Reduced Order Modeling and Analysis of Cellular Signal Transduction

Von der Fakultät für Konstruktions-, Produktions- und Fahrzeugtechnik der Universität Stuttgart zur Erlangung der Würde eines Doktor-Ingenieurs (Dr.-Ing.) genehmigte Abhandlung

Vorgelegt von

Markus Koschorreck

aus Mutlangen

Hauptberichter:	Prof. Dr. E. D. Gilles
Mitberichter:	Prof. Dr. O. Sawodny
	Prof. Dr. P. Scheurich
Tag der mündlichen Prüfung:	17.03.2009

Institut für Systemdynamik der Universität Stuttgart

2009

Vorwort

Die vorliegende Niederschrift entstand in den Jahren 2005 bis 2008 während meiner Tätigkeit am Institut für Systemdynamik (früher Institut für Systemdynamik und Regelungstechnik) der Universität Stuttgart und am Max-Planck-Institut für Dynamik komplexer technischer Systeme in Magdeburg.

Mein besonderer Dank gilt Herrn Prof. Dr.-Ing. Dr. h. c. mult. Ernst Dieter Gilles für die Überlassung des Themas, das mir entgegengebrachte Vertrauen und die Betreuung der Arbeit. Die großen eingeräumten Freiräume waren ein wesentlicher Beitrag zu dem Gelingen der Arbeit.

Ich möchte mich bei Herrn Prof. Dr.-Ing. Oliver Sawodny für die Schaffung von guten Arbeitsbedingungen und einer angenehmen Atmosphäre am Institut für Systemdynamik, die Unterstützung und die Übernahme des Koreferates bedanken.

Mein Dank gilt Herrn Prof. Dr. rer. nat. Peter Scheurich für die freundliche Übernahme des Koreferates.

Mein herzlicher Dank geht an die Systembiologie-Gruppen des Instituts für Systemdynamik sowie des Max-Planck-Instituts für Dynamik komplexer technischer Systeme, in besonderem Maße an Michael Ederer, Dr.-Ing. Holger Conzelmann und Johannes Witt für intensive und kritische Diskussionen sowie die sehr gute Zusammenarbeit. Insbesondere die signifikanten Beiträge von Holger Conzelmann und Michael Ederer zu den mathematischen Hintergründen der Modellreduktion waren von hoher Bedeutung.

Ebenfalls herzlich danken möchte ich Alexander Lutz und Peter Schumm für vielfältige Hilfen im EDV-Bereich sowie Duncan Payne und Benjamin Herzer für sprachliche Hilfe bei der vorliegenden Niederschrift.

Wertvolle Hinweise in Bezug auf die Physiologie von Ratten wurden von Prof. Dr. rer. nat. Gisela Drews, Dr. rer. nat. Jan-Michael Heinrich und Isabell Rode beigesteuert, hierfür vielen Dank.

Für die Förderung dieses im Rahmen von HepatoSys durchgeführten Projekts bedanke ich mich beim Bundesministerium für Bildung und Forschung (BMBF).

Ein besonderer Dank gilt meinen Eltern, meinem Bruder und meiner Frau, die mich immer unterstützt haben.

Stuttgart, im April 2009 Markus Koschorreck

Contents

Abbreviations **8**

Abstract **9**

Zusammenfassung **11**

1 Introduction **14**
1.1 Cellular signal transduction 14
1.1.1 The regulation of blood glucose 15
1.1.2 Insulin signaling 16
1.2 Modeling of chemical reaction networks 16
1.2.1 Underlying assumptions 17
1.2.2 Reactions 17
1.2.3 Generalized mass action kinetics 18
1.3 Combinatorial complexity 20
1.4 Limitations of conventional modeling 22
1.5 Rule-based modeling 22
1.6 Outline of the thesis 23

2 Insulin dynamics *in vivo* **26**
2.1 Biological background 26
2.1.1 Models of insulin receptor dynamics *in vitro* 27
2.1.2 Insulin dynamics and insulin receptor activation *in vivo* 28
2.2 The model 29
2.2.1 Important tissues and processes 29
2.2.2 Parameters, units and volumes 30
2.2.3 The liver 30
2.2.4 The kidney 36
2.2.5 Insulin secretion and injection 36
2.2.6 Insulin concentration in the plasma 37
2.3 Model validation 37
2.3.1 Insulin dynamics 37
2.3.2 Hepatic insulin receptor internalization 40
2.3.3 Insulin binding and receptor phosphorylation 40
2.4 Model analysis 42

2.4.1 Insulin degradation . . . 42
2.4.2 Insulin clearance . . . 43
2.4.3 Parameter estimation . . . 46
2.4.4 Sensitivity analysis . . . 47
2.5 Therapeutic insulin concentrations . . . 48
2.6 Conclusions . . . 50

3 Model reduction **52**
3.1 Intuitive model reduction . . . 52
3.1.1 Exemplification of intuitive model reduction . . . 53
3.2 General considerations about model reduction . . . 57
3.3 Matrix decomposition . . . 58
3.3.1 Eigenvalue decomposition (EVD) . . . 58
3.3.2 Singular value decomposition (SVD) . . . 58
3.4 Methods for linear systems . . . 59
3.5 Proper orthogonal decomposition (POD) . . . 60
3.6 Model reduction based on time hierarchy . . . 61
3.6.1 The quasi-steady-state approximation . . . 62
3.6.2 The rapid equilibrium assumption . . . 63
3.6.3 Enzyme kinetics . . . 64
3.7 Domain-oriented model reduction . . . 65
3.8 Conclusions . . . 67

4 Layer-based modeling **69**
4.1 Introduction to the layer-based approach . . . 69
4.2 Important concepts . . . 70
4.2.1 Definitions . . . 70
4.2.2 Processes and interactions . . . 72
4.2.3 The interaction graph and layers . . . 74
4.2.4 Mass flow and signal flow . . . 75
4.3 Modeling of layers . . . 76
4.3.1 Phosphorylation of binding sites . . . 76
4.3.2 Reactions . . . 77
4.3.3 Layer connections . . . 77
4.3.4 Reaction rates and ODEs . . . 78
4.3.5 Choice of initial conditions . . . 78
4.3.6 A small example system . . . 79
4.3.6.1 Approximation quality . . . 83
4.3.7 Synthesis, degradation and transport of proteins . . . 85
4.3.8 Step by step procedure: layer-based modeling . . . 85
4.4 Advanced strategies . . . 86
4.4.1 Equivalent binding sites . . . 86

4.4.2 One effector binding to different binding sites . . . 87
4.4.3 Additional signals between layers . . . 88
4.4.4 Domain-oriented reduction of layer-based models . . . 88
4.5 Mathematical background . . . 89
4.5.1 The nature of the approximation in layer-based modeling . . . 90
4.5.1.1 Independence of processes . . . 90
4.5.1.2 Diagonal reactions . . . 91
4.5.1.3 Non-uniqueness of diagonal reactions . . . 92
4.5.2 Reversible transformation . . . 93
4.5.3 Formal derivation of layer-based models . . . 95
4.5.4 Approximation quality . . . 95
4.5.4.1 Dynamics of the approximation error . . . 95
4.5.4.2 Input-to-state stability of the approximation error . . . 97
4.5.4.3 Evaluation of the approximation quality . . . 99
4.6 Modeling and analysis of larger systems . . . 101
4.6.1 An extended example system . . . 101
4.6.1.1 Approximation of variables . . . 102
4.6.2 Layer-based modeling of insulin signaling . . . 104
4.7 Reduction of the model size . . . 104
4.8 Conclusions . . . 107

5 Automation of layer-based modeling 109
5.1 Introduction to ALC . . . 109
5.1.1 Architecture and flow diagram . . . 110
5.2 Preparing the model definition . . . 110
5.2.1 Molecule definitions: the section '#molecules' . . . 112
5.2.2 Defining rules and reactions: the section '#reactions' . . . 114
5.2.2.1 Defining rules . . . 114
5.2.2.2 Symmetric reaction rules . . . 115
5.2.2.3 Generalized mass action kinetics for rules . . . 115
5.2.2.4 Advanced association rules . . . 116
5.2.3 Defining signal flows: the section '#layer connections' . . . 117
5.2.4 Defining parameters: the section '#parameters' . . . 117
5.2.5 Defining initial conditions: the section '#initial conditions' . . . 118
5.2.6 Constant concentrations: '#clamped concentrations' . . . 118
5.2.7 Defining conservation relations: '#molecular balances' . . . 118
5.2.8 Defining algebraic assignments: '#algebraic relations' . . . 119
5.2.9 Defining model outputs: the section '#output' . . . 119
5.2.10 Order of assignments . . . 119
5.3 Automated assignment of correction terms . . . 120
5.4 Using ALC . . . 121
5.4.1 Step by step procedure: using ALC . . . 121

5.4.2 Running ALC . . . 122
5.4.3 Resulting model files . . . 123
5.4.4 Computational aspects . . . 123
5.5 Conclusions . . . 124

6 Conclusions and outlook **125**

Bibliography **127**

A Appendix **133**
A.1 Scientific contribution of this thesis . . . 133
A.2 Software . . . 133
A.3 Insulin dynamics and insulin receptor activation . . . 133
A.3.1 Model parameters and initial conditions . . . 133
A.3.2 Dynamic model validation . . . 134
A.4 Measuring insulin clearance . . . 134
A.4.1 Analysis of plasma insulin concentration after insulin administration . 134
A.4.2 Continuous insulin infusion: euglycemic insulin clamp . . . 135
A.5 Approximation error $g(t)$ . . . 136
A.6 Layer-based model of the extended example system . . . 138
A.6.1 The receptor layer . . . 138
A.6.2 The E layer . . . 139
A.6.3 The F layer . . . 141
A.6.4 Conservation relations . . . 141
A.7 Command line parameters for ALC . . . 142
A.8 The 51 ODE model of insulin signaling . . . 144
List of tables . . . 148
List of figures . . . 149

Abbreviations

ass.	assumption
ALC	Automated Layer Construction
AUC	area under the curve
bpm	beats per minute
calc.	calculated
cpm	counts per minute
DNA	deoxyribonucleic acid
EN	in endosomes
g	gram
I1	one insulin molecule bound to the receptor
I2	two insulin molecules bound to the receptor
ID	administered insulin dose
l	liter
M	molar ($1\ M = 1\ mol \cdot l^{-1}$)
min	minute
ml	milliliter ($1\ ml = 10^{-3}\ l$)
mol	$6.02214179 \cdot 10^{23}$ molecules
nM	nanomolar ($1\ nM = 10^{-9}\ mol \cdot l^{-1}$)
$nmol$	nanomol ($1\ nmol = 10^{-9}\ mol$)
nsq	normalized sum of squares
ODE	ordinary differential equation
pH	potentia Hydrogenii (pH x $\widehat{=} 10^{-x}\ mol\ H_3O^+ \cdot l^{-1}$)
PM	at the plasma membrane
s	second
ssq	sum of squares
U	unit insulin ($1\ U = 6.17\ nmol\ insulin$)
μg	microgram ($1\ \mu g = 10^{-6}\ g$)
μM	micromolar ($1\ \mu M = 10^{-6}\ mol \cdot l^{-1}$)
μmol	micromol ($1\ \mu mol = 10^{-6}\ mol$)
μU	microunit insulin ($1\ \mu U = 6.17 \cdot 10^{-6}\ nmol\ insulin$)

Abstract

Cellular signal transduction is crucial for the regulation of many physiological processes. Understanding the signaling systems is of high medical interest because malfunctions can result in severe disorders such as cancer and diabetes. The behavior of these systems however, is often nonlinear and cannot be predicted intuitively. Therefore, mathematical modeling is necessary to understand and to analyze the system level properties of cellular signaling.

Insulin is a hormone that has a major role in the regulation of glucose concentration in the blood and the cellular energy metabolism. This thesis provides a mathematical model describing hepatic insulin receptor activation as well as insulin degradation and synthesis *in vivo*. Model analysis shows that insulin clearance and the relative contributions of the liver and the kidney to insulin degradation are highly dependent on insulin concentration. At low concentrations, insulin is mainly degraded by the liver, whereas renal insulin degradation is predominant at high insulin concentrations. Insulin clearance is therefore only a valid measure for the state of the insulin metabolism when corresponding insulin concentrations are taken into account, which is not the case in many experimental studies.

Building comprehensive models of complete signaling systems is in many cases impeded by combinatorial complexity. The association and modification of a few proteins can result in an enormous amount of feasible complexes and an equivalent amount of differential equations, when applying the conventional modeling approach. For example, $1.5 \cdot 10^8$ differential equations would be required to describe in detail the insulin signaling system, thereby establishing the need for a reduced order description.

This thesis introduces layer-based modeling, a new approximative method for the modeling of cellular signaling systems. Layer-based modeling provides high reduction of the model size and simultaneously a high quality of approximation. The errors introduced by the approximation are dynamically and ultimately bounded. In special cases, the reduced model is exact for steady states or even represents an exact minimal realization of the system. Layer-based models show a pronounced modularity and the state variables have a direct biochemical interpretation. Reduced order model equations can be generated directly employing a procedure quite similar to conventional modeling. The preceding generation of a potentially very large conventional model is not necessary, which allows for the modeling of systems not accessible previously.

Furthermore, the computer program Automated Layer Construction (ALC) is presented. Using ALC highly simplifies the generation of the model equations. The models are defined in terms of a rule-based model definition that utilizes a simple but powerful syntax. ALC allows the modeler to define layer-based models of very large systems with a relatively short and simple model definition. The output files of ALC are ready-to-run simulation files in

the formats C MEX, MATLAB, *Mathematica* and SBML. ALC also provides the model equations in LaTeX and plain text format to simplify their publication or presentation.

The application of ALC and layer-based modeling is demonstrated for a model definition for a layer-based model of insulin signaling with 51 ordinary differential equations (ODEs) approximating a conventional model with $1.5 \cdot 10^8$ ODEs.

Zusammenfassung

Alle Zellen eines Organismus sind durch Membranen von ihrer Umgebung getrennt. Dennoch müssen die Zellen auf Signale aus ihrer Umgebung reagieren, um ihre Aktivitäten an die Umweltbedingungen und den Bedarf des Organismus anzupassen. Diese Signale werden in vielen Fällen durch die Konzentration von Botenstoffen, zum Beispiel Hormonen wie Insulin, dargestellt. Die zelluläre Signalübertragung hat die Funktion, Signale aus der Umgebung in die Zelle zu übertragen, sie gegebenenfalls zu verstärken und die Aktivität von Zielmolekülen zu beeinflussen. Die ordnungsgemäße Funktion der Signalübertragungssysteme ist von höchster Wichtigkeit für das Überleben und die Gesundheit des Menschen. Fehlfunktionen können schwerwiegende Krankheiten wie Krebs und Diabetes mellitus verursachen. Obwohl viele Komponenten der Signalübertragungssysteme bekannt sind und charakterisiert wurden, sind die Funktionsweise und das Verhalten der Netzwerke in vielen Fällen noch nicht vollständig verstanden. Die Ursache dafür ist, dass die Interaktion von Komponenten mit gut charakterisierten Eigenschaften zu neuen, oft unerwarteten Eigenschaften auf der Systemebene wie zum Beispiel Oszillationen oder Bistabilitäten führen kann.

Die mathematische Modellierung ist eine systematische Herangehensweise, um diese versteckten Systemeigenschaften zu erkennen und zu verstehen. Ein Problem bei der Verwendung des konventionellen Modellierungsansatzes ist aber, dass die Assoziation von einigen wenigen Proteinen zu einer sehr hohen Anzahl von möglichen Komplexen und derselben Anzahl von für die Systembeschreibung notwendigen Differentialgleichungen führen kann. Diese kombinatorische Komplexität stellt ein enormes Problem bei der Modellierung von vielen Signalübertragungssystemen dar.

Die vorliegende Niederschrift gibt einen Einblick in die zelluläre Signalübertragung und zeigt die auftretende kombinatorische Komplexität am Beispiel des Insulinsignalsystems. Der konventionelle Modellierungsansatz wird vorgestellt, und seine durch die kombinatorische Komplexität bedingten Beschränkungen werden aufgezeigt.

Aufgrund der kombinatorischen Komplexität kann das Insulinsignalsystem nur dann auf konventionelle Weise modelliert werden, wenn die Komplexbildung stark vereinfacht dargestellt wird und sich die Modellierung auf wenige Prozesse beschränkt. Ein solches konventionelles Modell, das die Dynamik des Insulinrezeptors in Leberzellen und die daran gekoppelte, durch den Abbau und die Synthese von Insulin verursachte Dynamik der Insulinkonzentration im Blut beschreibt, wird in dieser Arbeit vorgestellt und analysiert. Es wird gezeigt, dass die relativen Beiträge von Leber und Nieren zum Insulinabbau in hohem Maße von der Insulinkonzentration abhängen.

Die Insulin Clearance ist ein häufig verwendeter Indikator für den Zustand des Insulinstoffwechsels. Es wird gezeigt, dass auch die Insulin Clearance aufgrund von Nichtlinearitäten

des Insulinabbaus von der Insulinkonzentration im Blut abhängt. Daraus folgt, dass die Verwendung der Insulin Clearance nur zu sinnvollen Aussagen führen kann, wenn die zugehörige Insulinkonzentration angegeben wird. Viele Publikationen erfassen die Insulinkonzentration nicht, daher ist die Aussagekraft der betreffenden Studien als begrenzt einzuschätzen.

Die konventionelle Modellierung von Signalsystemen mit auftretender kombinatorischer Komplexität führt zu extrem umfangreichen Modellen, wenn umfassende mathematische Beschreibungen der Systeme erwünscht sind. Eine sinnvolle und handhabbare Beschreibung dieser komplexen Systeme ist nur mit Modellen reduzierter Ordnung möglich.

Im mittleren Teil der Arbeit wird eine Einführung in die Modellreduktion gegeben. Da biologische Systeme im Allgemeinen nichtlinear sind, liegt der Schwerpunkt dabei auf Methoden für nichtlineare Systeme. Es werden allgemeine systemtheoretische Methoden für die Modellreduktion vorgestellt, die jedoch für große Systeme sehr rechenaufwändig sind und in den meisten Fällen zu Zustandsgrößen ohne physiologische Bedeutung führen. Des weiteren werden spezielle Modellierungs- und Modellreduktionsmethoden für chemische und biologische Systeme vorgestellt. Ein Schwerpunkt liegt auf der oft verwendeten intuitiven Modellreduktion, die auf der Annahme einer zeitlichen Reihenfolge der Prozesse basiert. Es wird aufgezeigt, dass diese Methode oft durch eine unzureichende Approximationsgenauigkeit gekennzeichnet ist. Im Gegensatz dazu erlauben die Annahmen von schnellem Gleichgewicht und Quasistationarität in vielen Fällen eine Reduktion der Systemordnung bei gleichzeitig hoher Approximationsqualität. Die diesen Annahmen zugrunde liegenden Konzepte und ihre Anwendung werden vorgestellt.

Ein schwerwiegender Nachteil aller systematischen, auf gewöhnlichen Differentialgleichungen beruhenden Methoden ist, dass ein konventionelles Modell generiert werden muss, bevor eine Reduktion der Modellgröße möglich ist.

In dieser Arbeit wird eine neue approximative Modellierungsmethode (layer-based modeling) beschrieben, die in hohem Maße für die Modellierung von Signalsystemen mit auftretender kombinatorischer Komplexität geeignet ist. Diese Methode zeichnet sich durch eine starke Verringerung der Systemordnung bei gleichzeitig hoher Approximationsgenauigkeit aus. Die resultierenden Modelle weisen eine ausgeprägte modulare Struktur auf, und die Zustandsgrößen haben eine direkte physiologische Bedeutung. Ein wesentlicher Unterschied zu anderen Modellreduktionsmethoden ist, dass das möglicherweise extrem große konventionelle Modell nicht erstellt werden muss, da das reduzierte Modell direkt generiert werden kann. Die für die Erstellung des reduzierten Modells notwendige Vorgehensweise weist große Ähnlichkeiten zur konventionellen Modellierung auf. Durch die direkte Erstellung der Modelle reduzierter Ordnung wird es möglich, Systeme zu modellieren, die zuvor nicht zugänglich waren.

Die Reduktion der Modellgröße gründet auf einer makroskopischen Zustandsbeschreibung und der impliziten systematischen Annahme von schnellem Gleichgewicht für virtuelle, zusätzlich eingeführte Reaktionen, die Diagonalreaktionen genannt werden. Die der Approximation zugrunde liegenden Gleichungen können alternativ auch aus Unabhängigkeitsannahmen für Prozesse aus verschiedenen Modulen, die nicht direkt interagieren, hergeleitet werden. Die Diagonalreaktionen müssen nicht explizit definiert werden, ihre Definition erfolgt

implizit bei der direkten Erstellung des reduzierten Modells. Die Approximationsfehler folgen Differentialgleichungen erster Ordnung, die die Entfernung der Diagonalreaktionen vom Gleichgewicht beschreiben, und sind dynamisch sowie stationär begrenzt. In Spezialfällen stellen die reduzierten Modelle eine stationär oder sogar dynamisch exakte Reduktion des entsprechenden konventionellen Modells dar.

Die Erstellung von reduzierten Modellen für komplexe kombinatorische Systeme wird durch das in dieser Arbeit vorgestellte Computerprogramm ALC (Automated Layer Construction) in hohem Maße vereinfacht. Die Modelle werden durch eine regelbasierte Modelldefinition in einer leicht verständlichen, aber mächtigen Syntax definiert. Die modulare Struktur der reduzierten Modelle spiegelt sich in den Modelldefinitionen wider. Diese werden einer Vielzahl von Konsistenzkontrollen unterzogen, was dazu führt, dass die meisten Fehler in den Modelldefinitionen leicht zu finden und zu beseitigen sind. ALC ist frei verfügbar und kann lokal oder über die ALC-Webseite ausgeführt werden. Die durch die Modelldefinitionen definierten Modelle werden in verschiedenen Formaten (C MEX, MATLAB, *Mathematica* und SBML) als direkt verwendbare Simulationsdateien ausgegeben. Des weiteren werden die Modellgleichungen auch in LaTeX und Textformat ausgegeben, um die Veröffentlichung und Präsentation der Modelle zu vereinfachen.

Durch die Verwendung von ALC ist es möglich, reduzierte Modelle von Systemen mit auftretender kombinatorischer Komplexität mittels einer relativ kurzen und einfachen Modelldefinition zu erstellen. Das Leistungsvermögen von ALC wird anhand verschiedener Modelldefinitionen demonstriert. Unter anderem wird eine Modelldefinition für ein Modell des Insulinsignalsystems vorgestellt, die die kombinatorische Komplexität berücksichtigt. Diese Modelldefinition führt zu einem reduzierten Modell mit 51 Differentialgleichungen, welches ein konventionelles Modell mit $1,5 \cdot 10^8$ Differentialgleichungen approximiert.

1 Introduction

Cellular signal transduction systems enable cells to adapt their physiological processes to the needs of the organism and to regulate many physiological processes. Though many of their components have been characterized, the signal transduction networks are not yet completely understood. An understanding of these systems is of high medical relevance because malfunctions can result in severe disorders such as cancer and diabetes. Mathematical models can be used to understand and to analyze cellular signaling systems.

Section 1.1 gives an introduction to cellular signal transduction and its biological relevance, while the conventional approach to model cellular signaling systems is described in Section 1.2. A challenge for conventional modeling of signal transduction networks is that the association of a few proteins may result in an enormous amount of complexes and an equivalent amount of necessary differential equations in the corresponding models. This phenomenon, called combinatorial complexity, is described in Section 1.3 and leads to enormous problems in the modeling of cellular signal transduction systems which are outlined in Section 1.4. A first step to overcome the problems associated with combinatorial complexity is rule-based modeling which is introduced in Section 1.5. At the end of this chapter, Section 1.6 gives an outline of this thesis.

1.1 Cellular signal transduction

The interior of eukaryotic cells is separated from the environment by the plasma membrane, a membrane consisting of a lipid bilayer and membrane proteins [1]. However, cells need to monitor their environment as they have to react to changing conditions and demands of the organism by adapting their activities, in particular the metabolism and the gene expression.

Many of these changes and demands are sensed by receptor proteins in the plasma membrane that initiate the intracellular transduction of extracellular signals. Receptor-mediated signal transduction plays an important role in biology since it influences many crucial physiological processes such as differentiation, cell division, cell death and the regulation of blood glucose [2]. Defects in signal transduction systems can result in severe implications for the human organism, e.g. diabetes mellitus, obesity and cancer.

Signaling systems sense extracellular signals, transmit them to the interior of the cell, amplify the signals and distribute them to their cellular targets [1,2]. Such a signal is in most cases the concentration of a relatively small molecule, for example a peptide hormone like insulin, in the cellular environment. In some cases, this signaling molecule can directly cross the cellular membrane, whereas in many other cases a receptor, a transmembrane protein, binds the signaling molecule (which is a ligand of the receptor) outside the cell. This ligand

binding induces changes, e.g. phosphorylation of binding sites, in the intracellular part of the receptor. A typical scenario is that other proteins (effectors) are recruited to activated binding sites and build a signaling complex at the receptor. These proteins can be scaffolds providing binding sites for other signaling proteins. Cellular signaling cascades which often strongly amplify the extracellular signal transmit it to its intracellular targets.

Though many biochemical details of several signal transduction systems are known today [1], an intuitive understanding of the entire signaling systems is rarely possible. The intuitive understanding is impeded by system level properties that are not apparent from the characteristics of the components [3]. This lack of understanding is the motivation for the emerging discipline of systems biology which aims at a holistic understanding of physiological processes using mathematical models [3, 4]. These models integrating the current state of biological knowledge are analyzed to reveal and understand the system properties. The models may also provide information about components of the system which are not yet described, but necessary to explain a certain observed system behavior.

Before describing the conventional modeling approach and its limitations, we give an introduction to blood glucose regulation and insulin signaling, a cellular system of tremendous physiological relevance that will be considered within several parts of this thesis.

1.1.1 The regulation of blood glucose

It is crucial for the organism that blood glucose concentration remains within relatively narrow bounds. Both, too high and too low glucose concentrations can cause serious problems up to death. Therefore, the glucose concentration is controlled by a complex regulation system, which is an interesting example for the importance of cellular signal transduction.

The β-cells of the pancreas secrete insulin if the glucose concentration in the blood is too high [5]. This usually occurs after a meal when the food is digested. Insulin induces glucose uptake of other cells, mainly in the fat and muscle tissue, which lowers the glucose concentration in the blood. The control signal of elevated insulin concentrations is turned off by insulin-sensitive cells taking up insulin and by the kidney which constitutively filters insulin out of the blood [6, 7]. Insulin removal and degradation is discussed in more detail in Section 2.1.

Over night or during exercise, the blood glucose concentration declines because muscles and other tissues consume glucose to maintain their function. If the glucose concentration in the blood is too low, glucagon is secreted by α-cells of the pancreas [5], which induces a reaction in other cells. Hepatocytes of the liver break down the intracellular glucose polymer glycogen and secrete glucose. The liver acts as a buffer for glucose as in the case of high glucose concentration in the blood, e.g. after a meal, glucose is taken up independently of insulin and partly stored as glycogen.

Both previously introduced mechanisms for sensing extracellular signals are present in the regulation of blood glucose. Pancreatic β-cells let the signal molecule glucose traverse the membrane and enter the cytoplasm. Transmembrane receptors for insulin and glucagon specifically bind their ligands outside the cell and transmit the signals to the interior of the

cells. The insulin signaling system is described in more detail below.

1.1.2 Insulin signaling

Insulin is a peptide hormone that is produced in the β-cells, which are specialized cells located in the islets of Langerhans in the pancreas. The insulin concentration in the blood is not only an important control signal for cellular glucose uptake but also for the energy metabolism [5, 8–10]. Insulin is also involved in gene expression, cell survival and differentiation [11].

The insulin signaling system transduces the information about extracellular insulin concentration inside the cell, amplifies it and routes it to different cellular targets [5,9,12,13]. A major compound of this signaling system is the insulin receptor. It is located at the plasma membrane and has an extracellular part that can bind insulin and an intracellular part that transduces the extracellular signal into the cell. Insulin binding to the receptor induces autophosphorylation of numerous intracellular sites on the receptor. The attachment of several signaling proteins to phosphorylated binding sites on the receptor is the first step in the intracellular amplification and transduction of the signal. This complex formation at the insulin receptor is described in Section 1.3 where the occurring combinatorial complexity is also shown. The activated receptor complexes induce several cellular signaling cascades (e.g. MAP kinase cascade and PI3K pathway) that result in the activation of cellular kinases such as PKB/Akt, PKC and ERK [2]. Finally, these signaling cascades lead to an adaptation of the cellular metabolism such that excessive glucose is broken down and for example used as a precursor for the synthesis of glycogen, fat and proteins [5].

Defects in the insulin signaling system can give rise to insulin resistance, obesity and type II diabetes mellitus [14–16], all of which provide severe implications for the organism. These diseases are widely spread in the western hemisphere and intense efforts are made to improve the corresponding therapies [17–20].

1.2 Modeling of chemical reaction networks

Biochemical systems can be considered as a special case of chemical reaction networks where the components, e.g. proteins, may have multiple binding sites and other complex functional properties. There is a well-established theory for the mathematical modeling of chemical reaction systems based on the use of ordinary differential equations (ODEs) [21], which can be applied to biochemical systems [22].

This section introduces the conventional approach for the modeling of chemical reaction networks which consists in describing the system by a set of chemical reaction equations and describing the dynamics of each component by a first order ordinary differential equation. This approach is also used and extended in this thesis to describe the dynamic behavior of cellular signaling systems.

1.2.1 Underlying assumptions

Cellular components show temporal and spatial dynamics [23] implying that partial differential equations should be used for an exact description. However, diffusion is quite fast for short distances [21]. Assuming that each cellular compartment is well stirred is therefore usually a good assumption that allows for the use of ODEs. Assuming fast diffusion is not appropriate when describing spatially distributed phenomena such as morphogenic gradients. Special solutions are required for these cases, however, the vast majority of systems can be adequately described by ODEs.

If the volumes of the compartments and the concentrations of the cellular substances are very low, the inherent stochastic behavior of chemical processes at the molecular level becomes important. This stochastic behavior is described by the chemical master equation, a set of first order differential equations describing the dynamics of the probabilities of all possible states of the system [24]. However, if the concentrations and the volumes are high enough, the system dynamics exhibits deterministic behavior which can be described by ODEs [24]. Additional requirements for the use of ODEs are equally distributed temperature and pressure within each compartment.

Altogether, using ODEs to describe the deterministic system dynamics requires the implicit assumptions that each compartment is well stirred and that the volumes and concentrations are high enough. This will always be assumed in this thesis.

1.2.2 Reactions

Chemical reactions are represented by chemical reaction equations [25]. In a general reaction equation

$$a\ A + b\ B... \leftrightharpoons c\ C + d\ D... \tag{1.1}$$

the stoichiometric coefficients (a, b, c and d) are natural numbers assigned to the species (A, B, C and D) participating in the reaction. The stoichiometric coefficients define how many molecules of the corresponding species are converted by the reaction. A species is a specific configuration of a molecule or a complex.

Notations for species consist of the molecule name followed by the comma-separated sequence of site configurations within one pair of squared brackets. As an example, the species $R[L, P]$ is a specified configuration of the molecule R where the ligand L is bound to the first site and the second site is phosphorylated. This indexing of species as configurations of a molecule is extremely valuable in the context of rules which are generalized reactions (see Section 1.5) and is widely used in Chapters 4 and 5. Another possibility for species notation is to use descriptive notations. As an example, the unphosphorylated and phosphorylated molecule E could be denoted as E and Ep, respectively. This descriptive notation is especially useful when describing systems with a relatively low number of species or reactions and is used in Chapter 2.

Reaction equations containing the reaction symbol '$\leftrightharpoons$' define reversible reactions. A reversible reaction formally consists of two irreversible reactions. The forward reaction pro-

ceeds from the left hand side to the right hand side of the reaction equation; the backward reaction proceeds from the right hand side to the left hand side of the reaction equation. Reactions may also be irreversible which can be indicated by a reaction symbol that indicates the sole direction of the reaction ('$\rightarrow$' or '$\leftarrow$' instead of '$\leftrightharpoons$').

