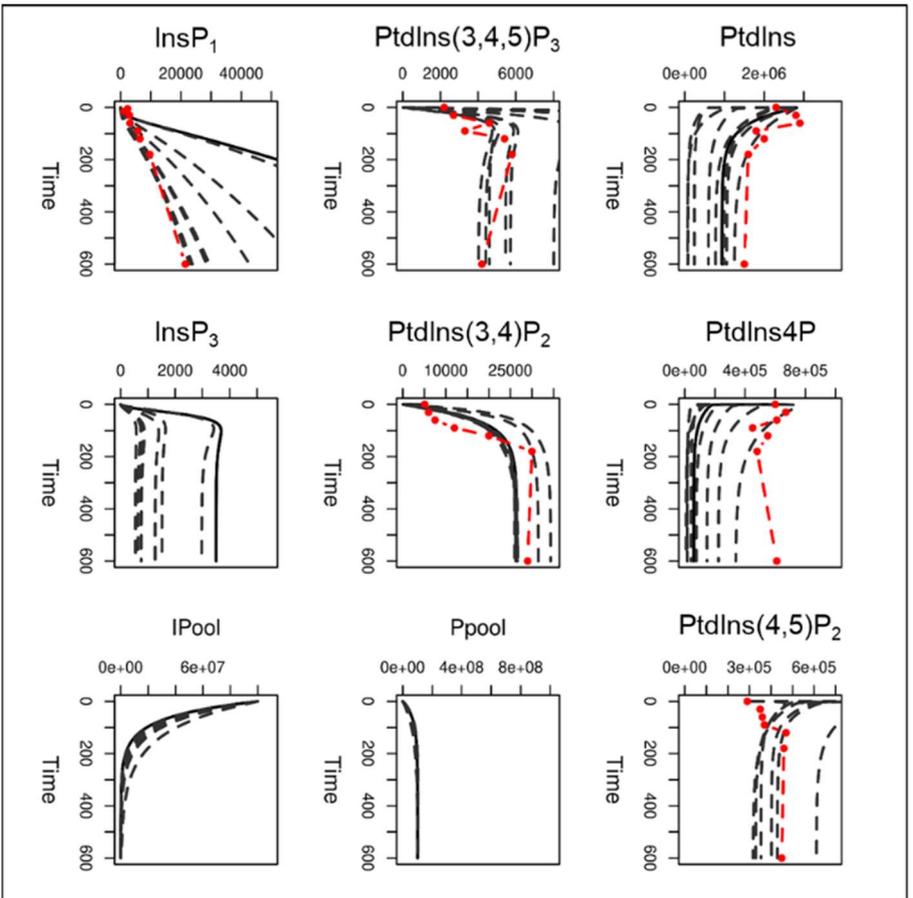
For the simulation of YU142670 treatment (Figure 6), the two Models 1 and 2 have similar profile trends and predicted a lower steady-state level of PtdIns4P than the original model because of the inhibition of PtdIns(4,5)P₂ phosphatase which reduces the amount of PtdIns4P formed by PtdIns(4,5)P₂ dephosphorylation. At the same time, the predicted levels of PtdIns(4,5)P₂, PtdIns(3,4)P₂ and PtdIns(3,4,5)P₃, InsP₃ and InsP₁ are higher compared to the original model, due to the lack of removal of PtdIns(4,5)P₂ through dephosphorylation, and those fluxes are instead transferred to other species that are produced from PtdIns(4,5)P₂. The effect of r_{-2} on Ipool and Ppool is minimal and similar to that observed in Figure 5. The simulation is consistent with Bura *et al.* who reported an increase of intracellular PtdIns(4,5)P₂ in activated platelet compared to the unstimulated,¹⁰ visualised by anti-PtdIns4P, anti-PtdIns(4,5)P₂ antibodies labelled platelets.

For the simulation of treatment with the PI4KA inhibitor GSK-A1 (Figure 7), both

models predicted that the steady-state level of PtdIns to be higher than the original model due to the inhibition of PI4K which reduces PtdIns removal through phosphorylation in both models. Consequently, the resynthesis of PtdIns(4,5)P₂ is reduced, and the predicted steady-state levels of PtdIns(4,5)P₂, PtdIns(3,4)P₂ and PtdIns(3,4,5)P₃ are 50% lower than the original model. Model 1 predicted a rapid drop in PtdIns4P and PtdIns(4,5)P₂ within 10 s of agonist stimulation followed by slow resynthesis of these species. This indicates the huge metabolic flux of r_1 (conversion of PtdIns to PtdIns4P) and r_2 (conversion of PtdIns4P to PtdIns(4,5)P₂) in Model 1, and with inhibition of r_1 these two species are rapidly depleted with slow replenishment.

Model 2 predicted that the level of PtdIns(4,5)P₂ PtdIns(3,4)P₂ PtdIns(3,4,5)P₃ and InsP₃ match the profile for the first 100 s, showing that inhibition of PI4K mostly has a less pronounced effect immediately affect agonist stimulation. Model 2 also predicted that GSK-A1 led to an increase in P_{pool} , possibly because of increased production of PtdIns3P and PtdIns5P due to more PtdIns being present. This increase is not observed in Model 1. The simulation is consistent with Bura *et al.*, who reported a decrease of both PtdIns(4,5)P₂ and PtdIns4P levels in activated platelets compared to the control.¹⁰

Model 1



Model 2

Figure 6. Simulation of the effect of YU142670 on phosphoinositides metabolism. The model is simulated by adjusting r_2 to 10% of its original value, while keeping all other parameters the same as in previous simulations. The ten best-fitting model profiles (black dotted lines) were compared to the original experimental data (red dotted lines).

Model 1

Model 2

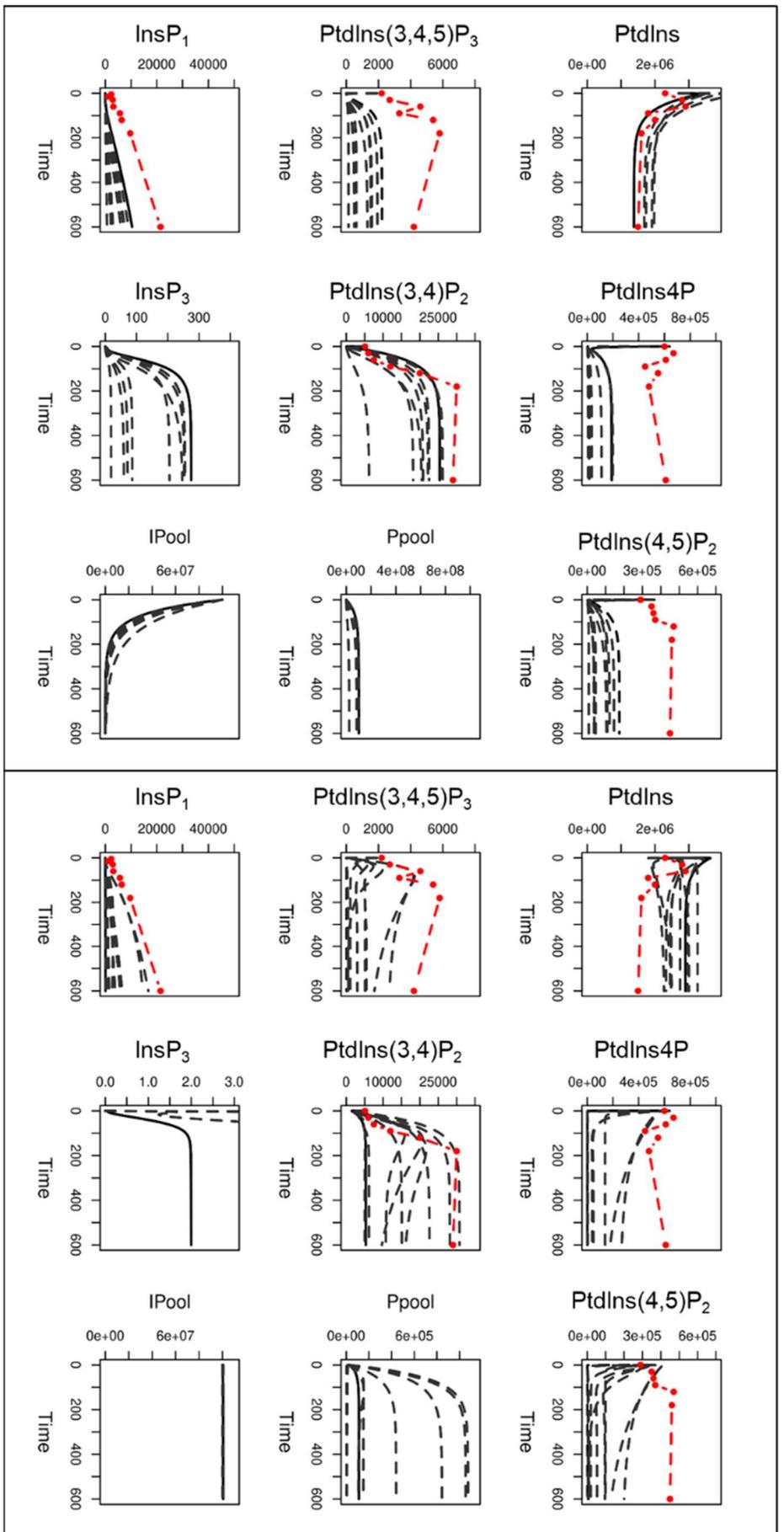


Figure 7. Simulation of the effect of GSK-A1 on phosphoinositides metabolism. The model simulated by adjusting r_1 to 10% of its original value, while keeping all other parameters the same as in previous simulations. The ten best-fitting model profiles (black dotted lines) were compared to the original experimental data (red dotted lines).

Effect of phosphoinositide kinase/phosphatase inhibitors on Ca²⁺ mobilisation

To validate the predictions of the models, functional experiments were conducted to investigate the effect of GSK-A1 and YU142770 on the Ca²⁺ mobilisation. For this purpose, Fura-2 loaded platelets were pre-treated for 10 min with GSK-A1 or YU142670 at the listed concentration, again in the presence of apyrase and indomethacin. The platelets were then stimulated with CRP in the presence of CaCl₂.

As shown in Figure 8A-B, 1 μM YU142670 has no effect on Ca²⁺ mobilisation in platelet. YU142670-treated platelets behaved similarly to that of the vehicle, with cytosolic Ca²⁺ increased rapidly for the first two min, from the basal level at 35 ± 13 to 233 ± 68 nM (vehicle control) and 262 ± 108 nM (YU142670). Afterwards, the cytosolic Ca²⁺ level remained sustained and it slightly decreased to 183 ± 70 nM (control) and 190 ± 88 nM (YU142670) over 10 min. On the other hand, for 1 μM GSK-A1 treated platelets, the Ca²⁺ also increased rapidly for the first two min to 233 ± 82 nM, but then it steadily decreased back to basal level for the next 8 min and reached 44 ± 9 nM, significantly lower than the vehicle (P < 0.01).

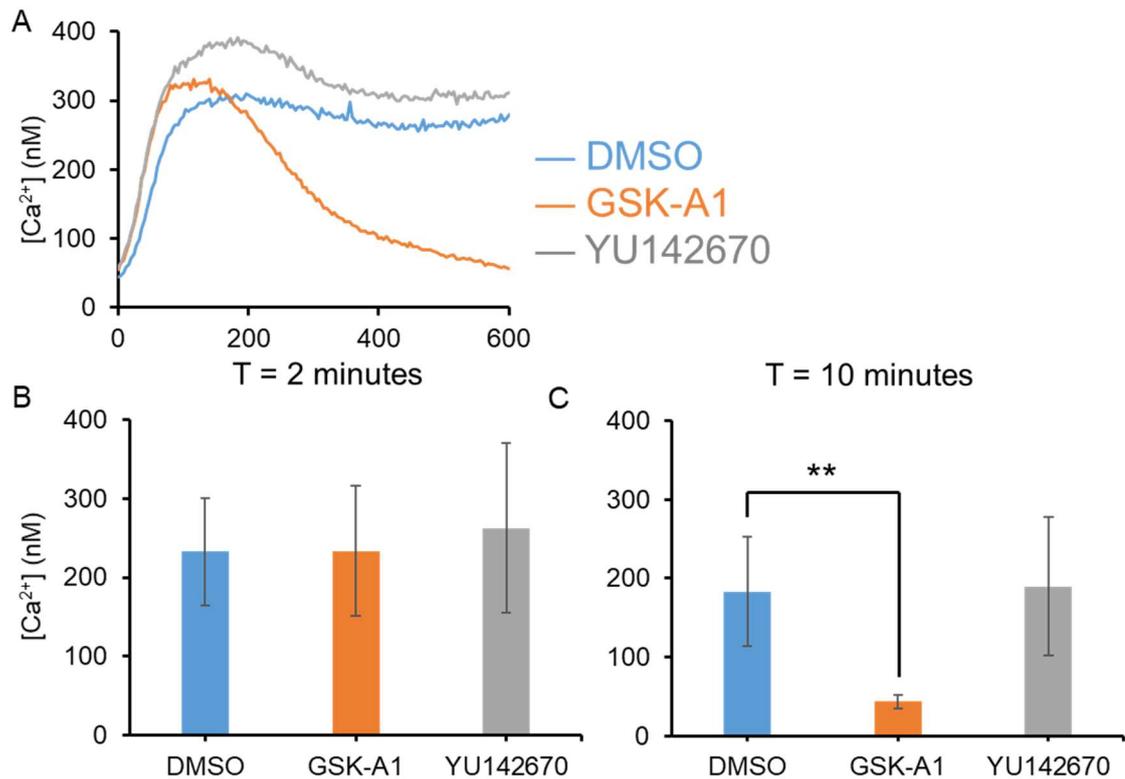


Figure 8. Effect of GSK-A1 and YU142670 on Ca^{2+} mobilisation. Fura 2-loaded washed platelets at 2×10^8 cells/mL were pre-treated for 10 min with DMSO vehicle, or at the stated concentration of GSK-A1 or YU142670 in the presence of apyrase (2.5 U/mL) and indomethacin (20 μ M). The platelets were then stimulated with 10 μ g/mL CRP in the presence of 1 mM $CaCl_2$. (A) Representative traces of changes in $[Ca^{2+}]_i$ with 1 μ M GSK-A1 or YU142670-treated platelets over 600 s were recorded. Quantification of increases in $[Ca^{2+}]_i$ for 2 min (B) or 10 min (C) after CRP stimulation. Data are means \pm SD ($n = 3$), ** $P < 0.01$, Welch's t-test.

Validation of model using InsP₁ measurement and phosphoinositide turnover inhibitors

Experiments were also conducted to determine the effect of GSK-A1 and YU142670 on CRP-induced accumulation of InsP₁ due to InsP₃ production and rapid hydrolysis. Washed platelets at 8×10^8 cells/mL were pre-treated for 10 min with vehicle, 1 μ M GSK-A1 or 1 μ M YU142670 in the presence of apyrase (2.5 U/mL), indomethacin (20 μ M) and 50 mM LiCl. The presence of Li⁺ inhibits inositol-phosphate phosphatase and prevents InsP₁ hydrolysis to inositol.²² The concentration of platelets, inhibitors and LiCl was based on similar studies and manufacturer's instructions.^{10,23}

As shown in Figure 9, similar to vehicle (from 30 ± 12 to 284 ± 112 nM), YU142670 pre-treatment led to the accumulation of InsP₁ to 334 ± 41 nM after 10 min of CRP stimulation, which was not significantly different from the vehicle control. On the contrary, 10 min pre-treatment of GSK-A1 was able to eliminate CRP-induced InsP₁ production, and remain at the basal level of 21 ± 3 nM after 10 min ($P < 0.05$ compared to vehicle). The results show that GSK-A1 inhibit PtdIns(4,5)P₂ resynthesis that prolonged the Ca²⁺ response, and that GSK-A1 treatment did not affect the initial Ca²⁺ mobilisation despite the lack of InsP₁ accumulation. This may suggest that the initial Ca²⁺ mobilisation is induced by InsP₃ is produced from the starting PtdIns(4,5)P₂.

The model predictions were compared with the experimental data in Figure 9B. The predictions were produced by simulating the model with the same parameters as the original model except for 10% θ_1 (the rate of conversion of InsP₁ to I_{pool}) only for the control condition, or together with 10% r_1 for GSK-A1 or 10% r_2 for YU142670. Both models predicted the inhibition of InsP₁ accumulation for GSK-A1, with most of the best-fitting model profiles (black dotted lines) matching the experimental data (orange line in Figure 9A), accumulating to 1,000 to 10,000 molecules after 10 min (equivalent to around 13 to 130 nM for 8×10^8 platelets/mL).

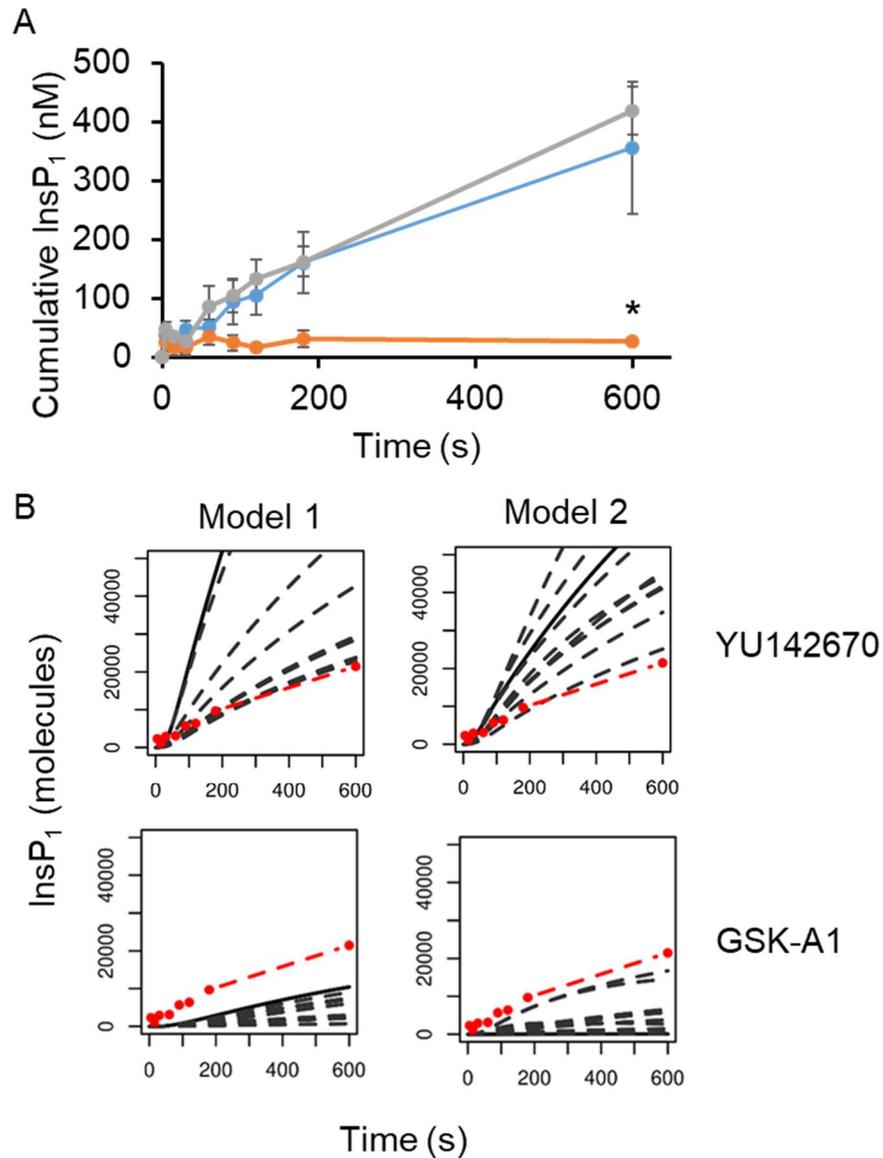


Figure 9. Comparison of experimental data and model predictions on the effect of GSK-A1 and YU142670 on InsP₁ accumulation. (A) Washed platelets at 8×10^8 cells/mL were pre-treated for 10 min with DMSO vehicle (grey), $1 \mu\text{M}$ GSK-A1 (orange) or $1 \mu\text{M}$ YU142670 (blue) in the presence of apyrase (2.5 U/mL), indomethacin ($20 \mu\text{M}$), and 50 mM LiCl to prevent InsP₁ hydrolysis. The platelets were stimulated with $10 \mu\text{g/mL}$ CRP, and the stimulation was stopped at the stated time with lysis buffer. Cumulative InsP₁ production was quantified using ELISA according to the manufacturer's instructions. * $P < 0.05$, Welch's t-test ($n = 3$ donors). (B) Model 1 (left) and 2 (right) were simulated at 10% r_2 to represent the effect of YU142670 (top) or 10% r_1 to represent the effect of GSK-A1 (bottom). The 6 best-fitting model profiles (black dotted lines) were compared to the observation in (A) (red dotted lines). Both simulated with ϑ_1 equal to 10% of its original value.

For YU142670, the model predicted an enhancement in InsP_1 accumulation which was not observed in the experiment. Comparing both models, in general simulation profiles of Model 1 deviates more from the red lines, compared to Model 2. In addition, the best-fitting model profiles failed to converge or reach steady states, and the predicted InsP_1 level after 10 min ranges from 20,000 to 50,000 molecules (equivalent to 260 to 650 nM for 8×10^8 platelets/mL).

Discussion

In this chapter, we created a biological model of the phosphoinositides metabolism that uses kinetic rate characteristics determined in platelets downstream of GPVI, allowing a more accurate representation of platelets. It has long been known that $\text{PtdIns}(4,5)\text{P}_2$ has a high turnover rate, which has been linked to so-called *futile cycles* of dephosphorylation and rephosphorylation that are thought to occur on the plasma membrane (PM).²⁴ This is shown by the rapid labelling kinetics of $\text{PtdIns}(4,5)\text{P}_2$ and $\text{PtdIns}4\text{P}$ compared to the much slower labelling kinetics of PtdIns and other phospholipids.²⁵ In our model, we are able to simulate the high turnover mathematically between $\text{PtdIns}(4,5)\text{P}_2$, $\text{PtdIns}4\text{P}$ and PtdIns which provided the kinetic basis for the futile cycle and demonstrated the system can reach a steady state in the simulated platelets system.

Controlling the amount of $\text{PtdIns}(4,5)\text{P}_2$ in the PM is crucial for regulating signalling and membrane dynamics. Platelets must replenish this pool, since $\text{PtdIns}(4,5)\text{P}_2$ only make up a tiny portion of all cellular PtdIns , particularly during prolonged PLC and PI3K activity. Early research in rat hepatocytes had shown that even a 10 min stimulation causes huge turnover in $\text{PtdIns}(4,5)\text{P}_2$ pool.²⁶ This can be explained by the sequential phosphorylation by PI4K and PIP5K,⁸ and the ER-PM membrane contact site (MCS), which can form in response to agonist stimulation and facilitate lipid-transfer proteins to shuttle phosphatidic acid and PtdIns back

and forth across membranes, to sustain PtdIns(4,5)P₂ level.²⁷ In the model, we integrated this concept by reducing the platelet environment into PM and Cyt compartments, with Cyt representing the cytosol and the inner membranes at the opposite side of MCS. This makes it possible for PtdIns to be transferred from the Cyt to the PM compartment by PITP in MCS in activated platelets without the need for vesicle transport, which would impede the replenishment of PtdIns and is unable to support continuous secondary mediators generation. And as a result, the model is constructed based on an activated platelet and is unable to simulate the resting condition in an unstimulated platelet.

The Ca²⁺ assay shows that GSK-A1 treatment did not affect the initial Ca²⁺ mobilisation despite the lack of InsP₁ accumulation, but it inhibited prolonged Ca²⁺ response. The discrepancy is probably because platelet Ca²⁺ response is regulated InsP₃ receptor activity, extracellular Ca²⁺ entry and Ca²⁺ back-pumping. In the first two min, InsP₃ is mainly produced by the initial pool of PtdIns(4,5)P₂ that is present in the PM, causing the opening of InsP₃ receptor and Ca²⁺ release from the internal stores. The Ca²⁺ spiking is caused by the co-stimulatory effect of InsP₃ and released Ca²⁺.²⁸ However, in the presence of GSK-A1 and inhibition of PI4KA, PtdIns(4,5)P₂ gradually depletes without replenishment from PtdIns. Together with the continuous hydrolysis of InsP₃ and its dissociation from InsP₃ receptors, the ion channels are eventually closed, and the elevated cytosolic Ca²⁺ is pumped back to the stores by Ca²⁺-ATPases until returning to the basal level.

Model 1 predicted a rapid depletion of I_{pool} from 10⁸ molecules to 10⁴ molecules in the first 200 s, which is converted to P_{pool}, increasing from 2 x 10⁴ molecules to 10⁸ molecules in the first 200 s through PtdIns as an intermediate. These predictions are too drastic to be physiologically possible within 200 s. While for the predictions in Model 2, which predicted a stable pool of I_{pool} and a tiny increase in P_{pool}, are more physiologically feasible. The reason

Model 1 produces such a simulation is that all 10^8 molecules of I_{pool} can participate in the reaction to synthesise PtdIns. By the law of mass action, with such a high amount of reactant, the conversion rate of I_{pool} to PtdIns is huge, and these molecules are subsequently converted to P_{pool} instead of PtdIns4P, as its level is constrained by experimental data. Model 2 prevents such a large flux by having half of the PtdIns and its synthesis located in a separate intracellular compartment. Therefore PtdIns (PM) can only be converted from PtdIns (Cyt), which is two orders of magnitude lower than I_{pool} , resulting in a much lower conversion rate compared to Model 1. This highlights the importance of spatial regulation, trafficking on maintaining the futile cycle and equilibrium of interconversion of phosphoinositides.

The model predicted a slight enhancement in InsP_1 accumulation for YU142670 which was not observed in the experiment. In addition, YU142670 produced a similar Ca^{2+} mobilisation trace as vehicle. The lack of effect of YU142670 can be because OCRL is localised in the trans-Golgi network instead of the PM compartment.²⁹ On the contrary, GSK-A1 inhibits PI4KA which localises in PM.³⁰ Therefore, inhibition of OCRL may assert minimal short-term effect on the pool of high turnover phosphoinositides that localise in PM in activated platelets.