All processes of a chemical or biochemical system can be described by a set of chemical reaction equations (or more shortly: reactions) which is called reaction network. The reaction network gives structural information about the corresponding system. The following section describes how information about reaction dynamics can be included in the reaction network to get a representation that uniquely defines the ODEs describing the dynamic behavior of the corresponding system.

1.2.3 Generalized mass action kinetics

If chemical systems are modeled at the molecular level, mass action kinetics are often a good description of the chemical processes [22] and are therefore frequently used to assign reaction rates to reaction equations. To include dynamic information, we rewrite Equation 1.1.

$$a\ A + b\ B... \leftrightharpoons c\ C + d\ D... \qquad k_i \qquad k_{-i} \tag{1.2}$$

The kinetic parameters of the forward and backward reactions are k_i and k_{-i}, respectively. According to the law of mass action [22], the reaction rate r_i for this reaction (Equation 1.2) has the unit $M \cdot s^{-1}$ and is given as

$$r_i = k_i \cdot A^a \cdot B^b ... - k_{-i} \cdot C^c \cdot D^d ... \tag{1.3}$$

where A, B, C and D are the concentrations of the corresponding species. For the sake of simplicity, the same notations are used for species and their concentrations. The standard unit for concentrations is molar (M). A concentration of 1 $M = 1\ mol \cdot l^{-1}$ means that there are one mol ($\approx 6.022 \cdot 10^{23}$) molecules of this substance per liter [25]. Concentrations of cellular signaling components are often in the nanomolar ($1\ nM = 10^{-9}\ mol \cdot l^{-1}$) or micromolar ($1\ \mu M = 10^{-6}\ mol \cdot l^{-1}$) range.

The units of the rate constants k_i and k_{-i} (Equations 1.2 and 1.3) depend on the sum of the stoichiometric coefficients of the corresponding side of the reaction. The units of k_i and k_{-i} are given as $M^{-fr} \cdot s^{-1}$ and $M^{-br} \cdot s^{-1}$, respectively. The exponents fr and br (associated to the forward and backward reactions, respectively) are defined as $fr = a + b + ... - 1$ and $br = c + d + ... - 1$.

One side of a reaction equation may be empty ($fr = -1$ or $br = -1$). This allows for the realization of synthesis and degradation reactions. Irreversible reactions are indicated by a reaction parameter that equals zero or a reaction symbol that indicates the direction of the reaction ('$\rightarrow$' or '$\leftarrow$' instead of '$\leftrightharpoons$'). In the latter case, only one kinetic parameter has to be given.

The information given in the parameterized reaction equations R_i defines the differential equations describing the dynamics of the concentrations of the species S_j. For the formal discussion, we denote the stoichiometric coefficient of the species S_j in the reaction R_i as s_{ij}.

The reaction rates r_i have to be considered with the correct factor f_{ij} when balancing the species S_j. This factor is the negative stoichiometric coefficient ($f_{ij} = -s_{ij}$) for reactants (species at the left hand side of the chemical reaction equation) and the positive stoichiometric coefficient ($f_{ij} = s_{ij}$) for products (species at the right hand side). If the species S_j occurs on both sides of the reaction equation, it holds that $f_{ij} = s_{ij,right} - s_{ij,left}$. The factor f_{ij} equals zero if the species S_j does not occur in the reaction R_i or shows no net turnover. In the case of constant volume the ODE for the concentration of each species S_j is given as

$$\frac{d}{dt}S_j = \sum_{R_i} f_{ij} r_i. \tag{1.4}$$

The following simple example illustrates how the ODEs are generated from parameterized reactions using the law of mass action.

$$\begin{array}{rcll} A + B & \leftrightharpoons & 2\ C & k_0 \quad k_{-0} \\ A & \leftrightharpoons & & k_1 \quad k_{-1} \\ C & \rightarrow & & k_2 \end{array} \tag{1.5}$$

These reactions proceed with the reaction rates

$$\begin{aligned} r_0 &= k_0 \cdot A \cdot B - k_{-0} \cdot C^2 \\ r_1 &= k_1 \cdot A - k_{-1} \\ r_2 &= k_2 \cdot C \end{aligned} \tag{1.6}$$

and define the differential equations

$$\begin{aligned} \frac{d}{dt}A &= -r_0 - r_1 \\ \frac{d}{dt}B &= -r_0 \\ \frac{d}{dt}C &= 2 \cdot r_0 - r_2. \end{aligned} \tag{1.7}$$

In the classical law of mass action for ideal mixtures, the kinetic parameters are constants [22] whose values only depend on the temperature and on the pressure [21]. The generalized law of mass action reviewed by Heinrich and Schuster [22] allows for a wide range of possible kinetics because both constant kinetic parameters of a reaction are multiplied by a common positive nonlinear function.

The generalized law of mass action used in this thesis is more general, since each reaction parameter may be an arbitrary nonlinear function. Note that these nonlinear functions may depend on the concentrations of species. A frequently occurring case in which a single reaction parameter depends on concentrations of reacting substances is shown in Section 3.6.3 for enzyme kinetics, which can be described by the generalized law of mass action.

1.3 Combinatorial complexity

The association and modification of a few proteins can result in an enormous amount of possible protein complexes [26,27]. This phenomenon is referred to as combinatorial complexity. There are four possible sources of combinatorial complexity in signaling systems that can also occur in combination with each other.

- Binding of several molecules to scaffold proteins with more than one binding site
- Modification (e.g. phosphorylation) of proteins with several sites
- Oligomerization (as a special case: dimerization) of proteins
- Chain formation

The complexity of the insulin signaling system results from a combination of the first three items listed above.

The insulin receptor is constitutively dimerized [28–30]. For the sake of simplicity however, we analyze the complexity for a virtual monomer. The receptor monomer can bind an insulin molecule and has intracellular binding sites for the signaling proteins IRS and Shc. Both intracellular sites of the receptor become phosphorylated before effector binding [12]. Shc becomes phosphorylated and binds Grb2. Grb2 can bind SOS, which in turn can be phosphorylated. IRS has four binding sites for PI3K (in fact it has at least nine binding sites for PI3K; each p85 subunit of PI3K occupies two binding sites), one for Grb2 and one for SHP2 [31]. All these binding sites can be phosphorylated. The number of feasible species is 17,038 for the receptor monomer and 145,156,469 in the total network including receptor dimers. These numbers can be calculated as shown in Figure 1.1. The facts that the insulin receptor and IRS (in fact there exist several different IRS molecules) can be phosphorylated on several regulatory sites and have additional binding partners [12,32,33] are not considered. This accounts for a further dramatic increase in complexity.

Many other signaling systems, e.g. the EGF signaling system, also show combinatorial complexity [26]. In some cases the complexity results from the association of a few proteins with a very limited amount of binding sites each. An extreme example for this is chain formation where a ligand with two binding sites for the receptor binds to a receptor with two binding sites for the ligand. Theoretically, the number of feasible complexes is infinitely high if there are no spatial constraints [27]. However, an upper bound for the theoretical complex size is given by the (bounded) number of cellular receptor molecules. There are also cases in which the possible modifications of only one protein can reach extremely high numbers, e.g. if this protein has many regulatory sites that can be phosphorylated.

As outlined below, one ODE is required for the balance of each possible species. This leads to enormous problems in the modeling of systems with inherent combinatorial complexity.

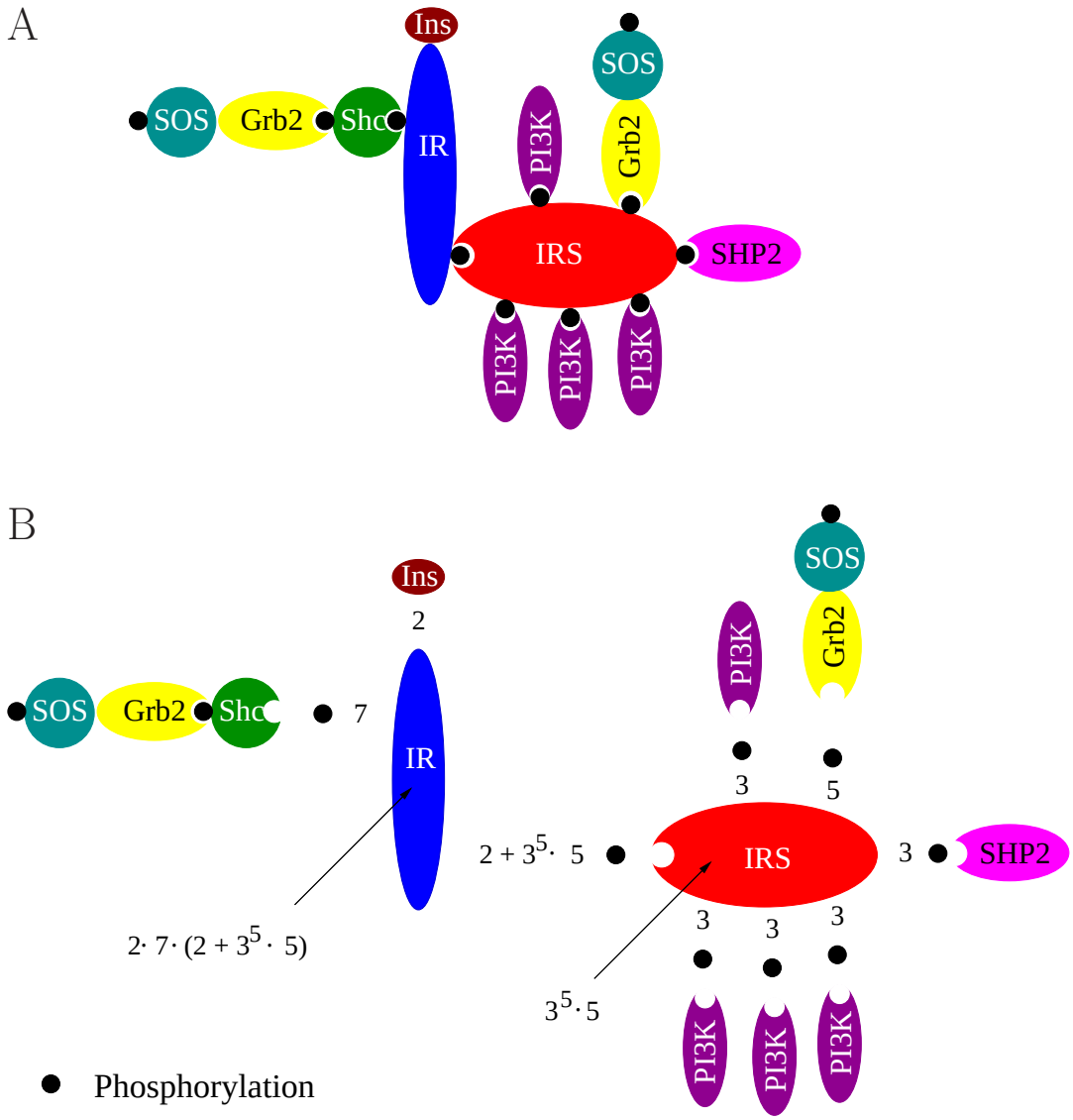

Figure 1.1: Combinatorial complexity in the insulin signaling system
A) The insulin receptor (IR) can bind insulin (Ins), Shc and IRS. IRS can bind four PI3K molecules, SHP2 and Grb2. Grb2 can bind SOS and phosphorylated SOS. Note that each binding process (except for the bindings of insulin to the receptor, receptor dimerization and the binding of SOS to Grb2) requires the phosphorylation of a binding site.
B) Without considering the insulin receptor, the described associations result in $3^5 \cdot 5 = 1215$ different complexes of IRS. For Shc binding to the receptor monomer, there are seven possible states of the corresponding site on the receptor, for insulin binding there are two possible states. For IRS binding to the receptor, there are $2 + 3^5 \cdot 5$ possibilities for the state of the binding site (unphosphorylated, phosphorylated but unoccupied and $3^5 \cdot 5$ possibilities for the occupied binding site). Altogether, there are $n = 2 \cdot 7 \cdot (2 + 3^5 \cdot 5) = 17,038$ different complexes of the receptor monomer. As the receptor is always a dimer ($k = 2$), there are $\binom{n+k-1}{k} = \binom{17039}{2} = 145,155,241 \approx 1.5 \cdot 10^8$ different combinations. Free species contribute another $1215 + 10 + 1 + 1 + 1 = 1228$ possible species: 1215 for free IRS complexes, 10 for all combinations of Shc, Grb2 and SOS and one for insulin, PI3K and SHP2 each.

1.4 Limitations of conventional modeling

The conventional modeling approach for biochemical networks was introduced above and is frequently applied in systems biology [34–43]. It is based upon defining a reaction equation with corresponding rate law for each chemical reaction. When the reaction network is set up, the balances of all species can be derived [22]. A basic feature of this approach is that the dynamics of the concentration of each species is described by one ODE.

Conventional modeling is very powerful for small reaction systems. However, its limitations become clearly visible when considering systems with inherent combinatorial complexity. For these systems, the number of necessary reactions is very high [27]. In the example of insulin signaling (Figure 1.1) there are $3^5 \cdot 5$ modifications of IRS that can bind to $2 \cdot 7$ modifications of the phosphorylated receptor monomer with unoccupied binding site. This results in $3^5 \cdot 5 \cdot 2 \cdot 7 = 17,010$ reactions describing IRS binding to the receptor monomer. Considering receptor dimerization there are $3^5 \cdot 5$ modifications of IRS that can bind to $2^2 \cdot 7^2 \cdot (2 + 3^5 \cdot 5) = 238,532$ modifications of the receptor dimer which may have bound two insulin molecules, two Shc molecules and one IRS molecule. Altogether, there are $2^2 \cdot 7^2 \cdot (2 + 3^5 \cdot 5) \cdot 3^5 \cdot 5 = 289,816,380 \approx 2.9 \cdot 10^8$ reactions for IRS binding to the receptor.

The generation of such a long list of reactions is not practical. In addition, a long list of reactions provides no simple representation of the signaling system and makes the understanding and modification of the model difficult. The number of necessary ODEs is also tremendously high for systems with inherent combinatorial complexity [27]. As an example, there are $1.5 \cdot 10^8$ different complexes in the insulin signaling system (see Section 1.3). To get a conventional model of this system, $1.5 \cdot 10^8$ ODEs would be necessary.

For this reason, the modeling of systems with inherent combinatorial complexity is very difficult, or even impossible, using the conventional modeling approach.

1.5 Rule-based modeling

A single molecular event, e.g the binding of an effector to a phosphorylated binding site of a receptor or the phosphorylation of a site, is often described by a long list of reactions. In the example of effector binding, a binding reaction is necessary for each modification of the unbound effector to each modification of the receptor with phosphorylated and unoccupied binding site (see Section 1.4).

Large subsets of these reactions or even all reactions are usually parameterized by the same kinetic constants. Such subsets can be represented using generalized reactions called rules which contain patterns instead of species [27, 44, 45]. Each rule represents a class of reactions that are parameterized by the same kinetic constants. This representation simplifies the description and the understanding of large reaction systems. In many cases, a relatively short list of rules is sufficient to describe the reactions of a signaling system. Software tools exist which automatically generate the simulation files from a rule-based model definition [45–51]. Thus, using rules highly simplifies the model representation and generation.

The usage of rules is demonstrated with the help of the following simple example, where

the reactions represent ligand binding to a receptor that also has a phosphorylation site.

$$\begin{array}{rcll} R[0,0]+L & \leftrightharpoons & R[L,0] & k_1 \quad k_{-1} \\ R[0,P]+L & \leftrightharpoons & R[L,P] & k_1 \quad k_{-1} \end{array} \tag{1.8}$$

These reactions can be represented by a single rule

$$R[0,X]+L \leftrightharpoons R[L,X] \qquad k_1 \quad k_{-1} \tag{1.9}$$

where patterns (characterized by at least one site modification 'X') occur on both sides of the rule. The notation of patterns is analogous to the notation of species, with the only difference, that the modification 'X' at a specific site represents all distinct modifications of this site individually. Note that in all other cases, e.g. in the reaction parameters, a site modification 'X' is interpreted as the sum of all possible modifications of this site (macrostate, see Section 4.2).

If there are several site modifications 'X', it is important to note that sites with a modification 'X' at the same position on both sides of the rule correspond to each other. When rules are used to generate the corresponding reactions, the species in each reaction have the same modification at these sites on both sides of the reaction equation.

The benefit of rules becomes apparent when considering larger reaction systems. As an example, consider a signaling protein with ten phosphorylation sites. If the phosphorylation of each site is not influenced by the phosphorylation state of all other sites, only 10 rules are sufficient to describe the reaction network defining $2^{10} = 1024$ ODEs. Each of these 10 rules corresponds to $2^9 = 512$ reactions that are parameterized by the same kinetic constants.

Rule-based modeling yields the same ODEs as conventional modeling of the system. Using rules therefore simplifies the model generation and representation but does not lower the order of the model. Note that rules can also be parameterized by generalized mass action kinetics in the same way as reactions. An example for this is given in Section 5.2.2.3.

1.6 Outline of the thesis

An increasing amount of biochemical knowledge about cellular signaling systems is being revealed. However, the behavior of these systems is often not apparent given the properties of their components. This results in an increasing need for mathematical modeling of signaling systems to address system level properties systematically. However, the association and modification of a few proteins can result in an enormous amount of feasible complexes. This phenomenon is called combinatorial complexity and leads to severe problems in the modeling of many cellular signaling systems because one ODE is required for the balance of each possible species.

Chapter 1 gives an introduction to cellular signal transduction and shows the occurring combinatorial complexity. A focus is on insulin signaling and the regulation of glucose concentration in the blood. We discuss the conventional modeling approach for biochemical systems and its limitations. Modeling by generalized reactions, called rules, simplifies the

model generation and representation but does not lower the number of necessary ODEs. Therefore, rule-based modeling is merely a first step to overcome the problems resulting from combinatorial complexity.

The insulin signaling system is characterized by combinatorial complexity. It can only be modeled in the conventional way if many known details about complex formation at the receptor are not included in the model. A conventional model focussing on the insulin receptor and also describing insulin degradation and synthesis *in vivo* is built and analyzed in Chapter 2. It is shown that the relative contributions of the liver and the kidney to insulin degradation and insulin clearance are highly dependent on insulin concentration. The widely used quantity of insulin clearance is therefore only meaningful when the corresponding insulin concentrations are taken into account. This however, is not the case in many experimental studies.

As discussed in Chapter 1, combinatorial complexity results in extremely large models if conventional mathematical descriptions of signaling systems are desired. This demonstrates the need for a reduced order description of such systems. Chapter 3 gives an overview of model reduction techniques. Since biochemical systems are highly nonlinear, we only shortly introduce methods for linear systems and focus on techniques that can be applied to nonlinear systems. Being a widely used technique, proper orthogonal decomposition (POD) is presented in more detail. However, using POD for models of signaling systems has several drawbacks including a high computational effort and state variables without direct physical interpretation.

The focus of Chapter 3 is on specialized model reduction methods for chemical and biological systems. The most important methods for chemical systems are the quasi-steady-state approximation and the rapid equilibrium assumption that both allow for the reduction of reaction systems when knowing qualitative features of reaction velocities. Apart from that, we discuss the frequently used intuitive model reduction approach for signaling systems that is the only non-systematic reduction technique considered. A focus of this discussion is the often insufficient approximation quality of intuitive model reduction. Chapter 3 also reviews domain-oriented model reduction, a special approach for signaling systems, which is based upon the concept of rules and is in some cases able to provide a strongly reduced description of the system. A major drawback of all systematic ODE-based approaches is that a potentially very large conventional model has to be built before the reduced order model can be generated.

Chapters 4 and 5 provide a framework for the modeling of signaling systems with inherent combinatorial complexity. Chapter 4 introduces layer-based modeling as a new approximative, but accurate method for the modeling of signaling systems with inherent combinatorial complexity. The method is characterized by an extremely high potential to reduce the model size and a high approximation quality. Convenient features of layer-based modeling are that the resulting models show a pronounced modularity and that the state variables are macroscopic quantities with a direct biochemical interpretation. An important difference to most other model reduction techniques is that the generation of the potentially very large conventional model to be approximated is not necessary. The reduced model equations can be

directly generated by a procedure quite similar to conventional modeling. Therefore, layer-based modeling allows for the mathematical description of systems not accessible previously.

The reduction of the model size is based upon an implicit systematic application of the rapid equilibrium assumption for additionally introduced virtual reactions, called diagonal reactions, and a macroscopic system description. A detailed analysis shows that the approximation errors can be described by first order ODEs whose solutions are dynamically and ultimately bounded.

Chapter 5 presents the computer program ALC (Automated Layer Construction) which highly simplifies the generation of reduced order models of large systems. ALC allows the modeler to define layer-based models of very large systems by a relatively short and simple rule-based model definition. To demonstrate the potential of ALC, we provide several model definitions, including a model definition for a layer-based model of insulin signaling that accounts for combinatorial complexity. This intuitively understandable model definition results in a layer-based model with 51 ODEs approximating a conventional model with $1.5 \cdot 10^8$ ODEs. ALC provides ready-to-run simulation files in different formats (C MEX, MATLAB, *Mathematica* and SBML). The model equations are also given in LaTeX and plain text format simplifying publication and presentation of the model.

Chapter 6 summarizes the main results of this thesis and gives an outlook on potential future work.

2 Modeling and analysis of insulin dynamics *in vivo*

Insulin is a major regulator of the cellular energy metabolism and glucose concentration in the plasma. A detailed understanding of insulin dynamics and insulin receptor activation in the target tissues is of high medical relevance.

This chapter presents a mathematical model describing insulin receptor activation in hepatocytes as well as the related processes of insulin degradation and synthesis in vivo*. Section 2.1 describes the biological background necessary for the following discussions and the state of the art in the modeling of insulin dynamics and insulin receptor activation. A detailed model of insulin receptor activation and insulin dynamics in the blood is introduced in Section 2.2. Model validation with experimental data sets from the literature is performed in Section 2.3. The model analysis performed in Section 2.4 shows that the relative contributions of the liver and the kidney to insulin degradation as well as insulin clearance highly depend on the insulin concentration. We show that using insulin clearance, a widely used quantity to characterize the state of insulin metabolism, is only justified if the corresponding insulin concentration is given. The analysis of insulin degradation implies that there is an upper bound for reasonable therapeutic insulin concentrations. Section 2.5 discusses modern insulin therapy in the light of these results.*

2.1 Biological background

Insulin is synthesized in pancreatic β-cells and released into the blood if the glucose concentration in the blood is too high [5]. Several tissues, in particular liver, kidney, adipose tissue and muscle contribute to its degradation, with the liver and the kidney being considered as the most important of them [7, 52]. The kidney mainly filters the insulin out of the blood, whereas liver cells (hepatocytes), fat cells (adipocytes) and muscle cells show a receptor-mediated uptake of insulin and intracellular insulin degradation [5, 12].

Insulin receptors bind insulin molecules from the space of Disse [53]. The space of Disse (perisinusoidal space) contains blood plasma and is an extracellular space between liver sinusoids (special blood vessels) and hepatocytes. In the space of Disse, there is also reversible nonspecific insulin binding to hepatocytes. The complex of insulin and the receptor is internalized to endosomes, where the receptor is located in endosomal membranes [12]. In the acidic endosomal compartment, insulin dissociates from the receptor and is degraded. The receptor then recycles to the cell surface (which is the usual case) or is degraded [8].

Receptors with bound insulin show a strongly elevated autophosphorylation activity. The receptor phosphorylates its intracellular binding sites for signaling molecules and also some regulatory sites modulating receptor activity [8, 12]. Dephosphorylation of phosphorylated sites is performed by protein phosphatases which in this way modulate receptor activity.

The complex formation at the activated receptor is described in Section 1.3. The resulting signaling complexes activate several cellular signaling cascades inducing changes in the cellular metabolism and in the gene expression [5, 8, 9, 12].

Defective insulin secretion or cellular insulin resistance may result in diabetes mellitus which is characterized by a decreased ability to regulate the glucose concentration in the blood [18]. Therapeutic insulin is injected to compensate for the insufficient amount of endogenously secreted insulin [16,18,19]. The aim of insulin therapy is to achieve sufficiently high insulin concentrations in the blood to regulate the glucose concentration. Too high insulin concentrations however have to be avoided because an overdose of insulin results in hypoglycemia which may cause serious complications.

Therapeutic insulin is injected into the subcutaneous tissue and traverses different compartments (e.g. the injection pocket and the interstitium) before entering the blood. Long acting insulins tend to form dimers or hexamers in the subcutaneous tissue which slows down the transition of insulin from the injection pocket to the blood. Fast acting insulin analogues however, have a decreased ability to form oligomers and enter the circulation rapidly. These different kinetic properties of different types of insulin are exploited in insulin therapy [18, 20]. The compensation for the glucose uptake after a meal is mainly performed by fast acting insulins, whereas the basal insulin level is adjusted by long acting insulins or continuous infusion of small amounts of fast acting insulins.

2.1.1 Models of insulin receptor dynamics *in vitro*

Several models in literature describe insulin receptor dynamics *in vitro*. Most of them [54–58] focus on a subset of the occurring processes and lump several processes into single reaction steps. The physiologic state of insulin-sensitive tissues, e.g. the activation of hepatic insulin receptors, cannot be precisely deduced from such models as their level of detail is too low. However, two recent *in vitro* models describe insulin receptor dynamics in more detail [35,43].

Sedaghat et al. combined models of insulin binding [58] and receptor internalization, recycling, synthesis and degradation [57] and extended them to a mathematical model of insulin signaling in adipocytes also describing receptor phosphorylation [35]. Model parameters were taken from other models and *in vitro* experiments. A very strong coupling between insulin binding and receptor phosphorylation is assumed. The second insulin molecule can only bind to phosphorylated receptors. Dephosphorylation of the receptor (with simultaneous insulin dissociation) is only possible if only one insulin molecule is bound to the receptor. Phosphorylated receptors without insulin are therefore not part of the model.

Hori et al. described receptor phosphorylation, internalization and recycling in Fao hepatoma cells at 100 nM insulin [43]. They analyzed several models corresponding to different model assumptions and different levels of detail. Model parameters were estimated using

experimental data sets from literature. The main limitations of the models of Hori et al. [43] are that they are only valid at 100 nM insulin (which is a supraphysiological insulin concentration [52]) and that insulin binding is not explicitly included. Due to the high insulin concentration of 100 nM, all receptors at the plasma membrane are assumed to be liganded. Hori et al. also provide a general model structure without parameterization that includes the binding of one insulin molecule to the receptor and is intended for variable insulin concentrations [43]. As in the model of Sedaghat et al. [35], receptor dephosphorylation and insulin dissociation are coupled in all models since insulin dissociation is assumed to be a prerequisite for receptor dephosphorylation or the processes are lumped into a single step. In addition, most processes are assumed to be irreversible.

Thus, there are many couplings between different processes in all detailed receptor models [35, 43]. Furthermore, one of the models is only valid for adipocytes [35], the other ones are not valid for other insulin concentrations than 100 nM [43]. This demonstrates the need for a detailed mathematical model for hepatocytes that allows for analyzing the physiological state of these cells.

2.1.2 Insulin dynamics and insulin receptor activation *in vivo*

Much work has been done in past decades to study insulin kinetics in the blood [6, 59, 60]. In the last few years, efforts have been focused on analyzing the dynamics of insulin concentration after the subcutaneous insulin injection [61–63]. The resulting models describe insulin removal from the blood in a highly reduced way [6, 63], whereas the subcutaneous tissue is usually modeled in more detail. Oligomer formation of insulin is included in some models [63]. In other studies, insulin dynamics are linked with glucose dynamics [64–69]. The corresponding models describe all involved processes in a highly reduced way.

In the last few decades, different kinetics for insulin removal from the blood were proposed. The most frequently used kinetics are linear first order kinetics, Michaelis-Menten kinetics or a combination of both [59]. Due to the experimental investigation of narrow concentration intervals, nonlinearity was difficult to demonstrate [70].

The presence of nonlinearities due to saturable processes is widely accepted now [7, 18]. However, insulin degradation is described as a linear first order process in most models and insulin receptor activation is not covered. Allocation of insulin degradation to specific tissues is not performed in the models of insulin dynamics [63]. Therefore, no model-based analysis of the contributions of the liver and the kidney to the degradation process has been done. In addition, the physiological state of the insulin-responsive tissues, e.g. the activation of insulin receptors, cannot be obtained from the existing *in vivo* models, as their level of detail is quite low [6, 59–63]. Consequently, there has also been no detailed analysis of the interactions between the highly related processes of insulin turnover and insulin receptor activation. A first step in this direction was taken by Hovorka et al. [60]. However, the receptor part of their model only distinguishes between free receptors and receptors with bound insulin. In addition, the focus of their study is clearly on insulin kinetics.

An *in vivo* model describing hepatic processes in such a detailed way as the existing *in vitro*

models [35,43] while simultaneously avoiding the deficiencies of those models will contribute to a deeper understanding and a model-based control of insulin and glucose dynamics. Additionally, a detailed model can serve as a starting point for modeling and analysis of the signaling cascades emerging from the hepatic insulin receptor *in vivo* which will further contribute to the understanding of insulin signaling.

2.2 The model

This section presents a literature-based mathematical model of insulin dynamics and hepatic insulin receptor activation in rats. The model consists of ordinary differential equations (ODEs) and describes the dynamic behavior of radioactively labeled and unlabeled insulin in the blood and the physiological state of hepatic insulin receptors. It can also be used for the injection of only labeled or only unlabeled insulin. Compared to other models [35,43,54–58], we describe receptor processes in more detail to get insights into the processes and into the connections between insulin dynamics and insulin receptor activation in hepatocytes. We decouple insulin binding and dissociation from receptor phosphorylation, as there is experimental evidence that receptor phosphorylation does not affect insulin binding [71]. In addition, the phosphorylation of liganded receptors is described as a reversible process.

This study uses the rat as model organism because much more parameters are known for rats than for humans.

2.2.1 Important tissues and processes

The liver and the kidney are the most important insulin degrading tissues [7,18]. Fat and muscle tissues contribute to insulin degradation as well, but all peripheral tissues (excluding the liver and the kidneys) together contribute with only about 13 % to total insulin degradation [52]. By some simple calculations it can be shown that the contribution of the fat tissue is small compared to that of the liver [72]. Since the contribution of the muscle tissue to insulin degradation is also relatively low [52] and a qualitatively similar behavior like that of the liver is expected, it is appropriate to focus on hepatic and renal insulin degradation.

Our model explicitly describes dynamic insulin receptor activation in hepatocytes of the liver. Processes considered are insulin binding to the receptor, receptor autophosphorylation, internalization and recycling. Compared to other models of the insulin receptor which also include these processes [35,43], we provide an extended description, model the *in vivo* situation and include reversible nonspecific insulin binding in the space of Disse.

The kidney's contribution to insulin degradation mainly consists in the filtering of insulin from the blood [7]. The filtering function of the kidney is modeled as a degradation rate that, according to experimental data [73], does not saturate and is proportional to the insulin concentration in the plasma.

Pancreatic insulin secretion is mainly induced by plasma glucose [5]. As we focus on insulin degradation, glucose is not included in the model describing pancreatic insulin secretion in a highly simplified way as a function of insulin concentration. Due to a high robustness to

changes in the parameters for insulin secretion [72, Additional file 3], this simplification does not lead to significant approximation errors in dynamic simulations. In addition, insulin secretion is irrelevant for stationary model analysis at constant insulin concentrations.

Altogether, the presented *in vivo* model describes the following processes: intravenous injection of radioactively labeled and unlabeled insulin, pancreatic insulin secretion, hepatic and renal insulin degradation, hepatic insulin receptor activation and nonspecific insulin binding by the liver.

2.2.2 Parameters, units and volumes

In vivo model parameters cannot be measured directly in most cases. Taking parameters from *in vitro* experiments or models for *in vivo* processes is a promising alternative. It is the only possibility if there is not sufficient experimental data and a model structure that guarantees identifiability which is a frequently occurring situation in systems biology.

In this study, model parameters are taken from previously published *in vitro* experiments [73–80] as well as small models of insulin binding [58], receptor internalization [56] and nonspecific hepatic insulin binding [53]. The models from literature [53, 56, 58] were combined and kinetic parameters for the remaining processes were taken from *in vitro* data (Table 2.1 and Appendix A.3.1).

All volumes are assumed to be constant. In addition, all tissues are assumed to contact the same total insulin concentration, which is the sum of the concentrations of labeled and unlabeled insulin. The physiological justification of this assumption is the high heart rate of rats ($320 - 480$ *bpm* [79]) that guarantees a fast distribution of circulating insulin.

Almost all state variables in the model represent concentrations and are given in nM. Exceptions are the state variables Ins_{ub} and $Ins_{*,ub}$ that represent amounts of substances and are given in $nmol$. All rates are given in $nM \cdot s^{-1}$. The rates describing insulin receptor dynamics (r_j, i_j and f_j, $j \in \mathbb{N}$) refer to the hepatocyte volume v_{hep}. All other rates refer to the blood plasma volume v_p.

2.2.3 The liver

Insulin degradation and insulin receptor activation in hepatocytes are modeled in a very detailed way. Processes considered are: successive binding of two insulin molecules to the insulin receptor, receptor phosphorylation and receptor internalization (Figure 2.1). The following model assumptions are supported by studies from literature.

- *Insulin binding and dissociation are independent of the phosphorylation state of the receptor.* This is directly supported by experimental evidence [71].
- *Only receptors with bound insulin show autophosphorylation activity.* Autophosphorylation is induced by insulin binding [5], and autophosphorylation of receptors without bound insulin is considerably weaker.

- *Receptor dephosphorylation is independent of insulin binding.* Receptor dephosphorylation is performed by protein phosphatases [5]. It seems very unlikely that insulin binding to the extracellular α-chain of the receptor induces conformational changes in the intracellular β-chain that are large enough to significantly change the affinity of phosphatases for their phosphorylated substrate sites.

- *Insulin dissociation from endosomal receptors is irreversible.* Upon internalization, the pH in endosomes decreases rapidly, which promotes insulin dissociation from the receptor [7]. Free endosomal insulin is degraded by proteases [7].

- *Only receptors without insulin are recycled.* Receptor recycling is faster if there is no external insulin [56]. This leads to the assumption that an additional step for receptors with bound insulin is necessary before recycling is possible. A very promising candidate for this step is insulin dissociation from the receptor. In this case, a single rate constant for recycling, independent of insulin concentration, is sufficient to explain the observation.

- *Phosphorylated receptors are internalized faster than unphosphorylated receptors.* In the presence of higher insulin concentrations, more insulin receptors are phosphorylated [5]. Receptor internalization is faster at high insulin concentrations than without external insulin [56]. In addition, there are reports that receptor internalization depends on phosphorylation [7].

- *Labeled and unlabeled insulin show the same physiological characteristics.* Labeling of the insulin molecules was performed with ^{125}I [81–83]. The size of this modification is small compared to the size of the insulin molecule and should not change its binding characteristics, the effect on receptor phosphorylation, the rate of nonspecific insulin binding or the rate of renal insulin filtration.

- *Degradation and synthesis of the insulin receptor is negligible.* Insulin receptor degradation in 3T3-L1 mouse adipocytes has a half-live of 7.5 h [84]. If protein synthesis is inhibited, the decrease in insulin binding in foetal hepatocytes has a half-live of 13 h [85]. Assuming that receptor degradation is not significantly different in adult hepatocytes, receptor turnover is slow compared to all other considered processes and can be neglected.

- *All processes in hepatocytes obey mass action kinetics.* The processes that were adopted from other models obey mass action kinetics [53, 56, 58]. Mass action kinetics are a good and frequently used approximation for processes at the molecular level [22, 24].

For the following assumptions there is no experimental data in literature supporting them. These assumptions were made to keep the number of parameters as low as possible.

- *Receptors with one or two bound insulin molecules show the same autophosphorylation activity.*

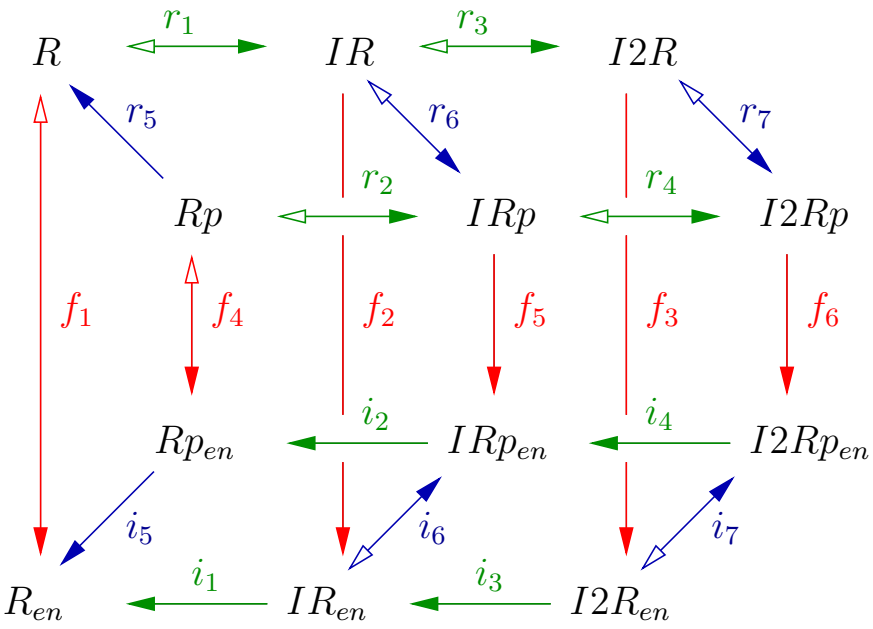

Figure 2.1: Insulin receptor activation in hepatocytes
The insulin receptor is denoted as R. One or two insulin molecules can bind to the receptor (green arrows). This is indicated by a prefix I or $I2$, respectively. Receptor phosphorylation (blue arrows) is indicated by a suffix p, receptor internalization to the endosomal compartment (red arrows) is indicated by a subscript en. Arrows with two heads indicate reversible reactions, whereas arrows with one head indicate irreversible reactions. Filled arrowheads indicate the positive direction of rates.

- *Receptor recycling is independent of receptor phosphorylation.*

Figure 2.1 shows the reaction scheme of processes in hepatocytes. In this figure and in the model equations, the insulin receptor is denoted as R. The binding of one or two insulin molecules is indicated by a prefix I or $I2$, respectively. A suffix p indicates receptor phosphorylation, a subscript en indicates internalization to the endosomal compartment. All concentrations of receptor species refer to v_{hep}, the total volume of hepatocytes.

Rates denoted by the standard notation r_j describe processes at the plasma membrane of hepatocytes ($j \in \mathbb{N}$) or outside the hepatocytes ($j \notin \mathbb{N}$). Rates denoted by i_j describe *internal* processes occurring in endosomes of hepatocytes, and rates denoted by f_j describe *flows* between the plasma membrane and endosomes of hepatocytes ($j \in \mathbb{N}$).

The hepatocyte part of the model does not distinguish between labeled and unlabeled insulin, which reduces the number of necessary ODEs. Hepatocytes have contact to the total insulin concentration Ins that is the sum of the concentrations of labeled and unlabeled insulin. The concentration of labeled insulin is denoted as Ins_*; unlabeled insulin ($Ins - Ins_*$) has no separate notation.

Rates $r_1 - r_4$ describe insulin binding to the insulin receptor at the plasma membrane. The values of the parameters $kins$, $kins1d$ and $kins2d$ were directly taken from the model

of Wanant et al. [58].

$$\begin{aligned} r_1 &= kins \cdot R \cdot Ins - kins1d \cdot IR \\ r_2 &= kins \cdot Rp \cdot Ins - kins1d \cdot IRp \\ r_3 &= kins \cdot IR \cdot Ins - kins2d \cdot I2R \\ r_4 &= kins \cdot IRp \cdot Ins - kins2d \cdot I2Rp \end{aligned} \tag{2.1}$$

The total contribution of the liver to plasma insulin degradation is given as

$$r_{liv} = (-r_1 - r_2 - r_3 - r_4) \cdot v_{hep}/v_p. \tag{2.2}$$

The plasma volume is denoted as v_p, the total hepatocyte volume is denoted as v_{hep}. Strictly speaking, r_{liv} defines insulin removal from the blood, whereas insulin degradation is performed in hepatic endosomes. However, r_{liv} is the contribution of the liver to insulin dynamics. In steady state, the values of the rates for insulin removal and insulin degradation are identical.

Rates $r_5 - r_7$ describe receptor phosphorylation at the plasma membrane.

$$\begin{aligned} r_5 &= kyd \cdot Rp \\ r_6 &= kyp \cdot IR - kyd \cdot IRp \\ r_7 &= kyp \cdot I2R - kyd \cdot I2Rp \end{aligned} \tag{2.3}$$

Rates $i_1 - i_4$ describe insulin dissociation from the receptor in endosomes.

$$\begin{aligned} i_1 &= kins1den \cdot IR_{en} \\ i_2 &= kins1den \cdot IRp_{en} \\ i_3 &= kins2den \cdot I2R_{en} \\ i_4 &= kins2den \cdot I2Rp_{en} \end{aligned} \tag{2.4}$$

Rates $i_5 - i_7$ describe receptor phosphorylation in endosomes.

$$\begin{aligned} i_5 &= kyden \cdot Rp_{en} \\ i_6 &= kyp \cdot IR_{en} - kyden \cdot IRp_{en} \\ i_7 &= kyp \cdot I2R_{en} - kyden \cdot I2Rp_{en} \end{aligned} \tag{2.5}$$

The parameter values in Equations 2.3-2.5 are taken from *in vitro* experiments (see Table 2.1 and Appendix A.3.1). According to our model assumptions, unphosphorylated receptors without insulin (R and R_{en}) have no autophosphorylation activity. Therefore, the reactions represented by the rates r_5 and i_5 are irreversible.

Rates $f_1 - f_6$ describe receptor internalization and recycling, where recycling is only possible for receptors without insulin.

$$\begin{aligned} f_1 &= intk2 \cdot R - reck1 \cdot R_{en} \\ f_2 &= intk2 \cdot IR \\ f_3 &= intk2 \cdot I2R \\ f_4 &= intk1 \cdot Rp - reck1 \cdot Rp_{en} \\ f_5 &= intk1 \cdot IRp \\ f_6 &= intk1 \cdot I2Rp \end{aligned} \tag{2.6}$$

Table 2.1: Model parameters and initial conditions

Note that the values of m_{body} (body weight in g), t_{in} (injection time in s), n_{in} and $n_{*,in}$ (amounts of injected unlabeled and labeled insulin in $nmol$) depend on the analyzed scenario. Initial conditions are: $Ins = 0.07$, $Ins^* = 0$, $R = 31.619$, $IR = 0.430007$, $I2R = 0.000696311$, $Rp = 0.227528$, $IRp = 2.07275$, $I2Rp = 0.00363012$, $R_{en} = 4.88528$, $IR_{en} = 0.145537$, $I2R_{en} = 0.000121295$, $Rp_{en} = 0.122602$, $IRp_{en} = 0.492464$, $I2Rp_{en} = 0.000433466$, $Ins_{ub} = 1.29948 \cdot 10^{-6} \cdot m_{body}$, $Ins_{*,ub} = 0$. The unit of Ins_{ub} and $Ins_{*,ub}$ is $nmol$, the unit of all other state variables is nM (see Appendix A.3.1). ass.: assumption, calc.: calculation.

Parameter	Value	Source	Meaning of the parameter
$kins$	$10^{-3}\ nM^{-1}s^{-1}$	[58]	insulin binding to the receptor
$kins1d$	$4 \cdot 10^{-4}\ s^{-1}$	[58]	insulin dissociation from the receptor (I1, PM)
$kins2d$	$4 \cdot 10^{-2}\ s^{-1}$	[58]	insulin dissociation from the receptor (I2, PM)
$kins1den$	$1.925 \cdot 10^{-3}\ s^{-1}$	[74]	insulin dissociation from the receptor (I1, EN)
$kins2den$	$3.85 \cdot 10^{-3}\ s^{-1}$	[75]	insulin dissociation from the receptor (I2, EN)
kyd	$3.85 \cdot 10^{-3}\ s^{-1}$	[76]	receptor dephosphorylation (PM)
$kyden$	$7.22 \cdot 10^{-3}\ s^{-1}$	[77]	receptor dephosphorylation (EN)
kyp	$0.0231\ s^{-1}$	[77]	autophosphorylation of the receptor (I1 and I2)
$intk1$	$5.5 \cdot 10^{-4}\ s^{-1}$	[56]	internalization of phosphorylated receptors
$intk2$	$2 \cdot 10^{-4}\ s^{-1}$	[56]	internalization of unphosphorylated receptors
$reck1$	$1.533 \cdot 10^{-3}\ s^{-1}$	[56]	recycling of receptors without insulin
$k1ub$	$0.35\ s^{-1}$	[53]	nonspecific insulin binding in the liver
$k2ub$	$0.2\ s^{-1}$	[53]	dissociation of nonspecifically bound insulin
$pansec$	$0.0016976\ nM \cdot s^{-1}$	calc.: [72]	pancreatic insulin secretion
$Kpan$	$0.5\ nM$	ass.	concentration of half-maximal insulin secretion
m_{liver}	$0.05 \cdot m_{body}$	[53]	mass of the liver
v_p	$0.03375 \cdot 10^{-3}\ l \cdot g^{-1} \cdot m_{body}$	[79]	plasma volume
ρ_{liver}	$1.051 \cdot 10^{3}\ g \cdot l^{-1}$	[78]	density of the liver
v_{hep}	$(m_{liver}/\rho_{liver}) \cdot 0.78$	[86]	total hepatocyte volume
v_d	$0.272 \cdot 10^{-3}\ l \cdot g^{-1} \cdot v_{hep} \cdot \rho_{liver}$	[53]	volume of the space of Disse
m_{kidney}	$2 \cdot 0.85\ g \cdot m_{body}/(230\ g)$	[80]	mass of the kidney
$Kkidney$	$0.0225 \cdot 10^{-3}\ l \cdot (s \cdot g)^{-1} \cdot m_{kidney}$	[73]	clearance of the kidney

The value of the parameter $intk1$ was directly taken from a model of receptor internalization and recycling at high insulin concentrations [56]. The values of the parameters $intk2$ and $reck1$ are from a model of receptor internalization and recycling without insulin [56].

Altogether, the described processes (Figure 2.1) result in the following balance equations for hepatic insulin receptor species.

$$\begin{aligned}
\dot{R} &= -r_1 + r_5 - f_1 \\
\dot{IR} &= r_1 - r_3 - r_6 - f_2 \\
\dot{I2R} &= r_3 - r_7 - f_3 \\
\dot{Rp} &= -r_2 - r_5 - f_4 \\
\dot{IRp} &= r_2 - r_4 + r_6 - f_5 \\
\dot{I2Rp} &= r_4 + r_7 - f_6 \\
\dot{R}_{en} &= i_1 + i_5 + f_1 \\
\dot{IR}_{en} &= -i_1 + i_3 - i_6 + f_2 \\
\dot{I2R}_{en} &= -i_3 - i_7 + f_3 \\
\dot{Rp}_{en} &= i_2 - i_5 + f_4 \\
\dot{IRp}_{en} &= -i_2 + i_4 + i_6 + f_5 \\
\dot{I2Rp}_{en} &= -i_4 + i_7 + f_6
\end{aligned} \tag{2.7}$$

The executable receptor model is given in MATLAB format in [72, Additional file 2] and can be used for the simulation of *in vitro* experiments.

The liver also performs nonspecific insulin binding. This reversible process does not saturate [53] and dampens rapid variations in insulin concentration. The rates r_{ub} and $r_{*,ub}$ define nonspecific binding of unlabeled and labeled insulin, respectively.

$$\begin{aligned}
r_{ub} &= (k1ub \cdot (\overbrace{Ins - Ins_*}^{\text{unlabeled insulin}}) \cdot v_d - k2ub \cdot Ins_{ub})/v_p \\
r_{*,ub} &= (k1ub \cdot Ins_* \cdot v_d - k2ub \cdot Ins_{*,ub})/v_p
\end{aligned} \tag{2.8}$$

The values of the parameters $k1ub$ and $k2ub$ were directly taken from the model of Hammond et al. [53]. The volume of the space of Disse, in which nonspecific insulin binding takes place, is denoted as v_d. The concentration of unlabeled insulin is $Ins - Ins_*$ (unit: nM), while $Ins_{*,ub}$ and Ins_{ub} are the amounts of substance (unit: $nmol$) of nonspecifically bound labeled and unlabeled insulin, respectively. Only the expressions for the forward reactions of the rates r_{ub} and $r_{*,ub}$ are multiplied by v_d (unit: l) because Ins and Ins_* (unit: nM) are concentrations, whereas Ins_{ub} and $Ins_{*,ub}$ are amounts of substance (unit: $nmol$).

The balances of the amounts of nonspecifically bound unlabeled and labeled insulin are given as

$$\begin{aligned}
\dot{Ins}_{ub} &= r_{ub} \cdot v_p \\
\dot{Ins}_{*,ub} &= r_{*,ub} \cdot v_p.
\end{aligned} \tag{2.9}$$

In order to obtain the unit $nM \cdot s^{-1}$ for all rates, we divide by v_p within the rates r_{ub} and $r_{*,ub}$ (Equation 2.8) and multiply the rates by v_p in the ODEs for Ins_{ub} and $Ins_{*,ub}$ (Equation 2.9), emphasizing the need for v_p.

Note that species notations with an asterisk indicate radioactively labeled insulin species. Species with insulin whose notations do not contain an asterisk can contain labeled or unlabeled insulin, except for Ins_{ub}, which only represents unlabeled nonspecifically bound insulin.

2.2.4 The kidney

The kidney performs insulin degradation by filtering insulin from the blood [18]. The degradation rate r_{kid} is proportional to the insulin concentration [73].

$$r_{kid} = -Kkidney \cdot Ins/v_p \tag{2.10}$$

Insulin clearance is defined as the quotient of the degradation rate and the insulin concentration [59]. Therefore, $Kkidney$ is the clearance of the kidney.

There are also reports that receptor-mediated transport in man contributes about one third to total renal insulin removal [7]. This may result in a slightly nonlinear behavior of renal insulin degradation. However, the nonlinearity resulting from receptor saturation is not visible in the experimentally examined concentration interval [73]. Therefore, linear first order kinetics are a good approximation of renal insulin degradation.

2.2.5 Insulin secretion and injection

Pancreatic insulin secretion is induced by plasma glucose [5], which is not included in the model. In the model, pancreatic insulin secretion r_{pan} is described as a function of insulin concentration and turned off at high insulin concentrations. This corresponds to the implicit assumption that glucose dynamics are faster than insulin dynamics. Peak concentrations in insulin therapy are about $60 - 80\ \mu U \cdot ml^{-1}$ [62, 63], which is about $0.35 - 0.5\ nM$. Insulin secretion is assumed to be cut off smoothly at $Kpan = 0.5\ nM$. The physiological basal insulin concentration ($0.07\ nM$) is guaranteed by adjusting the parameter *pansec* such that the secretion rate r_{pan} equals the sum of the stationary insulin degradation rates of the liver and the kidney at $0.07\ nM$ insulin [72, Additional files 4 and 5].

$$r_{pan} = pansec \cdot \left(1 - \frac{Ins^{10}}{Ins^{10} + Kpan^{10}}\right) \tag{2.11}$$

Intravenous injection of labeled and unlabeled insulin ($u_{*,in}$ and u_{in}) is performed during the injection time t_{in} with a constant injection rate that is sharply, but smoothly cut off.

$$\begin{aligned} u_{in} &= \frac{n_{in}}{t_{in}} \cdot \left(1 - \frac{t^{50}}{t^{50} + t_{in}^{50}}\right) / v_p \\ u_{*,in} &= \frac{n_{*,in}}{t_{in}} \cdot \left(1 - \frac{t^{50}}{t^{50} + t_{in}^{50}}\right) / v_p \end{aligned} \tag{2.12}$$

The amounts of injected unlabeled and labeled insulin are n_{in} and $n_{*,in}$ (unit: $nmol$), respectively. Each of these parameters can be set to zero if no injection of labeled or unlabeled insulin is desired. The corresponding input functions u_{in} or $u_{*,in}$ then equal zero.

Note that u_{in} and $u_{*,in}$ are not defined for $t_{in} = 0\ s$ which corresponds to a very rapid bolus injection of insulin. If $t_{in} = 0\ s$ is nevertheless desired, this infinitely small injection time can be realized by setting the initial conditions directly to the corresponding values.

$$\begin{aligned} Ins &= 0.07 + \frac{n_{in} + n_{*,in}}{v_p} \\ Ins^* &= \frac{n_{*,in}}{v_p} \end{aligned} \tag{2.13}$$

The rates u_{in} and $u_{*,in}$ then have to be set to zero. Note that 0.07 nM is the basal concentration of insulin.

2.2.6 Insulin concentration in the plasma

The balances of the concentrations of labeled (Ins_*) and total insulin (Ins) are given by:

$$\begin{aligned} \dot{Ins} &= r_{liv} + r_{kid} + r_{pan} + u_{in} + u_{*,in} - r_{*,ub} - r_{ub} \\ \dot{Ins}_* &= (r_{liv} + r_{kid}) \cdot \frac{Ins_*}{Ins} + u_{*,in} - r_{*,ub}. \end{aligned} \tag{2.14}$$

Note that r_{liv} and r_{kid} refer to total insulin. Therefore, only their fractions Ins_*/Ins that correspond to labeled insulin have to be considered in the balance of Ins_*. The distinction between labeled and unlabeled insulin is necessary because unlabeled insulin is synthesized in the pancreas whereas labeled insulin is not. Therefore, in experiments with labeled insulin, the fraction of labeled insulin changes over time.

The executable complete model is given in MATLAB format in [72, Additional file 1]. All parameter values are resumed in Table 2.1.

2.3 Model validation

Model validation is performed with experimental data sets from literature. The experimental data sets are for rats as the parameter values of the model are also for rats. We emphasize that the data sets used for the model validation are not used for parameter estimation which corresponds to a strict separation of model construction and model validation. A very remarkable result of the model validation is that the model with parameters from literature is able to match experimental data sets.

2.3.1 Insulin dynamics

The dynamic insulin degradation behavior of the model was compared to experimentally determined time courses of insulin concentration in plasma after intravenous insulin injections.

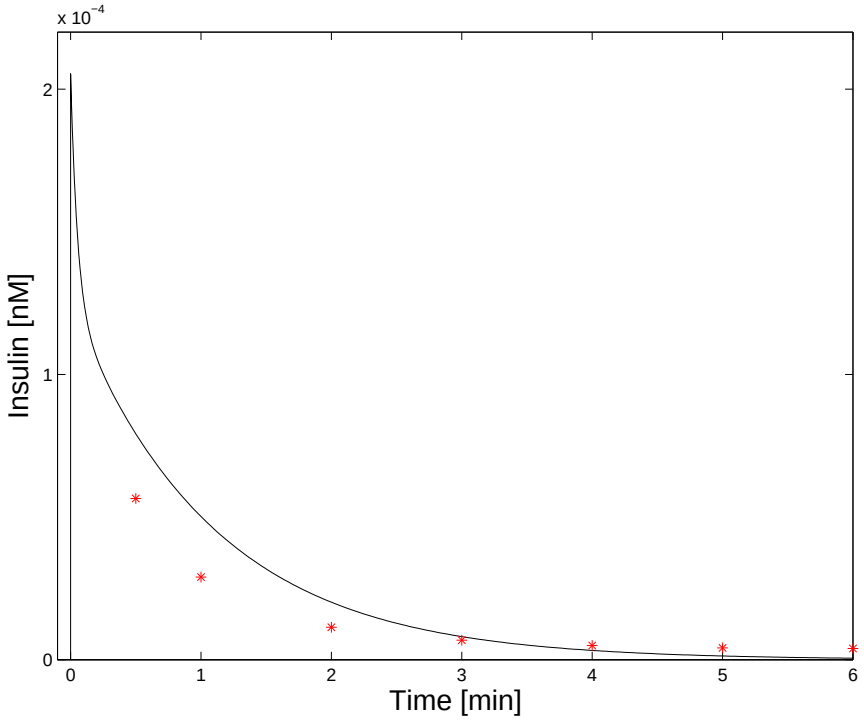

Figure 2.2: Dynamic model validation: physiological insulin concentrations
Simulation of the concentration of radioactively labeled insulin in plasma after the injection of a very low amount of radioactively labeled insulin is shown and compared to experimental data [83].

Experimental data sets with extremely high [82] and extremely low [83] amounts of injected insulin were used. The extremely low amount ($1.65 \cdot 10^{-6}$ $nmol$) was radioactively labeled. Therefore, the dynamics of injected insulin can be tracked though the amount of injected insulin is small compared to endogenous insulin. Measured concentrations of labeled insulin were about $0.5 \cdot 10^{-4}$ nM [83] (see Figure 2.2). In the other experiment, the extremely high amount of injected insulin (47.5 $nmol$) resulted in measured insulin concentrations above 1600 nM [82] (see Figure 2.3). Note that 1600 nM insulin are highly unphysiological because the basal concentration of insulin is 0.07 nM and peak concentrations in insulin therapy and physiological scenarios are below 1 nM [52, 62, 63].

Simulated insulin concentrations for low amounts of injected insulin [83] are relatively close to the experimental data set (Figure 2.2). As the absolute errors are moderate, Figure 2.2 is regarded as a semi-quantitative or at least qualitative validation of the dynamic model at physiological insulin concentrations.

Simulation results for the injection of unphysiologically high amounts of insulin [82] are not very close to the experimental data set (Figure 2.3). The measured insulin concentration of 1600 nM is four orders of magnitude higher than peak concentrations in insulin therapy ($0.35 - 0.5$ nM [62, 63]) and three orders of magnitude above physiological peak values [52]. In this concentration range new unmodeled effects occur. As an example,

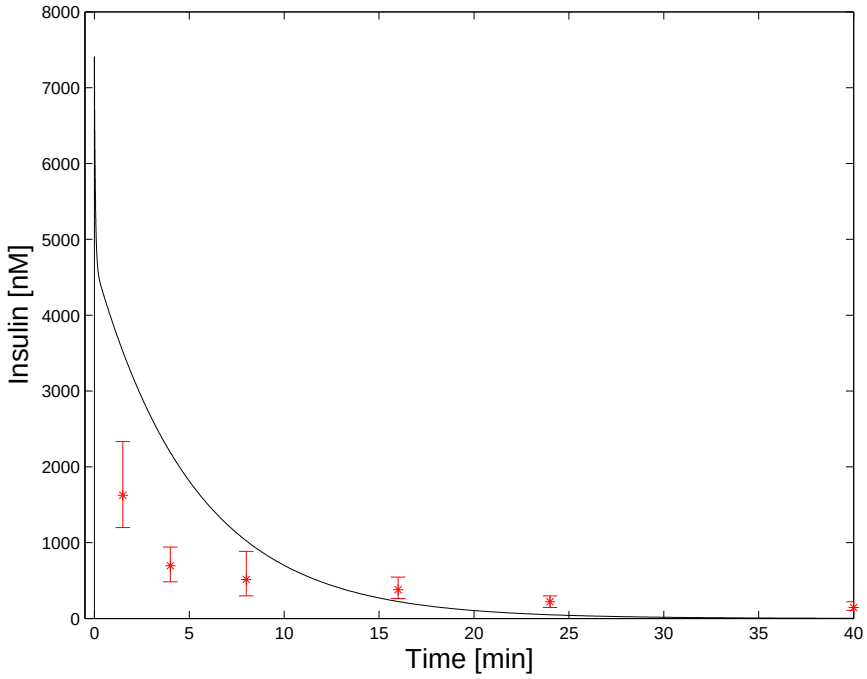

Figure 2.3: Dynamic model validation: extremely high insulin concentrations
Simulation of plasma insulin concentration after the injection of a large amount of insulin is shown and compared to experimental data [82]. The simulation does not match the experimental data set. This results from the presence of unmodeled effects at highly supraphysiological insulin concentrations and limitations in the detection quality of the experiment. Therefore, the model is not valid at these extremely high insulin concentrations.

pinocytosis (fluid-phase endocytosis) significantly contributes to hepatic insulin uptake at high insulin concentrations [7, 87]. In correspondence to nonspecific insulin binding by the liver, nonspecific insulin binding could also occur in other tissues. A result of this additional nonspecific binding would be reversible insulin removal at high insulin concentrations and subsequent insulin release at lower insulin concentrations. Furthermore, the assay of Desbuquois et al. [82] is not able to distinguish between insulin fragments and native insulin. After a few minutes however, insulin fragments contribute significantly to total insulin, as shown for the injection of small amounts of labeled insulin [83, Figure 3]. Assuming that this also holds for the injection of high amounts of insulin, the assay of Desbuquois et al. overestimates insulin concentrations at later points in time.

The effects of pinocytosis and additional nonspecific insulin binding at high insulin concentrations are not quantified in literature and not included in the model. Neglecting these processes (and maybe others that are important at high insulin concentrations) leads to an incorrect model structure for high insulin concentrations. Therefore, the model is not valid at extremely high insulin concentrations.

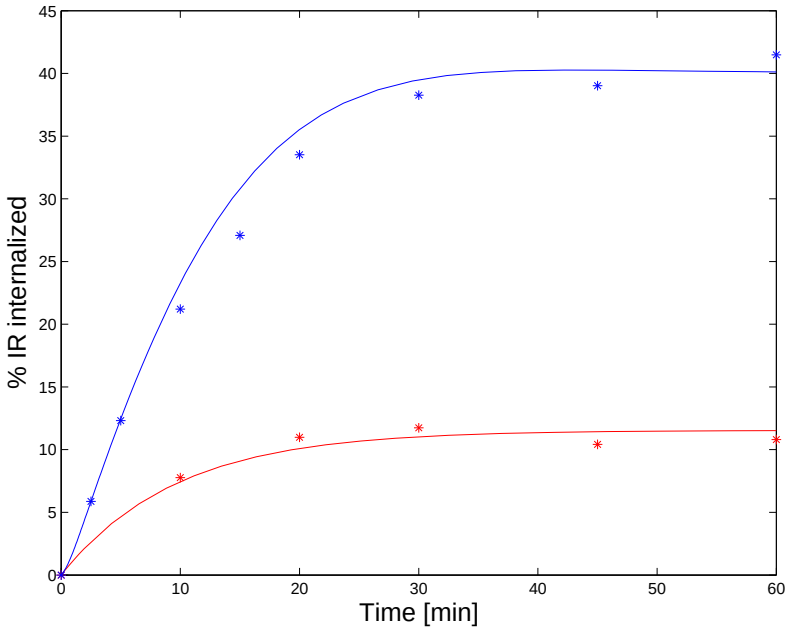

Figure 2.4: Dynamic model validation: receptor internalization
Simulation results for receptor internalization at 100 nM insulin (blue) as well as those without insulin (red) are shown and compared to experimental data [56]. Surface receptors were radioactively labeled at $t = 0$ s. This was simulated by taking the receptor model (Equation 2.7) and setting the initial conditions such that all receptors are in the state R at the plasma membrane. The assay can be simulated with this choice of initial conditions because the receptor model is linear for a constant insulin concentration.

2.3.2 Hepatic insulin receptor internalization

Simulation results for hepatic insulin receptor internalization at 100 nM insulin as well as those without insulin were compared to experimental data from literature [56]. As it can be seen in Figure 2.4, simulation results match the experimental data sets very well. Assuming that simulated insulin binding to the receptor mirrors physiological processes well (we show in Section 2.3.3 that it does), receptor internalization can be used as a direct indicator for hepatic insulin degradation.