There are several limitations of the model. First, the model does not account for spatial effects, combining cytosol and the inner membranes into the same Cyt compartment. This is not the full representation of the platelet system, ignoring the cytosol-inner membrane interface which would, for example, limit the amount of inositol available to be incorporated into PtdIns, and prevent the huge flux from happening in Model 1. Second, the posterior parameter approximations generated in this study (Figure 5B) span two orders of magnitude, showing that there are insufficient experimental data to constrain the parameter values. Acquiring extra data under different experimental conditions, such as the use of different levels of agonist and other phosphoinositides kinase/phosphatase inhibitors would help in

constraining the range of approximated parameters. In addition, the model combined PtdIns3P, PtdIns5P, and PtdIns(3,5)P₂ into a single variable Ppool, because I was unable to quantitate their abundance by the developed IC-MS method due to their peak overlap.

In conclusion, this chapter develops an experimentally calibrated dynamic mathematical model of phosphoinositide metabolism in platelets. Despite advances in mass spectrometry-based profiling or imaging techniques, our quantitative knowledge of these transient and unstable phosphoinositides is still limited. This is because there are no tools available to real-time measure the absolute concentrations and subcellular resolution of lipids in living cells at the same time. The mathematical model developed in this chapter can circumvent these problems and shed light on the interconnectedness of phosphoinositide metabolism. The model is also able to predict the effect of GSK-A1 on phosphoinositides metabolism and platelet activation, proving the hypothesis it would inhibit the resynthesis of PtdIns(4,5)P₂ level thereby reducing CRP-induced InsP₃ production and sustained Ca²⁺ mobilisation.

Authors contributions

HYFC designed research, analysed and interpreted data. HYFC and JLD developed the model and wrote the paper. CT and JLD wrote the code, ran and analysed model simulations. AS, SPW, JWMH and JMG revised the manuscript.

Acknowledgments

HYFC, CT, AS, SPW, JWMH, JMG and JLD have received funding from the European Union's Horizon 2020 Marie Skłodowska-Curie (#766118). HYFC is enrolled in a joint PhD program with the Universities of Maastricht (The Netherlands) and Birmingham (United

Kingdom). CT is enrolled in a joint PhD program with the universities of Reading (United Kingdom) and Maastricht. SPW is supported by the Wellcome Trust (204951/Z/16/Z). SPW holds a BHF Chair (CH03/003). J.L.D thanks the British Heart Foundation for funding (RG/20/7/34866 and RG/15/2/31224). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Declaration of Interests

The authors declare no relevant conflicts of interest.

Materials and Methods

Materials

Chemicals and reagents were obtained from the following sources: Cross-linked collagen-related peptide (CRP) came from the University of Cambridge (Cambridge, United Kingdom); Fura-2 acetoxymethyl ester and human fibrinogen were obtained from Invitrogen (Carlsbad, CA, USA); Pluronic F-127 from Molecular Probes (Eugene, OR, USA). MS-grade MeOH from Biosolve (Valkenswaard, The Netherlands); formic acid, 37% HCl, CHCl₃ and methylamine in MeOH from Sigma-Aldrich (Steinheim, Germany); NaCl, 1-butanol and isopropanol (IPA) from Merck (Darmstadt, Germany); Tris(hydroxymethyl)-aminomethane (Tris) was purchased from Applichem (Darmstadt, Germany); sodium dodecyl sulfate (SDS) from Roth (Karlsruhe, Germany); 16:0/16:0 PtdIns4P and 16:0/16:0 PtdIns(4,5)P₂ α -fluorovinylphosphonate (PtdIns(4,5)P₂-FP) from Echelon Biosciences (Salt Lake City, UT, USA); and 17:0/20:4 PtdIns3P, 18:1/18:1 PtdIns(3,4)P₂, 18:1/18:1 PtdIns(4,5)P₂, 18:1/18:1

PtdIns(3,5)P₂ and 17:0/20:4 PtdIns(3,4,5)P₃ from Avanti Polar Lipids (Alabaster, AL, USA). Ultrapure water (18 MΩ cm at 25°C) was obtained from an Elga Labwater system (Lane End, United Kingdom). The bicinchoninic acid (BCA) assay was purchased from Thermo Scientific (Schwerte, Germany). Platelets were activated using collagen-related peptide (CRP, Richard Farndale, University of Cambridge, United Kingdom) or thrombin from human plasma (Roche, Germany).

Subjects and blood collection

Blood was taken by venipuncture from healthy male and female volunteers who had not taken anti-platelets in the previous ten days, after full informed consent according to the Helsinki declaration. The study was approved by the Medical Ethics Committee of Maastricht University. According to the approval, blood donor age and sex were not recorded. Blood was collected into 3.2% sodium citrate (Vacuette tubes, Greiner Bio-One, Alphen a/d Rijn, The Netherlands). All blood donors had platelet counts within the reference range, as measured with a Sysmex XN-9000 analyser (Sysmex, Kobe, Japan).

Platelet isolation

Platelet-rich plasma (PRP) and washed platelet were obtained from citrated blood samples, using an earlier described protocol with slight modifications.³¹ The PRP was obtained through centrifugation at 260 g for 10 min, and supplemented with 1:10 vol/vol acid citrate dextrose (ACD; 80 mM trisodium citrate, 183 mM glucose, 52 mM citric acid). After transferring into eppendorf tubes, the PRP was centrifuged at 2360 g for 2 min. The pelleted platelets were resuspended into HEPES buffer pH 6.6 (10 mM HEPES, 136 mM NaCl, 2.7 mM KCl, 2 mM MgCl₂, 5.5 mM glucose, and 0.1% bovine serum albumin). After the addition of apyrase (1 U/mL) and

1:15 vol/vol ACD, a second similar centrifugation step was performed to obtain washed platelets.

Platelet stimulation experiment

Washed platelets at a concentration of 1×10^9 platelets/mL were stimulated with 30 $\mu\text{g/mL}$ CRP in the presence of apyrase (2.5 U/mL) and indomethacin (10 μM), before the addition of 10x volume of cold HCl to stop the reaction. The pellets were then shock frozen in liquid nitrogen and kept at -80°C after the platelets had been centrifuged for 5 min at 640 g at 25°C .

Lipid extraction

Acidified chloroform/methanol ($\text{CHCl}_3/\text{MeOH}$) extraction was carried out following the protocol of Clark *et al.*³² For platelet samples, after the addition of 242 μL CHCl_3 , 484 μL MeOH, 23.6 μL 1M HCl, 170 μL water and internal standard [100 pmol of PtdIns(4,5) P_2 -FP] to the cell pellets containing 1×10^8 platelets, the mixture was allowed to stand at room temperature for 5 min with occasional vortexing. Next, 725 μL of CHCl_3 and 170 μL 2M HCl was added to induce phase separation and the samples were centrifuged at 1,500 g for 5 min at room temperature (Eppendorf, Hamburg, Germany). This created a two-phase system with an upper aqueous layer and a protein interface. Then, the lower organic layer was transferred to another tube and dried under a continuous stream of nitrogen (1 L/min N_2 at 25°C).

The lipid extracts were then deacylated following the protocol of Jeschke *et al.*³³ The dried lipid extracts were resuspended in 50 μL methylamine in methanol/water/1-butanol (46:43:11) and incubated at 53°C for 50 min in a thermomixer at 1 000 rpm (Thermomixer Comfort; Eppendorf, Hamburg, Germany). Then 25 μL cold IPA was added to the mixture, and

the mixture was dried under a continuous stream of nitrogen to obtain dried lipid extracts (1 L/min N₂ at 25°C). The dried and deacylated lipid extract was resuspended in 50 µL water and stored at -80°C prior to further analysis.

Protein concentration determination

1200 µL methanol was added to the remaining protein interphase and aqueous upper phase, and the mixture was incubated at -80°C for 3 hours. Then the mixture was centrifuged at 19,000 g for 30 min at 4°C, the supernatant removed, and the remaining protein pellet was dried under the fume hood. The resulting protein pellet was then resuspended in 1% SDS, 150 mM NaCl, 50 mM Tris (pH 7.8) and the protein concentration was determined using a BCA assay.

IC-MS/MS

IC-MS/MS was conducted using a Dionex ICS-5000 instrument (Thermo Fischer Scientific, Darmstadt, Germany) connected to a QTRAP 6500 instrument (AB Sciex, Darmstadt, Germany) that was equipped with an electrospray ion source (Turbo V ion source). Chromatographic separation was accomplished with a Dionex IonPac AS11-HC column (250 mm × 2 mm, 4 µm; Thermo Fischer Scientific) fitted with a guard column (50 mm × 2 mm, 4 µm; Thermo Fischer Scientific). A segmented linear gradient was used for separation of GroPIInsP: initial 15 mM potassium hydroxide (KOH), then held at 15 mM KOH from 0.0 to 5.0 min, 15 to 25 mM KOH from 5.0 to 15.0 min, 50 to 65 mM KOH from 15.0 to 30.0 min, 100 mM KOH from 30.0 to 34.0 min, 10 mM KOH from 34.0 to 38.0 min, 100 mM KOH from 38.0 to 42.0 min, and 15 mM KOH from 42.0 to 45.0 min. The IC flow rate was 0.38 mL/min, supplemented postcolumn with 0.15 mL/min makeup flow of 0.01% FA in MeOH. The temperatures of the

autosampler, column oven and ion suppressor were set at 10, 30 and 20°C, respectively. The injector needle was automatically washed with water and 5 µL of each sample were loaded onto the column.

The following ESI source settings were used: curtain gas, 20 arbitrary units; temperature, 400°C; ion source gas I, 60 arbitrary units; ion source gas II, 40 arbitrary units; collision gas, medium; ion spray voltage, -4500 V; declustering potential, -150 V; entrance potential, -10 V; and exit potential, -10 V. For scheduled selected reaction monitoring (SRM), Q1 and Q3 were set to unit resolution. The collision energy was optimized for each GroPIInsP by direct infusion of the corresponding deacylated standard. The scheduled SRM detection window was set to 3 min, and the cycle time was set to 1.5 s. Data were acquired with Analyst version 1.6.2 (AB Sciex). Skyline (64-bit, 3.5.0.9319) was used to visualize results, integrate signals over the time, and quantify all lipids that were detected by MS.³⁴

Preparation of Fura-2 loaded platelets

The washed platelets were resuspended into HEPES buffer pH 7.45 at a count of 2×10^8 /mL, and then loaded with Fura-2 acetoxymethyl ester (3 µM) and pluronic (0.4 µg/mL) for 40 min at room temperature. The Fura-2-loaded platelets were then centrifuged again in the presence of apyrase (1 U/mL) and 1:15 vol/vol ACD. For all inhibitor experiments, extra apyrase (1 U/mL) was added during labelling followed by centrifugation in the presence of 1:15 vol/vol ACD. The final count after resuspension into HEPES buffer pH 7.45 was 2×10^8 /mL.

Calibrated platelet $[Ca^{2+}]_i$ measurements

Changes in $[Ca^{2+}]_i$ of Fura-2-loaded platelets were measured in 96-well plates using a FlexStation 3 robot (Molecular Devices, San Jose, CA, USA), basically as indicated before.^{35,36}

In brief, 200 μL samples of platelets ($2 \times 10^8/\text{mL}$) per well were left untreated or were pretreated with apyrase (0.1 U/mL) and indomethacin (20 μM) for 10 min. Where indicated, pharmacological inhibitors to block Ca^{2+} entry were added as well (see Table 1). After the addition of either 0.1 mM EGTA or 2 mM CaCl_2 , the disposed platelets in wells were temperature adjusted (37°C) and fluorescence at two excitation wavelengths were recorded for 10 min. During recording, 20 μL of agonist solution was added by automated pipetting. Note that the mixing of agonist with Fura-2-loaded platelets was diffusion-limited, and occurred by high-speed injection of 10% volume of the agonist solution. Prior to default use, injection volume and speed (125 $\mu\text{L}/\text{s}$) had been optimized to obtain maximal platelet responses.³⁶

Changes in Fura-2 fluorescence (37°C) were measured per row by ratiometric fluorometry, using excitation wavelengths of 340 and 380 nm and a single emission wavelength of 510 nm.³⁶ Fura-2 fluorescence ratio values per well were obtained every 4 s. Separate calibration wells contained Fura-2-loaded platelets that were lysed with 0.1% Triton-X-100 in the presence of either 1 mM CaCl_2 or 1 mM EGTA/Tris for determining R_{max} and R_{min} values.³⁷ After the correction for 340 and 380 nm background fluorescence levels, nanomolar changes in $[\text{Ca}^{2+}]_i$ were calculated according to the Grynkiewicz equation with a K_D of 224 nM.³⁸ All measurements were completed within 2-3 hours of isolation of cells. Dye leakage during measurements appeared to be negligible.

Platelet stimulation and InsP_1 ELISA

Washed platelets were resuspended in Tyrode's buffer at a concentration of 8×10^8 platelets/mL supplemented with 50 mM LiCl which prevents InsP_1 hydrolysis. Before stimulation, washed platelets were pre-treated with GSK-A1, YU142670, or DMSO as vehicle

together with 2.5 U/mL apyrase, 10 uM indomethacin for 10 min (presence of 2.5 U/mL apyrase, 10 uM indomethacin). Afterwards, the washed platelets were stimulated by 10 µg/mL CRP, and lysed with 1% lysis buffer (IP-ONE ELISA kit, Cisbio, Bedford, MA, USA). After cell lysis, the lysed platelets were snap-frozen in liquid nitrogen for future analysis. The InsP₁ ELISA was conducted according to IP-ONE ELISA kit manufacturer's instructions, and the final result was measured in 96-well plates using a FlexStation 3 (Molecular Devices, San Jose, CA, USA).

Model calibration

The mathematical models were calibrated to the experimental data utilising a Bayesian approach. We sampled parameter values via a Latin Hypercube from log-normal fitting constraints between 10^{-4} and 10^2 for 20,000 times, as summarised in Table 2. The constraints for s_1 , θ_1 , θ_6 and q_1 are different from the listed range and adjusted according to literature which shows the hydrolysis rate of inositol phosphates,^{39,40} or by the examination of posteriors distribution after the initial fitting. Based on the constraint, a gradient-based method (fmincon, MATLAB) was used to find the local minimum, minimising the distance between the mathematical model's simulations and the experimental data through the cost function:

$$SSE = (y_{i,x}(a) - Data_{i,x})^2,$$

where SSE denotes sum squared error, $y_{i,x}(a)$ indicates the model simulations for species x at time point i and $Data_{i,x}$ represents the experimental profile of species x at time point i . As best parameter values were defined those that allow the mathematical model to produce simulations with the lowest SSE and are closest to the experimental data.

References

1. Marée AF, Grieneisen VA, Edelstein-Keshet L. How cells integrate complex stimuli: the effect of feedback from phosphoinositides and cell shape on cell polarization and motility. *PLoS Comp Biol.* 2012;8:e1002402.
2. Dawes AT, Edelstein-Keshet L. Phosphoinositides and Rho proteins spatially regulate actin polymerization to initiate and maintain directed movement in a one-dimensional model of a motile cell. *Biophys J.* 2007;92:744-768.
3. Olivença DV, Uliyakina I, Fonseca LL, Amaral MD, Voit EO, Pinto FR. A mathematical model of the phosphoinositide pathway. *Sci Rep.* 2018;8:1-12.
4. Purvis JE, Chatterjee MS, Brass LF, Diamond SL. A molecular signaling model of platelet phosphoinositide and calcium regulation during homeostasis and P2Y₁ activation. *Blood.* 2008;112:4069-4079.
5. Mazet F, Tindall MJ, Gibbins JM, Fry MJ. A model of the PI cycle reveals the regulating roles of lipid-binding proteins and pitfalls of using mosaic biological data. *Sci Rep.* 2020;10:1-14.
6. Smejkal GB, Kakumanu S. Enzymes and their turnover numbers. *Expert Rev Proteomics.* 2019;16:543-544.
7. Eremin A, Moroz I, Mikhailova R. Use of cadmium hydroxide gel for isolation of extracellular catalases from *Penicillium piceum* and characterization of purified enzymes. *Appl Biochem Microbiol.* 2008;44:590-599.
8. Balla T. Phosphoinositides: tiny lipids with giant impact on cell regulation. *Physiol Rev.* 2013;93:1019-1137.
9. Burkhart JM, Vaudel M, Gambaryan S, et al. The first comprehensive and quantitative analysis of human platelet protein composition allows the comparative analysis of structural and functional pathways. *Blood.* 2012;120:e73-e82.
10. Bura A, Jurak Begonja A. Imaging of intracellular and plasma membrane pools of PI(4,5)P₂ and PI4P in human platelets. *Life.* 2021;11:1331.
11. Zhao L, Thorsheim CL, Suzuki A, et al. Phosphatidylinositol transfer protein- α in platelets is inconsequential for thrombosis yet is utilized for tumor metastasis. *Nature Commun.* 2017;8:1-12.
12. Lagarde M, Guichardant M, Menashi S, Crawford N. The phospholipid and fatty acid composition of human platelet surface and intracellular membranes isolated by high voltage free flow electrophoresis. *J Biol Chem.* 1982;257:3100-3104.
13. Morris JB, Hinchliffe KA, Ciruela A, Letcher AJ, Irvine RF. Thrombin stimulation of platelets causes an increase in phosphatidylinositol 5-phosphate revealed by mass assay. *FEBS Lett.* 2000;475:57-60.

14. Chicanne G, Severin S, Boscheron C, et al. A novel mass assay to quantify the bioactive lipid PtdIns3P in various biological samples. *Biochem J*. 2012;447:17-23.
15. Binder H, Weber PC, Siess W. Separation of inositol phosphates and glycerophosphoinositol phosphates by high-performance liquid chromatography. *Anal Biochem*. 1985;148:220-227.
16. Ragab A, Séverin S, Gratacap M-P, et al. Roles of the C-terminal tyrosine residues of LAT in GPVI-induced platelet activation: insights into the mechanism of PLC γ 2 activation. *Blood*. 2007;110:2466-2474.
17. Watson S, Auger J, McCarty O, Pearce A. GPVI and integrin α IIb β 3 signaling in platelets. *J Thromb Haemost*. 2005;3:1752-1762.
18. Kielkowska A, Niewczas I, Anderson KE, et al. A new approach to measuring phosphoinositides in cells by mass spectrometry. *Adv Biol Regul*. 2014; 54:131-141.
19. Valet C, Chicanne G, Severac C, et al. Essential role of class II PI3K-C2 α in platelet membrane morphology. *Blood*. 2015;126:1128-1137.
20. Bojjireddy N, Botyanski J, Hammond G, et al. Pharmacological and genetic targeting of the PI4KA enzyme reveals its important role in maintaining plasma membrane phosphatidylinositol 4-phosphate and phosphatidylinositol 4, 5-bisphosphate levels. *J Biol Chem*. 2014;289:6120-6132.
21. Pirruccello M, Nandez R, Idevall-Hagren O, et al. Identification of inhibitors of inositol 5-phosphatases through multiple screening strategies. *ACS Chem Biol*. 2014;9:1359-1368.
22. Leech AP, Baker GR, Shute JK, Cohen MA, Gani D. Chemical and kinetic mechanism of the inositol monophosphatase reaction and its inhibition by Li⁺. *Eur J Biochem*. 1993;212:693-704.
23. Gilio K, Munnix IC, Mangin P, et al. Non-redundant roles of phosphoinositide 3-kinase isoforms α and β in glycoprotein VI-induced platelet signaling and thrombus formation. *J Biol Chem*. 2009;284:33750-33762.
24. Hawkins PT, Michell RH, Kirk C. Analysis of the metabolic turnover of the individual phosphate groups of phosphatidylinositol 4-phosphate and phosphatidylinositol 4,5-bisphosphate. Validation of novel analytical techniques by using ³²P-labelled lipids from erythrocytes. *Biochem J*. 1984;218:785-793.
25. Downes P, Michell RH. Phosphatidylinositol 4-phosphate and phosphatidylinositol 4, 5-bisphosphate: lipids in search of a function. *Cell Calcium*. 1982;3:467-502.
26. Creba J, Downes CP, Hawkins PT, Brewster G, Michell RH, Kirk CJ. Rapid breakdown of phosphatidylinositol 4-phosphate and phosphatidylinositol 4, 5-bisphosphate in rat hepatocytes stimulated by vasopressin and other Ca²⁺-mobilizing hormones. *Biochem J*. 1983;212:733-747.
27. Posor Y, Jang W, Haucke V. Phosphoinositides as membrane organizers. *Nat Rev Mol Cell Biol*. 2022;23:797-816.

28. Clapham DE. Calcium signaling. *Cell*. 2007;131:1047-1058.
29. Faucherre A, Desbois P, Satre V, Lunardi J, Dorseuil O, Gacon G. Lowe syndrome protein OCRL1 interacts with Rac GTPase in the trans-Golgi network. *Hum Mol Genet*. 2003;12:2449-2456.
30. MJ A, Burke J. Molecular mechanisms of PI4K regulation and their involvement in viral replication. *Traffic (Copenhagen, Denmark)*. 2022.
31. Gilio K, van Kruchten R, Braun A, et al. Roles of platelet STIM1 and Orai1 in glycoprotein VI- and thrombin-dependent procoagulant activity and thrombus formation. *J Biol Chem*. 2010;285:23629-29638.
32. Clark J, Anderson KE, Juvin V, et al. Quantification of PtdInsP3 molecular species in cells and tissues by mass spectrometry. *Nat Methods*. 2011;8:267.
33. Jeschke A, Zehethofer N, Schwudke D, Haas A. Quantification of phosphatidylinositol phosphate species in purified membranes. *Meth Enzymol*. 587; 2017:271-291.
34. Peng B, Ahrends R. Adaptation of skyline for targeted lipidomics. *J Proteome Res*. 2015;15:291-301.
35. Jooss NJ, De Simone I, Provenzale I, et al. Role of platelet glycoprotein VI and tyrosine kinase Syk in thrombus formation on collagen-like surfaces. *Internatl J Mol Sci*. 2019;20:e2788.
36. Fernandez DI, Provenzale I, van Groningen J, et al. Ultra-high throughput Ca²⁺ response patterns in platelets to distinguish between ITAM-linked and G-protein coupled receptor activation. *iScience*. 2022;25:103718.
37. Feijge MAH, van Pampus ECM, Lacabartz-Porret C, et al. Inter-individual variability in Ca²⁺ signalling in platelets from healthy volunteers, relation with expression of endomembrane Ca²⁺-ATPases. *Br J Haematol*. 1998; 102:850-859.
38. Grynkiewicz G, Poenie M, Tsien RY. A new generation of Ca²⁺ indicators with greatly improved fluorescence properties. *J Biol Chem*. 1985;260:3440-3450.
39. Watson SP, Lapetina EG. 1,2-Diacylglycerol and phorbol ester inhibit agonist-induced formation of inositol phosphates in human platelets: possible implications for negative feedback regulation of inositol phospholipid hydrolysis. *Proc Natl Acad Sci USA*. 1985;82:2623-2626.
40. Watson SP, Reep B, McConnell RT, Lapetina EG. Collagen stimulates [³H] inositol trisphosphate formation in indomethacin-treated human platelets. *Biochem J*. 1985;226:831-837.

Supplemental

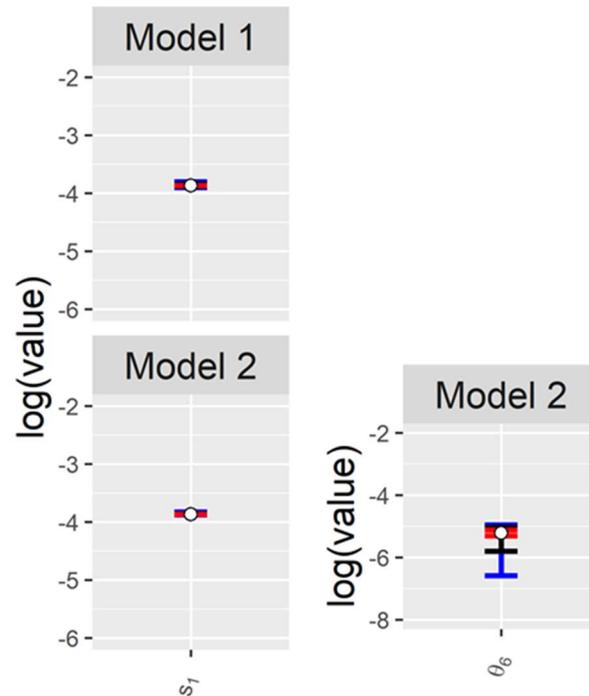


Figure S1. Posterior parameter ranges for s_1 and θ_6 . Posterior parameter ranges are shown for Models 1 (top) and 2 (bottom). The blue, black, and red lines and hollow circles show the full range, interquartile range, the inner 10% range and median of each estimated parameter.

Set	r_1	r_4	r_2	r_2	r_3	r_5	θ_2	θ_3	θ_3	θ_5	θ_5	θ_4	θ_4	s_1	s_2
1	1.4E-01	5.8E+01	1.7E+00	2.5E+00	2.4E+01	1.6E+00	1.7E-02	4.9E-01	5.5E-03	1.3E-03	7.9E-02	7.3E-03	2.9E+01	1.5E-04	3.3E-01
2	2.0E-01	3.0E+01	1.4E+01	1.9E+01	9.7E-01	7.5E+00	1.4E-02	3.1E-01	5.3E-03	2.0E-03	8.5E-01	2.6E-02	1.0E+02	1.6E-04	1.1E-01
3	2.0E-01	3.0E+01	1.4E+01	1.9E+01	9.7E-01	7.5E+00	1.4E-02	3.1E-01	5.3E-03	2.0E-03	8.5E-01	2.6E-02	1.0E+02	1.6E-04	1.1E-01
4	2.7E+00	9.2E+00	1.2E+01	1.5E+01	3.9E+01	2.9E+00	1.6E-02	1.9E-03	1.1E-03	1.0E-02	3.3E+00	2.3E-02	9.1E+01	1.5E-04	5.5E-01
5	3.4E-01	1.0E+00	2.4E+00	3.2E+00	2.7E+01	4.2E+00	1.2E-02	3.1E-01	4.2E-03	1.8E-03	5.6E-01	4.5E-04	2.5E+00	1.6E-04	4.1E-01
6	5.4E+00	1.8E+01	3.7E+00	3.3E+00	6.2E+01	8.4E-01	1.2E-02	2.7E-02	5.5E-03	2.0E-03	2.3E+00	1.7E-04	8.2E-01	1.6E-04	8.4E-01
7	1.3E+01	4.4E+01	1.3E+01	1.9E+01	2.3E+01	5.4E+00	1.6E-02	1.2E-01	2.3E-03	8.0E-03	2.5E+00	3.2E-03	1.4E+01	1.6E-04	3.9E-01
8	1.2E+00	4.7E+00	7.8E-01	9.8E-01	7.6E-01	1.2E+01	1.3E-02	3.8E-01	5.4E-03	1.3E-03	2.1E-01	2.0E-03	9.9E+00	1.6E-04	1.8E-01
9	7.1E+00	2.6E+01	1.5E+01	2.6E+01	3.0E+01	9.0E-01	1.5E-02	4.3E-01	7.0E-03	1.4E-03	5.3E-01	1.3E-02	5.2E+01	2.0E-04	5.1E-01
10	2.8E-01	1.8E-02	1.0E+00	1.3E-01	1.5E-02	1.0E+00	9.0E-03	4.4E-02	3.9E-03	1.8E-03	1.6E+00	1.1E-04	5.7E-01	1.6E-04	1.3E-02

Table S1. Results of parameter fitting Model 1 to experimental data. Parameter sets of the ten best results from a sample of $N = 20000$ fits.

Set	r_1	r_3	r_2	r_2	r_3	r_5	θ_2	θ_3	θ_{-3}	θ_5	θ_{-5}	θ_4	θ_{-4}	s_1	s_2
1	7.0E+00	1.3E+01	7.0E+00	9.8E+00	1.7E+00	2.7E+00	1.1E-04	1.4E-03	7.3E-03	2.2E-02	2.1E-03	1.1E+00	2.3E+01	1.5E-04	5.7E-02
2	1.1E-03	2.0E-03	3.4E-01	4.6E-01	1.0E+01	1.7E-01	6.3E-04	1.2E-03	6.2E-01	1.8E-01	3.0E-03	4.0E-03	3.3E-02	1.6E-04	1.4E-01
3	2.2E+01	4.8E+01	8.1E-03	1.1E-03	1.5E-03	8.5E-01	2.2E-02	1.0E-03	2.1E-02	1.0E-03	1.0E-03	4.8E-01	4.7E+00	1.6E-04	1.1E-02
4	4.3E-02	4.7E-02	7.1E-02	8.6E-02	4.6E-01	3.7E-02	2.3E-04	5.0E-03	6.0E-03	3.3E-03	1.8E-02	1.4E+00	5.2E+01	1.6E-04	6.4E-03
5	1.8E+00	3.0E+00	3.4E-01	4.8E-01	1.0E+01	9.6E-01	9.8E-01	3.0E-03	6.1E-03	7.2E-03	8.8E-03	1.2E-01	3.6E+00	1.6E-04	1.5E-01
6	6.9E+00	1.1E+01	3.5E-02	3.0E-02	5.2E+01	2.3E+00	2.5E-01	1.0E-03	1.1E-02	8.3E-03	6.5E-03	1.3E-01	4.0E+00	1.5E-04	6.8E-01
7	1.0E+00	1.4E+00	7.4E+00	1.0E+01	1.3E+01	1.3E-01	4.8E-01	8.2E-03	4.5E-03	4.9E-03	8.8E-03	2.2E-01	7.5E+00	1.6E-04	1.7E-01
8	3.6E-02	4.9E-02	3.7E-02	4.3E-02	2.2E-01	2.5E-01	2.8E-02	7.7E-03	5.9E-03	3.4E-03	1.0E-02	6.1E-02	2.3E+00	1.4E-04	6.1E-03
9	1.0E-03	2.7E-03	3.2E+00	4.3E+00	7.5E+01	9.9E-02	4.5E-02	1.0E-03	1.8E-01	3.6E+01	1.2E+00	1.9E-03	1.8E-02	1.6E-04	1.0E+00
10	7.0E+00	1.3E+01	7.0E+00	9.8E+00	1.7E+00	2.7E+00	1.1E-04	1.4E-03	7.3E-03	2.2E-02	2.1E-03	1.1E+00	2.3E+01	1.5E-04	5.7E-02

Table S2. Results of parameter fitting Model 2 to experimental data. Parameter sets of the ten best results from a sample of $N = 20000$

Chapter 5 (to be submitted)

Additive benefit of modelling of platelet calcium signalling by combined recurrent neural network and partial least squares analyses

The mechanisms that regulate calcium levels in the platelet cytosol following platelet activation via GPVI or other receptors/ligands is complex. This involves multiple ion channels in different subcellular locations. The relationship between these components is poorly understood. In this chapter, we developed two modelling techniques to investigate how different agonists and inhibitors impact intracellular Ca^{2+} dynamics in platelets. Using a fluorescence ratio probe, we stimulated human platelets with a panel of agonists under various conditions, including the presence of inhibitors of a range of ion channels that are implicated in Ca^{2+} homeostasis, and modulation of secondary mediators. The aim was to construct mathematical models that could effectively predict calcium response due to a combination of ligands and inhibitors, without overwhelming signalling complications.

Additive benefit of modelling of platelet calcium signalling by combined recurrent neural network and partial least squares analyses

Chukiat Tantiwong,^{1,2} Hilaire Yam Fung Cheung,^{2,3} Joanne L. Dunster,¹ Jonathan M. Gibbins,¹
Johan W. M. Heemskerk,^{2,4} Rachel Cavill^{5,*}

¹ School of Biological Sciences, University of Reading, Reading, United Kingdom

² Department of Biochemistry, Maastricht University, Maastricht, The Netherlands

³ Institute of Cardiovascular Sciences, College of Medical and Dental Sciences,
University of Birmingham, Birmingham, United Kingdom

⁴ Synapse Research Institute Maastricht, 6217 KD Maastricht, The Netherlands

⁵ Department of Data Science and Knowledge Engineering, Maastricht University,
Maastricht, The Netherlands

* Corresponding author: rachel.cavill@maastrichtuniversity.nl

Abstract

Platelets play critical roles in haemostasis and thrombosis. The platelet activation process is driven by agonist-induced rises in cytosolic $[Ca^{2+}]_i$, where the patterns of Ca^{2+} responses are still incompletely understood. In this study, we developed a number of techniques to model the $[Ca^{2+}]_i$ curves of platelets from a single blood donor. Using a fluorescence ratio probe, the platelets were stimulated with a panel of agonists, i.e. thrombin, collagen, or CRP under various conditions, preventing extracellular Ca^{2+} entry, secondary mediator effects or Ca^{2+} reuptake into intracellular stores. To analyse the data, we developed two non-linear models, a multilayer perceptron (MLP) network and an autoregressive network

with exogenous inputs (NARX). The trained networks accurately predicted the platelet $[Ca^{2+}]_i$ curves in the presence of combinations of agonists and inhibitors, with the NARX model achieving an R^2 up to 0.64 for trend prediction of unforeseen data. In addition, we used the same dataset for construction of a partial least square (PLS) linear regression model, which estimated the explained variance of each input. The NARX model demonstrated that good fits could be obtained for the calcium curves modeled, whereas the PLS model gave useful interpretable information on the importance of each variable. These modelling results can be used for the development of novel platelet $[Ca^{2+}]_i$ -inhibiting drugs.

Introduction

Blood platelets, derived from megakaryocytes, function in haemostasis and thrombosis via receptor-induced signalling responses¹⁻³. Important platelet-activating receptors are the protease-activated receptors (PAR1/4) for thrombin and the glycoprotein VI (GPVI) receptor for collagen, which signal as G-protein coupled receptors (GPCR) and as a protein tyrosine kinase-linked receptor (TKLR), respectively⁴. Given that arterial thrombosis is driven by the activation and aggregation of platelets⁵, and it is a prominent cause of death worldwide⁶, a clear understanding of the process of platelet activation is a must.

In platelets stimulated via GPCR or TKLR, a rise in cytosolic $[Ca^{2+}]_i$ is the common initial event, mediating all essential platelet functions^{7, 8}. The agonist-induced mobilisation of Ca^{2+} from intracellular stores in the endoplasmic reticulum (or dense tubular system) occurs via inositol 1,4,5-trisphosphate receptors (IP_3Rs), whereas sarcoplasmic/endoplasmic reticulum Ca^{2+} -ATPases (SERCAs) are responsible for Ca^{2+} back pumping into these stores (Figure S1)^{7, 8}. The IP_3R channels are triggered by IP_3 , which is produced as a result of activation of the GPCRs

for thrombin⁹ and ADP¹⁰, and of activation of the TLR GPVI by collagen or collagen-related peptide (CRP)⁸.

According to the mechanism of store-operated Ca^{2+} entry (SOCE), the store depletion is coupled to entry of Ca^{2+} from the extracellular medium, via Orai1 channels, which then interact with a Ca^{2+} sensor STIM1 (stromal interaction molecule 1) in the endoplasmic reticulum membrane⁷. The back pumping of Ca^{2+} over the plasma membrane occurs via plasma membrane Ca^{2+} -ATPases (PMCA). The primary agonists, thrombin and CRP, furthermore stimulate the release of autocrine agents that can enforce the Ca^{2+} signalling process. These are in particular autocrine-produced thromboxane A_2 (TxA_2) and ADP, both of which stimulate IP_3 production via GPCRs¹². Another paracrine-dependent Ca^{2+} entry mechanism is provided by ATP, which activates P2X_1 channels that specifically mediate Ca^{2+} entry¹¹.

Several pharmacological inhibitors can be used to interfere with the platelet Ca^{2+} responses. The entry of Ca^{2+} from the extracellular fluid is prevented by the Ca^{2+} chelator EGTA. Back pumping of Ca^{2+} from cytosol to intracellular stores is inhibited by the compound thapsigargin, which accordingly potentiates the Orai1-STIM1 dependent entry⁷. The effects of autocrine agents can be suppressed by the presence of apyrase (degrading ATP and ADP) and indomethacin (blocking TxA_2 formation). Figure S1 illustrates these platelet receptors, ligands, inhibitors and channels, relevant to the present study.

The high complexity of the Ca^{2+} -modulating process in platelets has triggered other authors to develop mathematical models, aiming not only to better understand the process but also to identify new therapeutic targets. Dolan and Diamond combined several models of Ca^{2+} fluxes in different platelet compartments into one single system, using a set of ordinary differential equations (ODEs)¹³. Although their system did not include ligand-receptor

interactions, it still consisted of 34 entities, 35 interactions and 86 parameters, thus reflecting the complexity of Ca^{2+} signalling process. An alternative approach presented by Chatterjee and Diamond¹⁴ was to create a neural network model, which was trained from the Ca^{2+} response patterns to specific agonists, using the platelets from a number of healthy donors. The neural network, acting as a black box, could predict synergistic effects on the Ca^{2+} responses of up to six receptor agonists. A trade-off of this network model was that all the parameters needed to be tuned and trained, which required extensive experimental data to achieve the desired predictive power. Another limitation was that the neural network approach did not provide information on the contribution of each type of Ca^{2+} channel and pump to the overall $[\text{Ca}^{2+}]_i$ levels. Similarly, it did not identify how the blockage of a given channel or (autocrine) process influenced the overall response.

In the present study, we constructed a computational model to predict the shapes of platelet $[\text{Ca}^{2+}]_i$ curves over time in response to thrombin or CRP for a given set of experimental conditions, with known agonists and inhibitors. We built several neural network-based models to better predict the agonist and inhibitor effects on the $[\text{Ca}^{2+}]_i$ time curves. We subsequently used a partial least square regression analysis to understand how specific curve variables contributed to obtained response. To exclude inter-individual variation, we used a coherent set of Ca^{2+} response curves in platelets, taken from a single healthy subject on one occasion.

Methodology

Materials

Human α -thrombin came from Kordia (Leiden, The Netherlands); cross-linked collagen-related peptide (CRP-XL) from the University of Cambridge (UK); Fura-2 acetoxymethyl ester from Invitrogen (Carlsbad CA, USA); Pluronic F-127 from Molecular Probes (Eugene OR, USA).

Horm-type collagen was obtained from Nycomed (Hoofddorp, The Netherlands). Other materials were from sources described before¹⁵.

Blood collection and platelet preparation

The study was approved by the Medical Ethics Committee of Maastricht University. Blood donor age and sex could not be recorded. Blood was taken into 3.2% sodium citrate (Vacurette tubes, Greiner Bio-One, Alphen a/d Rijn, The Netherlands) from consenting healthy volunteers who had not taken anti-platelet medication in the previous ten days. Platelet counts were within the reference range.

Platelet-rich plasma (PRP) was obtained from the citrated blood by centrifuging, after which the collected platelets were washed in the presence of apyrase (1 U/mL), and then loaded with Fura-2 acetoxymethyl ester (3 μ M) and Pluronic (0.4 μ g/mL) at a count of 2×10^8 /mL for 40 min at room temperature, such as described before¹⁶. The cells were finally resuspended at a concentration of 2×10^8 /mL in HEPES buffer pH 7.45 (10 mM HEPES, 136 mM NaCl, 2.7 mM KCl, 2 mM MgCl₂, 5.5 mM glucose, and 0.1% bovine serum albumin).

Calibrated cytosolic Ca²⁺ measurements

In the Fura-2-loaded platelets, changes in cytosolic [Ca²⁺]_i were measured in 96-well plates using a FlexStation 3 (Molecular Devices, San Jose, CA, USA), as described before¹⁶. When appropriate, the cells in wells were pretreated with apyrase (0.1 U/mL) plus indomethacin (20 μ M), or with thapsigargin (1 μ M) for 10 min. After the addition of 0.1 mM EGTA or 1 mM CaCl₂, the cells were stimulated by automated pipetting with one of the following agonists: CRP (1 or 10 μ g/mL), collagen (1, 3, 10 or 30 μ g/mL), thrombin (0.3, 1, 3 or 10 nM), or none of these (control). In all wells, changes in Fura-2 fluorescence were measured over time at 37 °C by ratiometric fluorometry, using appropriate control wells for calculating nM concentrations of [Ca²⁺]_i.¹⁶

Selection of platelet $[Ca^{2+}]_i$ traces for modelling

Calibrated agonist-induced time series of $[Ca^{2+}]_i$ with the various experimental conditions were performed with Fura-2-loaded platelets from 6 donors¹⁷. For the present modelling approach, a complete set of 72 time curves (Table 1) was chosen from one donor, which were representative for those all six donors. In the table, the validation and test conditions are highlighted in blue and red, respectively. The criteria for splitting these are indicated below.

Preparation of input data

The traces of nM changes in $[Ca^{2+}]_i$ in platelets for experiments involving CRP or collagen were measured every 4 s, while those for experiments with thrombin had an interval of 2-4 s. To be able to compare all 72 traces, all raw data (Figure S2) were linearly resampled and interpolated to generate 1 s time steps, from 0 s to 540 s (9 min). To minimize the noise in the dataset, the curves were smoothed with a Savitzky–Golay filter (Figure S3).

In cases where scaling of data was needed, a subset of interpolated curves was subjected to a standard min-max scaling algorithm to obtain values between 0 and 1. For the scaling of input conditions, experimental variables were set to have values in the range [-1, 1] (Table 1). Here, -1 meant no agonist or inhibitor present, while 1 indicated that the concentration of agonist or inhibitor was maximal across the samples.

For constructing a multilayer perceptron (MLP) network, a regression model was built using magnitudes of the $[Ca^{2+}]_i$ time series. The experimental variables were used as input values (Figure 1A). Herein we used the mean square error (MSE) as a cost function. This ensured a better fit for larger values (in the order of magnitude). To handle this complexity, we set the target (output) for the model as log-scaled values of the nM $[Ca^{2+}]_i$ range as $\log_{10}(\max - \min)$. This optimised the overall accuracy across log scales.

No.	Col	CRP	Thr	EG	Al	Tha
1	0	0	0	0	1	0
2	0	0	0	1	1	0
3	0	0	0	0	1	1
4	0	0	0	1	1	1
5	0	0	0	0	0	0
6	0	0	0	1	0	0
7	0	0	0	0	0	1
8	0	0	0	1	0	1
9	30	0	0	0	1	0
10	10	0	0	0	1	0
11	3	0	0	0	1	0
12	1	0	0	0	1	0
13	30	0	0	1	1	0
14	10	0	0	1	1	0
15	3	0	0	1	1	0
16	1	0	0	1	1	0
17	30	0	0	0	1	1
18	10	0	0	0	1	1
19	3	0	0	0	1	1
20	1	0	0	0	1	1
21	30	0	0	1	1	1
22	10	0	0	1	1	1
23	3	0	0	1	1	1
24	1	0	0	1	1	1
25	10	0	0	0	0	0
26	1	0	0	0	0	0
27	10	0	0	1	0	0
28	1	0	0	1	0	0
29	10	0	0	0	0	1
30	1	0	0	0	0	1
31	10	0	0	1	0	1
32	1	0	0	1	0	1
33	0	10	0	0	0	0
34	0	1	0	0	0	0
35	0	10	0	1	0	0
36	0	1	0	1	0	0
37	0	10	0	0	0	1
38	0	1	0	0	0	1
39	0	10	0	1	0	1
40	0	1	0	1	0	1
41	0	10	0	0	1	0
42	0	1	0	0	1	0
43	0	10	0	1	1	0
44	0	1	0	1	1	0
45	0	10	0	0	1	1
46	0	1	0	0	1	1
47	0	10	0	1	1	1
48	0	1	0	1	1	1
49	0	0	10	0	1	1
50	0	0	3	0	1	1
51	0	0	1	0	1	1
52	0	0	0.3	0	1	1
53	0	0	10	1	1	1
54	0	0	3	1	1	1
55	0	0	1	1	1	1
56	0	0	0.3	1	1	1
57	0	0	10	0	1	0
58	0	0	3	0	1	0
59	0	0	1	0	1	0
60	0	0	0.3	0	1	0
61	0	0	10	1	1	0
62	0	0	3	1	1	0
63	0	0	1	1	1	0
64	0	0	0.3	1	1	0
65	0	0	10	0	0	1
66	0	0	1	0	0	1
67	0	0	10	1	0	1
68	0	0	1	1	0	1
69	0	0	10	0	0	0
70	0	0	1	0	0	0
71	0	0	10	1	0	0
72	0	0	1	1	0	0

Table 1. Assignment matrix of variables of experimental conditions. Fura-2-loaded platelets from one single donor were used for Ca^{2+} response measurements on the same day. Conditions highlighted in blue (solid borders) were used as validation set, while those in red (dashed borders) were used as test set. Abbreviations: No., condition number; Col, collagen ($\mu\text{g/mL}$); CRP, collagen-related peptide ($\mu\text{g/mL}$); Thr, thrombin (nM); EG, EGTA: 0.1 mM if assigned to 1; or 1 mM $CaCl_2$ if assigned to 0; Al, apyrase (0.1 U/mL) plus indomethacin (20 μM); Thap, thapsigargin (1 μM).

Magnitude prediction neural network

Considering that the number of features was small with 6 experimental conditions (Table 1), we also generated polynomial features (quadratic feature combinations) to increase the feature number from 6 to 27. The MLP network was optimised by setting the number of hidden layers to 1, with the number of nodes randomly selected from 1 to 10. Given the relatively small training set that was available, network architecture options were chosen to train only a low number of parameters, thus preventing overfitting. Networks were trained 100 times, starting from random weights. The best structure was chosen as the one with a minimal score in the cost function of the validation set. Network training was performed using the Levenberg-Marquardt algorithm, containing a rectified linear unit (ReLU) as an activation function in each node. Modelling was conducted using Matlab R2022a and the Neural Network Toolbox.

Trend prediction of NARX network

Another type of neural network was constructed to predict the trends of smoothed and scaled $[Ca^{2+}]_i$ time curves. To better capture the time dynamics (*i.e.* the shape of the curve), we choose a non-linear autoregressive network with exogenous inputs (NARX) and parallel architecture^{18,19}, which is known as a closed-loop neural network. For this NARX network, the model's output $y(t)$ was used to fit the target (*i.e.* the smoothed and scaled $[Ca^{2+}]_i$ curves). The output then generated feedback as additional input to the network, when combined with the experimental condition (Figure 1B). The mathematical expression of $[Ca^{2+}](t)$ is written as follows:

$$y(t) = f(L_4 \times f(H_3 \times y_h + L_3 \times f(H_2 \times y_h + L_2 \times f(H_1 \times y_h + W \times I + b_1) + b_2) + b_3) + b_4)$$

where $y(t)$ is $[Ca^{2+}]_i(t)$, I is an input matrix of experimental conditions, y_h is the feedback delay (history) of y . Furthermore, W and H_n are the input matrix's weight and feedback delay of y ,

respectively; b_n are biases, and L_n are weights of each hidden layer; f is the activation (transfer) function. Note that the product of the matrix also is a matrix, meaning that the equation represents a summation of numerous parameters and functions.

For the feedback delays, we choose the values at the last 1, 3, 6, 10, 15, 21, 28, and 36 s prior to the current value of the $[Ca^{2+}]_i$ time series. These feedback delays hence kept the information about recent values, while preserving the long-term memory of the system. Initial values of the feedback delays were set to zero, as the system was assumed to be in a steady state prior to the agonist-induced activation of platelets. The use of MSE as a cost function allowed us to make predictions of the scaled min-max $[Ca^{2+}]_i$ time series. The scaling was performed per time series, implying that each series had the same range [0,1]. Polynomial features were also used in this network, thus expanding the number of inputs from 6 to 27.

The neural network architecture was optimised in a way to maximise the goodness of fit, but to prevent overfitting. We used three hidden layers, with each layer's size varying between 2 to 20 nodes (not including feedback delays). This gave approximately 7000 different architectures being trained. A randomised grid search was employed to find the best architecture. For training, the Levenberg-Marquardt algorithm was used with a hyperbolic tangent sigmoid (tansig) as an activation function. Since parameter fitting in the neural network depended on a random seed, each architecture was fitted 100 times, after which the best parameters were used for comparison. The networks were again built and trained in Matlab.

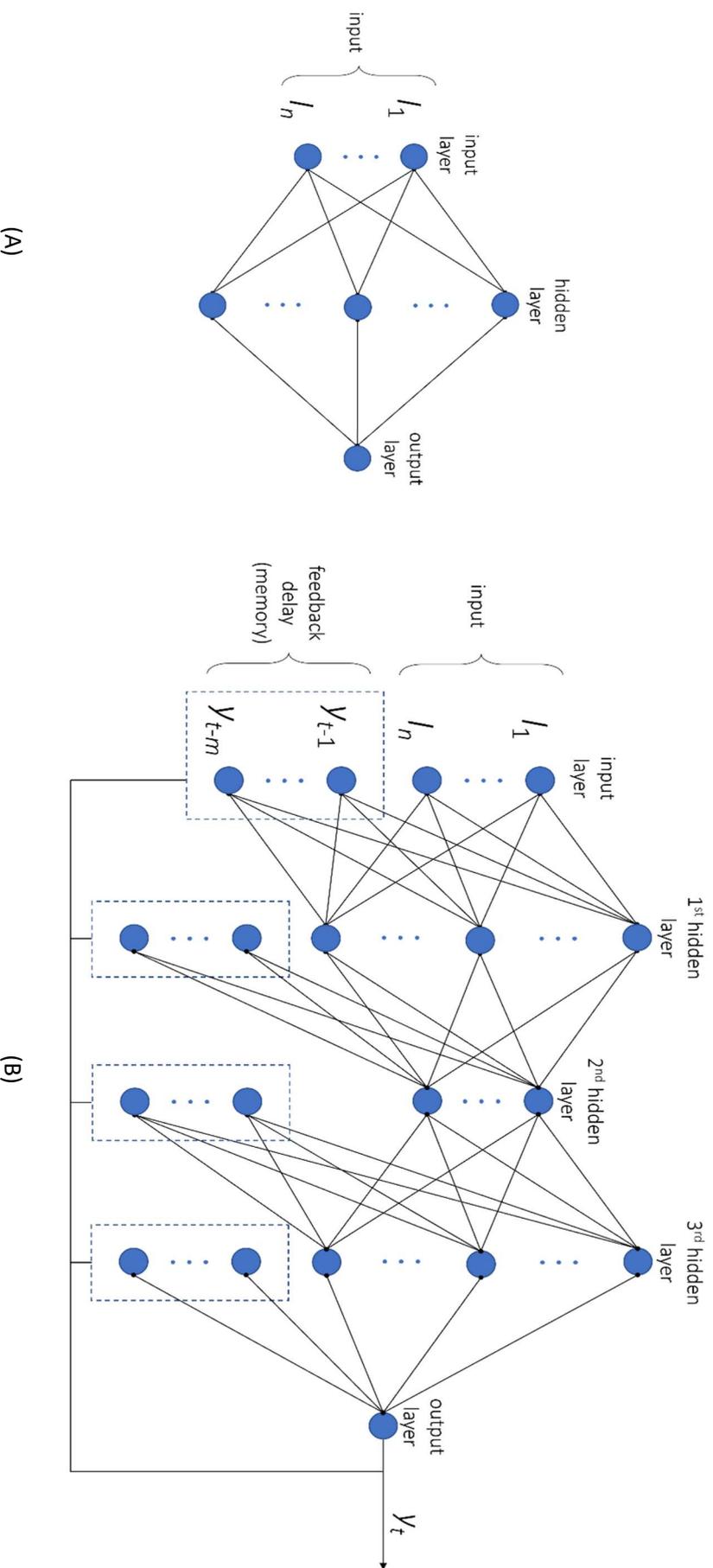


Figure 1. Construction of two neural networks. (A) Setup of MLP network, as a fully connected feedforward neural network, which was used for prediction of the magnitude of $[Ca^{2+}]_i$ time curves. (B) Closed-loop non-linear autoregressive network with exogenous inputs (NARX), which was used as a recurrent neural network. Herein, generated output served as input for a next time point. This network type was used for the prediction of trends in $[Ca^{2+}]_i$ time curves.

Parameter sensitivity analysis

To perform sensitivity analysis, the method of one-factor-at-a-time (OAT) was applied²⁰. This procedure keeps all other variables fixed to the central or baseline values while changing one variable at a time. Since all effects were computed with reference to the same central point in space, this improved comparability of the outcomes. As a default, we set the conditions of EGTA or CaCl₂, autocrine inhibitors (AI) or not, and thapsigargin or not as 1 or 0 ($2^3 = 8$ possible combinations). Furthermore, the agonist concentration was scaled from 0 to 10% of the maximal concentration, *i.e.* 30 µg/mL collagen, 10 µg/mL CRP or 10 nM thrombin. The shape of each $[Ca^{2+}]_i$ time curves was defined according to their scalar characteristics, namely the magnitude of the response, the peak time, the relative terminal level, and the mean deviation from a straight line, such as indicated in Figure 2.