The experimental data sets for receptor internalization result from experiments with Fao cells that are tumor cells of hepatic origin. Though simulations match the experimental data sets almost quantitatively, this can only be regarded as a qualitative model validation for hepatocytes.

2.3.3 Insulin binding and receptor phosphorylation

Simulation results for stationary insulin receptor phosphorylation and insulin binding were compared to experimental data sets [81] to validate the receptor part of the model.

Klein et al. determined cell-associated radioactively labeled insulin as a function of the

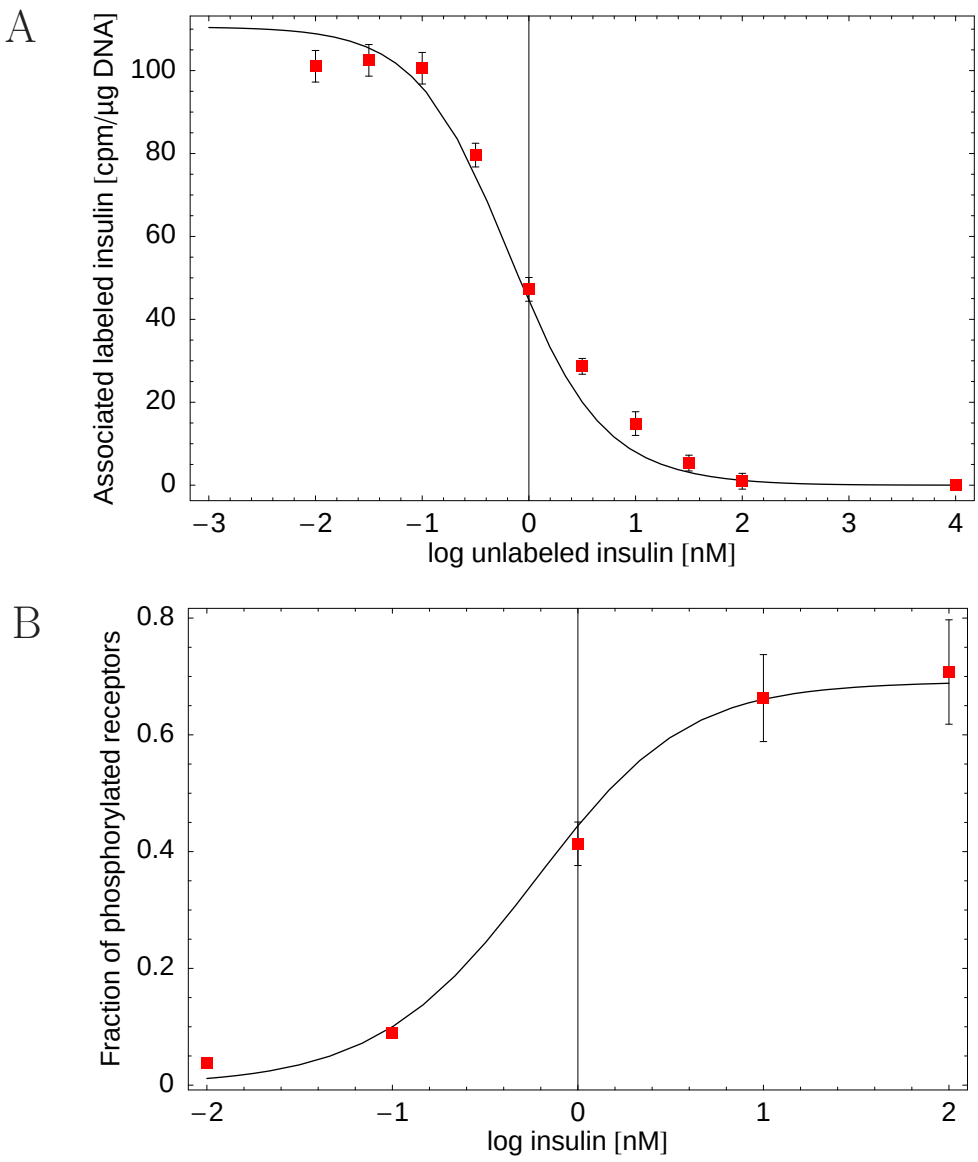

Figure 2.5: Stationary model validation: insulin binding and receptor phosphorylation
A) Cell-associated radioactively labeled insulin is shown as a function of the stationary concentration of unlabeled insulin and compared to experimental data [81, Figure 4 A]. Almost no labeled insulin should bind to receptors at maximal concentrations of unlabeled insulin. Therefore, the experimentally determined value for the highest concentration of unlabeled insulin was treated as background and subtracted from all values.
B) The fraction of phosphorylated receptors is shown as a function of the stationary insulin concentration and compared to experimental data for receptor activation [81, Figure 4 B]. We regard receptor phosphorylation as a good indicator for receptor activity.

concentration of unlabeled insulin [81, Figure 4 A]. This was done in a competition assay with a constant concentration of radioactively labeled insulin (0.01 nM) and variable concentrations of unlabeled insulin Ins_{unlab}. In steady state, cell-associated labeled insulin Ins_{cell} is given as

$$Ins_{cell} = \frac{0.01\ nM}{0.01\ nM + Ins_{unlab}} \cdot (IR + 2 \cdot I2R + IRp + 2 \cdot I2Rp) \tag{2.15}$$

if one assumes the same binding characteristics for labeled and unlabeled insulin. Only the labeled insulin attached to the surface of the cells was measured. Labeled insulin bound to internalized receptors is neither considered in the experiment nor in Equation 2.15.

The stationary model equations (Equation 2.7 with one ODE being replaced by a conservation relation for the insulin receptor and all derivatives set to zero) were solved for the concentrations of the species under the assumption of a constant insulin concentration. Inserting the solution into Equation 2.15 leads to a (cumbersome) expression for Ins_{cell} as a function of the concentration of unlabeled insulin in steady state. This *in silico* reproduction of the assay and projection of the result on the experimental data set results in a high accordance (Figure 2.5 A).

Klein et al. also determined the stationary dependency of receptor activity on the insulin concentration [81, Figure 4 B]. We regard receptor activity as an indicator for receptor phosphorylation. The stationary fraction of phosphorylated receptors in the model is given as

$$F_p = (Rp + IRp + I2Rp + Rp_{en} + IRp_{en} + I2Rp_{en})/R_{tot} \tag{2.16}$$

where $R_{tot} = 40\ nM$ is the constant total concentration of the receptor [88] (see Appendix A.3.1). The projection of the experimental data set for receptor activity on F_p shows a high accordance (Figure 2.5 B).

Altogether, the model is able to match the experimental data sets for receptor activity and insulin binding very well. The model parameters were not estimated to get these results.

2.4 Model analysis

2.4.1 Insulin degradation

The fractions of insulin that are degraded by the liver and the kidney were investigated in several studies. Values for the relative contribution of the liver to insulin degradation in man range from below 50 % to 70 %, and those for the kidney from 30 % to above 50 % [6, 7, 18]. We investigate the reason for this diversity by stationary model analysis.

Renal insulin degradation does not saturate [7, 73], whereas hepatic insulin degradation saturates [7, 18, 89]. The physiological situation is mirrored by the model, where hepatic insulin degradation ($-r_{liv}$) saturates, whereas renal insulin degradation ($-r_{kid}$) does not (Figure 2.6 A). It can be clearly seen that the relative contributions of the liver and the kidney to total insulin degradation strongly depend on the insulin concentration (Figure 2.6 B).

In stationary model analysis, the relative contribution of the liver to overall insulin degradation ranges between 81 % for insulin concentration tending to zero and 0 % for insulin concentration tending to infinity. The relative contribution of the kidney ranges between 19 % and 100 %. A significant part of these changes happens beyond physiological insulin levels. However, the fractions vary strongly in the physiological range of insulin concentration. Between 0 nM and 1 nM insulin, the relative contribution of the liver is between 81 % and 63 %, while the contribution of the kidney is between 19 % and 37 % (Figure 2.6 B). Only the liver and the kidney are considered in the analysis of insulin degradation. Other insulin degrading tissues, in particular fat and muscle, are neglected. Therefore, the sum of the relative contributions of the liver and the kidney to insulin degradation is one (100 %).

Note that changes in t_{in} or in the parameters for the pancreas do not affect the results of stationary model analysis, as the system is analyzed at constant insulin concentrations. The rate of nonspecific insulin binding equals zero in the unique steady state. Therefore, it also has no influence on stationary insulin degradation. The stationary analysis of degradation rates and relative contributions to insulin degradation is also independent of the parameter m_{body}. All stationary calculations are given in [72, Additional file 4].

Altogether, the relative contributions of the liver and the kidney to insulin degradation depend on the insulin concentration. At low insulin concentrations, hepatic insulin degradation is predominant, whereas at high insulin concentrations overall insulin degradation is mainly performed by the kidney. Therefore, different results for the relative contributions of the liver and the kidney to insulin degradation are expected for different experimental settings.

2.4.2 Insulin clearance

The quotient of insulin degradation rate and insulin concentration is denoted as insulin clearance c [59], which is a widely used quantity to characterize the state of insulin metabolism.

$$c = \frac{(-r_{liv} - r_{kid}) \cdot v_p}{Ins} = c_{liv} + c_{kid} \tag{2.17}$$

The value of insulin clearance corresponds to the volume of plasma from which insulin is completely removed (or cleared) per unit of time [52]. The physiological range of insulin clearance in man (70 kg) is $700-3350\ ml \cdot min^{-1}$ [6,59]. Most determined values from healthy individuals are in the interval $840-1050\ ml \cdot min^{-1}$, but also values as low as $140\ ml \cdot min^{-1}$ are sometimes measured [52]. Two frequently used methods to measure insulin clearance are described in Appendix A.4. One of the methods guarantees a constant insulin concentration during the measurement whereas the insulin concentration is not constant when using the other one. For nonlinear insulin dynamics, applying both methods results in different values for insulin clearance (Appendix A.4).

We use stationary model analysis to investigate why the measured values of insulin clearance vary in such wide ranges.

Renal insulin clearance $c_{kid} = Kkidney$ is independent of the insulin concentration, as the degradation rate r_{kid} is proportional to the insulin concentration (Equation 2.10). Therefore,

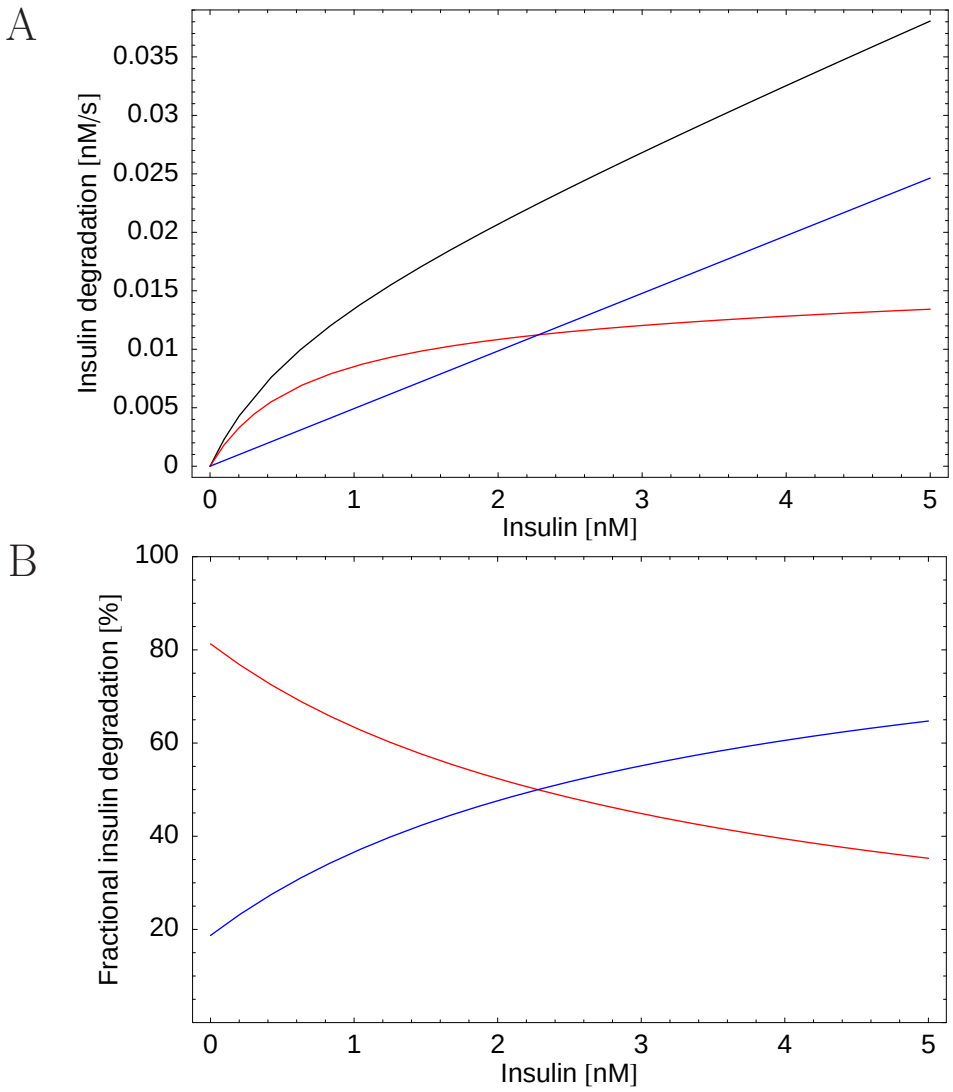

Figure 2.6: Stationary analysis of renal and hepatic insulin degradation
A) Stationary insulin degradation rates of the liver (red) and the kidney (blue) and the total insulin degradation rate (black) are shown as functions of insulin concentration.
B) Stationary relative contributions of the liver (red) and the kidney (blue) to total insulin degradation depend on the insulin concentration. Note that these fractions are slightly lower in reality because other tissues, in particular fat and muscle, also contribute to insulin degradation but are not analyzed here. The fractions in this plot refer to the sum of the degradation rates of liver and kidney.

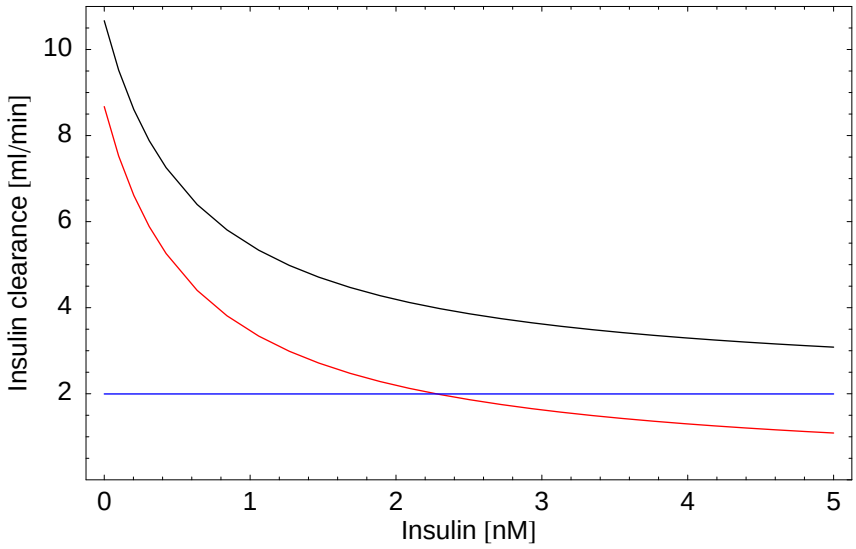

Figure 2.7: Stationary analysis of renal and hepatic insulin clearance
Insulin clearance is defined as the quotient of insulin degradation rate and insulin concentration (Equation 2.17). Total stationary insulin clearance (black) is a function of insulin concentration because hepatic insulin clearance (liver, red) depends on the insulin concentration, whereas renal insulin clearance (kidney, blue) is independent of the insulin concentration. A body weight of $m_{body} = 200\ g$ was used in the computations.

the quotient of rate and concentration is constant. A different situation occurs in the case of hepatic insulin clearance $c_{liv} = -r_{liv} \cdot v_p \cdot Ins^{-1}$ where the insulin degradation rate r_{liv} is nonlinear in the insulin concentration. The effect of Ins^{-1} strongly dominates the effect of the saturating degradation rate r_{liv} (compare Figures 2.6 A and 2.7). Therefore, hepatic insulin clearance decreases for increasing insulin concentrations and tends to zero for insulin concentration tending to infinity.

Altogether, due not nonlinearities in insulin degradation, insulin clearance strongly depends on the insulin concentration (Figure 2.7). As an example, in a rat whose body weight is 200 g, insulin clearance ranges between 10.7 $ml \cdot min^{-1}$ for insulin concentration tending to zero and 2.0 $ml \cdot min^{-1}$ for insulin concentration tending to infinity [72, Additional file 4].

Insulin clearance is often used to characterize the state of insulin metabolism. This is justified in the analysis of linear first order processes such as renal insulin degradation, where insulin clearance is independent of insulin concentration (Figure 2.7). However, the usefulness of insulin clearance for the analysis of processes that are dominated by saturable components, in particular hepatic insulin degradation, is very limited. A strong dependence on the insulin concentration hampers precise analyses, especially if the insulin concentration is not constant during the experiment.

An implication from this analysis is that insulin clearance should always be measured with methods that keep the insulin concentration constant during the experiment. In addition, the corresponding insulin concentration is necessary to asses the meaning of the value. Unfortunately, the corresponding insulin concentration is not given in most experimental

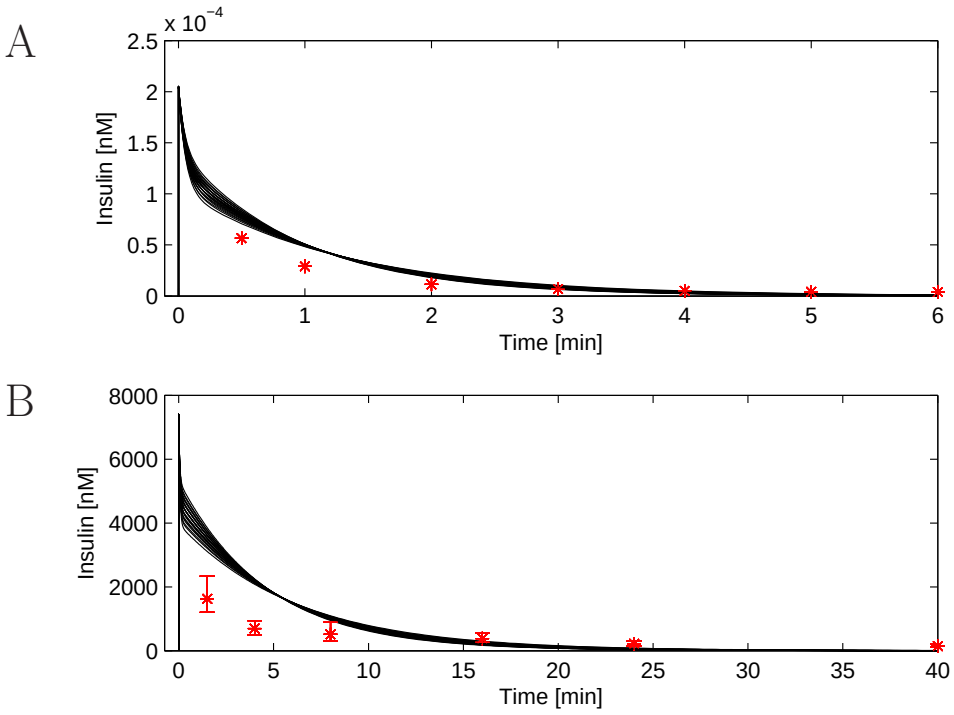

Figure 2.8: Sensitivity of simulation results to changes in $k1ub$ and $k2ub$
Simulations for each combination of $(0.8, 0.9, 1, 1.1, 1.2) \cdot k1ub$ and $(0.8, 0.9, 1, 1.1, 1.2) \cdot k2ub$ are shown. **A)** Simulation of the concentration of radioactively labeled insulin in plasma after the injection of a very low amount of radioactively labeled insulin is shown and compared to experimental data [83]. **B)** Simulation of plasma insulin concentration after the injection of a large amount of insulin is shown and compared to experimental data [82].

studies from literature.

Altogether, the strong dependency of insulin clearance on the insulin concentration is able to explain the wide range of reported values.

2.4.3 Parameter estimation

We investigate whether the model structure is able to reproduce the experimental data set for high amounts of injected insulin (Figure 2.3). The parameters estimated to match the experimental data set for high amounts of injected insulin [82] were chosen by the following considerations. Insulin degradation at high insulin concentrations is mainly performed by the kidney (Figure 2.6). Nonspecific insulin binding dampens rapid variations in insulin concentration at all insulin concentrations (Figure 2.8). Therefore, the most important parameters at high insulin concentrations are those for the kidney and nonspecific insulin binding.

It was possible to get simulation results much closer to the experimental data set for high amounts of injected insulin by changing the values of $k1ub$, $k2ub$ and $Kkidney$ simultaneously or only the value of $Kkidney$ [72, Additional file 3]. From a physiological point of view

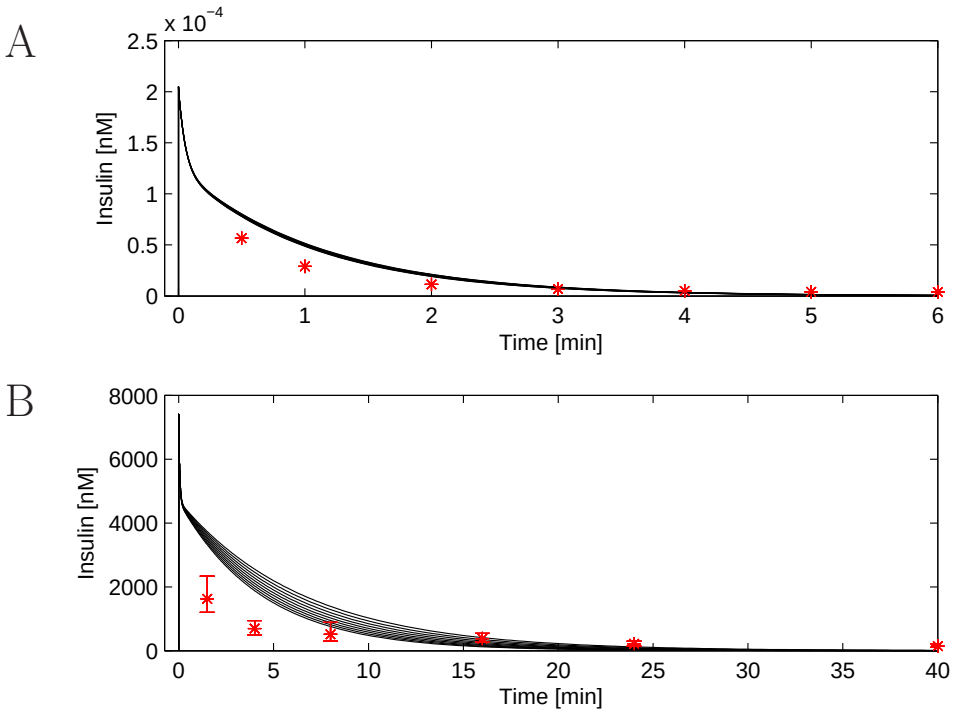

Figure 2.9: Sensitivity of simulation results to changes in $Kkidney$
The value of $Kkidney$ was varied in the interval $0.8 \cdot Kkidney \leq Kkidney \leq 1.2 \cdot Kkidney$. **A)** Simulation of the concentration of radioactively labeled insulin in plasma after the injection of a very low amount of radioactively labeled insulin is shown and compared to experimental data [83]. The differences between the distinct simulations vanish due to the thickness of lines. **B)** Simulation of plasma insulin concentration after the injection of a large amount of insulin is shown and compared to experimental data [82].

however, the better matching of experimental data and simulation results using estimated parameter values does not reflect increased model quality. As discussed above, in the presence of high insulin concentrations there are unmodeled effects which are not important at physiological insulin concentrations. Additionally, the assay of Desbuquois et al. [82] also detects insulin fragments. As a result, the estimated parameter values also include the effects of processes not explicitly described in the model and detection errors of the assay. Therefore, the model analysis was performed using the parameter values from literature (Table 2.1).

2.4.4 Sensitivity analysis

Since the model fails to match the experimental data set for high amounts of injected insulin (Figure 2.3), we investigate if small changes in the parameter values have a significant impact on the simulation results of the dynamic model validation. The focus is on the parameters for the kidney and nonspecific insulin binding which are most important at high insulin concentrations.

Increasing or decreasing the values of $k1ub$ and $k2ub$ by 20% results in moderate differences

in the simulation results for the injection of high and low amounts of insulin (Figure 2.8). Changing $Kkidney$ by 20% has practically no effect on simulation results for the injection of small amounts of insulin, whereas it results in moderate differences in the simulation results for the injection of high amounts of insulin (Figure 2.9). All simulation results with changed parameter values match the experimental data set for low amounts of injected insulin in an acceptable way (Figures 2.8 and 2.9). On the other hand, the values of $k1ub$, $k2ub$ and $Kkidney$ cannot be arbitrarily chosen. A nice example for this is that simulation results using an estimated parameter set (nsq, [72, Additional file 3]) fail to match the experimental data set for small amounts of injected insulin. Therefore, the values from literature are at least acceptable estimates for the parameterization of the processes they represent.

The parameters for the pancreas ($pansec$, $Kpan$, Hill coefficient) are chosen to guarantee that insulin secretion is turned off at peak concentrations in insulin therapy. They have negligible influence on the simulation results for insulin dynamics, as long as the physiological basal level of insulin (0.07 nM) is guaranteed by adjusting $pansec$. For detailed descriptions and simulation results see [72, Additional file 3].

Simulation results for high amounts of injected insulin are robust to changes in the parameters for the liver because the fractional contribution of the liver to insulin degradation is almost zero for such high insulin concentrations. In contrast to that, simulation results for insulin receptor internalization *in vitro* (Figure 2.4) only depend on the parameters for the liver.

Changes in the parameter t_{in} have little influence on the simulation results for high amounts of injected insulin. Simulation results for small amounts of injected insulin however, are sensitive to changes in t_{in} (Figure 2.10). If one assumes that $t = 0$ s corresponds to the end of insulin injection, simulation results are in each case relatively close to the experimental data set. However, if one assumes that insulin injection starts at $t = 0$ s, simulation results for slow injections (high values of t_{in}) are not close to the experimental data set any more (Figure 2.10). Unfortunately, the exact procedure of injection is not described in either study [82, 83]. We assume a bolus injection at $t = 0$ s for both experiments.

Altogether, simulation results for insulin dynamics are sensitive to changes in t_{in} but robust to changes in the other parameters. The deviations of the simulation results for high amounts of injected insulin from experimental data cannot be significantly reduced by small changes in the parameter values. As discussed above, these deviations result from an incorrect model structure at extremely high insulin concentrations.

2.5 Therapeutic insulin concentrations

The aim of insulin therapy is to achieve sufficient glucose uptake with minimal amounts of insulin [18, 19, 90]. An interesting question is whether an upper bound for reasonable insulin concentrations exists. We investigate this by combining the results of stationary model analysis and experimental studies from literature.

At about 10 nM insulin, the insulin receptor in hepatocytes of rats is almost maximally phosphorylated (Figure 2.5). Stationary model analysis shows that almost 80 % of insulin is

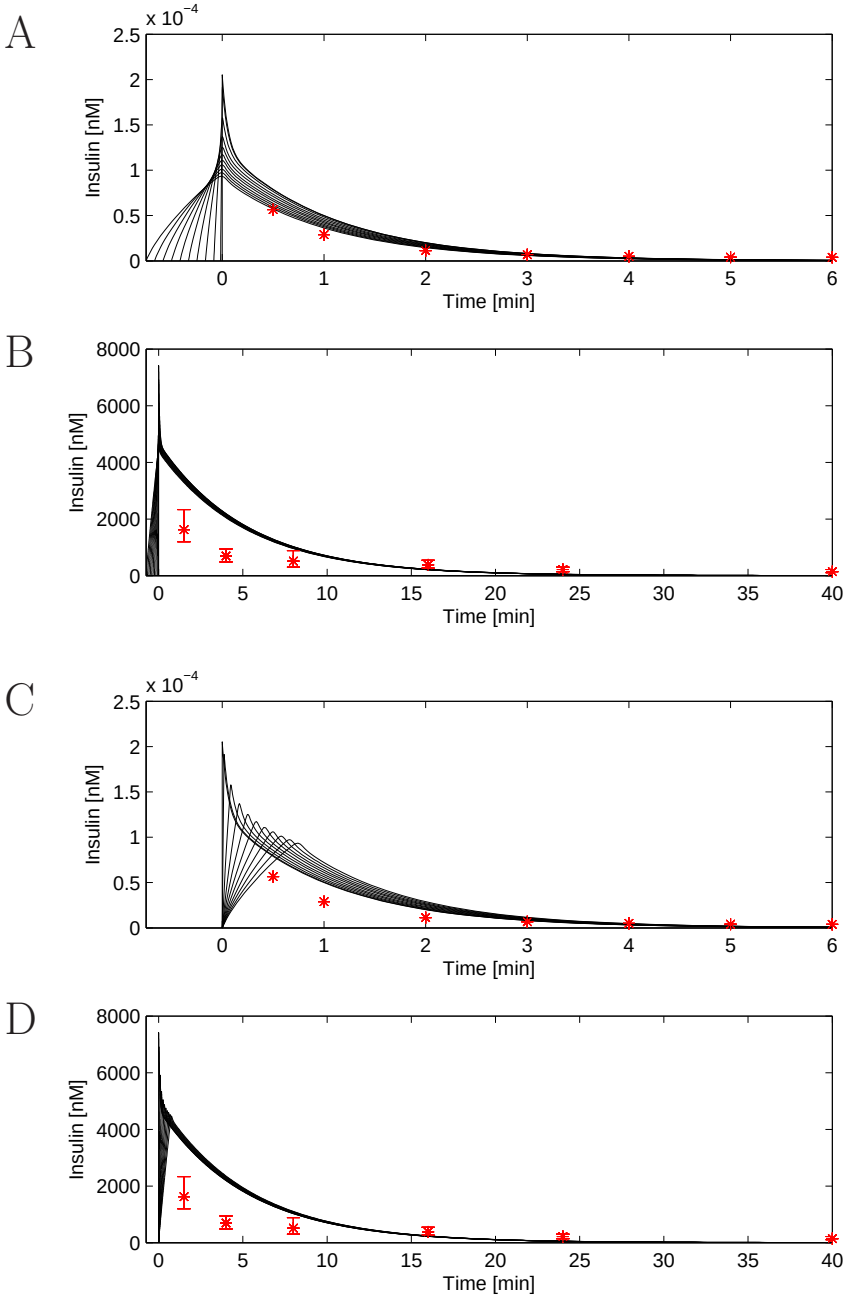

Figure 2.10: Sensitivity of simulation results to changes in t_{in}
The value of t_{in} (duration of the injection) was varied in the interval $0\ s \leq t_{in} \leq 45\ s$. In the plots A and B, the point in time $t = 0\ s$ denotes the end of insulin injection. Therefore, simulations start with the beginning of insulin injection at $t = -t_{in}$. In the plots C and D, the point in time $t = 0\ s$ denotes the beginning of insulin injection. **A,C)** Simulation of the concentration of radioactively labeled insulin in plasma after the injection of a very low amount of radioactively labeled insulin is shown and compared to experimental data [83]. **B,D)** Simulation of plasma insulin concentration after the injection of a large amount of insulin is shown and compared to experimental data [82].

degraded by the kidney and does not contribute to insulin receptor activation. The fraction of insulin that is degraded by the kidney further increases with increasing insulin concentration.

Half-maximal insulin receptor phosphorylation in rat adipocytes is at $7 \pm 1\ nM$ insulin [91]. Glucose uptake in adipocytes is half-maximal at $170\ pM$ insulin and saturates at about $3\ nM$ insulin [91]. These findings are expected to hold qualitatively also for human adipocytes. Therefore, an upper bound for therapeutic insulin concentrations in man seems to exist. This upper bound is the insulin concentration above which a higher insulin concentration does not result in a higher glucose uptake but only leads to increased insulin degradation. Characteristics of glucose uptake in rat adipocytes imply that this upper bound is at about $3\ nM$ and not at about $10\ nM$ as implied by the characteristics of hepatic insulin receptor phosphorylation.

We consider therapeutic insulin levels in the light of these findings. Overnight control of glucose concentration is performed with basal insulins that show slow absorption kinetics or with continuous injection of short acting insulins [18]. In both cases, only insulin concentrations close to the physiological basal concentration are expected, which is far below the upper bound. In contrast to this, relatively large amounts of insulin or insulin analogues are injected or infused in postprandial glucose control. This mimics the physiological response of healthy individuals to rising glucose concentrations in the blood [20, 92]. Insulin concentrations in physiological scenarios are usually lower than $150\ \mu U \cdot ml^{-1}$ ($\approx 0.93\ nM$) [52]. Postprandial plasma insulin concentration after a standard meal peaks at $60 - 80\ \mu U \cdot ml^{-1}$ [62,63], which is about $0.35 - 0.5\ nM$ and also below the proposed upper bound. Additionally, hepatic insulin degradation is predominant below $1\ nM$ insulin (Figure 2.6) and glucose uptake of adipocytes is strongly, but not fully activated [91].

Therefore, the theoretical upper bound for reasonable therapeutic insulin concentrations in rats (about $3\ nM$) lies significantly above therapeutic insulin levels in humans (about $0.5\ nM$). We suppose that the upper bound for reasonable therapeutic insulin concentrations in man is relatively close to the value postulated for rats.

Altogether, mathematical analysis and experimental results indicate that peak concentrations in insulin therapy are below the upper bound where a higher insulin concentration does not result in a stronger physiological effect.

2.6 Conclusions

The presented dynamic model describes *in vivo* insulin dynamics and hepatic insulin receptor activation in the rat. Model parameters are taken from *in vitro* experiments and other models. Using these parameter values, the model is able to reproduce experimental data sets from literature without parameter estimation.

The vast majority of statements about insulin degradation and insulin clearance in the literature is given without explicitly defining the corresponding insulin concentration, and the reported values widely vary. Mathematical analysis shows that relative contributions of the liver and the kidney to total insulin degradation highly depend on the insulin concentration. At low insulin concentrations, insulin is mainly degraded by the liver, whereas renal

insulin degradation is predominant at high insulin concentrations. This explains variations in reported values of relative contributions to insulin degradation.

Mathematical analysis also shows that insulin clearance strongly depends on the insulin concentration, which explains variations in literature data. Due to the concentration dependence of insulin clearance, measurements should always be performed at a constant insulin concentration and the insulin concentration during the measurement should also be given. This however is uncared-for in many studies from the literature.

The analysis of the relative contributions to insulin degradation and the dose-response characteristics of insulin receptor activation and glucose uptake imply the existence of an upper bound for reasonable therapeutic insulin concentrations. Higher insulin concentrations do not result in higher glucose uptake and additional insulin is degraded without having therapeutic effect. However, the upper bound for reasonable therapeutic insulin concentrations is above peak concentrations in insulin therapy.

The model presented here can be used as a starting point for *in vivo* modeling and analysis of the signaling cascades emerging from the hepatic insulin receptor (e.g. MAP kinase cascade and PI3K pathway). This will significantly contribute to understanding the effect of insulin on hepatocytes. In particular, insulin signaling is modulated by feedback phosphorylation of several sites on the receptor and IRS [33]. To understand these effects *in vivo*, much more detailed models reflecting complex formation at the receptor and regulatory phosphorylation are necessary. However, due to combinatorial complexity in signal transduction (Section 1.3), it is only possible to model the insulin signaling system in such detail if a reduced order description is performed.

The next chapter gives an introduction to model reduction which aims at replacing large models by smaller ones while preserving a high approximation quality. However, the presented methods are only partly applicable to biological signal transduction systems and even the existing specialized techniques cannot address all systems. In particular, a major drawback of most model reduction methods is that a potentially very large conventional model is required before the reduced order model can be generated.

Layer-based modeling, a new approximative but accurate technique, was developed to overcome the limitations resulting from combinatorial complexity. After the introduction to model reduction (Chapter 3), layer-based modeling is introduced in Chapter 4.

3 Model reduction

Conventional deterministic models of biological systems often consist of many ODEs. As shown in Section 1.3, this holds in particular for models of signaling systems with inherent combinatorial complexity. Using such large models to perform many numerical simulations, e.g. for the estimation of parameter values or for model analysis, is extremely time-consuming and rarely possible. This demonstrates the need for reduced order models with a high approximation quality allowing the modeler to perform simulations with acceptable computational effort.

This chapter gives an introduction to model reduction, i.e. the approximation of a large-scale dynamic system by a reduced order model, and introduces the most important techniques for the reduced order modeling of cellular signaling systems.

A frequently applied method for the modeling of signaling systems is that of intuitive model reduction. Section 3.1 discusses this heuristic approach and its often insufficient approximation quality. Formal considerations about systematic model reduction are given in Section 3.2. This introduction is followed by Section 3.3 containing a short review of matrix decomposition methods which are the basis of many model reduction approaches.

Some methods for the reduction of linear systems are shortly introduced in Section 3.4. As biological signal transduction systems are usually nonlinear, these techniques can only be applied to linearized systems, which may have lost interesting dynamic properties and are not a valid global description. Section 3.5 introduces proper orthogonal decomposition (POD), which can also be applied to nonlinear systems, and shows that it is in most cases not suited for the modeling of signaling systems. Model reduction methods basing on time hierarchy are introduced in Section 3.6. A focus in these descriptions is on the assumptions of rapid equilibrium and quasi-steady-state that are of high importance in the modeling of biochemical systems.

Domain-oriented model reduction allows for an exact description of macroscopic quantities with a reduced order model. This specialized model reduction technique for biological signal transduction systems is introduced in Section 3.7.

3.1 Intuitive model reduction

A frequently used approach for the modeling of cellular signaling systems is to simplify the combinatorial variety by focusing on small subsets of the possible reactions and complexes [34–41]. In most cases, a temporal order of the considered processes is assumed which leads to highly branched signaling networks being simplified to ordered pathways.

This method can work well, but there are also cases in which the simulations show a poor accuracy when compared to those of a model considering all possible complexes and reactions. As apparently well-founded assumptions may result in large approximation errors [93], it is very difficult to decide which reactions and complexes can be neglected.

In intuitively reduced models, most processes only occur if certain preconditions are fulfilled. As an example, it is often assumed that a binding site on a receptor can only be phosphorylated if a ligand is bound to the receptor or if the receptor is dimerized [39–41]. This strongly reduces the number of necessary state variables and reactions, since a highly branched reaction network is replaced by a reaction chain. However, there may be dramatic consequences for the approximation quality (see below).

Kholodenko et al. [39] applied the intuitive model reduction approach to EGF signaling. They replaced parts of this highly branched reaction network by a chain of reversible reactions. Receptor activation is described by the sequence: EGF binding to the receptor monomer, dimerization of liganded receptors, receptor dimer phosphorylation, effector binding. At the first glance, this sequence corresponds to biological knowledge. However, implications of this reaction chain are that EGF dissociation from the receptor is not possible as long as the receptor is dimerized and that dissociation of the receptor dimer is not possible as long as the receptor is phosphorylated. Some of the model assumptions are questionable, however, simulation results are comparable to those of a model that accounts for combinatorial complexity [94].

It was also shown for a model of FcεRI signaling [95] that for some parameter values (including the nominal ones) an intuitively reduced model can be a good approximation of a conventional model considering all possible protein complexes and reactions [96]. The structure of the reduced order model (i.e. the temporal order of the processes) however depends on the parameter values of the reaction network [96]. An enormous problem is that these parameter values are often not known and can only be estimated after the set-up of the model equations (for which they are necessary).

3.1.1 Exemplification of intuitive model reduction

In an example system also used in other chapters, a receptor can bind a ligand L and perform autophosphorylation. An effector E can bind to the phosphorylated binding site on the receptor. To simplify the discussion, the receptor is denoted as D in the conventional *detailed* model and as R in the *reduced* order model. A conventional model contains six receptor species connected by seven reactions.

$$\begin{array}{rcll} D[0,X]+L & \rightleftharpoons & D[L,X] & k_1 \quad k_{-1} \\ D[0,0] & \rightleftharpoons & D[0,p] & k_2 \quad k_{-2} \\ D[L,0] & \rightleftharpoons & D[L,p] & k_3 \quad k_{-3} \\ D[X,p]+E & \rightleftharpoons & D[X,E] & k_4 \quad k_{-4} \end{array} \tag{3.1}$$

The first rule expands to three reactions and the second rule corresponds to two reactions. Note that the second and the third reaction are no rules.

Applying the intuitive reduction approach, it is assumed that the processes only take place in the temporal order: ligand binding, receptor phosphorylation, effector binding.

$$\begin{array}{rcll} R+L & \leftrightharpoons & RL & k_1 \quad k_{-1} \\ RL & \leftrightharpoons & RLp & k_3 \quad k_{-3} \\ RLp+E & \leftrightharpoons & RLE & k_4 \quad k_{-4} \end{array} \tag{3.2}$$

Describing the system by this chain of three reactions including only four receptor species neglects ligand dissociation from phosphorylated receptors as well as phosphorylation and dephosphorylation of unliganded receptors. The problematic (implicit) assumption in this reaction chain is that receptor phosphorylation prevents ligand dissociation.

We perform a stationary model analysis to investigate the consequences of these simplifications. The stationary concentrations of the species can be obtained from the ODEs describing the dynamics of the systems which can be derived as discussed in Section 1.2.3. After replacing one ODE of each model by a conservation relation for the receptor and setting all derivatives to zero, the equations can be analytically solved for the concentrations of the species. For the sake of simplicity, we assume that the concentrations of L and E are constant.

As degrees of occupancy are important characteristics of such a system [97], we look at the fraction of liganded receptors F_L, whose stationary value is given by relatively simple analytic expressions in both models. In the stationary analysis of F_L, only the equilibrium constants K_i, which are the quotients of the kinetic parameters k_i and k_{-i}, occur. As the expressions for F_E (fraction of receptors with a bound effector) and F_p (fraction of phosphorylated receptors) in the conventional model are quite cumbersome and also depend on k_i and k_{-i}, we do not give the equations here and only discuss F_L.

The stationary fraction of liganded receptors $F_{L,c}$ in the conventional model is given as

$$F_{L,c} = \frac{D[L,0] + D[L,p] + D[L,E]}{R_{tot}} = \frac{K_1 L}{1 + K_1 L} = 1 - \frac{1}{1 + K_1 L} \tag{3.3}$$

where R_{tot} is the constant sum of all receptor species. The fraction of liganded receptors $F_{L,r}$ in the intuitively reduced model is

$$\begin{aligned} F_{L,r} = \frac{RL + RLp + RLE}{R_{tot}} &= \frac{K_1 L(1 + K_3 + K_3 K_4 E)}{1 + K_1 L(1 + K_3 + K_3 K_4 E)} \\ &= 1 - \frac{1}{1 + K_1 L(1 + K_3 + K_3 K_4 E)}. \end{aligned} \tag{3.4}$$

It is obvious that in most cases $F_{L,c} \neq F_{L,r}$. Nevertheless, $F_{L,c} = F_{L,r}$ holds in some scenarios. $F_{L,c} = F_{L,r} = 0$ holds for $K_1 = 0$ or $L = 0$. For L or K_1 going to infinity, the limit of the degree of occupancy F_L equals one in both models.

The conditions $K_1 = 0$ and $L = 0$ are trivial and indicate that F_L equals zero if there is no ligand binding or no ligand. It is also trivial that the limit of F_L is one for $L \to \infty$ or $K_1 \to \infty$, as in these cases ligand binding is infinitely strong.

For $K_3 = 0$, i.e. if there is no receptor phosphorylation, it holds that $F_{L,c} = F_{L,r}$, independently of the (finite) values of the other equilibrium constants. This results from

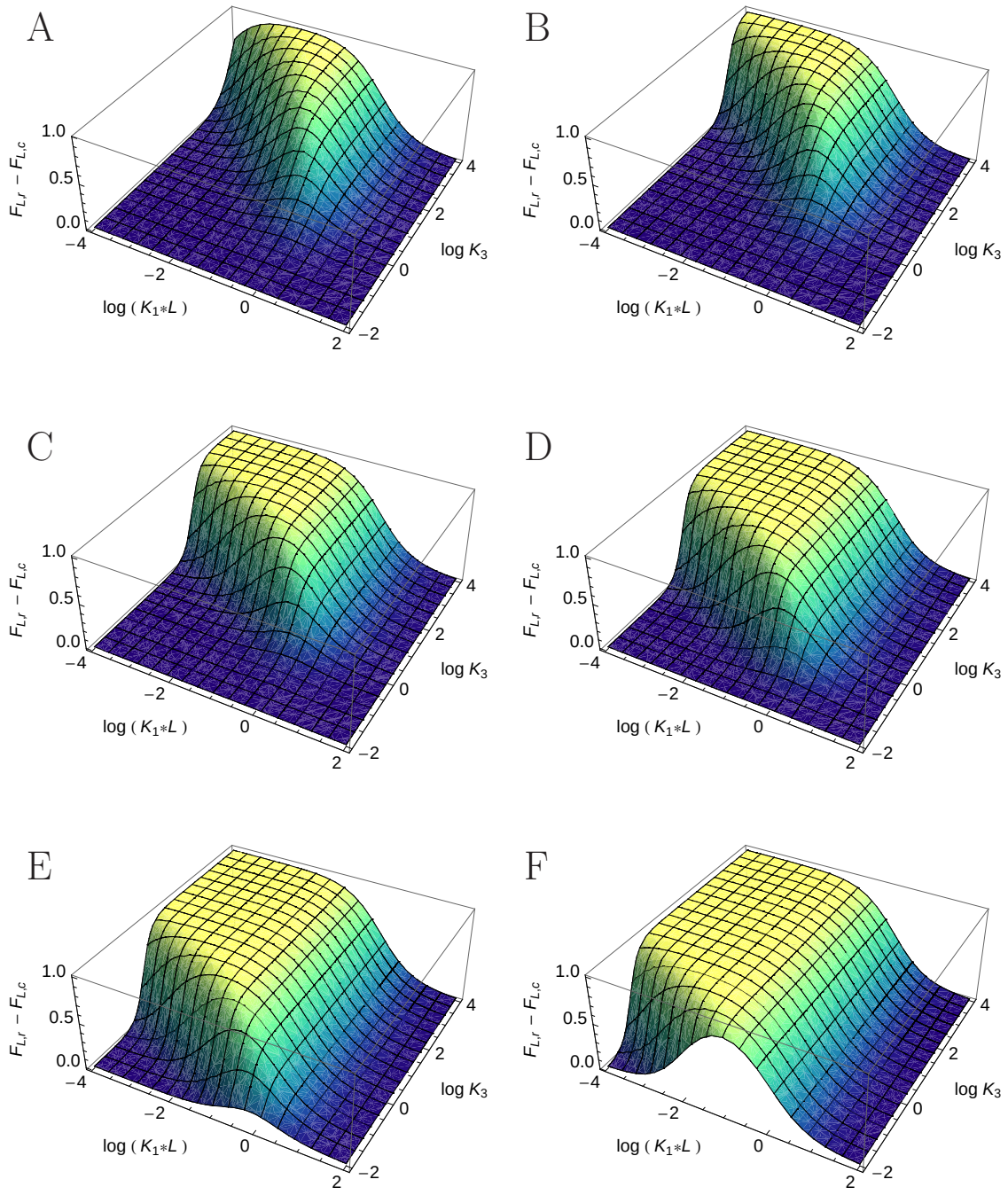

Figure 3.1: Stationary approximation error of intuitive model reduction
The stationary approximation error $F_{L,r} - F_{L,c}$ is shown as a function of $K_1 \cdot L$ and K_3. Note that L and K_1 occur in $F_{L,r} - F_{L,c}$ only as their product (Equations 3.3 and 3.4). The same holds for K_4 and E. Several scenarios characterized by different assumptions for $E \cdot K_4$ are analyzed. **A)** Approximation error for $E \cdot K_4 = 10^{-4} \cdot K_3$. **B)** Approximation error for $E \cdot K_4 = 10^{-2} \cdot K_3$. **C)** Approximation error for $E \cdot K_4 = K_3$. **D)** Approximation error for $E \cdot K_4 = 10^{2} \cdot K_3$. **E)** Approximation error for $E \cdot K_4 = 10^{4} \cdot K_3$. **F)** Approximation error for $E \cdot K_4 = 10^{6} \cdot K_3$.

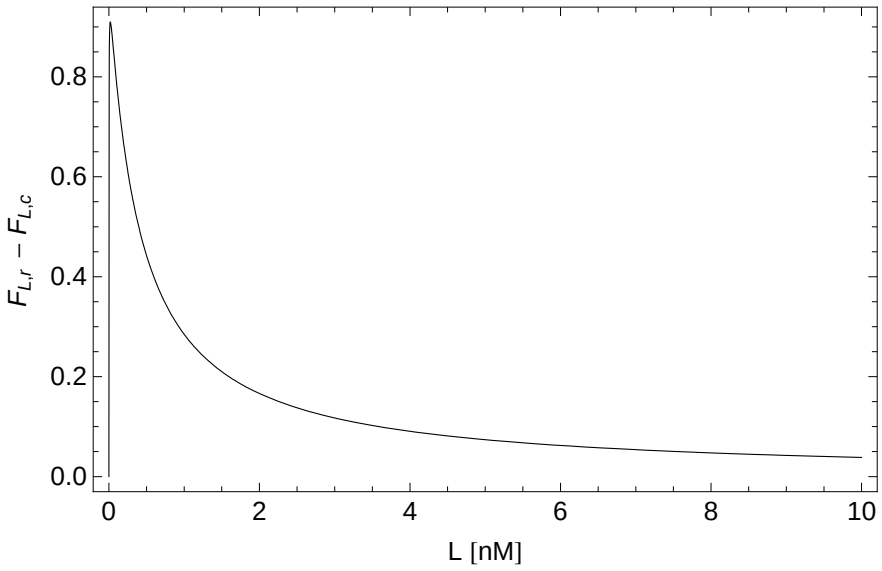

Figure 3.2: Stationary approximation error for parameter values from literature
Parameter values for insulin signaling were taken from literature (Table 4.1 on page 83) to investigate the stationary approximation error $F_{L,r} - F_{L,c}$ in a realistic situation. The relevant values are: $K_1 = 2.5\ nM^{-1}$, $K_3 = 6$, $K_4 = 0.29204\ nM^{-1}$ and $E = 250\ nM$.

the reaction chain in the reduced order model (Equation 3.2) where the dissociation of L is not possible for receptors that have a phosphorylated binding site. Therefore, receptor phosphorylation induces a potentially very large approximation error.

The difference between $F_{L,r}$ and $F_{L,c}$ is in the interval $0 \leq F_{L,r} - F_{L,c} < 1$ and increases with increasing values of K_3, K_4 and E. Furthermore, it holds that

$$\lim_{K_3 \to \infty} F_{L,r} = \lim_{K_4 \to \infty} F_{L,r} = \lim_{E \to \infty} F_{L,r} = 1 \tag{3.5}$$

independently of the (nonzero and finite) values of the other equilibrium constants and concentrations. Obviously, this does not hold for $F_{L,c}$ which is independent of K_3, K_4 and E (Equation 3.3).

A detailed analysis of the stationary approximation error $F_{L,r} - F_{L,c}$ is shown in Figure 3.1. If one looks at fixed values of K_3 and $K_4 \cdot E$, the approximation error is small for extremely low and extremely high values of $K_1 \cdot L$, which correspond to the trivial situations of almost no ligand binding and almost complete ligand binding. For intermediate values of $K_1 \cdot L$, the approximation error can be very large. For constant values of $K_1 \cdot L$, the approximation error increases with increasing values of K_3 and $K_4 \cdot E$. This results from the reaction chain in the reduced order model where receptor phosphorylation prevents ligand dissociation (Equation 3.2). For increasing values of K_3 and $K_4 \cdot E$, more receptors are trapped in states where ligand dissociation is not possible.

Peak values of the stationary approximation error $F_{L,r} - F_{L,c}$ almost equal 1 which means that almost all receptors in the reduced order model are liganded, whereas the vast majority of receptors in the conventional model is not. This is the theoretical worst case and occurs in many different scenarios (Figure 3.1).

The stationary approximation error $F_{L,r} - F_{L,c}$ was investigated for insulin signaling where L corresponds to insulin (Figure 3.2). Parameter values from literature were taken for all equilibrium constants and the concentration of E, corresponding to Shc or IRS (Table 4.1 on page 83). For physiological insulin concentrations (below 1 nM [52]), the stationary approximation error $F_{L,r} - F_{L,c}$ is quite high, whereas it is lower for supraphysiological insulin concentrations.

Altogether, the analysis of a small example system shows that the intuitive model reduction approach can have a very high approximation error. The approximation quality strongly depends on the kinetic parameters and the concentrations of the participating signaling proteins. No other satisfying method than comparing the simulation results with those of a conventional model has been proposed to decide if the intuitively reduced model is an adequate representation of the system [96]. Therefore, this approach should only be used after a careful preceding analysis requiring knowledge about the parameter values.

3.2 General considerations about model reduction

The starting point of common model reduction techniques is a conventional model

$$\dot{x} = f(x, u) \ , \qquad y = h(x, u) \tag{3.6}$$

where $x \in \mathbb{R}^n$ denotes the state vector, $u \in \mathbb{R}^m$ the system inputs and $y \in \mathbb{R}^q$ the system outputs. The objective of model reduction is to find another mathematical representation of the dynamic system which allows one to approximately describe the output variables by a reduced state vector [98]. In order to achieve this reduction, one has to transform the original dynamic system to new coordinates $z \in \mathbb{R}^n$ where $z = \phi(x)$. It is required that this generally nonlinear transformation is a diffeomorphism, which means that the function ϕ is invertible and smooth and that $\phi(0) = 0$. Additionally, we require that the transformation separates the states z_1 that have a strong impact on the output variables from the states z_2 that have only little influence on them. The transformed system is given as

$$\begin{aligned} \begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} &= \begin{bmatrix} g_1(z_1, z_2, u) \\ g_2(z_1, z_2, u) \end{bmatrix} \\ y &= h^*(z_1, z_2, u). \end{aligned} \tag{3.7}$$

If g_1 and h^* only depend on z_1 and not on z_2, the differential equations for z_2 can simply be omitted. In this case, it is guaranteed that the reduced order model has exactly the same input/output behavior as the original model (Equation 3.6). However, such an exact reduction is only possible in a restricted number of cases [93, 99]. Therefore, the states z_2 usually have to be approximated from the states z_1. If such an approximation $z_2 \approx \psi(z_1)$ is found, a reduced order model with $y^* \approx y$ is given as

$$\begin{aligned} \dot{z}_1 &= g_1(z_1, \psi(z_1), u) \\ y^* &= h^*(z_1, \psi(z_1), u). \end{aligned} \tag{3.8}$$

In the next section, we review matrix decomposition methods which are the basis for several model reduction techniques.

3.3 Matrix decomposition

Some of the model reduction techniques presented in the following sections base on matrix decomposition. We shortly review the eigenvalue decomposition (EVD) and the singular value decomposition (SVD) to simplify the understanding of these methods.

3.3.1 Eigenvalue decomposition (EVD)

A nonzero vector $v_i \in \mathbb{C}^n$ is an eigenvector of a matrix $A \in \mathbb{R}^{n\times n}$ if it satisfies the eigenvalue equation

$$Av_i = \lambda_i v_i \quad \Longleftrightarrow \quad (A - \lambda_i I_n)v_i = 0 \tag{3.9}$$

where $\lambda_i \in \mathbb{C}$ is the eigenvalue corresponding to v_i and $I_n \in \mathbb{R}^{n\times n}$ is the identity matrix [100]. Nontrivial solutions for the vectors v_i ($v_i \neq 0$) exist only if $(A - \lambda I_n)$ is singular. Therefore, the eigenvalues λ_i can be computed as solutions of the characteristic equation

$$det(A - \lambda I_n) = 0. \tag{3.10}$$

The corresponding eigenvectors can be computed by Equation 3.9.

If there are n linearly independent eigenvectors v_i, the matrix $T = [v_1 \dots v_n]$ can be used for a similarity transformation

$$\Lambda = T^{-1}AT \tag{3.11}$$

where $\Lambda \in \mathbb{R}^{n\times n}$ is a diagonal matrix with the diagonal elements λ_i [100]. This is widely used for the diagonalization of linear systems [98] (see Section 3.6).

3.3.2 Singular value decomposition (SVD)

The singular value decomposition (SVD) of a rectangular matrix $X \in \mathbb{R}^{n\times m}$, $n \leq m$ is given as

$$X = U\Sigma V^T = \sum_{i=1}^{n} \sigma_i u_i v_i^T \tag{3.12}$$

where $U = [u_1 \dots u_n] \in \mathbb{R}^{n\times n}$ and $V = [v_1 \dots v_m] \in \mathbb{R}^{m\times m}$ are orthogonal (or unitary in the case of complex matrices), i.e. $UU^T = I_n$ and $VV^T = I_m$ [98, 100].

The matrix $\Sigma \in \mathbb{R}^{n\times m}$ has the diagonal elements $\Sigma_{ii} = \sigma_i$, $i = 1, \dots, n$ and zeroes elsewhere. These diagonal elements σ_i ($\sigma_1 \geq \sigma_2 \geq \dots \geq 0$) are called singular values of X. The columns u_i and v_i of the matrices U and V are called left and right singular vectors of X, respectively.

The singular vectors u_i and v_i are the eigenvectors of XX^T and X^TX, respectively, and the singular values of X are the square roots of the corresponding eigenvalues of XX^T [98].

$$\sigma_i = \sqrt{\lambda_i(XX^T)} \tag{3.13}$$

The matrix X can be approximated by a matrix of rank $k < n$, optimally in the 2-induced norm, as

$$X = U\Sigma V^T \approx U_k \Sigma_k V_k^T = \sum_{i=1}^{k} \sigma_i u_i v_i^T \tag{3.14}$$

where $U_k \in \mathbb{R}^{n\times k}$ and $V_k \in \mathbb{R}^{m\times k}$ contain the first $k < n$ columns of U and V, respectively, and $\Sigma_k \in \mathbb{R}^{k\times k}$ is a diagonal matrix with σ_i, $i = 1, \ldots, k$ as the diagonal elements [98]. The error of this approximation is very small if the singular values σ_j for some $j > k$ are much smaller than the other ones.

3.4 Methods for linear systems

Several methods exist for the reduction of linear systems.

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned} \tag{3.15}$$

A detailed introduction is given in the textbook of A. C. Antoulas [98], a more condensed review is given in [101]. As biological signal transduction systems are nonlinear, we only provide a very short overview about these linear techniques.

SVD-based methods: Two SVD-based model reduction techniques, namely balanced truncation and Hankel norm approximation, base on the observation that there is a set of invariants belonging to every time-invariant stable linear system. These invariants are called Hankel singular values σ_i and determine the complexity of the reduced order system. Hankel singular values are a measure for the contribution of the states in a balanced representation to the energy transfer from the system inputs to the outputs. Their values are given as the square roots of the eigenvalues of $\mathcal{P}\mathcal{Q}$, where $\mathcal{P}$ and $\mathcal{Q}$ are the controllability and observability gramians, respectively.

$$\mathcal{P} = \int_0^\infty e^{At} B B^T e^{A^T t} dt \ , \qquad \mathcal{Q} = \int_0^\infty e^{A^T t} C^T C e^{At} dt \tag{3.16}$$

In balanced truncation, a state space transformation $x = Tz$ is used to get equal and diagonal gramians $\hat{\mathcal{P}} = \hat{\mathcal{Q}} = diag(\sigma_1 \ldots \sigma_n)$. The transformed system is a balanced representation of the original system (Equation 3.15) and is given as

$$\begin{aligned} \begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} &= \begin{bmatrix} \hat{A}_{1,1} & \hat{A}_{1,2} \\ \hat{A}_{2,1} & \hat{A}_{2,2} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} \hat{B}_1 \\ \hat{B}_2 \end{bmatrix} u \\ y &= \hat{C}_1 z_1 + \hat{C}_2 z_2 + Du. \end{aligned} \tag{3.17}$$

Model reduction is possible by neglecting the parts of the model that belong to the smallest Hankel singular values. This means that all states z_2 are neglected that both are difficult to control and difficult to observe. The states z_2 can be set to zero resulting in a reduced order model

$$\begin{aligned} \dot{z}_1 &= \hat{A}_{1,1} z_1 + \hat{B}_1 u \\ y &= \hat{C}_1 z_1 + Du. \end{aligned} \tag{3.18}$$

These SVD-based methods preserve stability and guarantee global error bounds. However, they are characterized by a high computational effort [98]. Balanced model reduction was modified for nonlinear systems leading to the method of empirical gramians which is characterized by an enormous computational complexity [98, 101].

Krylov-based methods: In contrast to the SVD-based methods, Krylov-based methods (or moment matching-based methods) do not rely on the computation of singular values. They are iterative in nature and base on moment matching of the impulse response of the system. Their aim is to find a reduced order model whose transfer function gives a series expansion around a specific frequency such that the first coefficients have the same values as those of the original system. Two widely used Krylov methods are the Lanczos and Arnoldi processes which can be applied to very large systems. Both methods have the drawbacks that stability of the reduced order model is not necessarily preserved and that there is no global error bound [98, 101]. An extension of these methods (implicit restart) however, preserves stability.

Another class of methods tries to combine the advantages of the SVD-based methods and the Krylov methods and is referred to as SVD-Krylov-based approximation methods [101].

3.5 Proper orthogonal decomposition (POD)

POD is an application of the SVD to the approximation of general dynamic systems [98]. The starting point of the model reduction procedure is a trajectory of the original system (Equation 3.6) subjected to an input signal $u(t)$. A matrix $X \in \mathbb{R}^{n \times N}$ ($N \gg n$) is constructed whose columns are samples of the state x at given points in time t_i which are referred to as snapshots $x(t_i)$.

$$X = [x(t_1) \;\; x(t_2) \;\; \ldots \;\; x(t_N)] \tag{3.19}$$

To approximate the system, the SVD of the snapshot matrix X (Equation 3.12) is computed. A reduced order system with $U_k^T U_k = I_k$, $U_k \in \mathbb{R}^{n \times k}$ and $z_1 \in \mathbb{R}^k$, $k < n$ is given as

$$\begin{aligned} \dot{z_1} &= U_k^T f(U_k z_1, u) \\ y &= g(U_k z_1, u) \end{aligned} \tag{3.20}$$

where the approximation z_1 evolves on a k-dimensional subspace spanned by the leading k left singular vectors of X [101].

The order k of the reduced system is chosen such that $\sigma_k \gg \sigma_{k+1}$. The (approximative) transformation between the lower-dimensional state z_1 and the original state x is given by

$$\begin{aligned} z_1 &\approx U_k^T x \\ x &\approx U_k z_1. \end{aligned} \tag{3.21}$$

Note that the SVD of X and the resulting reduced order model depend on the chosen input function $u(t)$ and the points in time at which the snapshots $x(t_i)$ are taken.

The reduced order model (Equation 3.20) is not always able to approximate other trajectories than the one taken for the generation of the snapshot matrix X. To improve the ability of the reduced order model to approximate the system behavior for different input functions, s matrices of time-snapshots X_i corresponding to different input functions can be combined to a larger matrix X.

$$X = [X_1 \dots X_s] \tag{3.22}$$

However, this improved approximation quality often comes at the price of a higher order of the reduced model.

A basic feature of the POD is that a model of the system (with all parameter values being known) has to be simulated before the reduced order model can be generated because the singular values and the matrix U_k depend on the parameters of the system. Therefore, it is not possible to use the reduced order model (Equation 3.20) for parameter identification. As it will be shown below, this is possible when applying other model reduction techniques. Another disadvantage of the POD is that outputs corresponding to concentrations which *per se* have positive values can become negative.