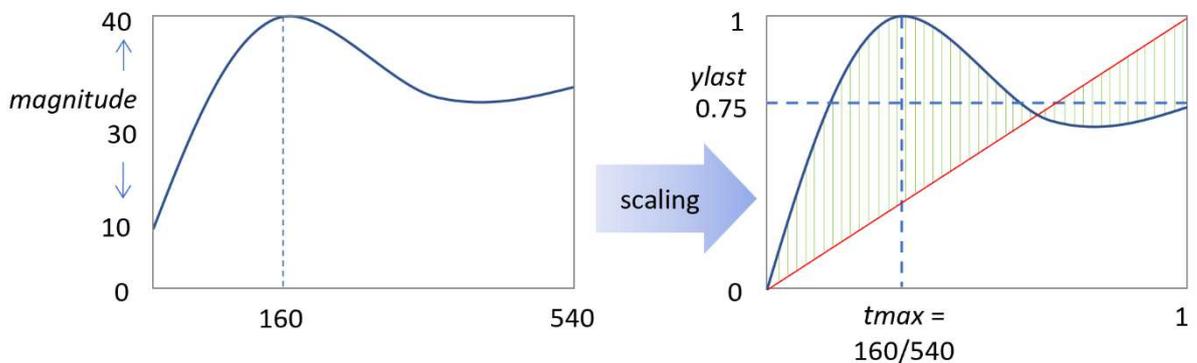


Figure 2. Defining the scalar characteristic of a $[Ca^{2+}]_i$ time curve. Scaling was performed using the following conversions: magnitude (nM) points to the maximal value minus the minimal value of a series. t_{max} refers to the time point that the curve reaches the maximal value, scaled by time range (540 s). Parameter y_{last} indicates the terminal value, scaled according to the magnitude. $absdev$ is a value indicating how much the time curve is deviating from a straight line (red line); calculated are at each time point the deviations from this line (green lines), and $absdev$ is the average of these deviations.

Partial least square regression analysis

Partial least square (PLS) regression analysis^{21,22} was used as an extension of principal component analysis, in which PLS instead of maximising the variance in each component, maximises the covariance between an input matrix X and an output matrix Y. Herein, each component has a latent variable t_i , while the linearly weighted combination of the latent variables generates the prediction of outcomes (Y matrix) as follows:

$$Y = C_1t_1 + C_2t_2 + \dots, \text{ where } C_i = a_{1i}x_1 + a_{2i}x_2$$

The experimental conditions of Table 1 were used as the X matrix, and the scalar characteristics of each $[Ca^{2+}]_i$ time series served as Y matrix. The number of components in the PLS analysis was obtained from the optimal variance achieved, when increasing the components. The loading weights depended on input variables that were most important for the predictions. By maximising the covariance between the explanatory variable X and the response variable Y, the most relevant components in X were obtained that affected the changes in Y. Stated otherwise, by examining the loading weights of first few latent variables that accounted for the majority of explained covariance, we could identify those experimental conditions with a most significant impact on $[Ca^{2+}]_i$ time curves.

Results

Comparing the input agonist-induced platelet $[Ca^{2+}]_i$ curves

Using a high-throughput method described before¹⁷, the Fura-2-loaded platelets from a single healthy donor were incubated in the presence of EGTA or $CaCl_2$ with or without secondary mediator inhibitors apyrase and indomethacin (AI); and then stimulated with collagen, CRP or thrombin. Under all these conditions, agonist-induced rises in $[Ca^{2+}]_i$ were measured as nM concentrations over a time period of 9 min. Altogether, by also varying the

agonist concentrations, resulted in a set 72 different experimental conditions (Table 1). For the present paper, the 72 different $[Ca^{2+}]_i$ time curves were obtained from the platelets of a single donor¹⁷.

Comparing the set of original traces (Figure S2), several characteristics can be observed, except for the expected agonist dose-dependency¹⁷. In general, the $[Ca^{2+}]_i$ curves induced by the weak GPVI agonist collagen showed steady increases with lower maximal amplitudes (Exp. 9-31), when compared to the higher amplitude and often biphasic $[Ca^{2+}]_i$ rises induced by the strong GPVI agonist CRP (Exp. 37-48). In particular, the curves with the PAR1/4 agonist thrombin (Exp. 49-72) had a transient shape, indicating high activity of the SERCA Ca^{2+} pumps. Other differences were 4-80 times higher amplitude curves (depending on other variables) in the presence of $CaCl_2$ than with EGTA, which in part was due to Orai1-dependent Ca^{2+} entry¹⁷. Furthermore, we observed potent $[Ca^{2+}]_i$ increase by adding the SERCA inhibitor thapsigargin, inhibiting SOCE and activating the Orai1 channels⁷. Effects of the autocrine inhibitors indomethacin and apyrase (IA) were a consistent lowering of most of the curves.

Workflow of the modelling approaches

In order to prepare the experimental data for further processing, we first interpolated and smoothed the 72 curves at 1 s time intervals (Figure S3), followed by a y-axis scaling per curve from 0-1 (Figure S4). The subsequent workflow (Figure 3) consisted of feature generation by combining and squaring of the experimental variables (see below), and split the curves into training, validation and test sets. The data were used as input for two types of modelling, *i.e.* neural network and PLS analyses. In the neural network analysis, we used the NARX procedure for trend prediction and the MLP procedure for magnitude prediction. A combined optimised network was tested on final performance. On the other hand, PLS was used to directly model

the scalar characteristics of the curves. The result from both approaches was interpreted and cross-checked with each other.

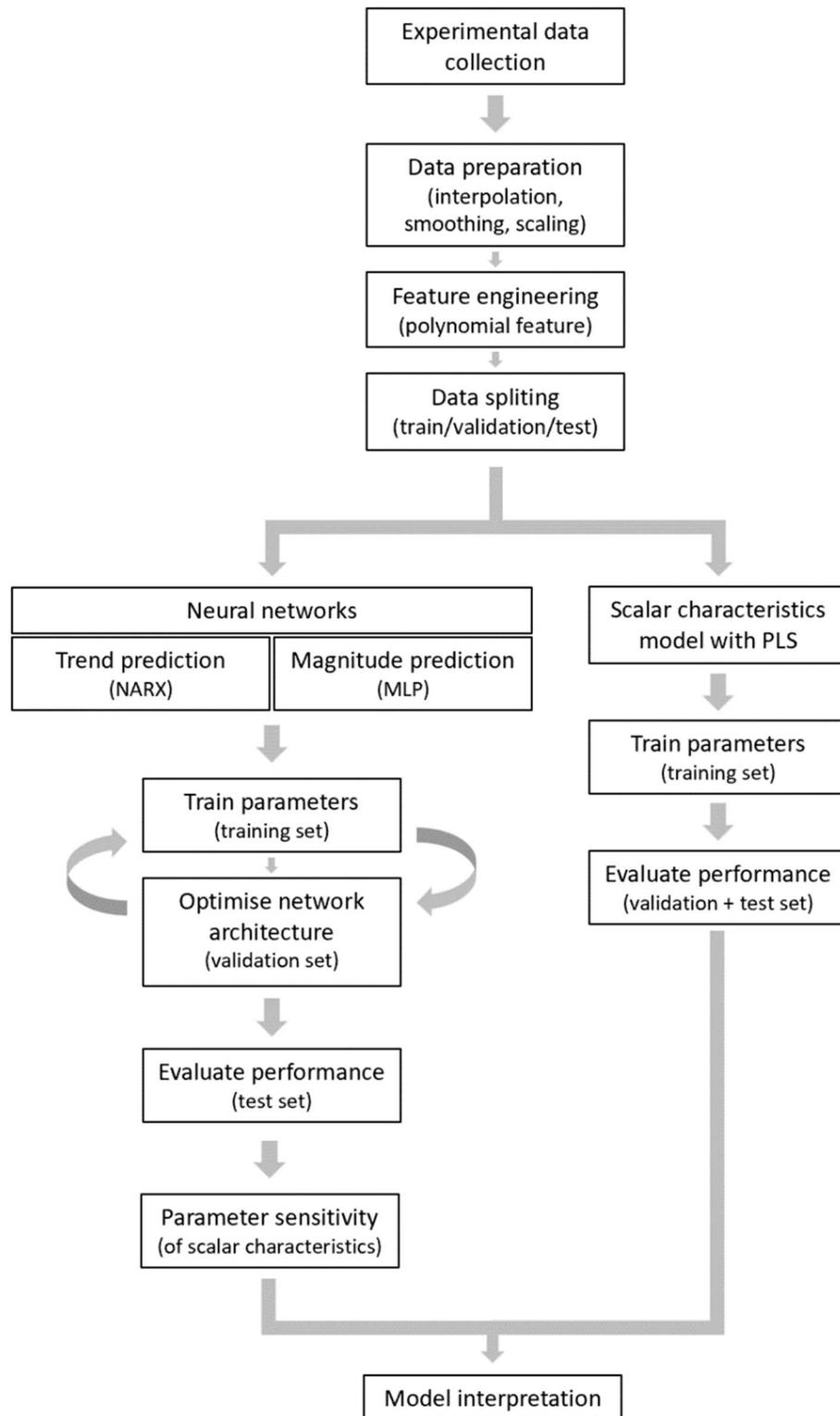


Figure 3. Workflow used for the data processing, neural network construction and scalar model development. For explanation, see text.

Neural MLP network for magnitude prediction

We first aimed to better understand how the smoothed $[Ca^{2+}]_i$ curves of platelets relied on the various experimental conditions (CaCl₂/EGTA, agonist dose, AI or thapsigargin). For this purpose, we generated a simple network able to predict the magnitude of the Ca²⁺ signal. The constructed multilayer perceptron (MLP) network was trained and validated, from which it appeared that the best MLP architecture had three nodes with a single hidden layer (Figure 1A). The results for the training, validation and test sets are shown in Figure S5. Plots were generated to compare the experimental data with the predictions in log scale and linear scale. Herein, each data point represents the experimental values and predicted magnitude values. These plots indicated an overall reasonable fitting, especially for the log-scale setting.

The obtained MLP parameters associated with each node are shown in Figure S6, with a colored way of the relative weights of the combined and squared parameters in the network. A limitation of this MLP approach is that only the curve magnitude is predicted and not the curve shape.

Neural NARX network for trend prediction

For prediction of the shape of trend of the $[Ca^{2+}]_i$ curves with all different amplitudes, uniform scaling is needed. Predictions modelling on these scaled time curves were made by constructing a recurrent, closed-loop neural network (NARX). For the training of the network, we used 58 scaled curves (Figure S7), which resulted in the best results for a network architecture with 3 hidden layers and 4 x 12 x 4 nodes (mean $R^2 = 0.84$) (Figure 1B). The validity of the network was overall confirmed for the validation set of 7 curves (mean $R^2 = 0.71$) (Figure S8). Fitting was less for the test set with 7 curves ($R^2 = 0.64$), in particular for the transient curve of Exp. 58 with thrombin (Figure 4A-G). For comparison, testing the same (unscaled)

amplitude curves with the MLP network resulted in a good prediction, especially for the high-magnitude curves (Figure 4H).

The NARX prediction trends also provided information on the non-linear shape of the $[Ca^{2+}]_i$ curves. Examining the trend values of R^2 , it appeared that these were negative for Exp. 58 (Figure S8) and Exp. 63 (Figure 4). This pointed to an explained variance worse than random, and hence inability of fitting. Furthermore, also other Exp. 67, 70 and 72 with thrombin as agonist gave an $R^2 < 0.4$. The likely explanation with this is the transiency of the thrombin-induced $[Ca^{2+}]_i$ rises. The above results prompted us to compare the neural network results of both magnitude and trend prediction.

Combining MLP and NARX networks

For a combined network curve prediction, we used the training set of 58 curves (Figure S9). The training with respect to magnitude and trend predictions was then validated and tested using the remaining 14 curves (Figure S10). The combined prediction resulted in a generally improved outcome. We also performed a one-factor-at-a-time (OAT) analysis by varying the agonist concentration at different inhibitor combinations, as shown in Figure 5A (as scaled variant curves) and Figure 5B (as scalar heatmaps). For additional visualization, also the unscaled curves are represented in Figure S11, which shows both the size and trend changes of the curves.

The OAT sensitivity analysis of Figure 5A shows how the predicted trend changed with the experimental input condition. It appeared that the predicted *magnitude* of the 'no inhibitor' (0 0 0) condition was mostly changed with the CRP concentration, when compared to collagen or thrombin. The presence of EGTA reduced the overall *magnitude* with to all agonists. Thrombin affected the *magnitude* most, while collagen and CRP had smaller effects. Furthermore, the presence of thapsigargin increased the overall *magnitude* prediction,

regardless of the type of agonist. Furthermore, the predicted *magnitude* increased less with the concentration of thrombin, than that of collagen or CRP.

As shown in the heatmap in Figure 5B, we also compared the three scalar characteristics (*tmax*, *ylast*, and *absdev*) of the scaled curves (see Figure 2). The $[Ca^{2+}]_i$ peak time (*tmax*) provided information on the curve transiency. If the *tmax* was equal to the final time point (540 s) usually indicated means that the Ca^{2+} response increased over the time range. In particular with thrombin the *tmax* was often <540 s, indicating a peaking and transient response. With CRP this was only seen to a limited extent at some inhibitor conditions.

The final level of $[Ca^{2+}]_i$, i.e. the parameter *ylast*, displayed similar trends as *tmax* (Figure 5B). In general, *ylast* is close to 1 under conditions of a continuous increase in $[Ca^{2+}]_i$, and <1 when the trend hit a peak before decreasing. Thus, thrombin inducing a non-linear curve pattern produced lower *ylast* values, even at higher agonist concentrations. The parameter *absdev* (absolute deviation from a straight line) indicated how the curve deviates from a linear response. Analysis of *absdev* showed that most of curves with thrombin were non-monotonic, except for conditions at which both thapsigargin and AI were present, i.e. resulting in more linear curves (Figure 5B). Accordingly, the three scaled curve characteristics provided additional information on the Ca^{2+} response patterns upon varying the $CaCl_2$, thapsigargin, AI and agonist concentrations.

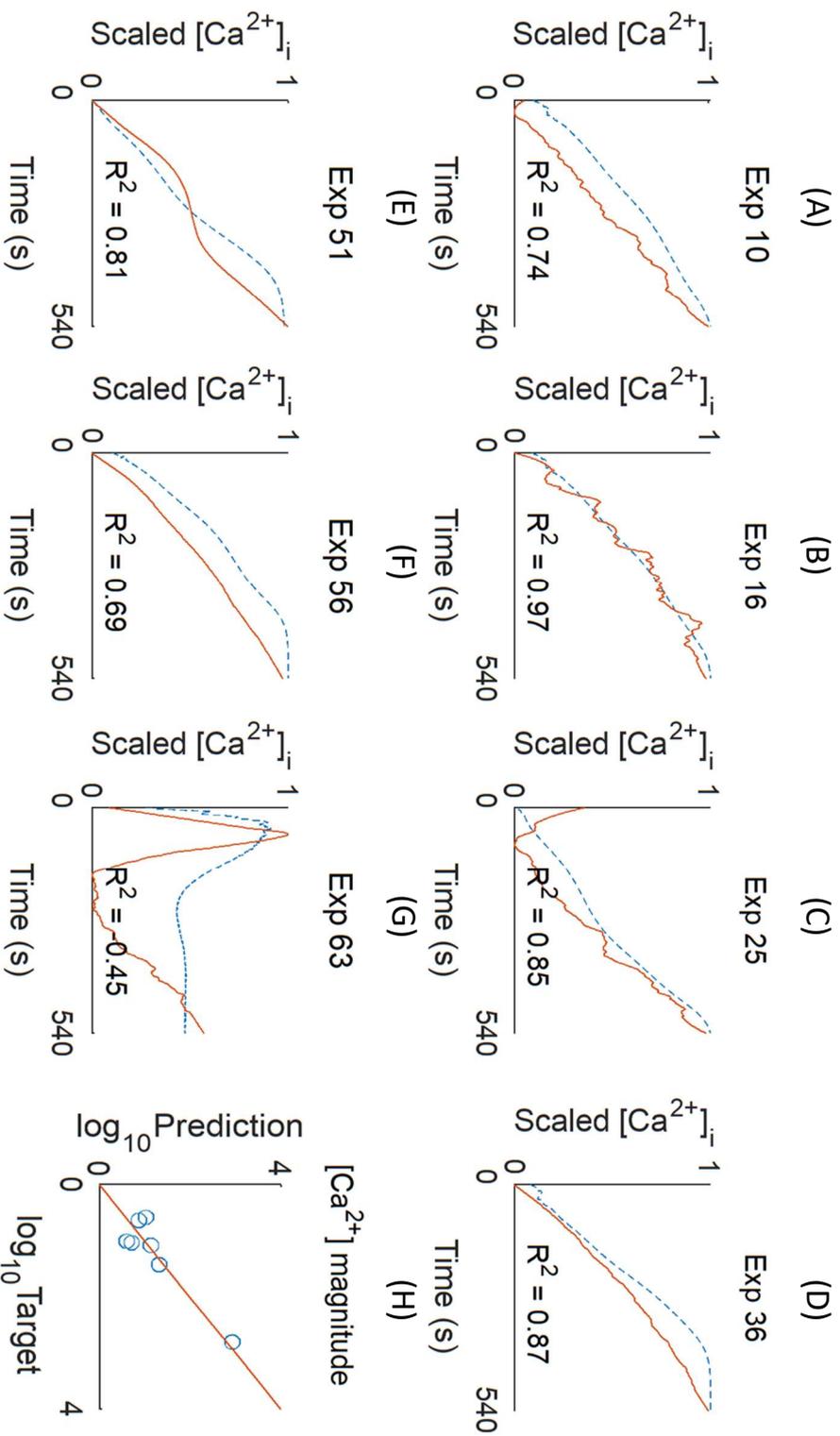
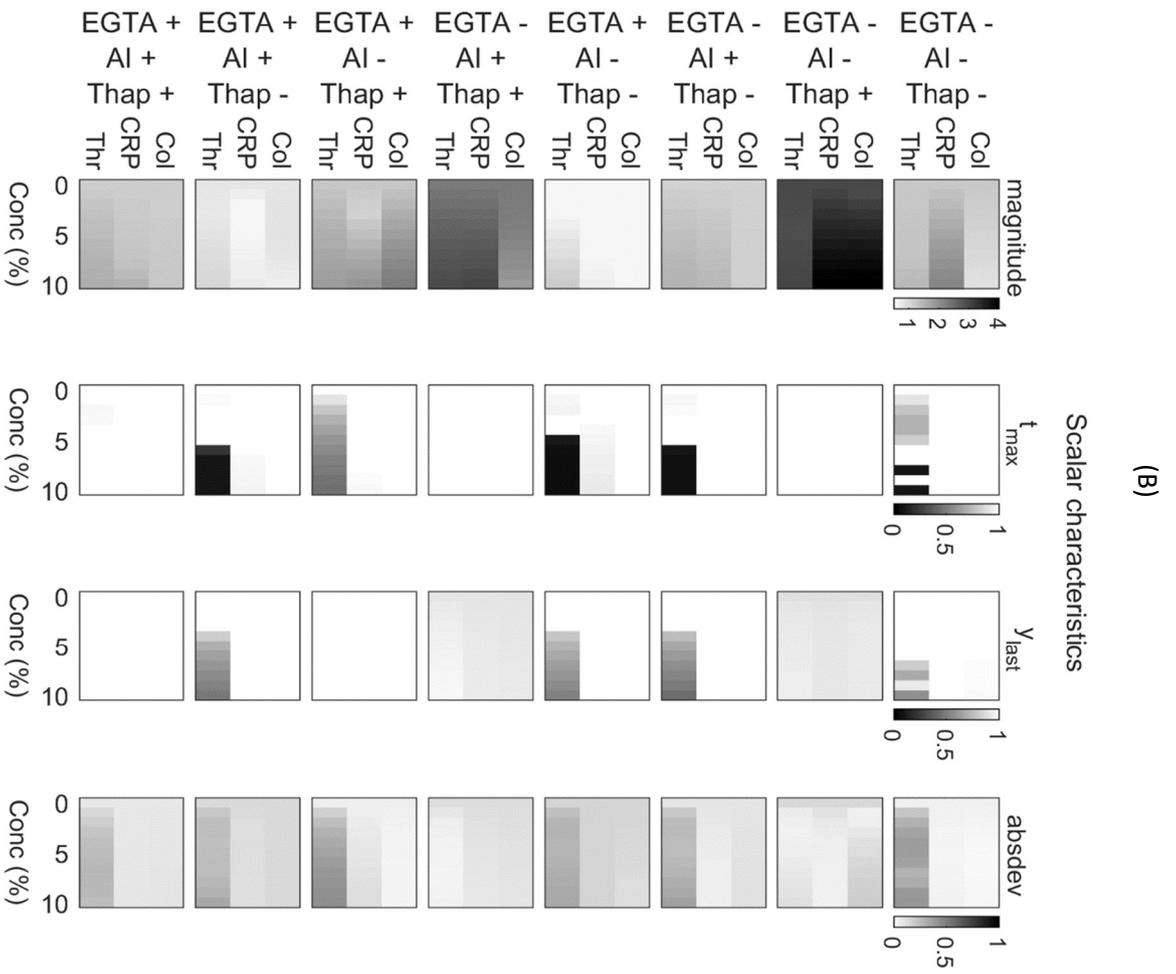
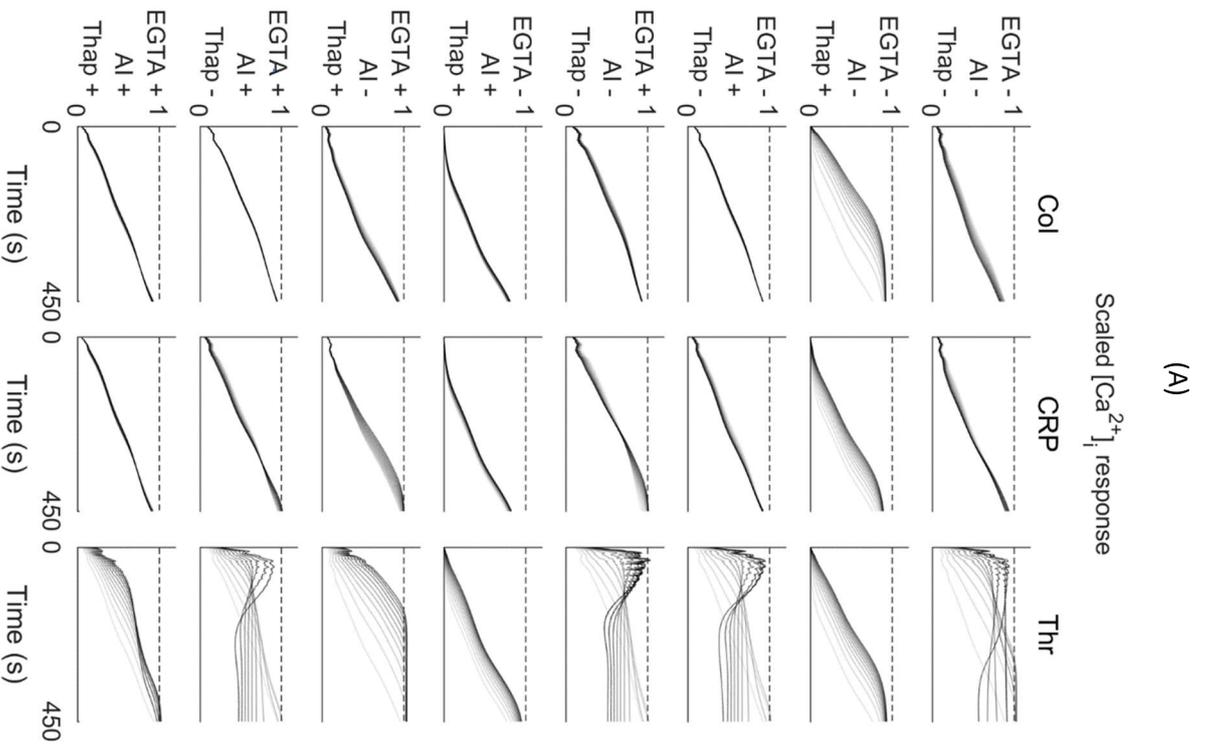


Figure 4. Test set results of NARX neural network to predict curve trends. (A-G) Testing of trend prediction of scaled $[Ca^{2+}]_i$ curves. The test set with experimental conditions was, as defined in Table 1. Red solid lines = actual values, blue dashed lines = predicted values. Calculated R^2 indicated per experimental condition (negative R^2 indicates an explained variance worse than random). (H) For comparison, results for the same test set are shown for curve magnitude predictions with the MLP network. Shown are the target and predicted nM levels of $[Ca^{2+}]_i$ in log scale.



◀ **Figure 5. Combined variation of trend prediction of scaled platelet $[Ca^{2+}]_i$ responses at varying agonist concentrations.** (A) Panels indicate prediction efficacy per agonist concentration. Lightest grey lines represent basal levels, while darker lines point to curve predictions in the presence of agonist by 1% from the basal level to 10% of the maximum concentration in the training set. Columns show conditions with indicated agonists (collagen, Col), CRP or thrombin (Thr). Rows represent different inhibitor conditions: + or - mean presence or not; from top to bottom: EGTA, apyrase plus indomethacin (AI), and thapsigargin (Thap). Dash lines at scale 1 indicate maximal trend per experiment. (B) Sensitivity of scalar characteristics of time curves generated by MLP and NARX model. Columns here indicate: $[Ca^{2+}]_i$ level at log10 base (magnitude), time of $[Ca^{2+}]_i$ (t_{max}), final $[Ca^{2+}]_i$ level (y_{last}), and mean absolute deviation from linear ($absdev$). All data shown are scaled at 0-1. For the unscaled $[Ca^{2+}]_i$ data, see Figure S7.

From combining the results of the two tested MLP and NARX networks, several conclusions can be drawn. The transient $[Ca^{2+}]_i$ responses with thrombin were harder to model than the non-transient responses with other agonists. For both the weak GPVI agonist collagen and the strong agonist CRP, the scaling approach showed a mostly monotonic curve increase, being close to linear at low agonist concentrations. Furthermore, the combined magnitude and trend modelling indicated for CRP additive effects of the absence of Ca^{2+} entry (EGTA, Exp. 36), absence of secondary mediators (AI, Exp. 42), of which the former was stronger (Exp. 44). However, in spite of these insights, the black-box nature of neural network approaches could have hidden other relevant relations between curves.

PLS regression analysis

As a more straightforward approach, we also directly investigated the contribution of each experimental variable (agonist dose, EGTA/ $CaCl_2$, IA, thapsigargin) to the scalar curve characteristic, *i.e.*, reducing the $[Ca^{2+}]_i$ time curves to t_{max} , y_{last} and $absdev$. For that purpose,

we used a PLS regression analysis to fit the relationships. As the PLS regression is a linear model, it is easier to investigate the impact of input variables on the output.

As input for the PLS model, we normalised all experimental conditions to separate variables of the concentrations of agonist and inhibitors (collagen dose, CRP dose, thrombin dose, EGTA/CaCl₂, AI, thapsigargin), all varying from 0 (none) to 1 (maximum). This resulted in a six-component model explaining the variance per component. As indicated in Figure S12, only the first two components contributed to the variance of the target. Accordingly, we fitted

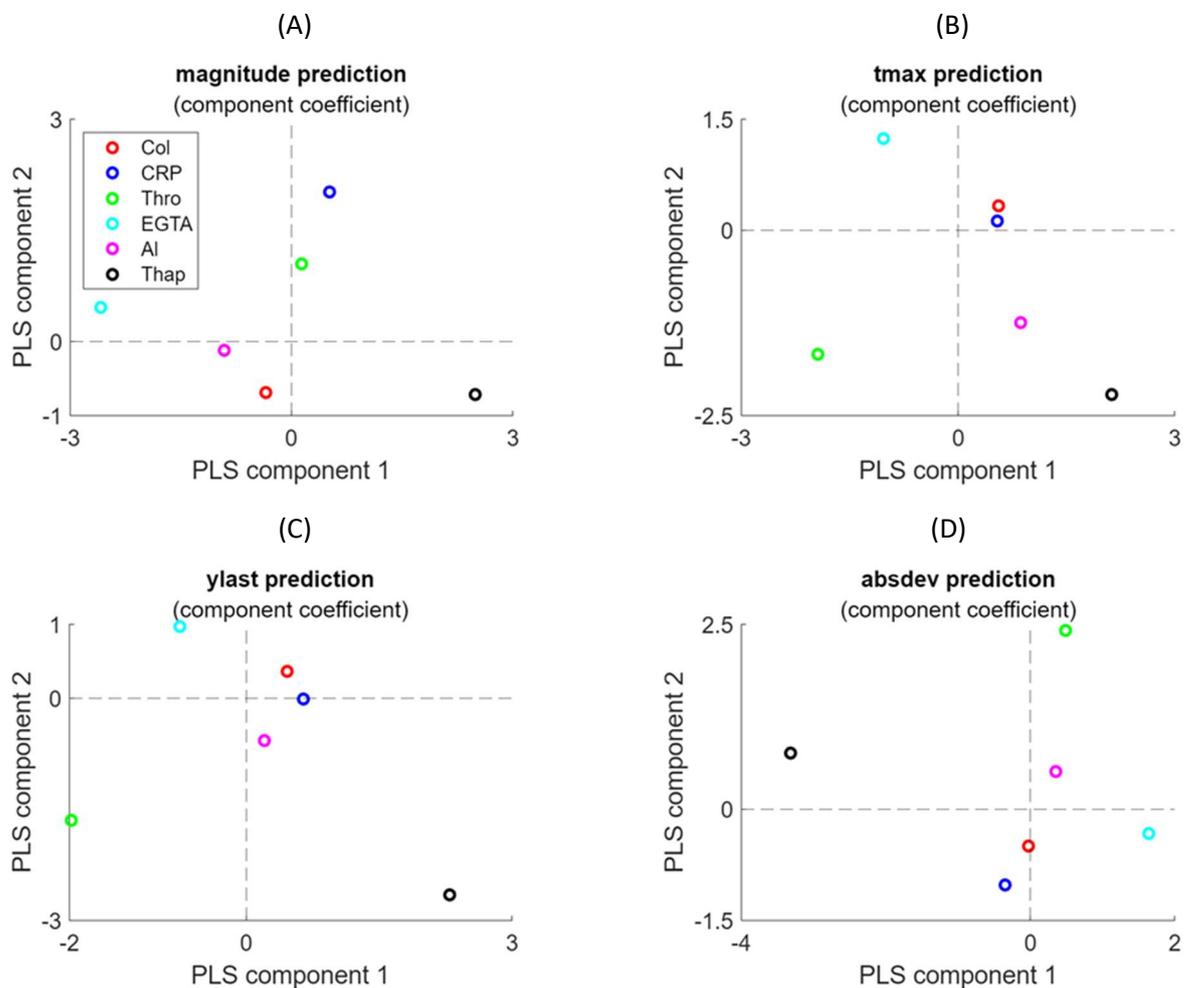


Figure 6. Loading coefficients of experimental variables in PLS regression analysis. Plots show for the first2 principal components the loading of the six experimental variables (collagen dose, CRP dose, thrombin dose, EGTA/CaCl₂, AI, thapsigargin). The PLS regression analysis was performed for prediction of curve magnitude (A), tmax (B), ylast (C) and absdev (D). Indicated in colours are the contributions per variable.

the 2-component PLS regression for curve *magnitude*, *tmax*, *y_{last}* and *absdev*, using the same training of 58 experimental conditions, while keeping the remaining 14 (previously validation and test sets) as test set of the PLS model. The variable loading coefficients of each PLS component are shown in Figure 6. This type of regression analysis was then used to predict the scalar characteristics of the test set. It also generated regression errors of both the training and test sets, which provided information on overfitting (Figure S13).

In agreement with the analysis above, it appeared that the first PLS component in the *magnitude* prediction had a negative loading in the presence of EGTA and/or AI, indicating a lower level of $[Ca^{2+}]_i$ (Figure 6A). On the other hand, the presence of thapsigargin resulted in a highly positive loading, due to an increased $[Ca^{2+}]_i$ level. Indeed, the presence of EGTA stopped the entry of extracellular Ca^{2+} , whereas thapsigargin increased this process by inhibiting the SERCA-type Ca^{2+} pumps controlling the STIM1-Orai1 entry pathway¹⁷. Furthermore, the *t_{last}* and *y_{max}* predictions showed an opposite loading in component 1 for thrombin (negative) and thapsigargin (positive) (Figure 6B-C). This reflected the more transient, non-linear Ca^{2+} responses with thrombin in comparison to those with thapsigargin. Regarding the *absdev* prediction, the thrombin condition showed a particularly high positive weight in component 2 (Figure 6D). The NARX model showed that the patterns of calcium curves modeled were accurately predicted, while the PLS model provided valuable understandable data regarding the significance of each variable.

Discussion

The combination of modelling approaches presented in this work introduces a new way to predict the response pattern of agonist-induced platelet Ca^{2+} responses under a great variety of conditions. The constructed neural networks by MLP and NARX were able to produce

mostly correct magnitude curves of $[Ca^{2+}]_i$, whereas the modelling by PLS regression captured the characteristic curve shape. Our work thereby adds to the idea of a platelet Ca^{2+} calculator introduced by Diamond and colleagues¹⁴, in that also curve patterns can be predicted without mathematical modelling. On the other hand, we did not consider the synergistic effects of agonist combinations such as presented in that study.

It is important to note that while the present machine-learning techniques were able to fit most of the input data, the obtained output does not give a direct biological interpretation, although sensitivity analysis was used to improve the interpretability. This contrast to other modelling approaches with a clear biological meaning, such as enzyme or receptor reaction rates in ODE-based kinetic models. However, the latter approaches cannot easily capture interactions between individual signalling steps, for instance due to combinations of agonists and inhibitors.

Both the NARX network and the PLS regression analysis yielded useful results for the $[Ca^{2+}]_i$ curve analysis. Thus, the up to magnitude differences between traces in the presence of $CaCl_2$ or EGTA and $CaCl_2$ (caused by Ca^{2+} entry into the platelets) were captured by both the MLP and PLS regression models. The prediction results - *i.e.* sensitivity analysis for MLP and PLS component analysis for PLS - were well interpretable for this case. On the other hand, NARX could better than PLS capture the curve effects by certain experimental variables. The curve magnitude and other characteristic effects (*tmax* and *absdev*) caused by thapsigargin, was also captured by NARX, but not by PLS regression. This illustrates that neural networks as NARX can easily handle non-linear effects function due to the complex activation functions, while PLS relies in linear regression analysis and hence cannot adapt to non-linearity.

A specific problem encountered was the different shapes of the $[Ca^{2+}]_i$ curves used for training by the various approaches, *i.e.* more often transient with thrombin and usually linear

with CRP or collagen. Although neural networks can capture any function, they need many data to train for such curve differences. In our case, only a limited number of curves with either agonist could be used for training, which caused a certain imbalance in the training set. One way to fix this problem is to use data augmentation, for example by the synthetic minority oversampling technique (SMOTE)²³, which is more often used for imbalanced datasets.

In the present paper, we trained all models using platelets derived from a single donor stimulated with a range of agonists and inhibitors, which thus resulted in a new tool for investigating the complex Ca^{2+} signalling pathways in single donor platelet activation. The models can now be used to generate hypotheses for additional experimentation and to provide insights that are otherwise not obtained by traditional analytical approaches. However, appropriate use of the models is important, ensuring that the data used for training are representative, while independent data are available for validation. The use of blood from a single donor can be seen as a limitation of the study, also because this reduced the number of variable experimental conditions and, accordingly, the machine learning models had a limited predictive power. Comparing the platelet responses from multiple donors will increase the number of samples available for model building, and may thereby decrease the accuracy of the predictions for each donor.

A solution to this issue is the approach of transfer learning²⁴, in which a generic model can be built for the samples from various donors, and then refine the model to obtain adjusted the weights for each donor separately. This approach is being used to build personalised models for drug development²⁵. An alternative is to train an auto-encoder in learning from a reduced part of the input, this to reproduce the output; this will also allow training on the data from several blood donors. Regardless of the approach followed, modelled analysis will be

interesting of the effects of additional inhibitors of relevant Ca^{2+} signalling pathways, such as P2X_1 Ca^{2+} channel antagonists²⁶ or STIM1-Orai1 pathways blockers²⁷.

Differently from the neural network models, the PLS regression analysis performed better with the relatively small sample size from one blood donor. The PLS approach is also less prone to overfitting. The present PLS regression analysis to predict the (scaled) $[\text{Ca}^{2+}]_i$ curve features easily allows for comparisons with the platelets from more donors. In work of the Diamond laboratory¹⁴, a NARX model was generalised by fitting multiple networks constructed from several donors, and the determining their average prediction. Our analysis indicates that this can be done more easily by PLS regression approaches.

Author Contributions

Conceptualisation, Methodology and Formal Analysis, CT, JLD, RC; Investigation, CT, HYFC; Resources and Supervision, JMG, JWMH, RC; Data Curation, CT, HYFC; Writing – Original Draft Preparation, CT, RC; Writing – Review & Editing, CT, JWMH, RC; Funding Acquisition, JMG, JWMH.

Acknowledgments and Funding

This work has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 766118.

Disclosures

JWMH is advisor of the Synapse Research Institute Maastricht. The other authors declare no relevant conflict of interest.

References

1. Patel, S. R., Hartwig, J. H. & Italiano, J. E. The biogenesis of platelets from megakaryocyte proplatelets. *J. Clin. Invest.* **115**, 3348–3354 (2005).
2. Chesterman, C., Owe-Young, R., Macpherson, J. & Krilis, S. Substrate for endothelial prostacyclin production in the presence of platelets exposed to collagen is derived from the platelets rather than the endothelium. *Blood* **67**, 1744–1750 (1986).
3. Davis, G. E. & Senger, D. R. Endothelial extracellular matrix: biosynthesis, remodelling, and functions during vascular morphogenesis and neovessel stabilisation. *Circ. Res.* **97**, 1093–1107 (2005).
4. Van der Meijden, P. E. & Heemskerk, J. W. Platelet biology and functions: new concepts and clinical perspectives. *Nat. Rev. Cardiol.* **16**, 166–179 (2019).
5. Jerjes-Sanchez, C. Venous and arterial thrombosis: a continuous spectrum of the same disease? *Eur. Heart J.* **26**, 3–4 (2005).
6. Jackson, S. P. Arterial thrombosis-insidious, unpredictable and deadly. *Nat. Med.* **17**, 1423–1436 (2011).
7. Mammadova-Bach, E., Nagy, M., Heemskerk, J. W., Nieswandt, N. & Braun, A. Store-operated calcium entry in blood cells in thrombo-inflammation. *Cell Calcium.* **77**, 39–48 (2019).
8. Versteeg, H. H., Heemskerk, J. W., Levi, M. & Reitsma, P. S. New fundamentals in hemostasis. *Physiol Rev.* **93**, 327–358 (2013).
9. Watson, S. P., McConnell, R. T. & Lapetina, E. G. The rapid formation of inositol phosphates in human platelets by thrombin is inhibited by prostacyclin. *J. Biol. Chem.* **259**, 13199–13203 (1984).
10. Daniel, J. L., Dangelmaier, C. A., Selak, M. & Smith, J. B. ADP stimulates IP₃ formation in human platelets. *FEBS Lett.* **206**, 299–303 (1986).
11. Oury, C., Toth-Zsamboki, E., Thys, C., Tytgat, J., Vermeylen, J. & Hoylaerts, M. F. The ATP-gated P2X₁ ion channel acts as a positive regulator of platelet responses to collagen. *Thromb. Haemost.* **86**, 1264–1271 (2001).
12. Capra, V., Bäck, M., Angiolillo, D. J., Cattaneo, M. & Sakariassen, K. S. Impact of vascular thromboxane prostanoid receptor activation on hemostasis, thrombosis, oxidative stress, and inflammation. *J. Thromb. Haemost.* **12**, 126–137 (2014).
13. Dolan, A. T. & Diamond, S. L. Systems modelling of Ca²⁺ homeostasis and mobilisation in platelets mediated by IP₃ and store-operated Ca²⁺ entry. *Biophys. J.* **106** (2014).
14. Chatterjee, M. S., Purvis, J. E., Brass, L. F. & Diamond, S. L. Pairwise agonist scanning predicts cellular signalling responses to combinatorial stimuli. *Nat. Biotechnol.* **28**, 727–732 (2010).
15. Gilio, K., Munnix, I. C., Mangin, P., Cosemans, J. M., Feijge, M. A., van der Meijden, P. E.,

- Olieslagers, S., Chrzanowska-Wodnicka, M. B., Lillian, R., Schoenwaelder, Koyasu, S., Sage, S. O., Jackson, S. P. & Heemskerk, J. W. Non-redundant roles of phosphoinositide 3-kinase isoforms alpha and beta in glycoprotein VI-induced platelet signalling and thrombus formation, *J. Biol. Chem.* **284**, 33750-33762 (2009).
16. Jooss, N. J., De Simone, I., Provenzale, I., Fernandez, D. I., Brouns, S. L., Farndale, R. W., Henskens, Y. M., Kuijpers, M. J., ten Cate, H., van der Meijden, P. E., Cavill, R. & Heemskerk, J. W. Role of platelet glycoprotein VI and tyrosine kinase Syk in thrombus formation on collagen-like surfaces. *Int. J. Mol. Sci.* **20**, e2788 (2019)
 17. Cheung, H. Y., Zou, J., Tantiwong, C., Fernandez, D. I., Huang, J., Ahrends, R., Roest, M., Cavill, R., Gibbins, J. M. & Heemskerk, J. W. High-throughput assessment identifying major platelet Ca²⁺ entry pathway via tyrosine kinase-linked and G protein-coupled receptors. *Cell Calcium*. In press (2023).
 18. Xie, H., Tang, H. & Liao, Y. Time series prediction based on NARX neural networks: an advance approach. *Computer Sci. Proc. Int. Conf. on Machine Learning and Cybernetics.* 38950256 (2009).
 19. Hewamalage, H., Bergmeir, C. & Bandara, K. Recurrent neural networks for time series forecasting: current status and future directions. *Int. J. Forecast.* **37**, 388–427 (2021).
 20. Razavi, S. & Gupta, H. V. What do we mean by sensitivity analysis? The need for comprehensive characterisation of "global" sensitivity in earth and environmental systems models. *Water Resour. Res.* **51**, 3070-3092 (2015).
 21. Wold, S., Sjöström, M. & Eriksson, L. PLS-regression: a basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **58**, 109–130 (2001).
 22. Abdi, H. Partial least squares regression and projection on latent structure regression (PLS regression). *Wiley Interdiscip. Rev. Comput. Stat.* **2**, 97–106 (2010).
 23. Chawla, N. V. et al. (2002) 'SMOTE: Synthetic Minority Over-sampling Technique', *Journal of Artificial Intelligence Research*, (16), pp. 321–357.
 24. Neyshabur, B., Sedghi, H. & Zhang, C. What is being transferred in transfer learning? *Adv. Neural Inf. Process. Syst.* **12**, 1–12 (2020).
 25. Dana, D., Gadhiya, S. V., Surin, L. G., Li, D., Naaz, F., Ali, Q., et al. Deep learning in drug discovery and medicine; scratching the surface. *Molecules* **23**, 1–15 (2018).
 26. Bennetts, F. M., Mobbs, J. I., Ventura, S. & Thal, D. M. The P2X₁ receptor as a therapeutic target. *Purinergic Signal.* **18**, 421–433 (2022).
 27. Shower, H., Norman, K., Cheng, C. W., Foster, R., Beech, D. J. & Bailey, M. A. ORAI1 Ca²⁺ channel as a therapeutic target in pathological vascular remodelling. *Front. Cell Dev. Biol.* **9**, 1–23 (2021).

Supplemental

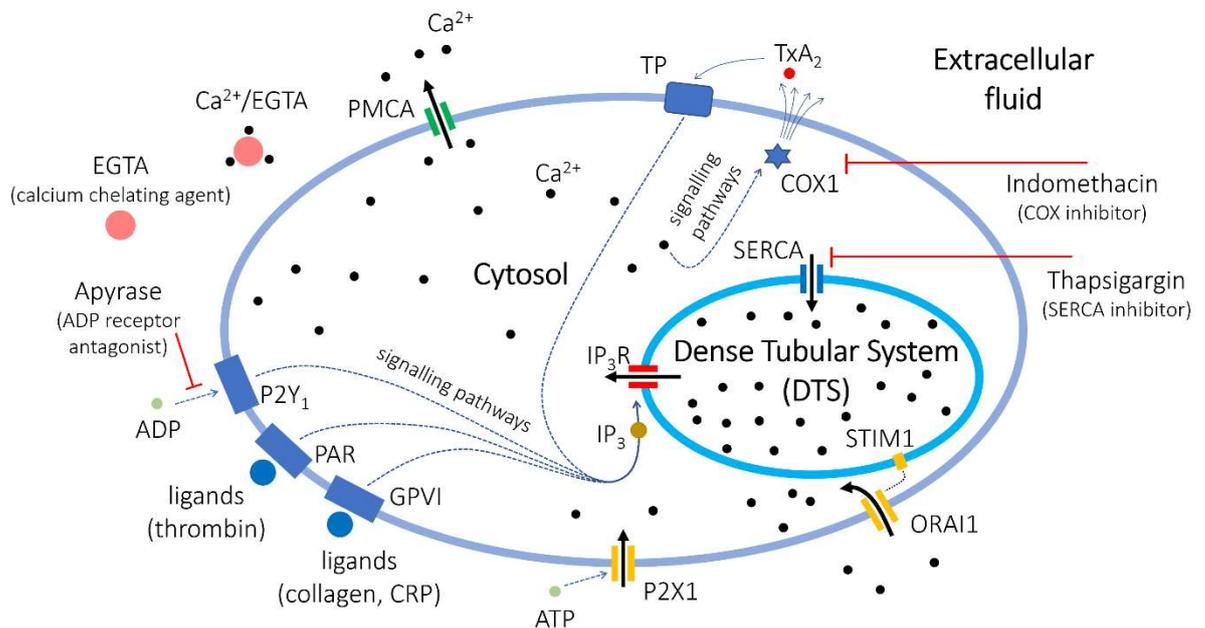


Figure S1. Overview of receptor-induced Ca^{2+} signalling mechanisms in platelets. Collagen and collagen-related peptide (CRP) activate platelet via glycoprotein GPVI (GPVI), while thrombin acts by cleaving proteinase-activated receptors (PAR). The ligands of these receptors induce the formation of inositol 1,4,5-trisphosphate (IP_3), which stimulates IP_3 receptors (IP_3R) in the membrane of the dense tubular system (DTS). This stimulation leads to discharge of Ca^{2+} from intracellular stores into the cytosol. Entry of extracellular Ca^{2+} is mediated by the Orai1 Ca^{2+} channels in the plasma membrane, which couple to STIM1 Ca^{2+} sensors in the DTS membrane. In addition, a fast and quickly desensitised entry of Ca^{2+} is mediated by ATP, activating the P2X_1 ion channels. Autocrine produced ADP and TxA_2 , via their receptors, potentiate the IP_3 production. Back pumping of released Ca^{2+} out of the cytosol occurs by the SERCA Ca^{2+} ATPases in the DTS, which are inhibited by thapsigargin. Back pumping out of the cells occurs by the PMCA Ca^{2+} ATPases. The presence of extracellular CaCl_2 or EGTA allows Ca^{2+} entry of not. The formation of TxA_2 is inhibited by indomethacin, whereas the effects of ADP are blocked by apyrase.

Raw $[Ca^{2+}]_i(t)$ traces

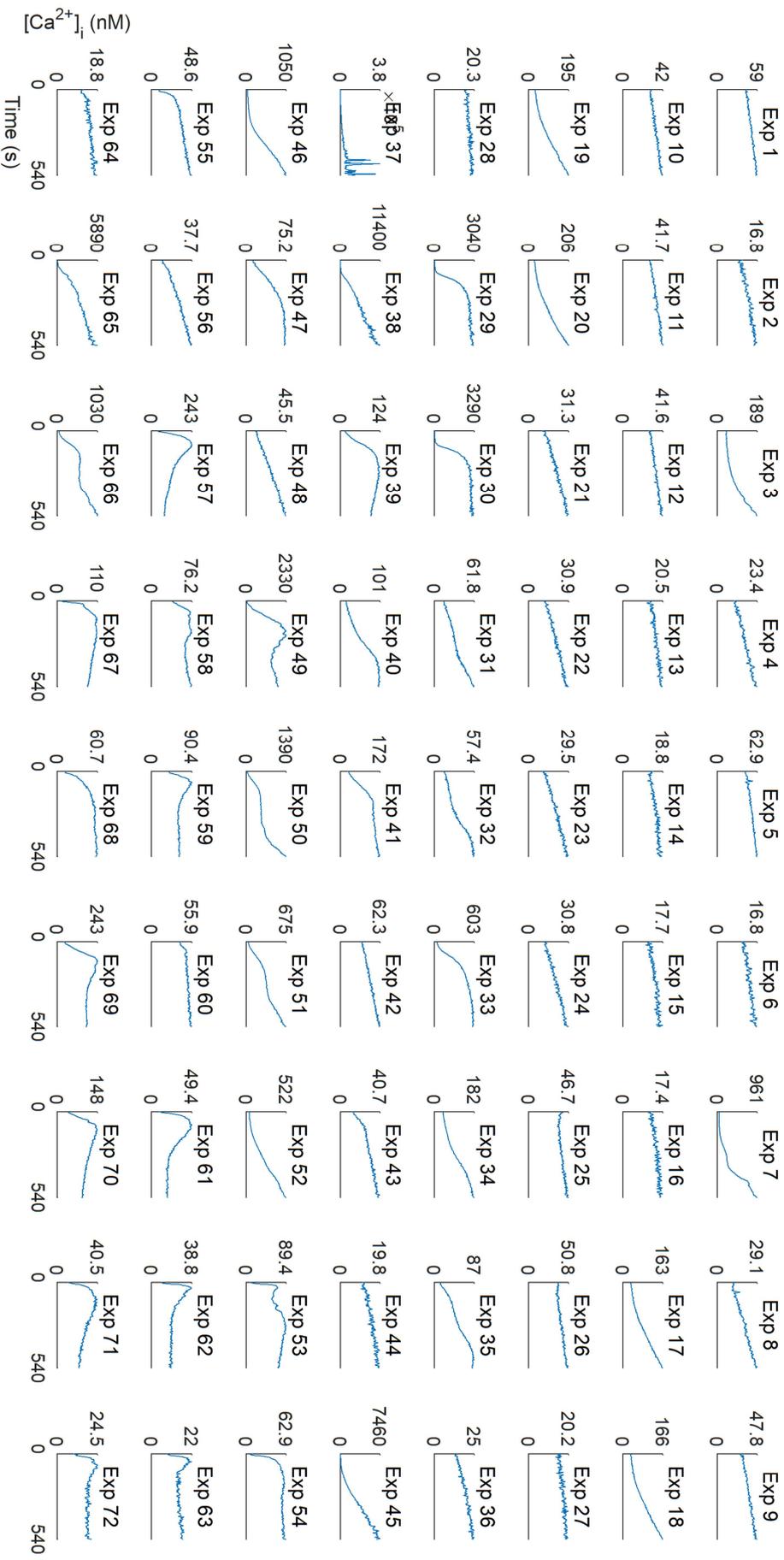


Figure S2. Raw data of agonist-induced $[Ca^{2+}]_i$ time traces of Fura-2-loaded platelets from a representative subject, used for the modelling studies. Note widely different ranges of nanomolar levels of $[Ca^{2+}]_i$ per condition. Y-axes represent linear ranges in nM $[Ca^{2+}]_i$. Time axes are 0 to 540 s.