Altogether, POD is highly valuable for many applications. However, it is usually not suited for the modeling of biochemical signaling systems, since in many cases parameter estimation is necessary and negative concentrations are physically impossible.

3.6 Model reduction based on time hierarchy

The processes defining the dynamics of a system can usually be clustered by their velocity into three classes [22], which can be exploited for model reduction.

The *slow* class comprises the processes that can be neglected because they have almost no impact on the system dynamics in the time frame of interest. For instance, if one is interested in the fast dynamics of the metabolism or cellular signaling which have time constants of minutes or seconds, the dynamics of gene expression having time constants of hours can be neglected. In the following discussions, we assume that all state variables that show very slow dynamics are already eliminated from the model by setting them to constant values.

The *fast* class comprises the processes that are so fast that their dynamics are not visible in the time scale of interest because their transient behavior declines instantaneously. This is used for model reduction by applying the quasi-steady-state approximation (Section 3.6.1) and the rapid equilibrium assumption (Section 3.6.2). The necessary stability conditions for this rapid decline are given by Tikhonov's theorem [22]. If they are not satisfied, there can also be very fast oscillations or other complex behavior which may get lost if one assumes a rapid decline of the transient behavior of the fast processes.

The *central* class comprises the processes moving on the time scale of interest which should be included in the model.

Simulation of a system is very difficult if there are processes on different time scales. During the simulation of such stiff systems, numerical instability can occur if the step size is not chosen to be sufficiently small. Removing the extremely fast dynamics does not only lower

the order of the model, but also reduces its stiffness and simplifies numerical simulations.

We exemplify model reduction based on time scales for linear systems, before we introduce the quasi-steady-state approximation and the rapid equilibrium assumption which both are highly suited for the reduction of chemical reaction systems.

A linear system (Equation 3.15) can be diagonalized by a similarity transformation $x = Tz$ if the dynamic matrix A has n linearly independent eigenvectors [98] (see Section 3.3.1).

$$\begin{aligned} \begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} &= \begin{bmatrix} \hat{A}_1 & 0 \\ 0 & \hat{A}_2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} \hat{B}_1 \\ \hat{B}_2 \end{bmatrix} u \\ y &= \hat{C}_1 z_1 + \hat{C}_2 z_2 + Du \end{aligned} \tag{3.23}$$

The diagonal matrices $\hat{A}_1$ and $\hat{A}_2$ have the eigenvalues of A as diagonal elements and it holds that $\hat{A} = T^{-1}AT$, $\hat{B} = T^{-1}B$, $\hat{C} = CT$ and $z(0) = T^{-1}x(0)$.

Let us assume that the transformation is chosen such that the states z_2 show much faster dynamics than the states z_1. If one is only interested in the dynamics of the central class of processes, the system can be approximated by setting $\dot{z}_2 = 0$. This corresponds to the assumption of a rapid decline of the (stable) dynamics of z_2 to a value given by $z_2 = -\hat{A}_2^{-1}\hat{B}_2 u$ and results in a reduced order system

$$\begin{aligned} \dot{z}_1 &= \hat{A}_1 z_1 + \hat{B}_1 u \\ y &= \hat{C}_1 z_1 - \hat{C}_2 \hat{A}_2^{-1} \hat{B}_2 u + Du. \end{aligned} \tag{3.24}$$

3.6.1 The quasi-steady-state approximation

A general nonlinear system is given as

$$\begin{aligned} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} &= \begin{bmatrix} f_1(x_1, x_2, u) \\ f_2(x_1, x_2, u) \end{bmatrix} \\ y &= h(x_1, x_2, u). \end{aligned} \tag{3.25}$$

Assume that changes in the values of the state variables x_2 are dominated by processes of the fast class whose transient behavior rapidly declines. Model reduction can be performed by assuming that the state variables x_2 are quasi-stationary. This is done via a singular perturbation, i.e. by setting $\dot{x}_2 = 0$. Setting $\dot{x}_2 = 0$ corresponds to the assumption of a rapid decline of the dynamics of x_2 and does not necessarily mean that the value of x_2 is constant. If the equation $f_2(x_1, x_2, u) = 0$ can be solved for $x_2 = \psi(x_1, u)$, the reduced order model is given as

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, \psi(x_1, u), u) \\ y &= h(x_1, \psi(x_1, u), u). \end{aligned} \tag{3.26}$$

In reaction networks, the quasi-steady-state approximation can be applied on components with low concentrations that are subjected to large fluxes [21]. This typically occurs if the production of a species is performed by processes that are significantly slower than

the consumption processes. In this case, the consumption of this species approximately equals its production and the concentration of this species is low. The quasi-steady-state approximation is often used to simplify rate laws of complex mechanisms [21], e.g. in enzyme kinetics (Section 3.6.3).

In a simple example, A is irreversibly converted to B which is converted to C. The first order constants of the reactions are k_1 and k_2.

$$A \overset{k_1}{\rightarrow} B \overset{k_2}{\rightarrow} C \tag{3.27}$$

The ODEs for the concentrations of the three compounds are given as

$$\begin{aligned} \frac{d}{dt}A &= -k_1 \cdot A \\ \frac{d}{dt}B &= k_1 \cdot A - k_2 \cdot B \\ \frac{d}{dt}C &= k_2 \cdot B. \end{aligned} \tag{3.28}$$

If $k_1 \ll k_2$ almost all molecules of B that are synthesized are immediately degraded.

$$k_1 \cdot A \approx k_2 \cdot B \tag{3.29}$$

The resulting approximation for the concentration of B shows that the concentration of B is much smaller than the concentration of A.

$$B \approx \frac{k_1}{k_2} \cdot A \tag{3.30}$$

Equations 3.28 and 3.29 imply $\frac{d}{dt}B \approx 0$. However, the concentration of B is only constant if the concentration of A is constant. In all other cases, the concentration of B very rapidly equals a value defined by Equation 3.30. Using the quasi-steady-state approximation for B (Equation 3.30), one can simplify Equation 3.28 to a reduced order system

$$\begin{aligned} \frac{d}{dt}A &= -k_1 \cdot A \\ \frac{d}{dt}C &= k_1 \cdot A \end{aligned} \tag{3.31}$$

where the dynamics of C only depends on k_1, the rate constant of the slow process.

Altogether, the quasi-steady-state approximation allows for an approximative model reduction by eliminating species with low concentrations that show a strong turnover.

3.6.2 The rapid equilibrium assumption

If the parameters k_i and k_{-i} of a reaction

$$\begin{aligned} &... \leftrightharpoons S \overset{r_i}{\leftrightharpoons} P \leftrightharpoons ... \\ &r_i = k_i \cdot S - k_{-i} \cdot P \end{aligned} \tag{3.32}$$

are much larger than the parameters of the remaining reactions, the quotient P/S will, after a short period of time, approximately equal the equilibrium constant [22].

$$\frac{P}{S} \approx \frac{k_i}{k_{-i}} = K_i \tag{3.33}$$

Due to the high values of k_i and k_{-i}, this does not mean that the flux r_i is small. Equation 3.33 can be used to replace one ODE of the system. The resulting reduced order model provides a good approximation of the dynamics after a short initial time span (or for all times if the initial conditions already fulfill Equation 3.33) [22]. Rapid equilibrium can be assumed for all fast reactions of a network, which may lead to a significant reduction of the system order.

In many cases, the rapid equilibrium assumption and the quasi-steady-state approximation lead to similar results. An example also showing a major difference between the two approximation methods is given in Section 3.6.3. For applying the quasi-steady-state approximation one needs to know the values of all kinetic parameters, whereas for the rapid equilibrium assumption it is sufficient to know the equilibrium constants of the fast reactions. If only qualitative features about the reaction velocities are known whereas the values of the kinetic parameters are not, model reduction can nevertheless be performed if certain reactions are fast compared to the others (rapid equilibrium) or if there are strong fluxes through components with low concentrations (quasi-steady-state). In this case, the reduced order model can be used for parameter estimation.

Both methods are highly suited for the reduction of cellular signaling systems. However, they have the drawbacks that a potentially very large conventional model is necessary and that each application of the rapid equilibrium assumption or the quasi-steady-state approximation only leads to an order reduction of one which necessitates many assumptions.

3.6.3 Enzyme kinetics

A simplified formal description of enzyme activity that bases on mass action kinetics was introduced by L. Michaelis and M. Menten [21, 22]. They assumed that a substrate S associates with an enzyme E and forms an enzyme-substrate complex ES. The enzyme converts S to the product P and releases P.

$$S + E \leftrightharpoons ES \qquad k_1 \quad k_{-1} \tag{3.34a}$$

$$ES \rightarrow E + P \qquad k_2 \tag{3.34b}$$

Assuming that ES is quasi-stationary and introducing the constant total enzyme concentration $E_{tot} = E + ES$ as well as the Michaelis constant $K_M = \frac{k_{-1}+k_2}{k_1}$ results in a reaction rate which is a saturable function of the substrate concentration S and only depends on the constants K_M and $r_{max} = k_2 \cdot E_{tot}$.

$$r = r_{max} \cdot \frac{S}{K_M + S} \tag{3.35}$$

The maximal rate of product formation is r_{max}, and K_M is the substrate concentration at which the reaction rate r equals $0.5 \cdot r_{max}$. A reaction following Michaelis-Menten kinetics can be formulated according to the generalized law of mass action (see Section 1.2.3)

$$S \to P \qquad \frac{r_{max}}{K_M + S} \tag{3.36}$$

which leads to a reaction rate defined by Equation 3.35. There exist several extensions of Michaelis-Menten kinetics that can also be formulated according to the generalized law of mass action [22].

Almost the same result can be obtained if one applies the rapid equilibrium assumption on the reaction describing substrate binding (Equation 3.34a), which for $K_M = \frac{k_{-1}}{k_1}$ also leads to Equation 3.35. Since it usually holds that $k_{-1} \gg k_2$, the difference in the resulting kinetics for both assumptions is small. Additionally, *in vitro* studies usually determine K_M directly so that it is not necessary to specify the underlying assumption.

3.7 Domain-oriented model reduction

Domain-oriented model reduction is a specialized technique for cellular signal transduction systems. State variables in domain-oriented models are macroscopic quantities, such as levels of occupancy or degrees of phosphorylation. They represent the state of domains (functional components of proteins) which are the fundamental elements of signal transduction [97]. Domain-oriented model reduction is also referred to as exact model reduction because the dynamics of macroscopic quantities are exactly preserved. In the last few years, two domain-oriented techniques have been developed [93, 99].

Borisov et al. proposed the first macroscopically exact approach to reduce the enormous amount of possible model equations [99]. The starting point is a rule-based model definition from which the reduced order model is directly generated. The method exploits the independence of distinct sites of a scaffold protein. It provides a strongly reduced model if combinatorial complexity results from complex formation at one large scaffold binding several proteins. Note that the method does not lead to a strongly reduced model if these proteins are again scaffolds. Therefore, the method of Borisov et al. [99] is very valuable for describing complex formation at one large scaffold protein, however, it cannot be applied to many real signaling systems.

Conzelmann et al. [93] extended and generalized this approach such that it can be applied to a large class of systems. However, for the application of this method it is necessary to evaluate the rule-based model definition and to use a conventional ODE model as the starting point. The conventional ODE model (that can be very large) is subjected to a hierarchically structured linear state space transformation which results in the domain-oriented model if the unobservable states are omitted. The formal procedure presented by Conzelmann et al. [93] has recently been extended and improved [102, 103]. The modified approach facilitates the generation of the exactly reduced model equations and does not require the former generation of a complete combinatorial model any more.

Borisov et al. automated a modification of the domain-oriented approach not requiring a state space transformation [104]. This advantage however comes at the price of incorrectly reduced models in some cases.

We exemplify domain-oriented model reduction according to Conzelmann et al. [93] on a small example system with four species and two rules.

$$\begin{aligned} R[0,X] &\leftrightharpoons R[1,X] \qquad k_1 \qquad k_{-1} \\ R[X,0] &\leftrightharpoons R[X,1] \qquad k_2 \qquad k_{-2} \end{aligned} \tag{3.37}$$

These two rules correspond to four reactions.

$$\begin{aligned} R[0,0] &\leftrightharpoons R[1,0] \qquad k_1 \qquad k_{-1} \\ R[0,1] &\leftrightharpoons R[1,1] \qquad k_1 \qquad k_{-1} \\ R[0,0] &\leftrightharpoons R[0,1] \qquad k_2 \qquad k_{-2} \\ R[1,0] &\leftrightharpoons R[1,1] \qquad k_2 \qquad k_{-2} \end{aligned} \tag{3.38}$$

The conventional model of this system is given as

$$\begin{aligned} r_0 &= k_1 \cdot R[0,0] - k_{-1} \cdot R[1,0] \\ r_1 &= k_1 \cdot R[0,1] - k_{-1} \cdot R[1,1] \\ r_2 &= k_2 \cdot R[0,0] - k_{-2} \cdot R[0,1] \\ r_3 &= k_2 \cdot R[1,0] - k_{-2} \cdot R[1,1] \end{aligned} \tag{3.39}$$

$$\begin{aligned} \frac{d}{dt}R[0,0] &= -r_0 - r_2 \\ \frac{d}{dt}R[0,1] &= -r_1 + r_2 \\ \frac{d}{dt}R[1,0] &= r_0 - r_3 \\ \frac{d}{dt}R[1,1] &= r_1 + r_3. \end{aligned} \tag{3.40}$$

As the total concentration of R is constant, one ODE could be replaced by a conservation relation for R. According to Conzelmann et al. [93], the hierarchical and reversible transformation is given as

$$\begin{aligned} z_{XX} &= R[0,0] + R[1,0] + R[0,1] + R[1,1] \\ z_{1X} &= R[1,0] + R[1,1] \\ z_{X1} &= R[0,1] + R[1,1] \\ z_{11} &= R[1,1] \end{aligned} \tag{3.41}$$

where z_{XX} is the overall constant concentration of R, z_{1X} and z_{X1} are the degrees of occupancy of the first and the second site, and z_{11} is a mixed degree of occupancy where both sites are occupied.

By differentiating the transformation equations, we get the transformed system

$$\begin{aligned}
\frac{d}{dt}z_{XX} &= 0 \\
\frac{d}{dt}z_{1X} &= r_0 + r_1 = k_1 \cdot (\underbrace{R[0,0] + R[0,1]}_{z_{XX}-z_{1X}}) - k_{-1} \cdot (\underbrace{R[1,0] + R[1,1]}_{z_{1X}}) \\
\frac{d}{dt}z_{X1} &= r_2 + r_3 = k_2 \cdot (\underbrace{R[0,0] + R[1,0]}_{z_{XX}-z_{X1}}) - k_{-2} \cdot (\underbrace{R[0,1] + R[1,1]}_{z_{X1}}) \\
\frac{d}{dt}z_{11} &= r_1 + r_3 = k_1 \cdot \underbrace{R[0,1]}_{z_{X1}-z_{11}} + k_2 \cdot \underbrace{R[1,0]}_{z_{1X}-z_{11}} - (k_{-1} + k_{-2}) \cdot \underbrace{R[1,1]}_{z_{11}}.
\end{aligned} \tag{3.42}$$

If one is only interested in the degrees of occupancy z_{1X} and z_{X1}, the ODE for z_{11} can be omitted because z_{11} has no influence on z_{1X} and z_{X1}. The ODE for z_{XX} is also not necessary as the total concentration of R is constant. Therefore, the domain-oriented model consists of only two ODEs.

As in the approach of Borisov et al. [99], the performance of the extended approach [93,103] strongly depends on the interactions in the network. Enormous reduction is possible in some cases whereas in other cases no reduction is possible [105]. In many cases, the resulting reduced order models still consist of too many ODEs for efficient simulation or analysis.

3.8 Conclusions

Model reduction aims at replacing a mathematical model by a reduced order model with approximately the same input/output behavior.

The intuitive model reduction approach is frequently used in systems biology. It allows for a strong reduction of the system order. However, as shown for an example system, it may result in a model providing a very low approximation quality.

Very powerful model reduction techniques are available for linear systems, however, biological systems are usually nonlinear. A frequently used technique for the reduction of nonlinear systems is proper orthogonal decomposition (POD). Using POD for the reduction of biological systems has several drawbacks which make the method unsuitable for this application. In particular, POD requires a conventional model and the knowledge of all parameter values. In addition, negative concentrations, which are physically impossible, cannot be excluded.

The quasi-steady-state approximation and the rapid equilibrium assumption are highly suited for the reduction of biochemical systems if there is at least qualitative knowledge about parameter values and if the necessary stability conditions are fulfilled. A drawback of these methods is that a conventional model of the system is necessary. In addition, each application of one of these assumptions only leads to an order reduction of one which makes many assumptions necessary to get sufficiently small models.

In contrast to the approximative techniques, domain-oriented model reduction allows for an exact description of macroscopic quantities. However, this method may require a high computational effort and the resulting models can still be very large.

Altogether, model reduction in biology is an emerging field and many problems are not yet solved. In particular, there are many cases in which reduced order models of cellular signaling systems showing a high approximation quality are difficult to obtain.

In the next chapter, we introduce layer-based modeling, an approximative but accurate technique providing a solution of the model reduction problem for a large class of biological signal transduction systems.

4 Layer-based modeling

Layer-based modeling is a reduced order modeling technique with a high approximation quality. It is well suited for the mathematical modeling of signaling systems with inherent combinatorial complexity. The number of state variables in the simulation equations is highly reduced and the resulting dynamic models show a pronounced modularity. Layer-based modeling allows for the modeling of systems not accessible previously.

Section 4.1 gives a brief overview of layer-based modeling. The detailed description of layer-based modeling starts with Section 4.2 providing some definitions and a formal introduction to general concepts. In particular, it provides a precise notation of the used variables and introduces the concepts of processes and interactions. These concepts are applied in the construction of interaction graphs, which define the modularity of layer-based models. Section 4.3 describes the modeling of layers and the definition of their connections. A small example system is discussed in detail and serves to illustrate the main concepts of layer-based modeling. Advanced strategies which often allow for a further reduction of the model size are discussed in Section 4.4.

The mathematical background of layer-based modeling and a general discussion of the approximation quality are given in Section 4.5. Layer-based modeling of two larger systems is described in Section 4.6. Section 4.7 describes the most common scenarios where a strong reduction of the model size is possible.

4.1 Introduction to the layer-based approach

Layer-based modeling allows for a macroscopic and highly reduced description of signaling systems with inherent combinatorial complexity. A graph visualizing all processes and interactions of the considered system serves to define the modules of the reduced order model which are called layers. Layers can be modeled separately from each other once their connections are defined. This strongly reduces combinatorial complexity because a significantly decreased number of processes has to be considered in each single layer. Processes belonging to the same layer in most cases describe the dynamics of one or two signaling proteins. The binding of effectors to phosphorylated binding sites of these proteins is most often not explicitly considered as these dynamics are described in other layers. The different layers are connected in a highly standardized way. Two layers either are not connected or exchange signals that correspond to macroscopic variables describing phosphorylation and occupancy of binding sites.

In conventional modeling, binding or modification events are often represented by a huge number of reactions, since the involved proteins can exist in a high number of feasible

configurations (see Section 1.3). Due to the modular structure of layer-based models and the use of macroscopic state variables, a highly reduced number of reactions defines the reaction network.

The number of necessary ODEs in layer-based models is dramatically reduced as well. As an example, modeling of the insulin signaling system is possible with only 214 ODEs, instead of the $1.5 \cdot 10^8$ ODEs resulting from conventional modeling (Section 1.3). Only 51 ODEs are sufficient if additional conditions concerning the equivalence of binding sites are fulfilled.

An advantageous feature of layer-based models is that their state variables have a direct physiological meaning. In addition, layer-based models can be directly obtained, without the preceding generation of potentially very large conventional models. However, there also exists a mathematical formalism to derive the reduced order model equations from a corresponding conventional model.

Layer-based models are characterized by a high approximation quality. For physiological parameter values, they provide a highly accurate description of the corresponding signaling networks. Though layer-based modeling is an approximative technique, macroscopic quantities are described stationarily or even dynamically exact in special cases.

4.2 Important concepts

4.2.1 Definitions

Molecule definitions: A molecule definition represents the class of possible modifications of a certain protein (or any other molecule). It consists of the molecule name that is optionally followed by the successive definition of all sites (if there are any). A site definition consists of a comma-separated sequence of all possible modifications of this site (e.g. unphosphorylated and phosphorylated), enclosed by curly brackets.

As an example for a molecule definition, `R{0,L}{0,P}` defines the receptor molecule R with a binding site for the ligand L and a phosphorylation site. The possible modifications (or configurations) of R are '0' and 'L' at the first site as well as '0' and 'P' at the second site. The modification 'L' of the first site indicates the binding of L, the modification 'P' of the second site indicates phosphorylation. The absence of bound L or a phosphate group is indicated by the modification '0' at the corresponding site. Note that we allow for the consideration of a bound ligand (e.g. L) as a site configuration of another molecule (e.g. R). There can also be molecules that exist in only one configuration. Their conversion performed by reactions results in new molecules.

Species and concentrations: Species correspond to a specific configuration of a molecule and can be considered as instances of the corresponding molecule definitions. Notations for species consist of the molecule name followed by the comma-separated sequence of site configurations within one pair of squared brackets. As an example, the species $R[L, P]$ is a specified configuration of the molecule R. The ligand L is bound to the first site, the second site is phosphorylated. The species $R[0, 0]$ defines the unphosphorylated receptor without

ligand. If a molecule A can occur in just one state (i.e. the molecule definition is `A`), this single species is also denoted as A. For the sake of simplicity, we use the same notation for species and their concentrations.

Macroscopic variables: In domain-oriented model reduction [93,99,103] (see Section 3.7) and layer-based modeling, variables can correspond to sums of species of a corresponding conventional model. We call such variables macroscopic variables. Note that species can be macroscopic variables (or macroscopic species).

Microscopic species: All species that are not macroscopic variables are microscopic species. Microscopic species correspond to a distinct modification of a molecule and do not represent a pool of chemical molecules or complexes with common properties. Microscopic species are species that could also occur in a conventional model.

Macrostates and patterns: Macrostates describe macroscopic variables, e.g. degrees of phosphorylation or occupancy, which correspond to sums of species. Their notation is analogous to that for species, with the sole difference being that the modification 'X' at a specific site indicates that all distinct modifications at this site are included. Therefore, a site modification 'X' of a molecule indicates the sum of all possible modifications of this site. As an example, the macrostate $ERB[L, X, X]$ defines the sum of all species of the molecule ERB that have a bound ligand at the first site, regardless of the state of the other sites.

Macrostates in rules are interpreted as patterns. Each species of the class defined by the pattern occurs in a separate reaction when the rule is evaluated. Therefore, patterns do not represent sums of species (as macrostates do) but each of the corresponding species individually (see Section 1.5). Like species, macrostates can also be considered as instances of the corresponding molecule definitions.

Note that the term macrostate is used with a slightly different meaning in other contributions [93, 99], where it represents what we call a macroscopic variable (or macroscopic species). We reserve the term macrostate for sums of species that are defined by at least one site modification 'X'.

Complexes: Complexes result from the association of species. A complex is represented by a list of comma-separated species within curly brackets where the binding partner is indicated at each occupied binding site. Indices have to be used if it is required to achieve uniqueness.

As an example, consider the binding of a molecule R with n sites and a molecule S with k sites. The complex of R and S is $\{R[m_1, .., S, ..., m_n], S[m_1, ..., R, ..., m_k]\}$, where m_i denotes the modification at the site position i. If we also consider another molecule T that has q sites, the complex of R, S and T, where S and T bind to R, is denoted as $\{R[m_1, .., S, ..., T, ..., m_n], S[m_1, ..., R, ..., m_k], T[m_1, ..., R, ..., m_q]\}$. This definition of complex notation is very general but cumbersome in many cases.

We introduce a simplified notation that will be used from now on. In many cases, this notation is less cumbersome and more intuitive than the general one. Consider the case that the molecules S and T both have only one site which is a binding site for R. The complex of R, S and T then is $R[m_1, .., S, ..., T, ..., m_n]$ and can be treated as a species of the molecule R. If S and T have more sites, the configurations of S and T in the complex with R can be indicated by introducing additional site modifications of the corresponding sites on R. Another possibility is to define the class of possible complexes by a new molecule definition which consists of a new molecule name and the sequence of site definitions of the corresponding molecule definitions for the monomers. We exemplify this for the general case where R has a binding site for S and n other sites, while S has a binding site for R and k other sites. A species representing a complex of R and S is $RS[m_1^R, ...m_n^R, m_1^S, ...m_k^S]$, where the superscripts indicate the origin of the sites. Dimerization can be treated analogously.

Due to the modular structure of layer-based models, this simplified notation is usually more convenient than the general one, even if the signaling network is highly branched and contains many interacting scaffold proteins.

Summary: Species and macrostates are instances of molecule definitions which in turn define classes of species. Species can be microscopic species that correspond to a distinct modification of a chemical molecule or complex. Species can also be macroscopic variables that correspond to sums of microscopic species. Macrostates are macroscopic variables that correspond to sums of species and whose hallmark is that they must contain at least one site modification 'X'. Patterns occurring in rules have the same notation as macrostates and define classes of species. Each species of such a class occurs in a separate reaction when the rule is evaluated. Following the simple notation of complexes, they are treated as species of one of the participating molecules or as species of newly defined molecules.

4.2.2 Processes and interactions

In combinatorial reaction networks, the same molecular event often occurs in many different reactions. The high number of reactions results from the high amount of different species that participate in the reactions. All reactions that describe the same molecular event define the corresponding process. For example, all reactions that change the level of occupancy or the degree of phosphorylation of a specific site belong to the same process. Note that the inverse molecular event also belongs to the same process, which means for example that the process of effector binding also contains all reactions describing effector dissociation.

The reactions of a process may either be parameterized by the same or by different kinetic constants depending on the presence of process interactions. Interactions between processes result in different parameter values for different reactions of the same process. This means that the reactions of a process that is influenced by other processes are not parameterized by the same kinetic constants. If the reaction parameters of a process depend on a modification that is performed by another process, these two processes interact with each other. If the reaction parameters of two processes do not depend on the state of the corresponding other

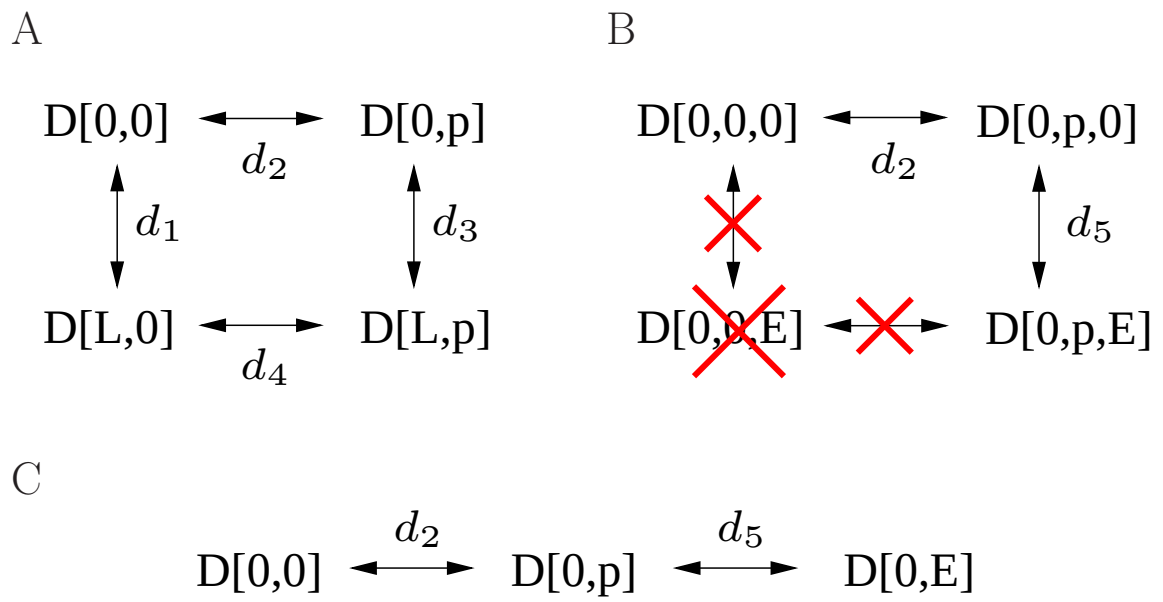

Figure 4.1: Types of interactions between processes
Reactions are indicated by arrows. The reaction rate for each reaction is according to the generalized law of mass action (Section 1.2.3). It is denoted as d_i and parameterized by k_i and k_{-i}. **A)** Visualization of a graded interaction or no interaction between the processes of L binding and D phosphorylation. A graded interaction is present if at least one of the conditions $k_1 = k_3$, $k_{-1} = k_{-3}$, $k_2 = k_4$ and $k_{-2} = k_{-4}$ is violated. If all conditions are fulfilled, the processes do not interact. **B)** Visualization of an all-or-none interaction between the processes of binding site phosphorylation and E binding. E can only bind to phosphorylated sites, while the dephosphorylation of D is only possible in the absence of a bound effector. **C)** The reaction cycle degenerates to a reaction chain, as the species $D[0, 0, E]$ does not exist. In this case, the notation can be simplified by describing binding site phosphorylation and E binding as modifications of the same site.

process, these two processes do not interact.

Identifying the processes of the system as well as their interactions is the first step in layer-based modeling and finally defines the modularity of the model (see Section 4.2.3). There exist three structurally different types of interactions between processes: *graded interaction*, *all-or-none interaction* or *no interaction*.

Graded interaction: Processes that undergo graded interactions can influence each other in arbitrary ways. A typical case is the binding of a ligand that influences the kinetic parameters for receptor phosphorylation (Figure 4.1 A). Graded interactions can be unidirectional or mutual. In the example shown in Figure 4.1 A, receptor phosphorylation may, but does not have to influence the kinetic parameters for ligand binding. If each process is defined by two reactions, these four reactions can be arranged as a reaction cycle. If the processes are defined by more reactions, there are also more cycles.

All-or-none interaction: All-or-none interactions are a limiting case of graded interactions, where both processes can only occur under certain preconditions, provided by the other process. A typical case is the interaction of the processes of binding site phosphorylation and effector binding (Figure 4.1 B). The effector can only bind if the binding site is phosphorylated. Additionally, the dephosphorylation of the binding site is only possible if the site is not occupied. A hallmark of all-or-none interactions is the degeneration of reaction cycles to reaction chains (Figure 4.1 C).

Phosphorylation and effector binding are the most common examples of processes that undergo an all-or-none interaction. To simplify the discussion, and without loss of generality, we will always assume that the processes that undergo all-or-none interactions are phosphorylations and effector bindings.

No interaction: The third type of interaction, also a limiting case of graded interactions, is that the processes do not interact. Kinetic parameters of each process are not influenced by the other process (Figure 4.1 A).

4.2.3 The interaction graph and layers

The interaction graph is a systematic visualization of the considered processes and their interactions. After identifying the processes and their interactions, building the interaction graph is the second step in layer-based modeling. In the interaction graph, boxes representing the processes are connected by lines indicating their interactions. All-or-none interactions are represented by green lines, graded interactions are represented by red lines. Figure 4.2 shows the interaction graphs of two example systems discussed in Sections 4.3.6 and 4.6.1.

The third step in layer-based modeling consists in determining the modularity of the model from the interaction graph. All processes that are directly or indirectly connected by graded interactions (represented by red lines) form a module that is called layer. The set of these processes can be directly obtained from the interaction graph. Note that layers may only contain a single process, and that there may be many layers.

The connectivity of the layers is defined by all-or-none interactions between the processes of different layers. After the definition of the connections, the layers are modeled independently from each other, as if all processes belonging to other layers do not exist.

In contrast to other modularization techniques [106–109], where the modularity is deduced from the entire model, layers can be modeled independently from each other once their connections are defined. This may give the impression that the modeler defines the modularity, however, the modularity is uniquely defined by the interaction graph. Once the modeler has defined which processes interact in which way, there is only one way of influencing the modularity, which is to merge two or more layers into a larger one.

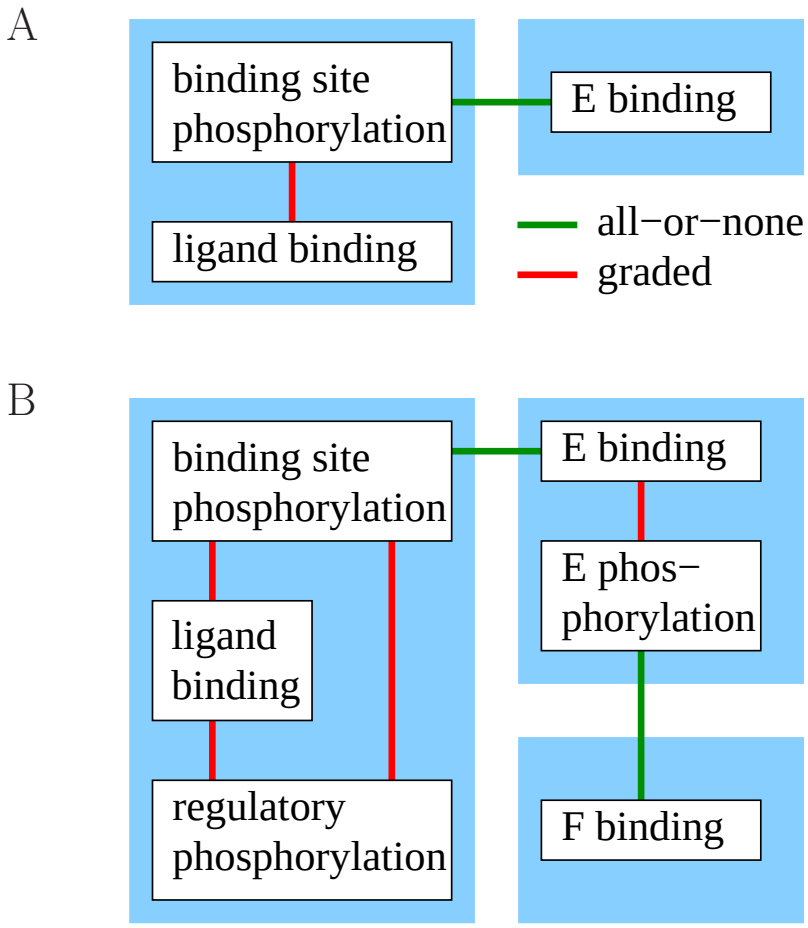

Figure 4.2: Interaction graphs of two example systems
All processes (white boxes) that are coupled by graded interactions (red lines) are described in the same layer (blue boxes). Processes of different layers are only connected by all-or-none interactions (green lines) or do not interact. **A)** Interaction graph of a small example system which is discussed in Section 4.3.6. Processes considered are ligand binding, binding site phosphorylation and effector binding. **B)** Interaction graph of an extended example system (discussed in Section 4.6.1) illustrating the successive arrangements of layers.

4.2.4 Mass flow and signal flow

In this subsection, we discuss the characteristics of mass flows and signal flows, and their occurrence in layer-based modeling. This helps in the understanding of basic principles of layer-based modeling, as the flows within layers and between layers are qualitatively different.

Mass flow: Each reaction defines a transition between species whose quantity is given by the reaction rate. We refer to these transitions as mass flows. Mass flows occur within layers and do not cross layer boundaries.

In a technical block diagram, a reaction is represented by a converter connected to storages whose potentials represent the concentrations of the reactants and products. The potentials

are transmitted to the converter where the reaction rate is determined. The converter transmits signals representing fluxes to the storages where the concentrations are balanced. Considering mass flows as transitions between species implicitly includes the signals transferring potentials and fluxes between converters and storages. Such signals are not considered as signal flows in the following discussions.

Signal flow: Signal flows define information transfers. In contrast to mass flow, signal flow is not associated to reactions and does not change the content of storages, i.e. the concentrations of species. In layer-based modeling, signal flows occur between layers and can be represented by macroscopic variables, usually sums of species with phosphorylated binding sites x_i and corresponding sums of species with occupied binding sites $x_i b$. These signals are exchanged between layers containing processes that interact via all-or-none interactions. This will be discussed in detail below.

Note that signal flow in a biological context is often interpreted as a sequence of molecular events transmitting a signal through a signaling system. Unless otherwise noted however, we will always refer to signal flow as an information transfer to which a value is associated.

Flows and layers: Layers are either only connected by signal flows or are not connected at all. No mass flows cross layer boundaries as no reactions transport substance from one layer to the other. Within layers, mass flows are defined by reaction equations and the corresponding rates as in conventional modeling.

The connections between layers show retroactivity [108] because the signal flow is bidirectional. However, changes in a layer (e.g. the introduction of an additional regulatory site) that do not affect the notation of x_i and $x_i b$ do usually not necessitate changes in other layers. Therefore, the layers can be modeled separately from each other as long as the notations of the signal flows between the layers are defined.

4.3 Modeling of layers

In the fourth step of layer-based modeling, the signal flows between the layers are defined and the layers are modeled separately from each other. Modeling within layers shows remarkable similarities to conventional modeling, however, there are differences that mainly result from the presence of all-or-none interactions between processes of different layers.

4.3.1 Phosphorylation of binding sites

All-or-none interactions between binding site phosphorylations and the corresponding effector bindings account for much of the reduction potential of layer-based modeling. An all-or-none interaction implies that a phosphorylated binding site remains phosphorylated while the effector is bound, and that the effector binds only to phosphorylated binding sites. The phosphorylation of a binding site is often described in a different layer than effector binding to this site. However, if there are graded interactions that connect the binding site

phosphorylation and effector binding processes indirectly, these two processes have to be described in the same layer. It is obvious that the phosphorylation of binding sites to which effector binding is described in another layer has to be treated differently than the phosphorylation of regulatory sites or the phosphorylation of binding sites to which the effectors bind in the same layer.

To indicate the phosphorylation of a binding site to which effector binding is described in another layer, the site notation (uppercase) 'P' is used. The phosphorylation of other sites can be indicated by other notations, e.g. (lowercase) 'p'. Species with a site modification (uppercase) 'P' comprise microscopic species with a phosphorylated but unoccupied binding site as well as microscopic species with a phosphorylated and occupied binding site.

Effector binding to phosphorylated binding sites in other layers does not directly change the concentration of species with a site modification 'P', because 'P' represents both unoccupied and occupied binding sites. The signal flows between the layers are used to guarantee that only unoccupied binding sites are dephosphorylated (see below).

4.3.2 Reactions

The processes of each layer are described by reactions as if there were no other layers. This also holds for the phosphorylation of binding sites to which effector binding is described in another layer. The sole difference is that the phosphorylation of binding sites to which the effectors bind in other layers has to be indicated by a site modification (uppercase) 'P'.

Effectors can only bind to a subset of phosphorylated binding sites, namely those that are not occupied. Effector binding to such a binding site in another layer is performed by introducing a new species that represents the sum of all microscopic species with phosphorylated but unoccupied binding sites. This new species acts as a binding partner for the effector and is defined by an algebraic equation. It is defined as the difference of the sum of species with phosphorylated binding sites x_i and the corresponding sum of species with occupied binding sites $x_i b$. These sums define the signal flows between the layers and are discussed in the next subsection.

4.3.3 Layer connections

For the realization of all-or-none interactions between processes of different layers, signal flows have to be exchanged between the layers. These signal flows are typically defined by the sums of species with phosphorylated binding sites x_i and the corresponding sums of species with occupied binding sites $x_i b$. Note that the signal flows x_i and $x_i b$ often correspond to experimental readouts. These signals are used to compute correction terms assuring that only unoccupied binding sites are dephosphorylated (see below, e.g. Equation 4.1). Their differences define the concentrations of binding partners for effectors whose binding site phosphorylations are described in other layers.

4.3.4 Reaction rates and ODEs

Reaction rates are assigned to reactions as in conventional modeling. In most cases generalized mass action kinetics are used, however, the layer-based approach also allows for other kinetics. The sole exception from this analogy is the dephosphorylation of binding sites with a site modification 'P'. Without a special treatment of these dephosphorylation reactions, generalized mass action kinetics result in an overestimation of the rates for the dephosphorylation, as occupied binding sites are also dephosphorylated. Remember that the site modification 'P' represents the phosphorylation of a binding site without distinguishing if an effector is bound or not and that only unoccupied binding sites can be dephosphorylated. Therefore, a correction term c_i has to be introduced for the dephosphorylation of each binding site with a modification 'P'.

$$c_i = \frac{x_i - x_i b}{x_i} \tag{4.1}$$

It approximates the fraction of species with phosphorylated binding sites that are not occupied. The need for these correction terms results from the loss of information and the corresponding reduction of the model size that is due to the modularization of the network. The correction terms always have the same structure, however, different x_i and the corresponding $x_i b$ are taken for each binding site.

The rate for each dephosphorylation reaction of a binding site with a site modification 'P' is multiplied by the appropriate c_i which guarantees that only unoccupied binding sites are dephosphorylated.

All dephosphorylation rates for a specific site with a modification 'P' are multiplied by the same c_i. The implicit assumption behind this is that all species with this phosphorylated binding site (that all count to the same x_i) have the same fraction of unoccupied phosphorylated binding sites. This is the core of the approximation and the reduction and is formally discussed in Section 4.5.

The ODEs are generated as they would be in conventional modeling. Note that no ODE is necessary for the binding partners in binding reactions that are defined by algebraic equations as differences of x_i and $x_i b$. The sums x_i and $x_i b$ are also defined by algebraic equations.

4.3.5 Choice of initial conditions

During the assignment of initial conditions, one has to assure that $x_i \geq x_i b$ and $x_i > 0$ hold for all i. The condition $x_i > 0 \; \forall \; i$ has to be fulfilled because negative concentrations are not possible and initial conditions leading to $x_i = 0$ result in division by zero in the correction term c_i (Equation 4.1). A value very close to zero can be taken if zero is the desired initial condition for x_i. The condition $x_i \geq x_i b \; \forall \; i$ has to be fulfilled because phosphorylation is a necessary precondition for effector binding. Therefore, it is not possible that more binding sites are occupied than phosphorylated. Violating a condition $x_i \geq x_i b$ results in a negative concentration of the binding partner defined by $x_i - x_i b$.

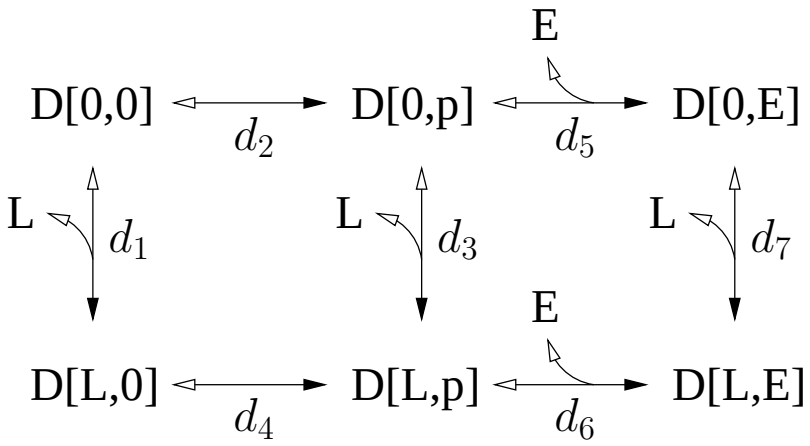

Figure 4.3: Conventional reaction network of the small example system
A receptor D has two sites and is defined as `D{0,L}{0,p,E}`. The first site is a binding site for the ligand L, the second site is a binding site for the effector E that has to be phosphorylated before E can bind. Reactions are indicated by arrows. The reaction rate for each reaction is denoted as d_i. All reactions are reversible, filled arrowheads indicate positive directions of rates.

4.3.6 A small example system

Layer-based modeling is demonstrated on a small example system where a receptor (e.g. the insulin receptor) has two sites. The first is a binding site for the ligand L (e.g. insulin), the second is a binding site for the effector E (e.g. Shc or IRS). This binding site has to be phosphorylated before E can bind and the binding of E protects the binding site from dephosphorylation. To simplify the discussion, the receptor is denoted as D in the conventional *detailed* model and is denoted as R in the layer-based *reduced* order model.

A conventional model of these processes comprises eight microscopic species and can be described by five ODEs due to conservation relations for the receptor D and the effector E. The concentration of the (extracellular) ligand L is considered as the input of the system. Figure 4.3 shows the reaction network of the conventional model.

The conventional model is given as

$$\begin{aligned}
d_1 &= k_1 \cdot L \cdot D[0,0] - k_{-1} \cdot D[L,0] \\
d_2 &= k_2 \cdot D[0,0] - k_{-2} \cdot D[0,p] \\
d_3 &= k_1 \cdot L \cdot D[0,p] - k_{-1} \cdot D[L,p] \\
d_4 &= k_3 \cdot D[L,0] - k_{-3} \cdot D[L,p] \\
d_5 &= k_4 \cdot E \cdot D[0,p] - k_{-4} \cdot D[0,E] \\
d_6 &= k_4 \cdot E \cdot D[L,p] - k_{-4} \cdot D[L,E] \\
d_7 &= k_1 \cdot L \cdot D[0,E] - k_{-1} \cdot D[L,E]
\end{aligned} \tag{4.2a}$$

$$\begin{aligned}
\frac{d}{dt}D[0,0] &= -d_1 - d_2 \\
\frac{d}{dt}D[L,0] &= d_1 - d_4 \\
\frac{d}{dt}D[0,p] &= d_2 - d_3 - d_5 \\
\frac{d}{dt}D[0,E] &= d_5 - d_7 \\
\frac{d}{dt}D[L,E] &= d_6 + d_7
\end{aligned} \tag{4.2b}$$

$$\begin{aligned}
D[L,p] &= totR - (D[0,0] + D[L,0] + D[0,p] + D[0,E] + D[L,E]) \\
E &= totE - (D[0,E] + D[L,E])
\end{aligned} \tag{4.2c}$$

where $totR$ and $totE$ are the total concentrations of the receptor and the effector, respectively.

Building the layer-based model requires to identify the processes and their interactions (see Section 4.2.2). In the small example system, there are three processes: ligand binding, binding site phosphorylation and effector binding. The processes of ligand binding and binding site phosphorylation perform a graded interaction (Figure 4.1 A). The processes of binding site phosphorylation and effector binding undergo an all-or-none interaction (Figure 4.1 B) since receptor phosphorylation is a precondition for effector binding and dephosphorylation is only possible if the effector is not bound. This knowledge about processes and interactions allows one to build the interaction graph (see Section 4.2.3) from which the modularity of the layer-based model is derived. The interaction graph of the small example system is given in Figure 4.2 A. As all processes that are coupled by graded interactions belong to the same layer, the layer-based model consists of two layers. The receptor layer describes the processes of ligand binding and binding site phosphorylation, the effector layer describes the process of effector binding. Once the signal flows between the layers (given by the sum of phosphorylated binding sites x and the sum of occupied binding sites xb) are defined, both layers can be modeled as if the corresponding other layer does not exist. Figure 4.4 shows the reaction network of the layer-based model.

The model equations of the receptor layer are given as

$$\begin{aligned}
r_0 &= k_1 \cdot L \cdot R[0,0] - k_{-1} \cdot R[L,0] \\
r_1 &= k_1 \cdot L \cdot R[0,P] - k_{-1} \cdot R[L,P] \\
r_2 &= k_2 \cdot R[0,0] - k_{-2} \cdot ((x - xb)/x) \cdot R[0,P] \\
r_3 &= k_3 \cdot R[L,0] - k_{-3} \cdot ((x - xb)/x) \cdot R[L,P]
\end{aligned} \tag{4.3a}$$

$$\begin{aligned}
\frac{d}{dt}R[0,P] &= -r_1 + r_2 \\
\frac{d}{dt}R[L,0] &= r_0 - r_3 \\
\frac{d}{dt}R[L,P] &= r_1 + r_3
\end{aligned} \tag{4.3b}$$

$$R[0,0] = totR - (R[0,P] + R[L,0] + R[L,P]). \tag{4.3c}$$

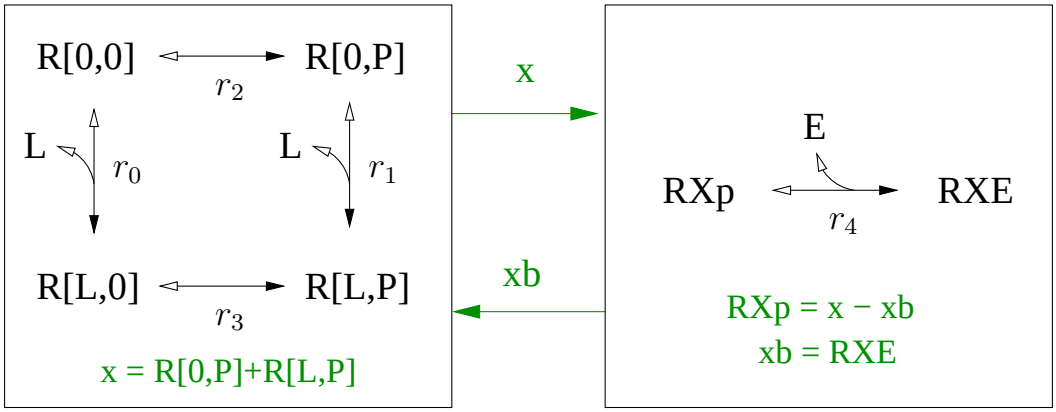

Figure 4.4: Layer-based reaction network of the small example system
The receptor layer (left) describes the processes of ligand binding and binding site phosphorylation by two reactions each. Note that the receptor is defined as `R{0,L}{0,P}`. The effector layer (right) describes the effector binding process by one reaction. The signals between the layers are the sum of species with a phosphorylated binding site x and the sum of species with an occupied binding site xb (which in this case is only the species RXE). Black arrows indicate reactions, green arrows indicate signal flows. All reactions are reversible, filled arrowheads indicate positive directions of rates.

In this layer we see two important characteristics of layer-based models that do not occur in conventional models. The first characteristic is the occurrence of sites with a modification (uppercase) 'P'. As described in Section 4.3.1, a site modification 'P' indicates the phosphorylated state of a binding site to which effector binding is described in a different layer. There is no distinction if the effector is bound to this site or not, both possibilities are included. Such a macroscopic description of phosphorylation however requires to approximate the fraction of phosphorylated binding sites that are not occupied, since only unoccupied binding sites can be dephosphorylated. This approximation is given by the correction term $(x - xb)/x$ occurring in the rates r_2 and r_3.

Describing effector binding in a different layer and using a correction term in the dephosphorylation rates necessitates signal flow between the layers given by x and xb, where x is the sum of species with phosphorylated binding sites and xb is the sum of species with occupied binding sites.

$$\begin{aligned} x &= R[X,P] = R[0,P] + R[L,P] \\ xb &= RXE \end{aligned} \tag{4.4}$$

Omitting the factor $(x - xb)/x$ in the dephosphorylation rates would result in too high dephosphorylation rates since the site modification 'P' represents both unoccupied and occupied phosphorylated binding sites. The microscopic receptor species $D[0,p]$ and $D[L,p]$ (see Figure 4.3) represent receptor molecules with phosphorylated but unoccupied binding sites. They are approximated from the species $R[0,P]$ and $R[L,P]$ that due to the site modification 'P' also comprise microscopic receptor species with bound E. Using the correction

term $(x - xb)/x$ in the rates r_2 and r_3 therefore guarantees that only unoccupied binding sites are dephosphorylated.

$$\begin{aligned} D[0,p] &\approx \frac{x - xb}{x} \cdot R[0,P] \\ D[L,p] &\approx \frac{x - xb}{x} \cdot R[L,P] \end{aligned} \tag{4.5}$$

The underlying assumption is that the processes of ligand binding and effector binding do not interact, which results in the same fractions of $R[0,P]$ and $R[L,P]$ having a bound effector E. According to the interaction graph, this (approximative) assumption is justified because E binding only depends on binding site phosphorylation (Figure 4.2 A), but not on ligand binding. However, there is an indirect interaction between these processes since ligand binding influences the phosphorylation of the binding site for the effector which may lead to an approximation error. A formal discussion about the nature and the quality of the approximation in layer-based modeling is given in Section 4.5.

Altogether, it is assumed that effector binding only depends on binding site phosphorylation but not on other modifications of the receptor. This leads to the same fractions of occupied phosphorylated binding sites for $R[0,P]$ and $R[L,P]$.

The aggregation of $R[0,P]$ and $R[L,P]$ into the signal x allows for the modularization of the system since the only information necessary for effector binding in the effector layer is the sum of phosphorylated binding sites. Transferring the signal xb, which represents the sum of species with occupied binding sites, from the effector layer to the receptor layer is necessary for building the correction terms in dephosphorylation rates guaranteeing that only unoccupied binding sites are dephosphorylated.

The model equations of the effector layer are

$$r_4 = k_4 \cdot RXp \cdot E - k_{-4} \cdot RXE \tag{4.6a}$$

$$\frac{d}{dt} RXE = r_4 \tag{4.6b}$$

$$\begin{aligned} RXp &= x - xb \\ E &= totE - RXE \end{aligned} \tag{4.6c}$$

where RXE represents the sum of all microscopic receptor species with bound E. The binding partner for an effector whose binding site phosphorylation is described in a different layer is defined by signals exchanged between the layers (see Section 4.3.2). RXp is such a binding partner representing the sum of microscopic receptor species with phosphorylated but unoccupied binding sites.

The analogy of layer-based modeling and conventional modeling can be clearly seen in the effector layer as the reaction rate, the ODE and the conservation relation for the effector (Equation 4.6) could also occur in a conventional model describing the association of two species.

Note that the letter 'X' only indicates a macrostate if it is a site configuration (as is the case for $R[X,P]$). In the species RXE and RXp, the 'X' is a part of the molecule name.

Table 4.1: Kinetic parameters and initial conditions for the small example system
Initial conditions were 40 nM for $D[0,0]$ [88] (see Appendix A.3.1), 250 nM for E and 10^{-20} nM for all other species of the conventional model. Initial conditions for the layer-based model were determined according to the transformation equations (Equation 4.25). The ligand concentration was set to 100 nM, which is a typical concentration for *in vitro* insulin stimulation. The parameters k_4 and k_{-4} originally describe p85 (a PI3K subunit) binding to its phosphorylated binding site on IRS [110], and the concentration of E (250 nM) is the value for Shc in a model of EGFR signaling in hepatocytes [40].

Parameter	Literature value	Unit	Source
k_1	0.001	$nM^{-1}s^{-1}$	[58]
k_{-1}	$4 \cdot 10^{-4}$	s^{-1}	[58]
k_2	0	s^{-1}	assumption
k_{-2}	0.00385	s^{-1}	[76]
k_3	0.0231	s^{-1}	[77]
k_{-3}	0.00385	s^{-1}	[76]
k_4	0.033	$nM^{-1}s^{-1}$	[110]
k_{-4}	0.113	s^{-1}	[110]

Therefore, RXE and RXp are not macrostates. However, both are macroscopic species since they correspond to sums of microscopic species. Note that $R[0,P]$ and $R[L,P]$ are macroscopic species as well.

If we compare the reactions of the layer-based model (Figure 4.4) and the reactions of the conventional model (Figure 4.3), we find

$$\begin{array}{lll} r_0 \mathrel{\widehat{=}} d_1 & r_1 \mathrel{\widehat{=}} d_3 + d_7 & r_2 \mathrel{\widehat{=}} d_2 \\ r_3 \mathrel{\widehat{=}} d_4 & r_4 \mathrel{\widehat{=}} d_5 + d_6. & \end{array} \tag{4.7}$$

The corresponding reaction rates and in particular the reaction rates that are merged together (d_3 and d_7 as well as d_5 and d_6) are parameterized by the same kinetic constants.

Altogether, the small example system shows that layer-based modeling provides a reduced order description of the system where macroscopic species are connected by macroscopic fluxes. The resulting model has a modular structure defined by the interaction graph. The modules (called layers) are connected by signal flows representing phosphorylation and occupancy of binding sites which are macroscopic characteristics of the system. All layers can be modeled independently from each other once their connections are defined.

4.3.6.1 Approximation quality

The transformation equations between the layer-based model and the conventional model (see Section 4.5.2) allow for the comparison of the simulation results. Simulation with parameter values from literature (Table 4.1) shows that the approximation error of the

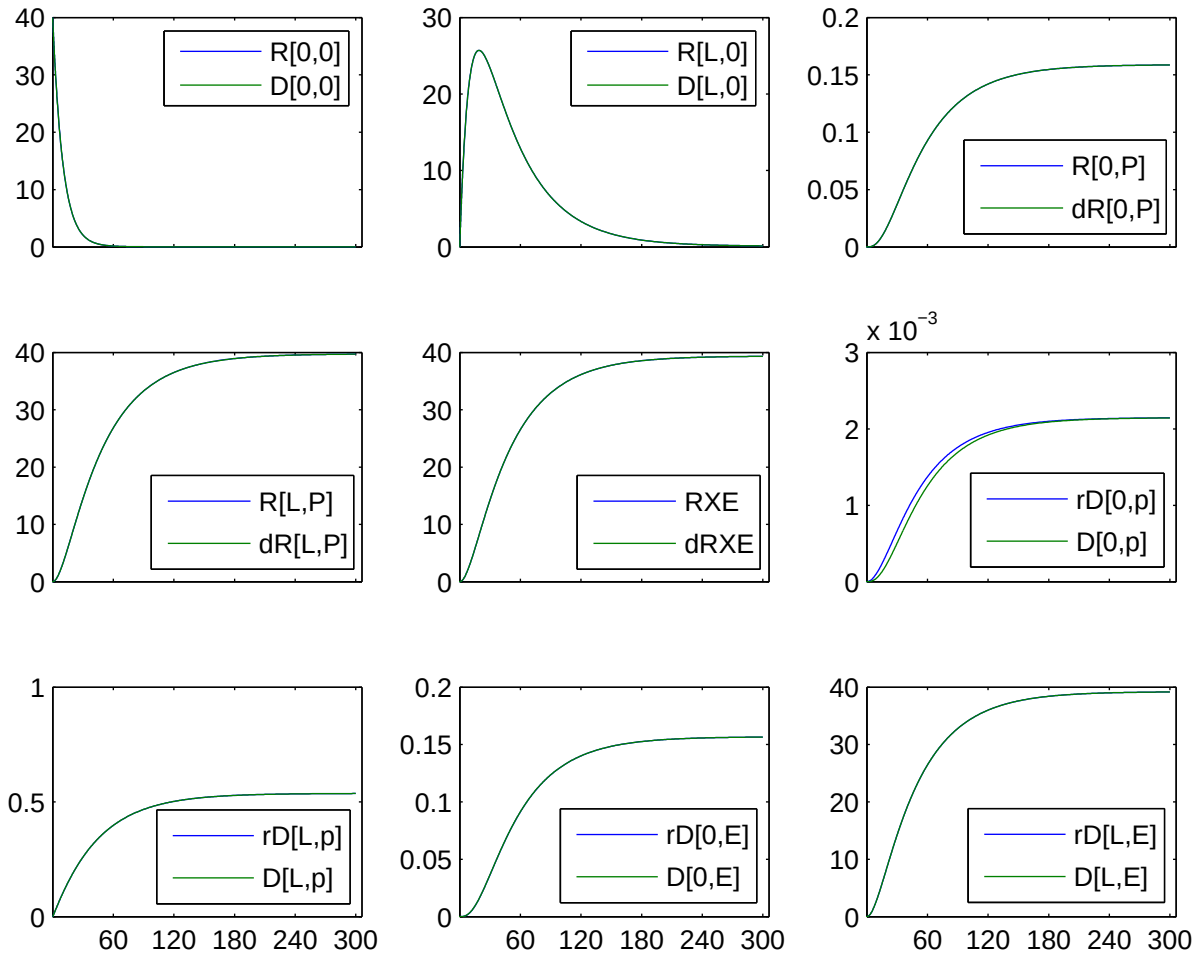

Figure 4.5: Simulation results for the small example system
Simulation results of the layer-based model and the conventional model are shown for parameter values from literature (Table 4.1). The differences vanish due to the thickness of lines, showing that the approximation error of the layer-based model is negligible. A prefix r or d of the species notation indicates that this species is computed from the states of the (reduced) layer-based or conventional (detailed) model, respectively, according to the transformation equations (see Section 4.5.2). The axis of abscissae indicates time in s, the axis of ordinates indicates concentration in nM.

layer-based model is negligible (Figure 4.5). Even the approximation of the state variables of the conventional model is possible with a high accuracy. This also holds for the variation of each parameter in a wide range (not shown).

In contrast to the insufficient approximation quality of the intuitively reduced model (Section 3.1.1, Figures 3.1 and 3.2), the stationary fraction of liganded receptors in the layer-based model $F_{L,l} = (R[L,0] + R[L,P])/totR$ equals $F_{L,c}$, the value of the conventional model (Equation 3.3), independently of the parameter values. Since receptor phosphorylation does not influence ligand binding, $F_{L,l}$ is even dynamically exact (not shown).

4.3.7 Synthesis, degradation and transport of proteins

Up to now it has been assumed that there is no protein synthesis or degradation. However, synthesis and degradation of free unmodified proteins are easy to handle as they are performed like in conventional modeling.

Degradation of complexes often requires additional effort. If a scaffold protein or a receptor with phosphorylated binding sites is to be degraded, one has to observe that there may exist species in several layers that correspond to complexes including this scaffold or receptor. One has to decide if only the specific protein is to be degraded (which leads to the liberation of all binding partners) or if all binding partners attached to the corresponding protein are also degraded. In both cases, the degradation of a protein in a certain layer also has implications on other layers which demands additional signals between the layers. Therefore, the same rates multiplied by different factors (to reflect complex composition) occur in distinct layers. Transport between different compartments can be handled as degradation of species in one compartment and synthesis of the same species with the same rates (corrected for the volume) in the other.

Altogether, degradation and transport of species can be realized in layer-based models.

4.3.8 Step by step procedure: layer-based modeling

This step by step procedure can be used for efficient and standardized layer-based modeling.

1. Identify all processes and their interactions.
2. Draw the interaction graph.
3. Deduce from the interaction graph how many layers are in the model and which processes form part of which layer. All processes that are coupled by graded interactions are part of the same layer. Processes of different layers are coupled by all-or-none interactions or do not interact.
4. Model each layer separately as if the processes of all other layers do not exist.
   a) Define rules and reactions, as in conventional modeling. The site modification (uppercase) ‘P’ is reserved for indicating the phosphorylation of binding sites to which effectors bind in other layers. A different notation (e.g. lowercase ‘p’) has to be used to indicate the phosphorylation of other sites.
   b) Define all sums of species with phosphorylated binding sites x_i and all sums of species with occupied binding sites x_ib. This only has to be done for binding sites with a site modification ‘P’. Note that x_ib has the same notation as the corresponding x_i, followed by (lowercase) ‘b’.
   c) Use algebraic equations to define the concentrations of the species that act as binding partners in effector binding to binding sites whose phosphorylation is described in other layers. These species are defined as differences of x_i and x_ib.

d) Assign a rate to each reaction, as in conventional modeling. A special case is the dephosphorylation of binding sites with a site modification (uppercase) 'P'. The rate expressions for these dephosphorylation reactions have to be multiplied by the appropriate correction term $c_i = (x_i - x_i b)/x_i$. Using the correction terms c_i guarantees that dephosphorylation is only possible for unoccupied binding sites.

e) Conservation relations can be used to replace one ODE for each conserved moiety (e.g. a protein that is not degraded or synthesized).

f) Use reactions and rates, as in conventional modeling, to derive the ODEs for all species that are not defined by an algebraic equation.

g) Assign initial conditions guaranteeing $x_i > 0$ (prevents division by zero) and $x_i \geq x_i b \; \forall \; i$ (prevents negative concentrations of binding partners).