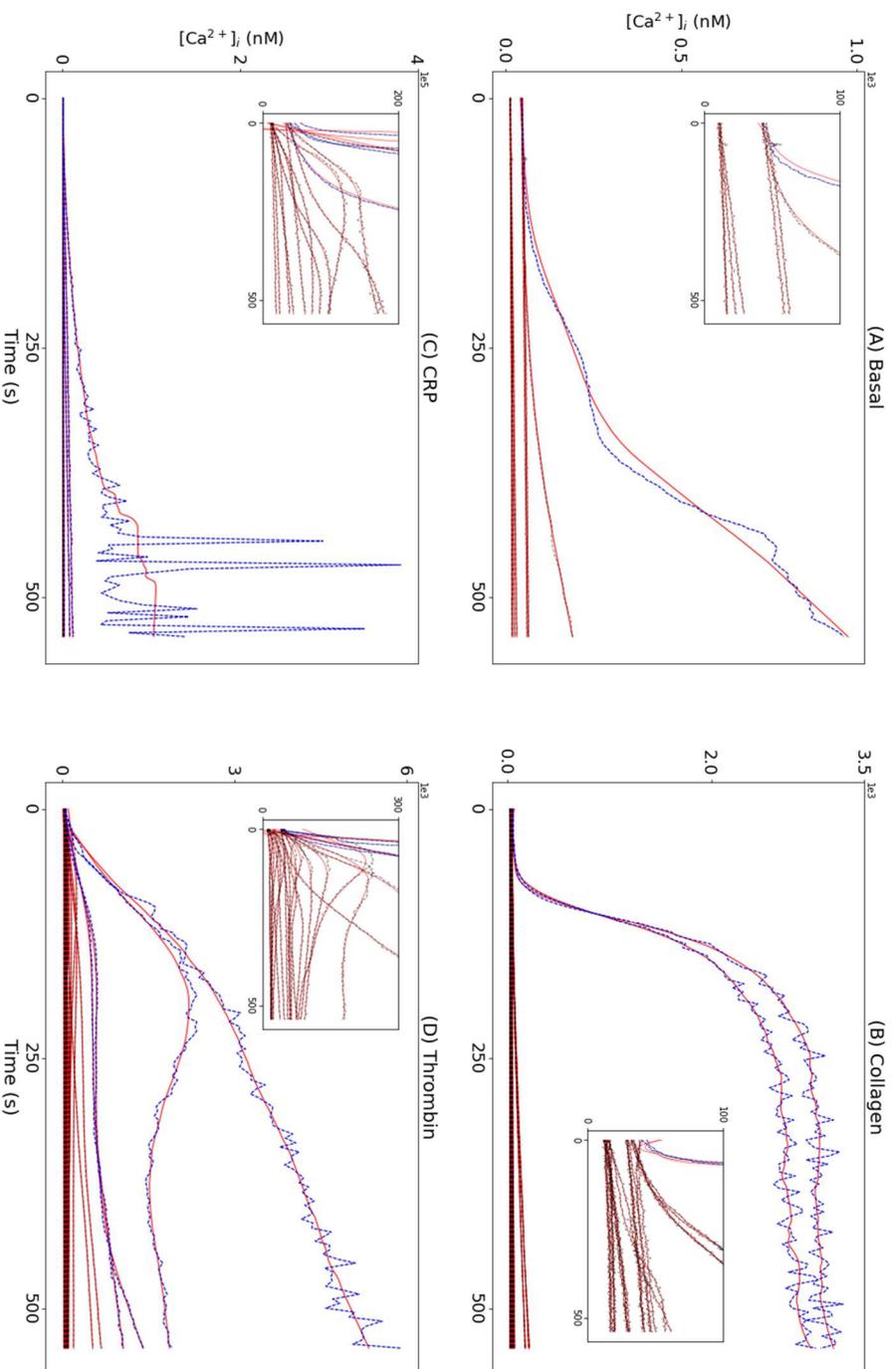


Figure S3. Raw and smoothed agonist-induced $[Ca^{2+}]_i$ curves (in nM) of Fura-2-loaded platelets from a representative subject. The 72 experimental conditions (Table 1) were grouped into four panels according to the agonist used: (A) basal (no agonist), (B) collagen, (C) CRP, or (D) thrombin. Resampled and interpolated curves were smoothed with a Savitzky-Golay filter. Shown are the original curves (dash black lines) and the filtered curves (red lines). Note that highest supra-micromolar rises in $[Ca^{2+}]_i$ were obtained in the presence of thapsigargin and $CaCl_2$ (dashed blue lines).

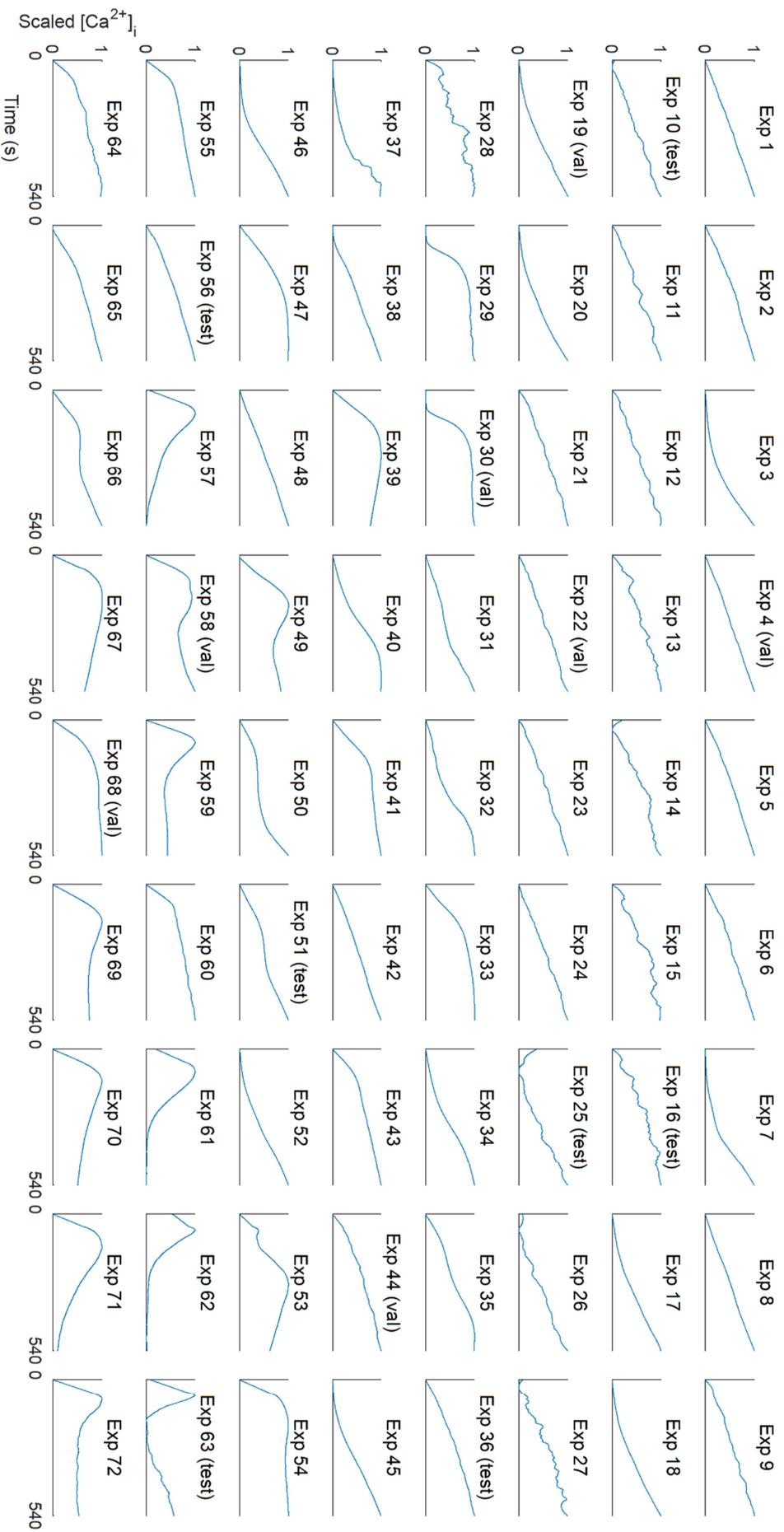


Figure S4. Scaled agonist-induced $[Ca^{2+}]_i$ time traces of Fura-2-loaded platelets (in nM) from one healthy representative donor. Numbers for experiments and agonist/treatment conditions are indicated in Table 1. Raw input data in nM were curve interpolated, smoothened and linearly scaled 0-1. Time axes are from 0 to 540 s.

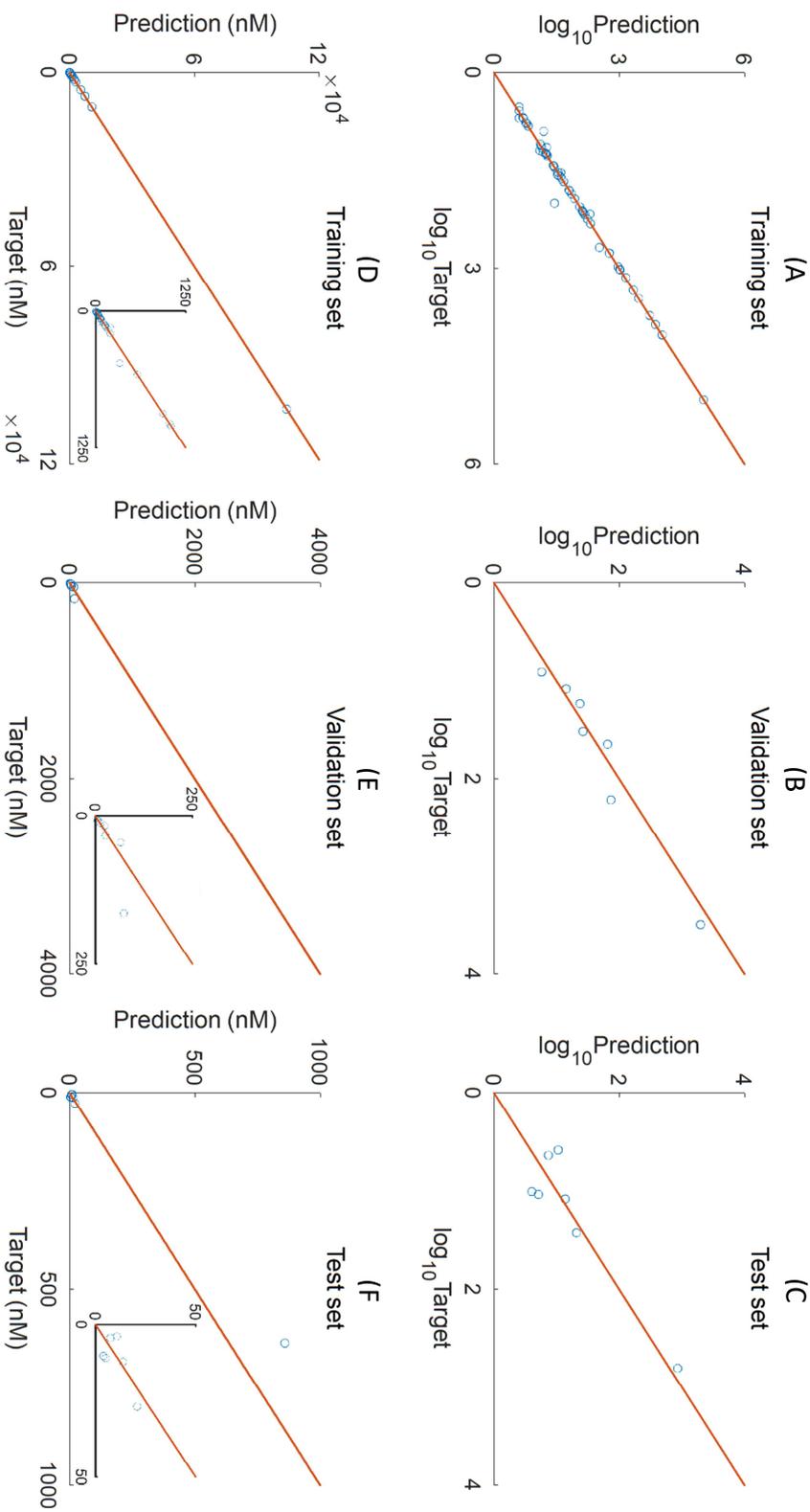


Figure S5. Magnitude prediction of agonist-induced platelet $[Ca^{2+}]_i$ responses. Shown are for time traces of the selected training set (A, D), validation set (B, E) and test set (C, F), the relation between the measured target levels and the predicted levels (nM). (A-C) Log₁₀ scale, (D-F) linear scale. Note the used curve selection for validation set (Exp. 4,19,22,30,44,58,68), test (Expt. 10,16,25,36,51,56,63), and training set (all the rest). Red lines represent diagonals.

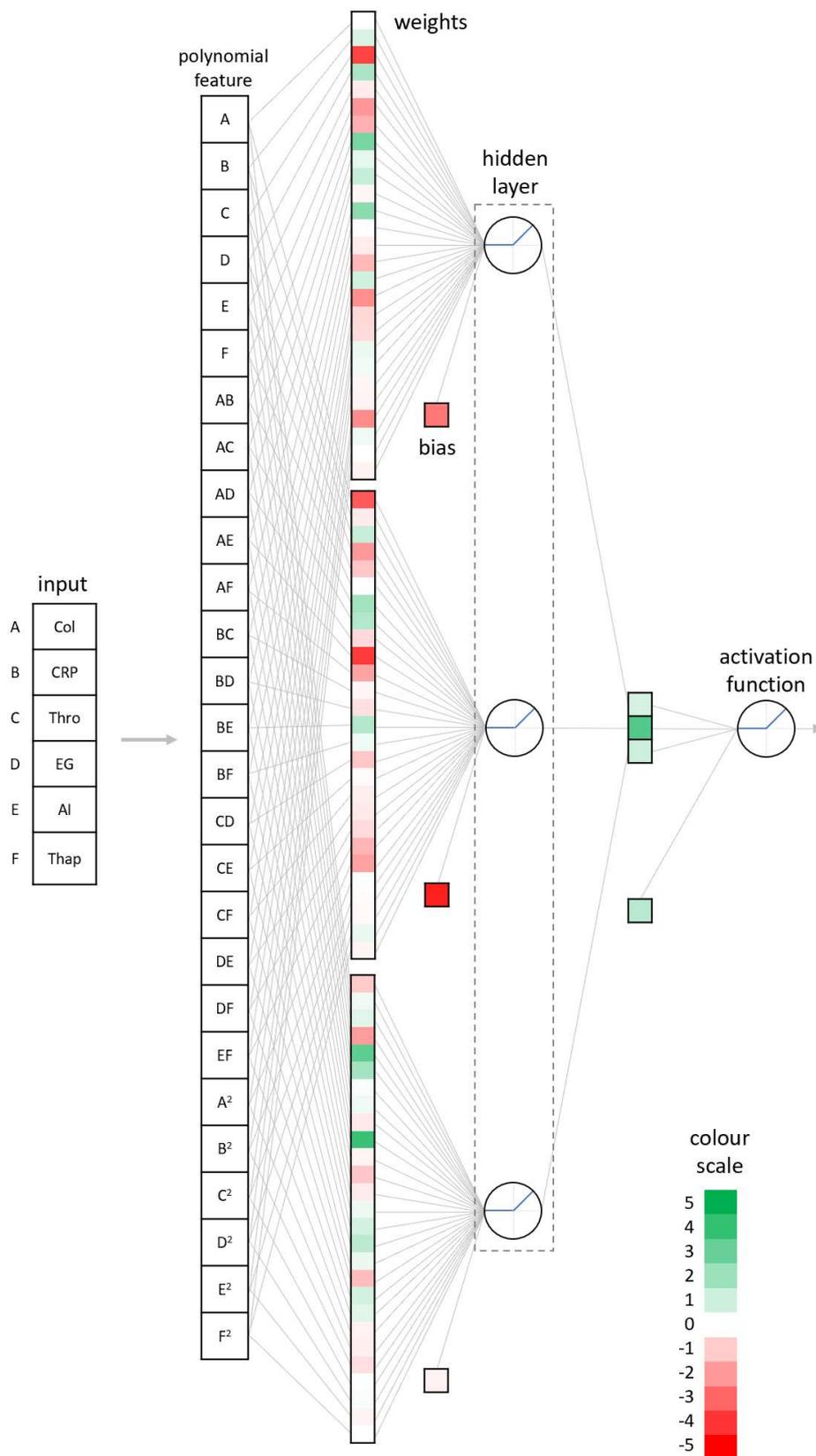


Figure S6. Parameter composition of polynomial multilayer perceptron (MLP) network associated with each node. Note the 27 combinations made from 6 input variables. Relative weights of per node are displayed in colour scale.

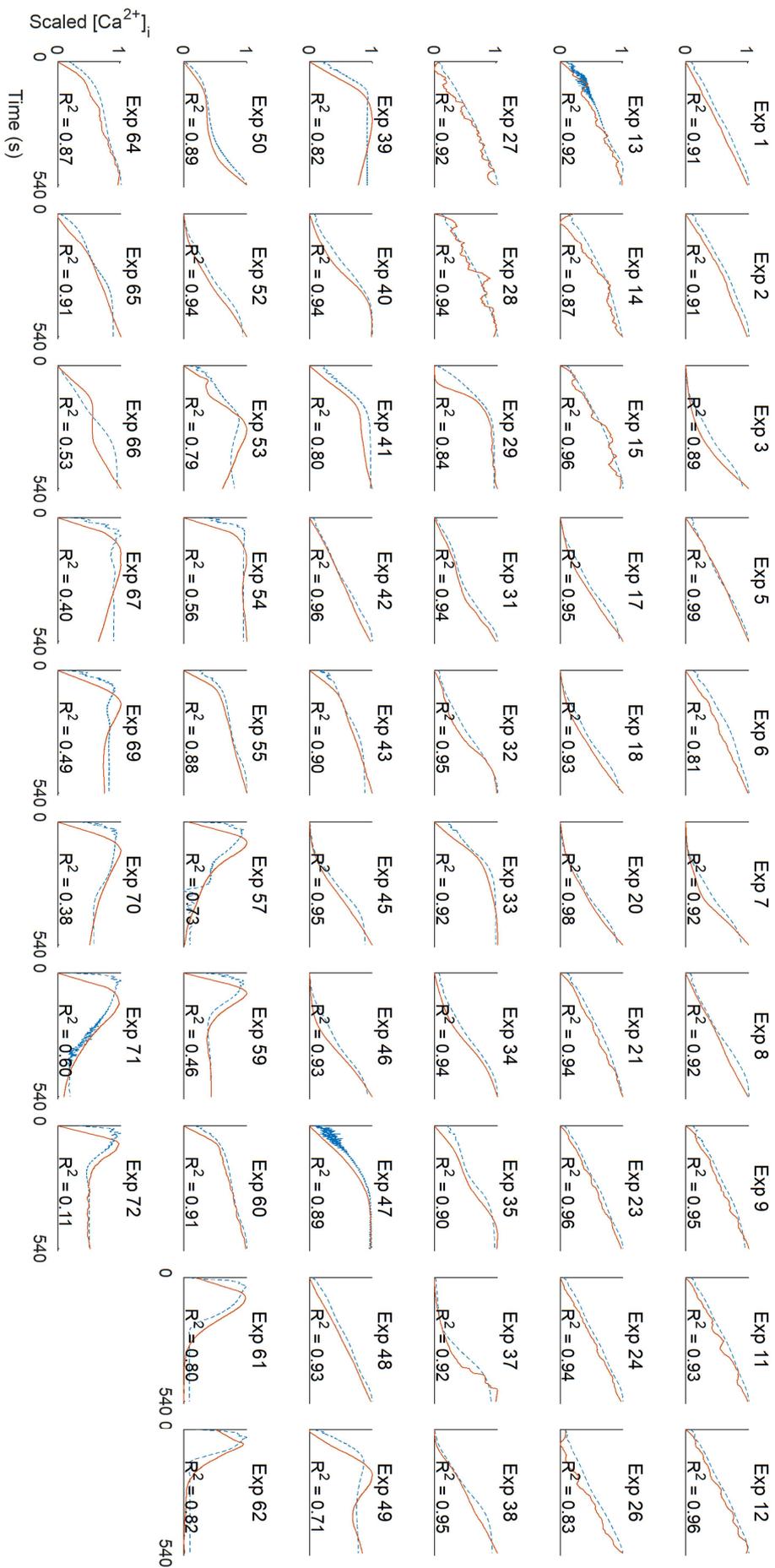


Figure S7. Trend prediction with NARX of agonist-induced platelet $[Ca^{2+}]_i$ responses in the training set. Numbers for experiments and agonist/treatment conditions are indicated in Table 1. Vertical axes represent scaled responses from 0 to 1. Horizontal axes represent measurement time from 0 to 540 s.

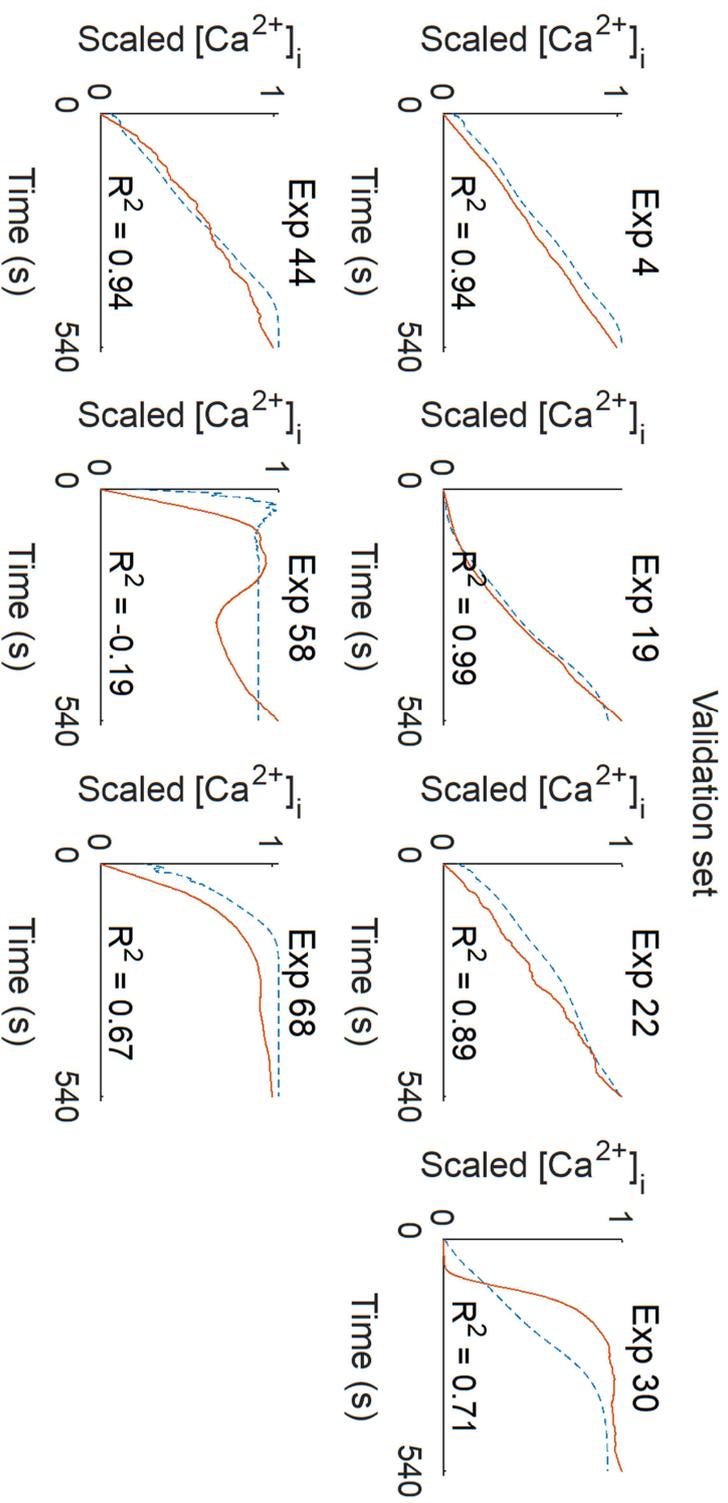


Figure S8. Trend prediction of agonist-induced platelet $[Ca^{2+}]_i$ responses in the validation set. Experimental conditions were numbered as in Table 1. Vertical axes represent scaled responses from 0 to 1. Horizontal axes represent experimental time from 0 to 540 s.

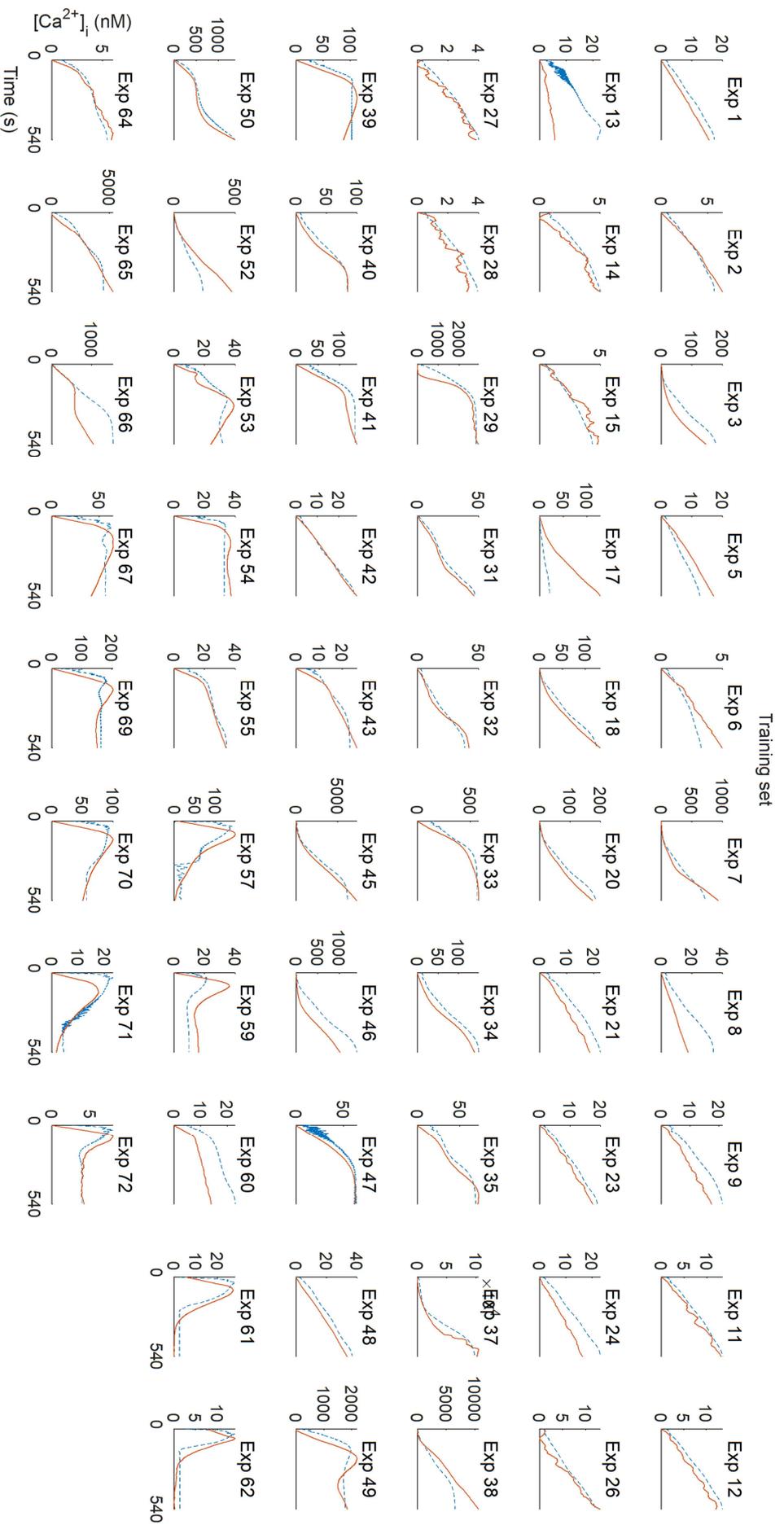
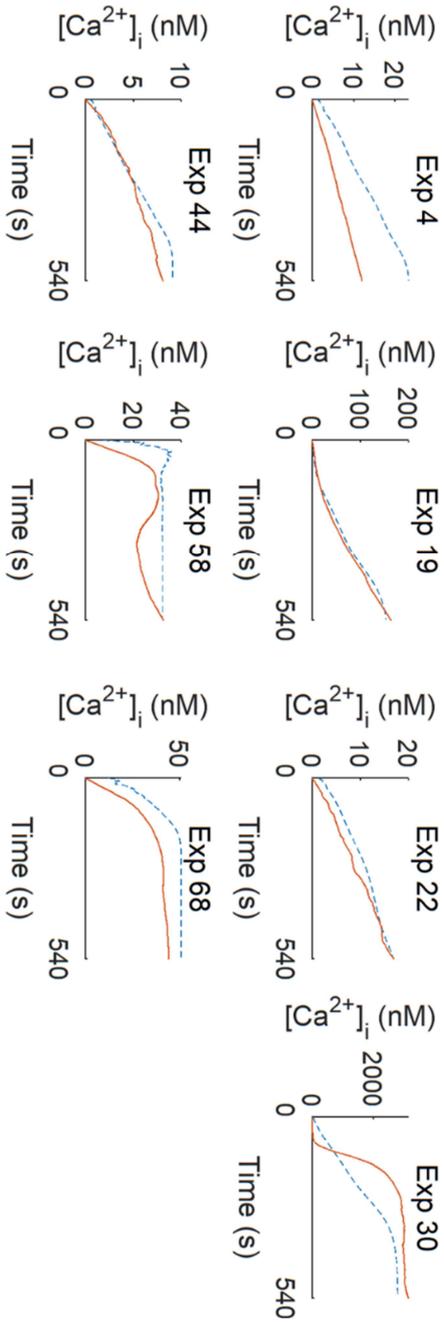


Figure S9. Magnitude and trend prediction of platelet $[Ca^{2+}]_i$ responses in training set. The predicted curves resulted from a combination of magnitude and trend predictions. Numbers for experiments and agonist/treatment conditions are indicated in Table 1. Panels indicate nanomolar $[Ca^{2+}]_i$ levels versus time (0 to 540 s).

Validation set



Test set

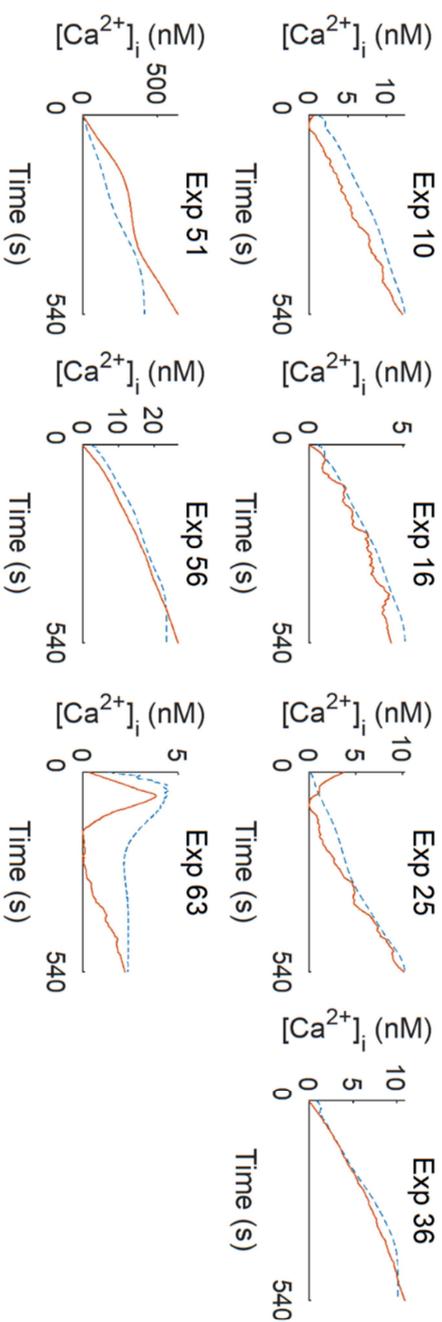


Figure S10. Combined prediction of platelet $[Ca^{2+}]_i$ responses with magnitude and trend predictions. Indicated are results from validation set (A) and from test set (B). The results from 2 models were combined (magnitude and trend prediction). Numbers for experiments and agonist/treatment conditions are indicated in Table 1. Vertical axes indicate $[Ca^{2+}]_i$ levels in nM, and horizontal axes the time.

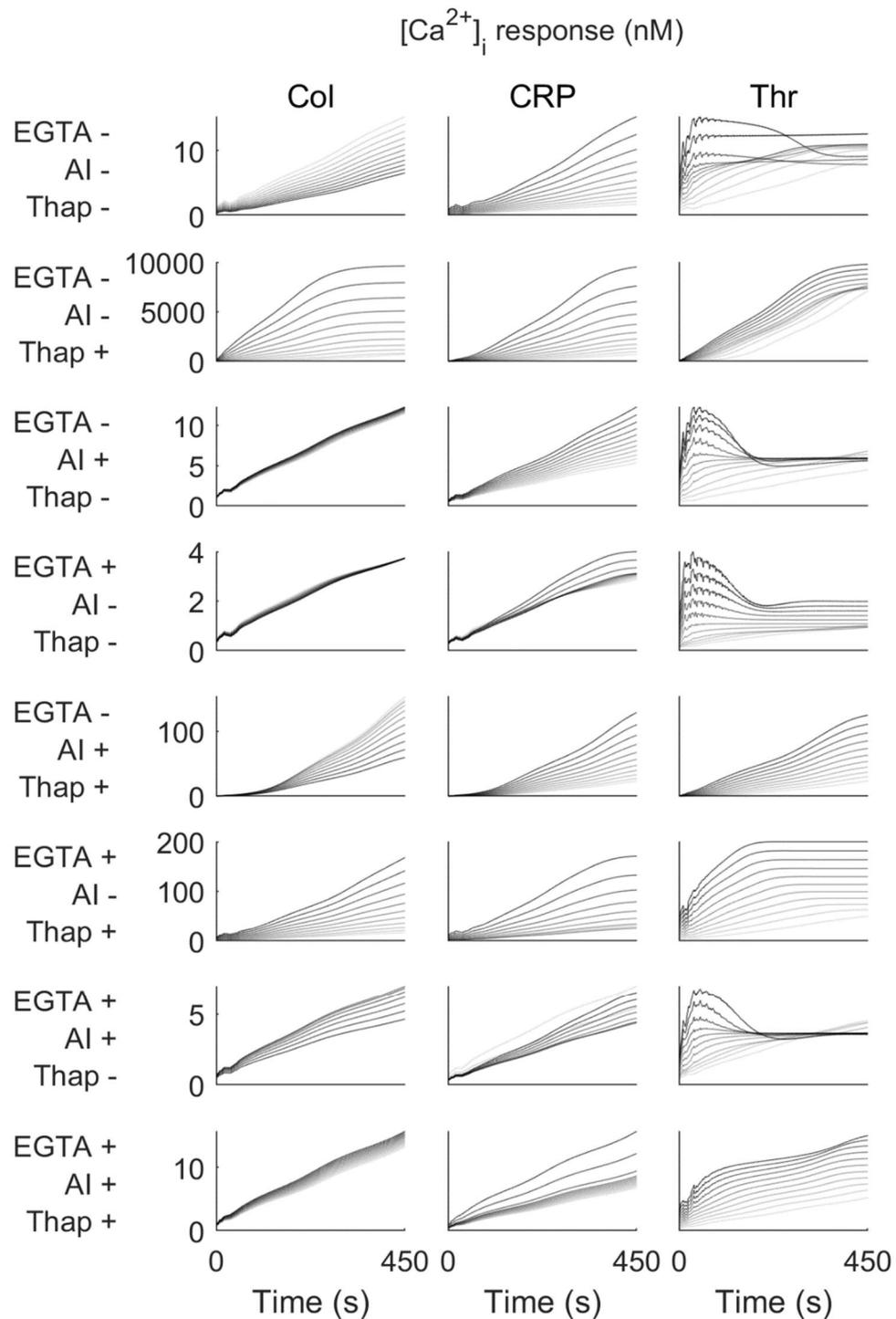


Figure S11. Variation of trend prediction of nanomolar $[Ca^{2+}]_i$ responses with increasing agonist concentrations. Panels indicate prediction efficacy per agonist concentration. Lightest grey lines represent basal levels, while darker lines represent a curve prediction due to an increment of ligand by 1% from basal level to 10% of the maximum concentration used in the training set. Columns show conditions with different agonists (collagen, Col), CRP or thrombin (Thr). Rows represent different inhibitor conditions: + or - mean presence or not; from top to bottom: EGTA, apyrase plus indomethacin (AI), and thapsigargin (Thap). Shown are unscaled levels of $[Ca^{2+}]_i$ (nM); for scaled data, see Figure 5.

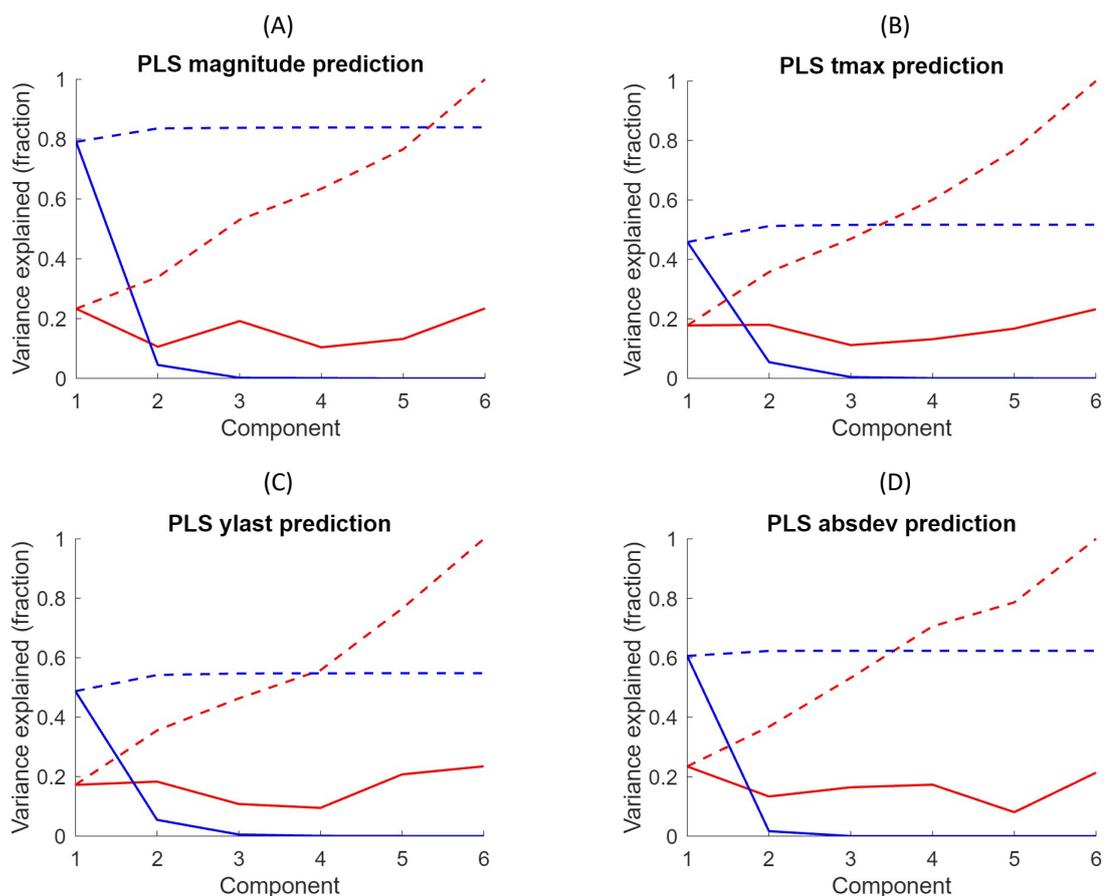


Figure S12. Variance in PLS regression model explained per component. Plots show the per principal component of PLS the fraction of explained variance of the dataset. PLS regression analysis was performed for prediction of curve magnitude (A), tmax (B), ylast (C) and absdev (D). For definition of the six included experimental variables, see Figure 6. Red lines indicate the explained variance of input (experimental condition); blue lines show explained variance of the target (curve's scalar characteristic). Dashed lines display cumulative sums for increasing components. Note that only the first 2 components contribute to most of the target variance.

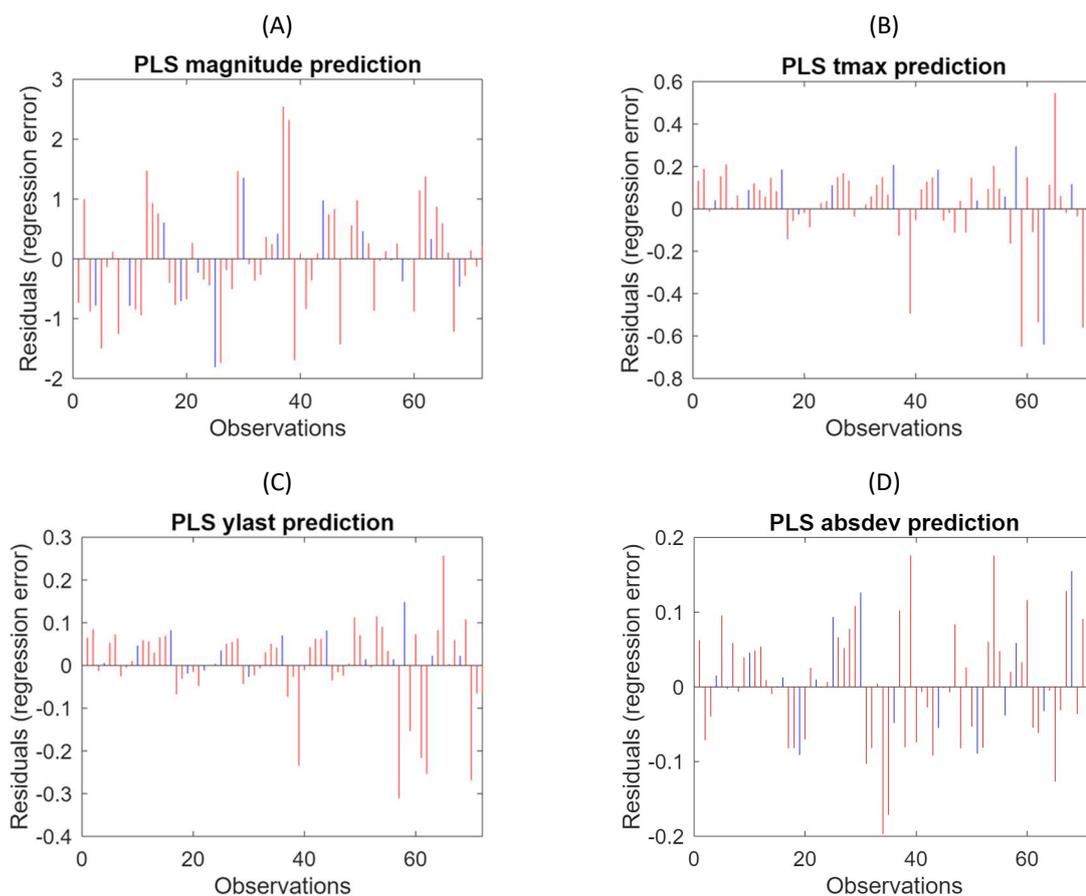


Figure S13. Residual deviations of the PLS regression model. Indicated are absolute regression errors upon increasing observations in the training set (red bars) and test set (blue bars). Regression errors were defined as the actual value minus the predicted value. Positive errors indicate an underestimate of the explained variance. PLS regression analysis was performed for prediction of curve magnitude (A), tmax (B), ylast (C) and absdev (D).

Chapter 6

General Discussion

In this chapter, a reflection is presented on the application of mathematical modelling approaches to address complex cell biology questions, using the platelet GPVI collagen receptor as a model system.

General Discussion

Overview

Transmembrane receptor clustering and signalling are crucial processes in many biological activities ranging from cell adhesion, migration, and cell-cell interactions, to control immune function and neurotransmission¹⁻³. By studying these processes, scientists can gain a better understanding of how cells interact with their environment and how different signalling pathways regulate cellular responses.

There are several reasons why the study of transmembrane receptor clustering and signalling is important. Dysregulation of receptor clustering and signalling is linked to many diseases, from chronic inflammation to cancer. In breast cancer, overexpression or mutation of human epidermal growth factor receptor 2 (HER2) lead to their constitutive activation, resulting in excessive clustering of HER2 receptors on the cell surface. This can further increase signalling cascades that promote cell proliferation and survival⁴. Understanding these processes can lead to the development of new therapies and treatments. Trastuzumab (Herceptin) is a targeted therapy for early-stage breast cancer that overexpresses HER2. It has been shown to improve outcomes by reducing the risk of recurrence and improving survival in HER2-positive early breast cancer patients when used in combination with chemotherapy and/or hormonal therapy⁵. Many drugs target these receptors and their signalling pathways. This is a recent development in the last few years. An increasing number of drugs target signalling processes rather than the traditional receptor antagonists, for example, targeting PI3K signalling pathway in HER2⁶. Studying these processes can help identify new drug targets and improve existing treatments. Understanding transmembrane receptors also provides insights into how cells function together to maintain tissue and organ function.

Platelet GPVI (glycoprotein VI) is primarily responsible for initiating platelet activation in response to damage to the blood vessel wall, which leads to the formation of a blood clot and the prevention of excessive bleeding. Biologists have studied GPVI's function and regulation and found that it is possibly becoming one of the promising therapeutic targets for thrombosis. This also serves as a model receptor and readily accessible cell to generally model receptor behaviour - in a way that could be applied to a range of cell systems. To understand the mechanisms and quantitative aspects of GPVI and platelet activation, mathematical/computational modelling was used in the present studies to help us fill these gaps and to provide a framework in which to optimise experimental design and for predicting the impact or consequences of GPVI function, its mechanism of action and the impact of perturbing these.

Our initial biological questions about GPVI and calcium regulation were focused on understanding how it occurs and which component contributes the most to its mechanism. We started by investigating GPVI's clustering on the platelet surface prior to ligand binding in Chapter 2. We asked how the clustering behaviour, such as dimerisation and localisation, depended on membrane properties and other proteins. With regards to membrane physics, diffusivity of transmembrane receptors depend on various factors, from membrane lipid composition to cytoskeleton elements. In addition to diffusivity, the presence of an inert protein on the membrane might block the movement of receptor, affecting receptor encounter rate. This raises the question of how lipid rafts and presence of inert protein could affect receptors localisation and dimerisation.

We then focused on the complexity of GPVI signalling immediately after ligand binding, which included the LAT signalosome (Chapter 3) and the phosphatidylinositol cycle (PI-cycle, Chapter 4), as we aimed to mathematically describe how these components interact. The idea

behind these models was that biological interactions and their signalling pathways might be known, but most related parameters have not been quantified experimentally. Successful modelling of signalling processes may reflect a complete understanding of the components of signalling pathways, while failure to model data may indicate important gaps – for example, missing proteins or feedback systems.

Finally, we have studied how calcium is released from an internal store of platelet due to these activations (chapter 5). To be specific, cytosolic calcium production can be exchanged through various channels, i.e. from/to internal store of platelet (DTS) via IP3 receptor and SERCA, or from/to extracellular matrix via PMCA and STIM/Orai1. These processes involve many signalling pathways and could lead to hundreds or thousands of parameters and variables. These parameters and variables may include the expression levels of different receptors and signalling molecules, the activation status of different signalling pathways, the downstream effects of these signals on cellular behaviour, and their reaction rates and thresholds. We want to investigate the calcium behaviour without constructing a complicated, overwhelming model (which easily leads to uncertainty and overfitting) while maintaining a useful prediction.

In this work, we have created computational models for the initial stages of the mechanisms that regulate the functions of platelets, from surface receptor mobilisation to internal calcium release. Combined, these models utilised several branches of mathematics and modelling techniques. They may seem unrelated at first sight and might raise the question of why each part of biology needs a different model, but they are actually a continuous process and easier to address with different approaches. In this chapter, we will discuss how those models are connected, or can be connected, in terms of biology and modelling. We will also see their applications, limitations, and possibilities to improve them in the future.

How models can be helpful for biological understanding

One important consideration for biologists is how mathematical modelling can enhance biological research, particularly in the field of platelet biology. By integrating mathematical and physical principles with experimental biology, modelling provides a powerful tool for studying platelets in various ways. Rather than being a separate entity from experimental biology, modelling complements and enhances it by allowing researchers to make predictions and test hypotheses *in silico* before carrying out costly and time-consuming experiments. This also sometimes allows complex questions to be asked which would be difficult or impossible to explore in the research laboratory. While it is true that biology and physics operate on different levels, and there may be philosophical questions regarding their complete representation of biological systems using mathematical approaches, although modelling provides a valuable framework for bridging the gap between these fields and gaining deeper insights into platelet behaviour. In the following section, we discuss examples of how the models presented in this work can be applied to platelet biology research.

Models are able to provide an estimation of how a system will behave in response to different experimental conditions, such as changes in ligand concentration, the presence of other signalling molecules, or the introduction of inhibitors or activators of platelet activation. Through this simulation, the model could predict how the system would respond to these changes and provide insight into the underlying mechanisms of platelet behaviour. The prediction might not be exact or fully quantitative, but it provides powerful information helpful for experimental design. Agent-based models allow users to experiment with a different range of parameters, i.e., artificial membrane/receptor, to observe receptor clustering behaviour. The research done by Chen et al.⁷ is a good example of how receptor dimerisation is studied. In their work, artificial transmembrane receptors and their signalling were designed by

modular DNA scaffolds and triggered the dimerisation by external signal input. The process induced the expression of a peroxidase-like enzyme and facilitated an output signal, which can be observed by fluorescence or absorbance. This work provides a generalisable method for an artificial transmembrane receptor design. Our agent-based model (ABM) can be used to design (and can be validated by) similar experiments.

In our other studies, the GPVI signalling and Ca^{2+} mobility models enabled us to investigate the magnitude of how each component affects the outcome. These could be helpful for a real-world experiment to estimate the range of related ligand/inhibitor concentrations to achieve the desired results. For example, we could design a set of experiments varying possible concentrations of CRP that could increase calcium response by the expected order of magnitude. Another similar scheme has been illustrated by Cho et al.⁸, where experiments estimated the model parameters, and the parameter sensitivity could be conversely used for experimental design. They used TNF α -mediated NF- κ B pathway, which is involved in immunity and inflammation and cell proliferation, differentiation, and apoptosis, as a case study. They constructed a set of ordinary differential equations (ODEs) and used the Monte-Carlo (MC) simulation to analyse the parametric sensitivity. Thus, the parameters were estimated and used for designing an experiment. This technique relies on the random initialisation of parameters. It makes reaching the global (or even local) optimum harder than our Bayesian technique, which utilised the posterior distribution of parameters. However, their developed framework is still helpful and can be a good example of how modelling could be used for experimental design.

In the context of drug target identification, computational models can be used to simulate the behaviour of a biological system in response to the presence of a drug. This can provide insights into drug's mechanism of action and help identify potential drug targets.

Simply decreasing (or increasing) a component in the model and observing how this change will affect the outcome can illustrate to us what the component's inhibition could lead to. The GPVI clustering model allowed us to test various effects of lipid raft or inert proteins on GPVI dimerisation, which leads to increased affinity for collagen thereby reinforcing activation signalling. It is possible to alter these components by interrupting some processes, thus effectively inhibiting GPVI activation. The GPVI signalling and calcium model will enable us to vary the effect of each element on a whole signalling pathway or outcome. These models can help identify potential drug targets by predicting which components of the system are most affected by the drug and which interactions are most important for the drug's activity, which could act as an inhibitor of a specific protein or channel. Many researchers have used this similar technique in drug targeting for cancer. Lebedeva et al. performed a sensitivity analysis of an ODE-based model to identify anti-cancer drug targets in ErbB2/3 network⁹. Zhu et al. developed a stochastic logical model, where the activation or inhibition were modelled using digital logic, with ON/OFF states, to analyse the vulnerability of signalling components in breast cancer¹⁰. They found that vulnerabilities of some components vary depending on sub-pathways, thus requiring more personalised treatment. Additionally, some components' vulnerabilities are high and invariant, making them a more universal target.

Comparison of modelling approaches

In these studies, we have used many modelling approaches for different systems or problems. The following section describes the rationale behind our decision to pick these specific models for each problem. How do the models suit their problems?

In this work, we decided to use ABM since we must consider a spatiotemporal effect of receptors' clustering on the membrane. We also need the discretisation so receptor

behaviour can be assigned individually. Diffusion models based on the partial differential equations (PDE) do not lend themselves to studying individual receptors and their physical processes; conversely, it assumes that the receptor is a continuum concentration that can not be discretised. The deterministic dynamic, such as solving the equation of motion of each atom, molecule, or receptor ($F = ma$), would consume substantial computational power and include random force, which adds uncertainty to the simulation. On the other hand, agent-based models can assign individual sets of rules to receptors/proteins, interaction, and randomness (via stochastic movement). Thus, ABM can be considered an option that offers a compromise between an overly simplistic and an overly complicated model. More protein diffusion modelling choices and their pros and cons have been discussed in Burrage et al.¹¹, starting from an overly simplistic model, i.e., mean-field kinetic, where the whole biological system is treated as a continuum media, to an overly complicated model, i.e., molecular dynamics (MD), where all atoms are simulated with their equation of motion and interactions.

Next, we used ODE-based technique to model signalling components as it could quantify the time series response. Although the main components in most signalling pathways are well known and documented in many textbooks, their interaction parameters frequently have not been measured experimentally. If a good amount of experimental data exists, then we can fit some of these parameters. A model should be able to describe the interaction of each component rationally. Thus, these interactions can be converted into a system of ordinary differential equations using mass-action kinetics. The unknown parameters can be determined by fitting the model prediction with experimental data. This idea is similar to what has been done by colleagues¹², where the early events of GPVI-ligand signalling have been modelled, and parameters quantified. The range for parameter values can be limited by proper and possible order of magnitude, and their uncertainty can be estimated by observing their

posterior distribution. The computational models of signalling networks have been discussed step-by-step by Rangamani and Iyengar¹³. They stated that the use of computational models could help to identify the behaviour of signalling that experimental data cannot achieve. Models provide the ability to explain what is happening to different modules and components of a network at different time scales. Even for a basic signalling cascade, modelling can lead to insights that are not easily comprehensible.

Some biological processes, such as complicated cell signalling, involve too many parameters and are almost impossible to be described by conventional mathematical models. It could be a better idea to treat this kind of problem as a black box, where we know only an input and can only probe the output signal. In platelet biology, the second messenger, like Ca^{2+} , can be easily measured and studied, Ca^{2+} model can be a useful tool for understanding the platelet response. Since Ca^{2+} fluxes involve many related components and signalling pathways, including all of these could lead to a highly complicated model that involves hundreds of parameters¹⁴. Fitting these parameters could easily lead to overfitting due to parameter redundancy. Thus, a machine learning model could be a powerful tool to handle this kind of problem. Neural networks (NN) can learn complex relationships between inputs and outputs without requiring explicit knowledge of the underlying process or system, and can provide accurate predictions even when the relationship is nonlinear or difficult to describe using traditional mathematical models. In addition, recurrent neural networks like NARX (nonlinear autoregressive network with exogenous inputs)¹⁵ not only use the input from the initial condition but also from output series. It can recycle the output of the previous time step as an input for the next time step; this work as a 'memory' of the network where other inputs, i.e. experimental conditions, are treated as a constant. Diamond et al.^{16,17} have performed similar work using NARX on calcium concentration modulation although this was approached

differently through study or the synergistic effect of ligands, while in the present study we explored the combinatorial effect of ligands and channel inhibitors.

Although we carefully modelled these problems and avoided unrealistic and unreasonable assumptions (as far as we can be sure), some limitations could still be inevitable due to limited available experimental data or biological information. The following are a few limitations of our modelling approaches with specific examples from these present studies and potential mitigating measures.

First, all models make assumptions, whether due to gaps in knowledge or underlying biological availability of data, or the modelling approach itself. Some parts of the constructed models in this study were oversimplified. For example, in receptor clustering ABM, we assumed that the receptor and other proteins equally distribute in all periodic simulation boxes; we also disregarded the effect of membrane curvature. This might not be neglectable since several enzymes and proteins display curvature-sensitivity¹⁸. In the signalling model, we have not included the spatial effect of components. This might result in inaccurate, biologically implausible predictions or a wide range of possible values that could be assigned to each parameter. These uncertainties can be improved by adding more experimental data. With more data, it is possible to estimate model parameters and assess the model's ability to predict new observations accurately. For example, in ABM, this added data could be the observed receptor distribution on platelet membrane, allowing us to assign a proper rule for diffusion. In ODE models, we can improve the model which lack spatial effects by adding compartmentalisation, i.e., separating the variable between the surface, cytosol, and nucleus, as we have done with PI in chapter 4. However, we still need to disregard diffusive and advective processes assuming they are fast compared to other reaction rates in the system. Using Virtual Cell computational modelling and simulation software (VCell), the simulation of

spatial PDE and compartmental ODE models were shown to be approximately identical in that case¹⁹.

Another idea is to convert an ODE to a PDE model. Levy²⁰ has highlighted how PDE-based model could add spatial terms to ODE-based model, and what additional parameters/variables are needed to generalise the model, for example, localisation of scaffold protein and spatial gradient of concentration. In his work, the author started considering the most simplistic case, the symmetric spherical cell with one degree of freedom – distance from the centre (1D PDE model). Even for this most straightforward case, solving an analytical solution was complicated.

A problem that arises from model assumptions, which often oversimplify the system, is there are missing parameters. In ODE model, we have disregarded some components or intermediate steps in signalling pathway. This might render the model incapable of capturing some behaviour that can be observed experimentally. For example, we have not included an intermediate step of IP₃ recycling pathway in our GPVI signalling model, which should be IP₃ → IP₂ → IP₁ → I or even more complex intermediate steps²¹. To overcome this problem, we would need to add at least 2 more parameters and, therefore 2 more differential equations. However, adding more parameters could lead to overfitting. This problem, again, can be solved by producing more experimental data for the fitting, especially with additional components. In addition, we can assess their relative quality using Akaike information criteria (AIC)²² to find the optimal candidate model. AIC is a mathematical method for determining how well a model fits the data expressed as a score. The AIC score is not only improved by the better fit of the model but also impaired by redundancy parameters. Thus, it can balance the goodness of fit and penalty for overwhelming parameters and, therefore, can be used for model selection.

Although we tried to base our developed models in this work on experimental data as much as possible, data availability could limit our models from being generalisable. The receptor clustering model used estimated data to construct a theoretical model that can only predict trends. In contrast, the GPVI signalling model used phosphorylation data from many donors to fit the ODE parameters. The latter, in principle, can predict a better representation of the population since it relies on actual experimental data compared with those models based on rough estimation. The calcium prediction model, on the other hand, relies on only the data from a single donor. Although it can give a useful and interpretable fitting, its result is preliminary as it cannot be generalised to other donors. On the other hand, Ca^{2+} model might not represent population well due to variability in data of human samples; it would probably be easier to utilise.

There are examples of machine learning models that are trained on data from a single individual and can generalise to a larger population, and their performance can be compared with models trained on data from many individuals. One such example is the use of transfer learning²³, where a model trained on data from a single person can be fine-tuned on a larger dataset from a population to leverage the learned representations. This approach has been applied in various domains, including healthcare, where models trained on data from individual patients can be fine-tuned on larger datasets from diverse populations to improve generalisation performance.

For instance, in a study by Miotto et al.²⁴, a deep learning model was trained on electronic health record (EHR) data from individual patients and then fine-tuned on a larger dataset from a population to predict disease outcomes. The study showed that the model trained on individual-level data followed by fine-tuning on population-level data outperformed

models trained only on individual-level or population-level data, indicating the potential benefits of leveraging both types of data in a sequential manner.

The connection between models

In this thesis, we have worked on four different sections of platelet biology. There is no clear boundary between models. The constructed models divided different aspects of biological complexities into pieces for the sake of simplicity. Although we have modelled all these problems with different approaches, they are biologically connected. We use the labels (A to E), as in Figure 1, to represent each model.

First, the key connection between GPVI clustering (A) and early signalling model (B) is the ligand binding in its dimeric form. As dimeric GPVI is able to bind its ligand, i.e. collagen or CRP, with higher affinity than in monomeric form, we are convinced²⁵⁻²⁷ that the dimerisation of GPVI is essential for its activation. Thus, the dimerisation study can be viewed as an early stage of GPVI signalling prior to the ligand binding.

Soon after the ligand binding, GPVI is activated and initiates a signalling cascade through Syk phosphorylation as modelled previously¹² (model B, Figure 1). This work continues expanding the last model by dividing the downstream signalling into 2 main parts regarding the availability of the data: LAT signalosome (C) and PI-cycle (D) studies. The first one relies on phosphorylation data of LAT, PLC, Btk, and PKC, with less detail of the PI-cycle, while the latter one used quantification of PI, PI4P, PI(3,4)P2, PI(4,5)P2, and PI(3,4,5)P3, which focus on PI-cycle and treat other parts of the signalling pathway as empirical input function obtained from experiment. These 2 parts used similar techniques and complimented each other. Thus, they are pieces of jigsaws that can be glued together to form a complete GPVI signalling pathway model.

In our GPVI signalling model (B and C), the last molecular specie that was included in the pathway is IP₃, which could lead to Ca²⁺ release via IP₃-receptor on the DTS. There are many IP₃R-Ca²⁺ models available. For example, Sneyd and Dufour²⁸ used saturating kinetic and system of ODEs to describe interaction between different states of IP₃R. However, as there are also many possible ligands and inhibitors that could alter the Ca²⁺ signal in addition to GPVI, we developed the model (E) that treats all complex signalling pathways as a black box, and shortcuts the ligand binding to Ca²⁺ production. This held the advantage of being able to assess calcium production and ligands/inhibitors conditions while minimising the number of variables/parameters. GPVI model is a small section of the black box with a detailed pathway which well constructed and parameterised.

What are the missing pieces?

Although it can be seen that the models presented in this thesis are biologically connected, combining all these models into a single complete model might not be plausible due to the lack of techniques for modelling connections. In the following section, the missing pieces of information encountered while modelling GPVI function and signalling are discussed.

Firstly, although models developed in this work are meaningful and able to predict interpretable results, they are far from perfect. The first piece we can easily spot is the absence of a ligand binding simulation in (A). Instead, we varied the binding and dissociation rate (k_b and k_d , respectively) to simulate the effect of ligand by assuming that the ligand could stabilise dimeric form of receptor (decrease k_d), and also induce dimerisation (increase k_b). However, no physical agent of ligand, i.e. collagen, has been introduced in the model. The idea of adding ligands which lie in a different medium, i.e. moving in 3-dimensional extracellular space instead of a 2-dimensional membrane, would lead to the introduction of another level of

complexity and associated limitations. The idea of ABM with compartmentalisation has been implemented in different cell lines. In studies by Das et al.²⁹, an ABM of ErbB signalling was implemented as 5 compartments: extracellular domain, plasma membrane, early endosome, late endosome, and lysosome. While ligand has been modelled as a moving agent in extracellular domain, receptors were moving on plasma membrane with assigned ligand affinity.

Secondly, while GPVI early signalling model (B, ODE-based) describes the signalling pathway of GPVI inside platelets, it simplifies ligand (collagen or CRP) binding to receptor (GPVI) using simple mass-action kinetics. Our available clustering model (A, ABM-based) does not yet connect to the signalling model. For this, we would need to connect this missing piece from ligand binding to GPVI dimerisation, which could include spatial effects of receptor clustering within the GPVI activation and signalling model.

Thirdly, the current work on GPVI model separates signalling pathways into 2 parts: PI-cycle (D) and LAT signalosome (C). The previously developed model (A)¹² focused on the connection between GPVI-ligand binding and Syk phosphorylation, and our LAT signalosome model continues that connection to IP₃. However, the PI part of our LAT model was simplified, and we have developed another separate model for this complexity. Although the input from LAT signalosome, feeds into the PI-cycle model as a mathematical function, we have not made a combined model that connects these 2 parts as a single large model. Once the models are completed and well-parameterised, this issue should be easily solved. To achieve this goal, more time course results of other components might be needed, for example, unphosphorylated species of LAT, BTK, PLC, and PKC, IP₃ and its degradation, and missing components in PI-model, i.e., PI3P, PI5P, and PI(3,5)P₂.

Finally, IP_3 production is a step that can be predicted by the GPVI signalling model developed in this study (C and D). By combining IP_3 production with IP_3 -receptor model, which has been developed by Sneyd et al. ³⁰, we should be able to generate the time course Ca^{2+} response to GPVI. Since GPVI is one of the explanatory variables used in our calcium black box model, it is possible to validate its result using the GPVI signalling model by varying the GPVI concentration. However, this method can only capture the variation in calcium concentrations due to GPVI, and cross-validate between 2 models (ODEs for GPVI and NN for Ca^{2+} (E)). It cannot be used to validate Ca^{2+} production model from other ligands/receptors, i.e. PAR1-model³¹. The proper way to assess these different models is to test them with experimental results.

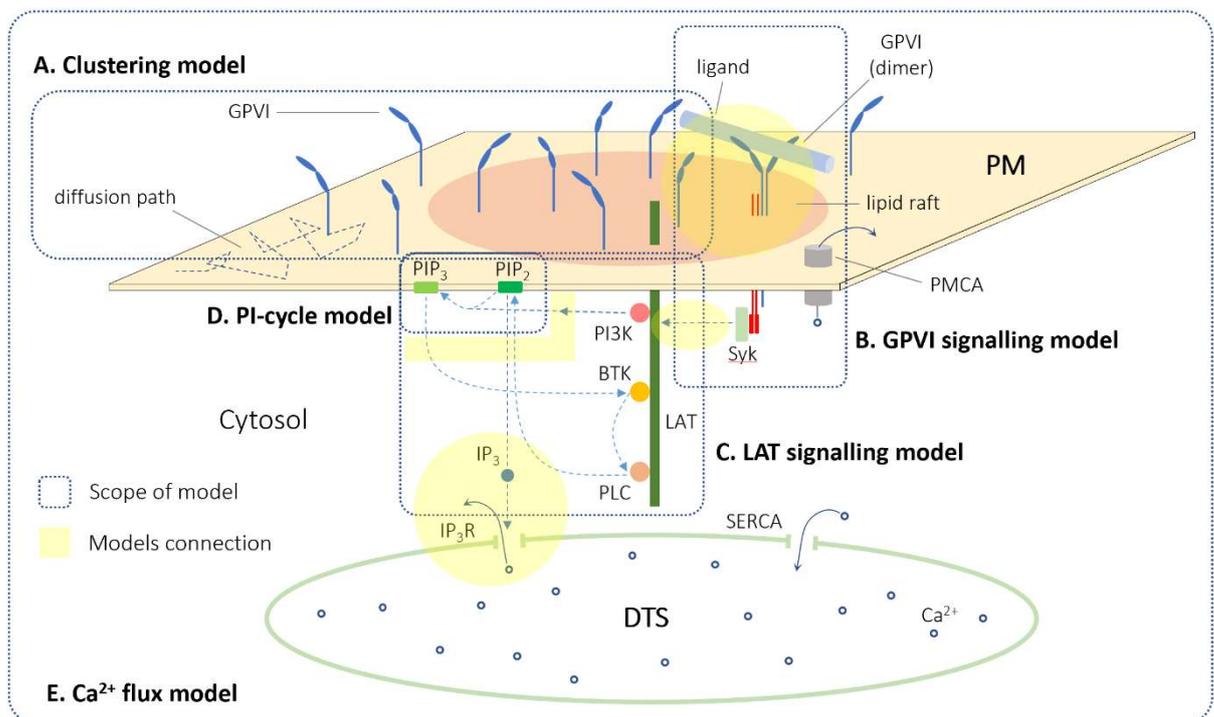


Figure 1. Summary of models in this thesis. The scope of each model is framed by a dash rectangle and labelled with bold text. The connection between models is highlighted with faded yellow.

Model	Approach	Chapter	Connections	Missing pieces
A. Receptor clustering	ABM	2	Dimerisation (A) → GPVI activation (B)	Ligand binding
B. GPVI early signalling	ODEs	previous work ¹²	Syk phosphorylation (B) → LAT activation (C)	Spatial effect of receptor clustering
C. LAT signalling	ODEs	3	Activated PI3K (C) → input signal for PI cycle (D)	Complete GPVI signalling model
D. PI cycle signalling	ODEs	4	IP ₃ production (C/D) → IP ₃ R on DTS (Ca ²⁺ release) (E)	
E. Ca ²⁺ flux	NN + PLSR	5	Ligands binding (A) → Ca ²⁺ flux (E)	Other receptors' model validation

Table 1 Summary of models in this thesis. The table shows the approaches used, the (possible) connection, and the missing pieces in each model. Abbreviations: ABM, agent-based modelling; ODEs, ordinary differential equations; NN, neural network; PLSR, partial least squares regression.

Future development

The use of mathematical and computational models has become increasingly popular in the study of biological systems. In this section, we discuss potential future developments in the validation, data requirements, and modelling approaches for the platelet models presented in this thesis. Specifically, we explore experimental data that could be used for model validation and examine whether more data are needed for modelling. Additionally, we discuss alternative modelling approaches, including agent-based modelling (ABM) for cellular signalling and multiple donor calcium modelling. These approaches have the potential to enhance our understanding of platelet behaviour and could contribute to the development of novel therapeutic strategies.

There are a few experiments that can be used to validate the models presented in this thesis. The receptor clustering model is theoretical and designed to show the effect of each parameter variation. It can be validated by experiment, i.e. synthesising an artificial membrane, to test the prediction by observing receptor behaviour with single-particle-tracking techniques. The GPVI signalling model can be validated by either variation of ligand concentration or by using inhibitors that alter specific pathways; phosphorylation outcomes may be investigated and compared to the predictions made by the model. There are many possibilities to validate the calcium model by varying the concentrations of each ligand, channel inhibitors, and their combinations.

Within the receptor clustering model, most parameters are arbitrary, and the framework of this model is theoretical rather than experiment-based. The data that could improve the model should be able to be parameterised as model input, i.e. receptor binding and dissociation rates. More phosphorylation data of various components in the signalling pathway could improve GPVI signalling model by narrowing down the posterior parameter range, thus reducing the error of parameter estimation. In calcium modelling, where all signalling complexity is treated as a black box, more data could lead to a more accurate model. For example, calcium time course due to all possible combinations of ligand and inhibitor concentration could improve the accuracy of model prediction due to synergistic effect of ligands/inhibitors. It is also possible to fit these models for several donors, and find the average to achieve the prediction of the pool.

Are there any other modelling approaches to explore? By extending the current ABM model for receptor clustering to GPVI signalling downstream, each component in signalling pathway can be converted into a model's agent with assigned interaction parameters obtained from the ODE model. This model might be able to capture the spatial behaviour of lipids (2D

movement in membrane) and proteins (3D movement in cytosol). The multiscale modelling technique can be used as a full model that connects GPVI clustering and signalling from end-to-end process. These 2 models can also be combined by interchanging their simulation outcome via connected components, i.e. activated receptors. This hybrid modelling idea has been implemented to predict pressure ulcer formation by Solovyev et al.³², where the ABM (used for skin injury, inflammation, and ulcer formation modelling) and ODE model (used for blood flow modelling) are connected via specific variables (blood flow and external pressure).

For current calcium model, the limitation is that it is based on data from a single donor. This might not capture the overall trend of calcium response for the average population. The next challenge is to extend the model to cover multiple donors. One straightforward way to do this is using the data from multiple donors, constructing the model for each donor with personalised parameters, and making a prediction from their average. In addition to transfer learning, which we mentioned earlier, Bernardo-Faura et al.³³ have developed another approach using multiple individual patient models to identify the drug combination that could convert the signal in an 'unhealthy state' to a 'healthy-like' state. Although the idea is not to generalise the model to cover all donors, it can be used to obtain similar applications, i.e., to identify the therapeutic target.

Concluding remarks

The models in this work have many applications, from experimental design to drug targeting. In this thesis, we created ABM for transmembrane receptors, which can be generalised to other receptors/proteins or cell lines, not limited to GPVI on the platelet surface. The model provides a useful prediction for receptor localisation and dimerisation due to the various conditions of their diffusivity on membrane and the addition of inert

transmembrane proteins. We developed an equation-based model and quantified signalling parameters for LAT signalosome and PI-cycle, which lie in between the GPVI signalling after activation and IP3 production, which leads to platelet internal calcium release. Finally, we simplified the complicated platelet calcium mobilisation due to ligands and channel inhibitors/antagonists as a machine learning model such as NN and PLSR. These models can generate meaningful results and are useful for trend analysis. However, the ODE models developed relied on a small blood sample from a small group of donors, while the calcium work is developed from single donor data, and the ABM work is a theoretical approach based on estimated parameters obtained from many sources, which is hugely varied. This means that our models cannot be generalised to large populations. The scope of these works is also limited to the factor included in the models; they cannot be used for other purposes, i.e. predict the effect of other components.

The findings of this thesis highlight the potential of cross-disciplinary modelling approaches for understanding and predicting biological processes. At the same time, there is still room for improvement, such as incorporating more experimental data to enhance model performance and expanding the scope of prediction through spatiotemporal modelling or combining different modelling approaches. The future priorities for further development of this approach involve utilising modelling to guide biological data acquisition in order to provide the necessary data for parameterisation and testing of informative questions. This may apply greater openness to using models to ultimately advance our understanding of complex biological systems.

References

1. Alberts, B., Johnson, A., Lewis, J., Morgan, D., Raff, M., Roberts, K., *et al.* *Molecular biology of the cell.* (Garland Science, 2015).
2. Mellman, I. & Nelson, W. J. Coordinated protein sorting, targeting and distribution in polarized cells. *Nat. Rev. Mol. Cell Biol.* **9**, 833–845 (2008).
3. Tony, P. & John D., S. Protein phosphorylation in signaling - 50 years and counting. *Trends Biochem. Sci.* **30**, 283–286 (2005).
4. Moasser, M. M. The oncogene HER2; Its signaling and transforming functions and its role in human cancer pathogenesis. *Oncogene* **45**, 6469–6487 (2007).
5. Moja, L., Tagliabue, L., Balduzzi, S., Parmelli, E., Pistotti, V., Guarneri, V., *et al.* Trastuzumab containing regimens for early breast cancer. *Cochrane Database Syst. Rev.* **2021**, (2012).
6. Yip, H. Y. K. & Papa, A. Signaling pathways in cancer: Therapeutic targets, combinatorial treatments, and new developments. *Cells* **10**, 1–30 (2021).
7. Chen, H., Zhou, L., Li, C., He, X., Huang, J., Yang, X., *et al.* Controlled dimerization of artificial membrane receptors for transmembrane signal transduction. *Chem. Sci.* **12**, 8224–8230 (2021).
8. Cho, K. H., Shin, S. Y., Kolch, W. & Wolkenhauer, O. Experimental Design in Systems Biology, Based on Parameter Sensitivity Analysis Using a Monte Carlo Method: A Case Study for the TNF α -Mediated NF- κ B Signal Transduction Pathway. *Simulation* **79**, 726–739 (2003).
9. Lebedeva, G., Sorokin, A., Faratian, D., Mullen, P., Goltsov, A., Langdon, S. P., *et al.* Model-based global sensitivity analysis as applied to identification of anti-cancer drug targets and biomarkers of drug resistance in the ErbB2/3 network. *Eur. J. Pharm. Sci.* **46**, 244–258 (2012).
10. Zhu, P., Aliabadi, H. M., Uludag, H. & Han, J. Identification of Potential Drug Targets in Cancer Signaling Pathways using Stochastic Logical Models. *Sci. Rep.* **6**, 1–11 (2016).
11. Burrage, K., Hancock, J., Leier, A. & Nicolau, D. V. Modelling and simulation techniques for membrane biology. *Brief. Bioinform.* **8**, 234–244 (2007).
12. Dunster, J. L., Mazet, F., Fry, M. J., Gibbins, J. M. & Tindall, M. J. Regulation of early steps of GPVI signal transduction by phosphatases: A systems biology approach. *PLoS Comput. Biol.* **11**, 1–26 (2015).
13. Sible, J. C. & Tyson, J. J. Mathematical modeling as a tool for investigating cell cycle control networks. *Methods* **41**, 238–247 (2007).
14. Dolan, A. T. & Diamond, S. L. *Systems Modeling of Ca²⁺ Homeostasis and Mobilization in Platelets Mediated by IP₃ and Store-Operated Ca²⁺ Entry.* *Biophysical Journal* vol. 106 (2014).
15. Xie, H., Tang, H. & Liao, Y. Time Series Prediction Based on NARX Neural Networks: An Advance Approach. in *Proceedings - International Conference on Machine Learning and Cybernetics*

- (2009).
16. Chatterjee, M. S., Purvis, J. E., Brass, L. F. & Diamond, S. L. Pairwise agonist scanning predicts cellular signaling responses to combinatorial stimuli. *Nat. Biotechnol.* **28**, 727–732 (2010).
 17. Lee, M. Y. & Diamond, S. L. A human platelet calcium calculator trained by pairwise agonist scanning. *PLoS Comput. Biol.* **11**, 1–24 (2015).
 18. McMahon, H. T. & Boucrot, E. Membrane curvature at a glance. *J. Cell Sci.* **128**, 1065–1070 (2015).
 19. Cowan, A. E., Moraru, I. I., Schaff, J. C., Slepchenko, B. M. & Loew, L. M. Spatial Modeling of Cell Signaling Networks. *Methods Cell Biol.* **110**, 195–221 (2012).
 20. Levy, C. Modelling the spatial effects of the signal transduction process. (Dalhousie University, 2015).
 21. Majerus, P. W. Inositol phosphate biochemistry. *Annu. Rev. Biochem.* **61**, 225–250 (1992).
 22. Akaike, H. A New Look at the Statistical Model Identification. *IEEE Trans. Automat. Contr.* **19**, 716–723 (1974).
 23. Neyshabur, B., Sedghi, H. & Zhang, C. What is being transferred in transfer learning? *Adv. Neural Inf. Process. Syst.* **2020-Decem**, 1–12 (2020).
 24. Miotto, R., Li, L., Kidd, B. A. & Dudley, J. T. Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records. *Sci. Rep.* **6**, 1–10 (2016).
 25. Miura, Y., Takahashi, T., Jung, S. M. & Moroi, M. Analysis of the interaction of platelet collagen receptor glycoprotein VI (GPVI) with collagen: A dimeric form of GPVI, but not the monomeric form, shows affinity to fibrous collagen. *J. Biol. Chem.* **277**, 46197–46204 (2002).
 26. Jung, S. M., Tsuji, K. & Moroi, M. Glycoprotein (GP) VI dimer as a major collagen-binding site of native platelets: Direct evidence obtained with dimeric GPVI-specific Fabs. *J. Thromb. Haemost.* **7**, 1347–1355 (2009).
 27. Jung, S. M., Moroi, M., Soejima, K., Nakagaki, T., Miura, Y., Berndt, M. C., *et al.* Constitutive dimerization of glycoprotein VI (GPVI) in resting platelets is essential for binding to collagen and activation in flowing blood. *J. Biol. Chem.* **287**, 30000–30013 (2012).
 28. Sneyd, J. & Dufour, J.-F. A dynamic model of the type-2 inositol trisphosphate receptor. *Proc. Natl. Acad. Sci.* **99**, 2398–2403 (2002).
 29. Das, A. A., Darsana, T. A. & Jacob, E. Agent-based re-engineering of ErbB signaling: A modeling pipeline for integrative systems biology. *Bioinformatics* **33**, 726–732 (2017).
 30. Sneyd, J. & Falcke, M. Models of the inositol trisphosphate receptor. *Prog. Biophys. Mol. Biol.* **89**, 207–245 (2005).
 31. Sveshnikova, A. N., Ataulakhanov, F. I. & Pantelev, M. A. Compartmentalized calcium signaling

- triggers subpopulation formation upon platelet activation through PAR1. *Mol. Biosyst.* **11**, 1052–1060 (2015).
32. Solovyev, A., Mi, Q., Tzen, Y. T., Brienza, D. & Vodovotz, Y. Hybrid Equation/Agent-Based Model of Ischemia-Induced Hyperemia and Pressure Ulcer Formation Predicts Greater Propensity to Ulcerate in Subjects with Spinal Cord Injury. *PLoS Comput. Biol.* **9**, (2013).
 33. Bernardo-Faura, M., Rinas, M., Wirbel, J., Pertsovskaya, I., Pliaka, V., Messinis, D. E., *et al.* Prediction of combination therapies based on topological modeling of the immune signaling network in multiple sclerosis. *Genome Med.* **13**, 1–17 (2021).