4.4 Advanced strategies

Using the descriptions provided above, it is possible to model a large class of systems. Here we introduce advanced strategies that allow for an additional reduction of layer-based models. Examples for the application of these strategies can be found in [111, Additional file 2].

4.4.1 Equivalent binding sites

If n binding sites on a molecule are exactly equivalent and do not influence each other (and if this molecule is not an oligomer whose dissociation separates the sites) it is sufficient to model only one binding site. Note that this holds only if the initial conditions for these sites (i.e. the initial degrees of phosphorylation) are also identical. In this case, the x_i for this binding site results from multiplying the sum of species with phosphorylated binding sites by n, e.g. $x_i = n \cdot R[P]$. For linear phosphorylation or dephosphorylation kinetics, the rate law is equivalent to the case where each site is modeled separately. If the factor n is considered in the definition of x_i one needs not to consider this factor in any reaction. For nonlinear phosphorylation or dephosphorylation kinetics, e.g. Michaelis-Menten kinetics, n times the concentration of the corresponding species has to be used as the substrate concentration in the rate laws.

$$r_{tot} = \frac{r_{max} \cdot n \cdot R[0]}{n \cdot R[0] + K_M} \tag{4.8}$$

In addition, the total rate r_{tot} has to be divided by n because r_{tot} is the sum of the rates for all n sites and we only consider one of them.

$$r = \frac{r_{tot}}{n} \tag{4.9}$$

The rate r describes the kinetics of the phosphorylation of $R[0]$ to $R[P]$. According to the generalized law of mass action (Section 1.2.3), the reaction and its parameterization is given as

$$R[0] \rightarrow R[P] \qquad \frac{r_{max}}{n \cdot R[0] + K_M}. \tag{4.10}$$

If it is desired to have a model output representing the sum of all phosphorylated binding sites (for both linear and nonlinear kinetics), the factor n has to be considered and the output is $n \cdot R[P]$.

Equivalent sites that are not binding sites and that do not influence each other can be treated in an analogous way. The equivalence of binding sites can be exploited to exactly reduce the model of insulin signaling with 214 ODEs to 51 ODEs (see Section 4.6.2 and Appendix A.8).

4.4.2 One effector binding to different binding sites

Assume that m different binding sites (not necessarily belonging to the same molecule) are equivalent with respect to effector binding but may have different phosphorylation characteristics. Frequently occurring examples for this scenario are homodimerization and heterodimerization of receptors. In the example of homodimerization, phosphorylation of monomers could be parameterized by different kinetic constants than the phosphorylation of dimers, whereas the process of effector binding is not influenced by the dimerization. In these cases, a common x_i for all m binding sites can be used. This x_i equals the sum of all sums of species with phosphorylated binding sites for the effector, $x_i = x_{i_1} + ... + x_{i_m}$.

As an example, assume that a receptor $R1$ can dimerize to $R2$. If the binding of an effector E to a phosphorylated binding site on the receptor depends on binding site phosphorylation, but not on receptor dimerization, a common x_i can be used.

$$\begin{aligned} x_R &= R1[P] + R2[P,X] + R2[X,P] \\ x_R b &= E[R] \\ RXp &= x_R - x_R b \end{aligned} \tag{4.11}$$

RXp represents the sum of all phosphorylated but unoccupied binding sites on the dimers and monomers of the receptor. The binding of $E[0]$ to RXp leads to $E[R]$.

The correction term c_R for the dephosphorylation of all binding sites is $c_R = (x_R - x_R b)/x_R$. This corresponds to the assumption that the same fractions c_R of phosphorylated binding sites are occupied for all considered sites with a site modification 'P'.

This strategy also prevents negative concentrations in the effector layers that may occur if e.g. receptor dimerization is faster than effector dissociation and phosphorylated monomers rapidly 'vanish' due to dimerization. In an example (where the actual advanced strategy is not applied) $x_{R1} = R[P,X,...]$ is the sum of monomers with phosphorylated binding sites and the corresponding $x_{R1}b$ represents all effectors that are bound to receptor monomers. Let the system be in a state where a large fraction of the binding sites is occupied, i.e. $x_{R1} - x_{R1}b \approx 0$. Assume that an external stimulus induces the rapid dimerization of the receptor which rapidly lowers the value of x_{R1} because the dimers are not included in x_{R1}. Note that the dimerization has no direct influence on $x_{R1}b$. Therefore, it is possible that the concentration of the species $R1Xp = x_{R1} - x_{R1}b$ representing all receptor monomers with phosphorylated but unoccupied binding sites becomes smaller than zero. As mentioned

above, this problem can and should be avoided by using a common x_i for corresponding sites on the monomers and dimers.

4.4.3 Additional signals between layers

Additional signals between layers are allowed as long as no graded interactions are introduced by them. In a typical example, the processes of receptor activation and receptor phosphorylation are described in the receptor layer, whereas the binding and phosphorylation of effectors is described in different layers. The phosphorylation of effectors that are bound to receptors is performed by activated receptors. It is important that one specific activated receptor protein does not selectively phosphorylate effector molecules that are bound to it. This would introduce a graded interaction between the processes of receptor activation and effector phosphorylation and therefore disrupt the modular structure of the model. A different situation occurs if the pool of activated receptors acts as an enzyme and phosphorylates receptor-bound effectors. In this case, the signal representing receptor activity can be transferred to the effector layers as an additional signal.

In general, the additional coupling between processes of different layers must occur in the form of a mean-field assumption. This means that the rate of a certain event (e.g. effector phosphorylation) is proportional to the average value of some property computed from another layer (e.g. receptor activation).

4.4.4 Domain-oriented reduction of layer-based models

A special case of macroscopically exact reduction of layer-based models was introduced in Section 4.4.1 where the equivalence of sites is exploited. Macroscopically exact reduction is also possible for more general scenarios by applying the original domain-oriented model reduction approach [93] (see Section 3.7) to each layer of the layer-based model separately. It can also be applied to selected layers only. Natural candidates are layers with occurring combinatorial complexity that may for example describe large scaffold proteins with many binding sites. The procedure consists of two steps. In the first step, the complete layer-based model is generated, in the second step domain-oriented reduction is applied. Such a procedure is demonstrated in [105, Additional files 7 and 8] for the 214 ODE model of insulin signaling which results in only 56 ODEs providing a macroscopically exact reduction.

The recent extension of domain-oriented model reduction [102,103] can also be applied. It suggests a more simple procedure in which independent subsystems are modeled separately before the model equations are transformed and redundant information is omitted.

The ideal scenario for domain-oriented reduction of layer-based models is the occurrence of molecules with many sites that do not (or only scarcely) interact. In this case, domain-oriented model reduction allows for a strong reduction of the layer-based model.

4.5 Mathematical background

Layer-based modeling fits into the general procedure of model reduction (Section 3.2). Following this procedure, one has to define the state variables of the reduced order model z_1 and the state variables z_2 that have only little influence on the outputs. Algebraic equations can be used to approximate $z_2 \approx \psi(z_1)$ in the ODEs for z_1 (Equation 3.8 on page 57).

The definition of z_1 is quite simple, since the layer-based approach makes clear statements about the state variables of the reduced order model (see the step by step procedure in Section 4.3.8). In the small example system discussed in Section 4.3.6, the state vector z_1 includes $R[0,P]$, $R[L,0]$, $R[L,P]$ and RXE (Equations 4.3 and 4.6). As shown in Section 4.5.2, this definition of z_1 leads to a linear transformation $z_1 = \phi_1(\hat{x})$ (Equation 4.25). To avoid confusion with x_i (signal flows representing sums of species with phosphorylated binding sites), we deviate from the standard notation x for state variables of the conventional model (used in Chapter 3) and refer to them as $\hat{x}$.

The state variables z_2 (in the small example system: one state variable z_2) have to be chosen such that the transformation $z = \phi(\hat{x})$ is invertible. Interestingly, the layer-based approach does not make any statement about how to choose z_2 and the layer-based model can be directly generated without considering z_2. As long as the transformation is invertible, the choice of the additional states z_2 has no influence on the model equations which can be proven by some simple considerations.

Assume that a certain z_2 was chosen and that there is a model as shown in Equation 3.7. There are some algebraic equations for the approximation step $z_2 = \psi(z_1)$ and a resulting layer-based model

$$\dot{z}_1 = g_1(z_1, \psi(z_1), u) \tag{4.12}$$

(see Equation 3.8). Assume that another representation of z_2, namely $\tilde{z}_2$, is chosen. This can be realized by using a linear transformation

$$\begin{bmatrix}\tilde{z}_1\\ \tilde{z}_2\end{bmatrix} = T \cdot \begin{bmatrix}z_1\\ z_2\end{bmatrix} \quad \text{where} \quad T = \begin{bmatrix}I & 0\\ T_1 & T_2\end{bmatrix} \quad \text{and} \quad T^{-1} = \begin{bmatrix}I & 0\\ -T_2^{-1}T_1 & T_2^{-1}\end{bmatrix} \tag{4.13}$$

which transforms the states z_2 of Equation 3.7 to the new states $\tilde{z}_2$. Since I is the identity matrix, it holds that $\tilde{z}_1 = z_1$ and $z_2 = -T_2^{-1}T_1\tilde{z}_1 + T_2^{-1}\tilde{z}_2$ which leads to the ODEs of the transformed system.

$$\begin{bmatrix}\dot{\tilde{z}}_1\\ \dot{\tilde{z}}_2\end{bmatrix} = \begin{bmatrix}g_1(\tilde{z}_1, -T_2^{-1}T_1\tilde{z}_1 + T_2^{-1}\tilde{z}_2, u)\\ \tilde{g}_2(\tilde{z}_1, \tilde{z}_2, u)\end{bmatrix} \tag{4.14}$$

The algebraic equations $z_2 = \psi(z_1)$ transformed to the new coordinates are given by

$$\tilde{z}_2 = T_1\tilde{z}_1 + T_2\psi(\tilde{z}_1). \tag{4.15}$$

If one uses Equation 4.15 to replace $\tilde{z}_2$ in Equation 4.14 and considers $\tilde{z}_1 = z_1$, the resulting model is again Equation 4.12. This means that the model structure is completely determined by defining the state variables z_1 and $dim(z_2)$ algebraic equations for the approximation step.

For building the layer-based model it is not necessary to postulate the algebraic equations explicitly as they are implicitly contained in the model equations. The derivation of these equations which are the basis of the approximation is shown in the next section. They can be used for a formal derivation of the layer-based model which is described in Section 4.5.3.

4.5.1 The nature of the approximation in layer-based modeling

The approximation $z_2 = \psi(z_1)$ results from algebraic equations $\Psi(\hat{x}) = 0$ transformed to z coordinates. These equations can be derived by two considerations, described below, which both lead to the same result. $\Psi(\hat{x}) = 0$ is also used to derive the inverse transformation $\hat{x} = \phi^{-1}(z_1)$ which allows one to approximate the state variables of the conventional model from the state variables z_1 of the layer-based model. The dimension of $\Psi(\hat{x})$ is given by $dim(\Psi(\hat{x})) = dim(z_2) = dim(\hat{x}) - dim(z_1)$.

4.5.1.1 Independence of processes

Borisov et al. [99, 112] discussed algebraic constraint equations in combinatorial reaction networks that can be used to derive the equations $\Psi(\hat{x}) = 0$. Consider a scaffold protein with a large number of independent binding sites where the binding of a ligand to the scaffold has no influence on all other binding sites. In this case, the independent binding sites can be described separately in a strongly reduced model that only provides information about the levels of occupancy of the different domains [99, 112]. The information about the detailed (microscopic) complex composition can be reconstructed using the calculus of probability. This reconstruction is exemplified for a scaffold `S{0,1}{0,1}` with two independent binding sites where a site modification '1' indicates an occupied binding site. Due to the independence of binding sites, the calculus of probability suggests that the concentration of the scaffold with both domains occupied can be calculated as

$$S[1,1] = \frac{S[1,X] \cdot S[X,1]}{S[X,X]}. \tag{4.16}$$

Borisov et al. showed that if this equation is fulfilled at a time point t_0, it will hold for all times $t > t_0$ [99, 112].

We show that the findings of Borisov et al. can be used to derive the equations $\Psi(\hat{x}) = 0$ which define the approximation in layer-based modeling. Equation 4.16 can be simplified by elementary transformations to

$$S[0,0] \cdot S[1,1] - S[1,0] \cdot S[0,1] = 0. \tag{4.17}$$

Analogous considerations for the small example system, where it is assumed that the processes of L binding and E binding are completely independent, lead to

$$\frac{D[0,p]}{D[L,p]} = \frac{D[0,E]}{D[L,E]} \quad \Leftrightarrow D[0,p] \cdot D[L,E] - D[L,p] \cdot D[0,E] = 0 \tag{4.18}$$

which is the basis of the approximation in the small example system (Section 4.3.6). It is an approximation because there is an indirect interaction between these processes (Figure 4.2 A on page 75). The independence assumption given by Equation 4.18 leads to the same fractions of $R[0,P]$ and $R[L,P]$ having a bound effector E. In larger models, independence assumptions between processes of different layers that do not interact directly lead to the same correction term $(x_i - x_i b)/x_i$ being used in the dephosphorylation rates of all species with the same binding site phosphorylated. This approximation corresponds to the assumption that effector binding depends on binding site phosphorylation, but not on other modifications of the scaffold providing the binding site.

In both the example here and in Figure 4.3, the four species occurring in the equation representing the independence assumption (Equations 4.17 and 4.18) form a reaction cycle. These cycles consist of reactions of two processes of different layers which do not interact *directly*, e.g. ligand binding and effector binding. Each branched reaction network which is decomposed into layers includes such independent cycles. All corresponding pairs of processes are described in distinct layers and do not interact directly. Indirect interactions with at least one all-or-none interaction in-between are possible, but not necessary.

For each of these reaction cycles one can formulate an independence assumption like Equation 4.18. Each independent equation decreases the number of states of the reduced order model by one. All independent equations together define $\Psi(\hat{x}) = 0$ which is the basis of the approximation. The next section describes how the same equations can also be derived from the assumption of rapid equilibrium.

4.5.1.2 Diagonal reactions

Another possibility for deriving the equations $\Psi(\hat{x}) = 0$ is to assume rapid equilibrium [22] for virtual reactions connecting processes of different layers that do not interact directly. We demonstrate this for the small example system where a reaction cycle is formed by the reactions describing the bindings of L and E to the phosphorylated receptor (Figures 4.3 and 4.6). There are four equilibrium conditions for this reaction cycle.

$$\begin{aligned} \frac{D[0,E]}{D[0,p]\cdot E} &= K_4 & \qquad \frac{D[L,p]}{D[0,p]\cdot L} &= K_1 \\ \frac{D[L,E]}{D[L,p]\cdot E} &= K_4 & \qquad \frac{D[L,E]}{D[0,E]\cdot L} &= K_1 \end{aligned} \tag{4.19}$$

As the Wegscheider condition [22] is fulfilled for this reaction cycle, the equilibrium can be described by only three independent equations. Instead of using three of the four equations of Equation 4.19, one can also choose another representation including a parameter-free equation. This parameter-free equation is of special interest and represents $\Psi(\hat{x}) = 0$.

$$D[0,p]\cdot D[L,E] - D[L,p]\cdot D[0,E] = 0 \tag{4.20}$$

It corresponds to a rapid equilibrium assumption for the reaction

$$D[0,p] + D[L,E] \leftrightharpoons D[L,p] + D[0,E] \tag{4.21}$$

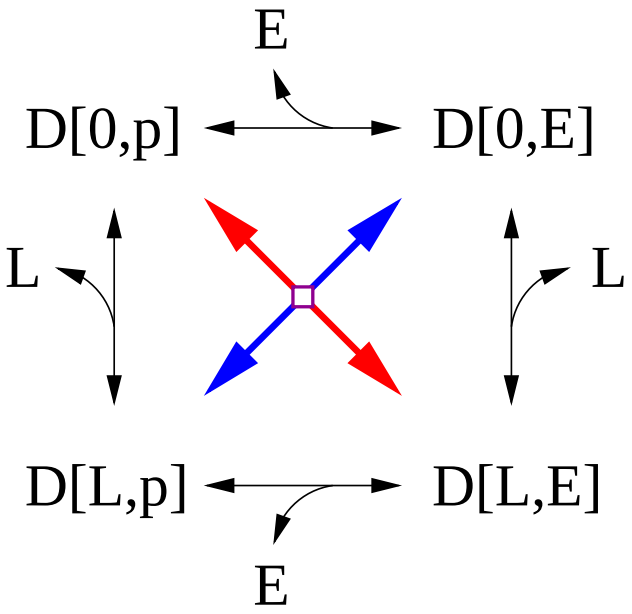

Figure 4.6: Reaction cycle and diagonal reaction
The diagonal reaction (Equation 4.21) whose equilibrium assumption is the basis for model reduction of the small example system is symbolized in the middle of the reaction cycle.

which is not part of the reaction network (Figure 4.3) and where the equilibrium constant equals one. This reaction fills the diagonals of the square representing the reaction cycle (Figure 4.6). Therefore, we refer to this kind of fast virtual reaction as a diagonal reaction.

As Equation 4.20 is only one of three independent equilibrium conditions, assuming equilibrium for the diagonal reaction is a weaker assumption than assuming equilibrium for the whole reaction cycle.

In a general system, each independent equilibrium assumption for a diagonal reaction like Equation 4.20 reduces the number of necessary ODEs by one. All independent equilibrium assumptions for diagonal reactions connecting processes of different layers that do not interact directly together define $\Psi(\hat{x}) = 0$, which is the basis of the approximation.

4.5.1.3 Non-uniqueness of diagonal reactions

The set of diagonal reactions (or independence assumptions) is not unique. In almost all larger models there are several conditions such as Equation 4.18, which lead to

$$q_1 = q_2 = ... = q_n \tag{4.22}$$

where q_i are quotients of microscopic species of the conventional model (for an example see [105, Additional file 4, p. 3]). Equation 4.22 provides $n - 1$ independent equations. It is obvious that the equations $\Psi(\hat{x}) = 0$ depend on the choice of the $n - 1$ independent equations resulting from Equation 4.22. However, the set of microscopic species connected by the diagonal reactions (or independence assumptions) is unique.

When choosing components of $\Psi(\hat{x}) = 0$, one should pay attention that the corresponding diagonal reactions belong to reaction cycles where all transitions are given by single reac-

tions. This highly simplifies the derivation of the equations for the approximation errors (Section 4.5.4) and the interpretation of the underlying independence assumptions.

4.5.2 Reversible transformation

There are two equivalent possibilities to approximate the state variables of a conventional model from the state variables of the corresponding layer-based model. The first is to set up the linear transformation equations

$$z_1 = \phi_1(\hat{x}). \tag{4.23}$$

This is done by summing up the corresponding species of the conventional model for each species of the layer-based model. When considering the equations $\Psi(\hat{x}) = 0$ and all conservation relations, Equation 4.23 can be analytically solved for the state variables of the conventional model resulting in

$$\hat{x} = \phi^{-1}(z_1). \tag{4.24}$$

Using the nonlinear equations $\Psi(\hat{x}) = 0$ makes the reverse transformation nonlinear. Note that Equations 4.23 and 4.24 define no diffeomorphism as one needs the equations $\Psi(\hat{x}) = 0$ and all conservation relations to solve Equation 4.23 for $\hat{x}$.

We demonstrate this method on the small example system. Calculating the variables of the layer-based model from variables of the conventional model follows a linear transformation

$$\begin{aligned} R[0,0] &= D[0,0] \\ R[L,0] &= D[L,0] \\ R[0,P] &= D[0,p] + D[0,E] \\ R[L,P] &= D[L,p] + D[L,E] \\ RXp &= D[0,p] + D[L,p] \\ RXE &= D[0,E] + D[L,E] \\ rE &= dE \end{aligned} \tag{4.25}$$

where rE and dE represent E in the layer-based model and the conventional model, respectively. Note that $z_1 = \phi_1(\hat{x})$ for the small example system is given by Equation 4.25 but does not contain equations for $R[0,0]$, E and RXp as these variables are not defined by ODEs (see Equations 4.3 and 4.6). Due to

$$R[0,P] + R[L,P] = RXp + RXE \tag{4.26}$$

the equations in Equation 4.25 are linearly dependent and Equation 4.25 cannot be inverted.

$\Psi(\hat{x}) = 0$ (Equation 4.20) is used to get the inverse transformation.

$$\begin{aligned} D[0,0] &= R[0,0] \\ D[L,0] &= R[L,0] \\ D[0,p] &= \frac{x - xb}{x} \cdot R[0,P] \\ D[L,p] &= \frac{x - xb}{x} \cdot R[L,P] \\ D[0,E] &= \frac{R[0,P]}{x} \cdot RXE \\ D[L,E] &= \frac{R[L,P]}{x} \cdot RXE \\ dE &= rE \end{aligned} \tag{4.27}$$

Note that $\hat{x} = \phi^{-1}(z_1)$ is given by Equation 4.27 but does not contain equations for $D[L,p]$ and E as these variables are not defined by ODEs (Equation 4.2). For the inverse transformation (Equation 4.27) the definition of the signal flows (Equation 4.4) was used. Note that the expressions for $D[0,p]$ and $D[L,p]$ are part of the rates r_2 and r_3 of the layer-based model (Equation 4.3), where they are used to approximate the corresponding microscopic species of the conventional model (Equation 4.5). The derivation of the transformation equations is also shown for an extended example system (Section 4.6.1) in [105, Additional files 1 and 4].

This method to approximate variables of the conventional model is a brute force possibility, where all equations $\Psi(\hat{x}) = 0$ (about 145 millions for insulin signaling) have to be explicitly formulated. Additionally, inversion is difficult for such highly nonlinear systems that may contain many equations.

The second possibility for the approximation of variables of the conventional model from variables of the layer-based model is more effective if only a few variables of the conventional model are to be approximated. The starting point of the approximation of a variable is an arbitrarily chosen species of the layer-based model containing this variable. The species of the layer-based model is multiplied by factors specifying the characteristics of the desired variable. For the small example system, we consider the approximation of $D[0,p]$, which is contained in $R[0,P]$. The fraction of phosphorylated binding sites that is not occupied is $(x - xb)/x$. The approximation is therefore given by $D[0,p] \approx R[0,P] \cdot (x - xb)/x$.

More detailed descriptions about the approximation of selected variables can be found in [105, Additional file 4]. This method is also demonstrated in Section 4.6.1.1 for the approximation of two variables of an extended example system. Note that this method is implicitly applied by using correction terms to guarantee that only unoccupied binding sites are dephosphorylated (see Equation 4.5).

The approximated variables can be used in layer-based models. Such variables may be of physiological importance and for example represent activated effectors. The approximation of selected variables provides additional information about the system not explicitly contained in the layer-based model. However, it demands additional signal flows between the layers necessary for the factors in the approximation.

4.5.3 Formal derivation of layer-based models

The step by step procedure for layer-based modeling (Section 4.3.8) describes how layer-based models can be directly generated. This is a comfortable possibility not requiring the preceding generation of a conventional model. Another procedure is according to formal considerations about model reduction (Section 3.2) and leads to the same layer-based model. It requires a conventional model $\dot{\hat{x}} = f(\hat{x}, u)$ and an interaction graph of the system.

1. Use the interaction graph to determine which processes form part of which layer (see Section 4.2.3) and specify the state variables of each layer.
2. Derive the transformation equations $z_1 = \phi_1(\hat{x})$ by assigning the corresponding sum of species of the conventional model to each state variable of the layer-based model.
3. The equations $\Psi(\hat{x}) = 0$ representing the approximation (see Section 4.5.1) follow from assuming that all processes of different layers not interacting directly are independent. $\Psi(\hat{x}) = 0$ is given by the set of resulting independent equations like Equation 4.18.
4. Solve $z_1 = \phi_1(\hat{x})$ and $\Psi(\hat{x}) = 0$ for $\hat{x}$ to obtain the inverse transformation $\hat{x} = \phi^{-1}(z_1)$.
5. Differentiate the transformation equations $z_1 = \phi_1(\hat{x})$ and insert $\hat{x} = \phi^{-1}(z_1)$ as well as $\dot{\hat{x}} = f(\hat{x}, u)$ to get the layer-based model $\dot{z}_1 = g_1(z_1, \psi(z_1), u)$ (Equation 4.12).

4.5.4 Approximation quality

4.5.4.1 Dynamics of the approximation error

In order to analyze the error of the approximation $\Psi(\hat{x}) = 0$, a reaction cycle with four different in-fluxes J_i is considered (see Figure 4.7). This is a general case because the effect of all reactions not considered in the cycle is represented by the in-fluxes J_i. According to Borisov et al. [99], the equation $\Psi(\hat{x}) = 0$ provides an exact approximation if the processes whose reactions form the cycle interact neither directly nor indirectly and if the initial conditions already fulfill the equation (see Section 4.5.1.1). In all other cases, $\Psi(\hat{x}) = 0$ is an approximation.

Each pair of processes of different layers that forms such reaction cycles leading to equations $\Psi_i(\hat{x}) = 0$ does not interact directly. However, there may be indirect interactions between the considered processes realized by all-or-none interactions between processes of different layers. The presence of other processes results in external in-fluxes J_i as shown in Figure 4.7. Therefore, the relations like Equation 4.18 may become erroneous.

We introduce a measure for the approximation error $g(t)$ as the deviation from Equation 4.18.

$$D[0, p] \cdot D[L, E] - D[L, p] \cdot D[0, E] = g(t) \tag{4.28}$$

The error $g(t)$ quantifies the distance of the diagonal reaction (Equation 4.21, Figure 4.6) from equilibrium. It follows a first order ODE

$$\dot{g} = -a(t) \cdot g + u(t) \ , \qquad g(0) = g_0 \tag{4.29}$$

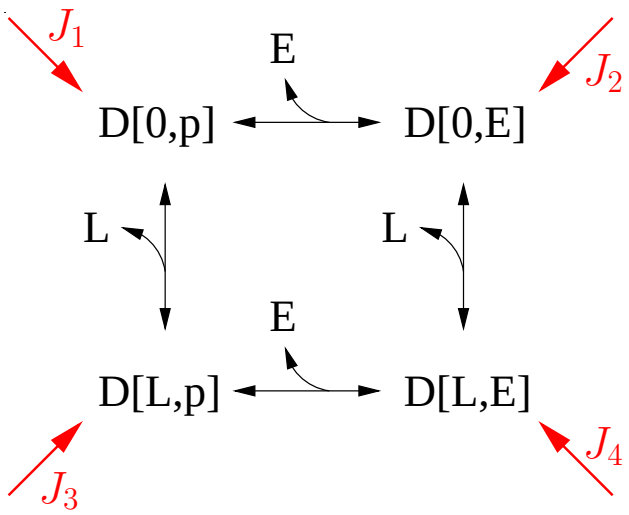

Figure 4.7: Typical reaction cycle with in-fluxes
The reaction cycle is composed of four reactions that belong to two processes of different layers that do not interact directly. As there is no direct interaction between the processes, the reactions belonging to the same processes are parameterized by the same kinetic constants. Like in the small example system, L binding is parameterized by k_1 and k_{-1}, and E binding is parameterized by k_4 and k_{-4}.

where $u(t) = J_1 \cdot D[L,E] - J_2 \cdot D[L,p] - J_3 \cdot D[0,E] + J_4 \cdot D[0,p]$ (see Figure 4.7). The coefficient $a(t)$ is given as $a(t) = k_1 \cdot L + k_{-1} + k_4 \cdot E + k_{-4}$ and it holds that $a(t) > 0 \; \forall \; t$. If there is no L or E entering the reaction cycle, the corresponding parameters in $a(t)$ (k_1 and k_4, respectively) are not multiplied by L or E, which may lead to $a(t) \equiv a$. A detailed derivation of Equation 4.29 is given in Appendix A.5.

In the small example system (Figure 4.3) it holds that $J_2 = J_4 = 0$ and therefore $u(t) = d_2 \cdot D[L,E] - d_4 \cdot D[0,E]$. Equation 4.29 gives the analytical solution

$$g(t) = e^{-A(t)} \left(g_0 + \int_0^t e^{A(\tau)} u(\tau) d\tau \right) , \qquad A(t) = \int_0^t a(\tau) d\tau. \tag{4.30}$$

The following calculations provide an upper bound for the approximation error $g(t)$. If one considers

$$\frac{d}{d\tau} A(\tau) \cdot \frac{1}{a_{min}} \geq 1 \tag{4.31}$$

where $a_{min} = \min\limits_{0 \leq \tau \leq t} (a(\tau))$ and assumes $u(t) \geq 0 \; \forall \; t$, one obtains

$$g(t) \leq e^{-A(t)} \left(g_0 + \int_0^t \frac{d}{d\tau} A(\tau) \cdot \frac{1}{a_{min}} \cdot e^{A(\tau)} u(\tau) d\tau \right) \tag{4.32}$$

where $e^{A(\tau)} \cdot \frac{d}{d\tau} A(\tau)$ is easy to integrate. For $u(t) \geq 0 \; \forall \; t$ and $g_0 \geq 0$, an upper bound for the dynamic error is given by

$$g(t) \leq g_0 \cdot e^{-a_{min} \cdot t} + \frac{u_{max}}{a_{min}} \cdot \left(1 - e^{-a_{max} \cdot t}\right) \tag{4.33}$$

where

$$\begin{aligned} u_{max} &= \max_{0\leq\tau\leq t}(u(\tau)) \\ a_{min} &= \min_{0\leq\tau\leq t}(a(\tau)) \geq k_1 \cdot \min_{0\leq\tau\leq t}(L(\tau)) + k_{-1} + k_4 \cdot \min_{0\leq\tau\leq t}(E(\tau)) + k_{-4} \\ a_{max} &= \max_{0\leq\tau\leq t}(a(\tau)) \leq k_1 \cdot \max_{0\leq\tau\leq t}(L(\tau)) + k_{-1} + k_4 \cdot \max_{0\leq\tau\leq t}(E(\tau)) + k_{-4}. \end{aligned} \tag{4.34}$$

Note that $A(t) = a \cdot t$ if $a(t) \equiv a = a_{min} = a_{max}$.

A very important result from Equation 4.33 is that the approximation error declines and is bounded because $u(t)$ is bounded from above and $a(t)$ is bounded from both below and above. Similar expressions to Equation 4.33 can be given for $u(t) \leq 0$ and $g_0 \leq 0$. Equation 4.33 shows that both the steady state error and the upper bound of the dynamic error decrease for increasing values of a, which corresponds to increasing values of the kinetic parameters k_1, k_{-1}, k_4 and k_{-4}. Both errors go to zero for one of these parameter values going to infinity.

For each diagonal reaction i, an error equation like Equation 4.29 (with g_i, a_i and u_i) can be given. In the most simple case, a_i equals the sum of the four kinetic parameters of the reaction cycle. In each case, a_i is large if the reaction parameters in the corresponding cycle are large, and u_i is a weighted sum of fluxes (that can be positive or negative) entering the reaction cycle. When unbalanced ($u_i(t) \neq 0$), these incoming fluxes can lead to virtual fluxes through the diagonal reactions that lower the approximation quality. However, there are also cases in which u_i may reduce the error g_i, in particular if u_i is positive while g_i is negative or vice versa.

4.5.4.2 Input-to-state stability of the approximation error

An important stability property is input-to-state stability which guarantees that the state $g(t)$ is bounded for any bounded input $u(t)$. Direct consequences of input-to-state stability are that the system is globally uniformly asymptotically stable for $u \equiv 0$ and that the state g finally goes to zero in the case of a vanishing input, i.e. $\lim_{t\to\infty} g = 0$ for $\lim_{t\to\infty} u = 0$ [113].

This section shortly reviews input-to-state stability and shows input-to-state stability of the approximation errors g_i in layer-based modeling. The following theorem is taken from the textbook of H. K. Khalil [113, Theorem 4.19] and gives a sufficient condition for input-to-state stability of the system

$$\dot{g} = f(t, g, u), \quad g \in \mathbb{R}^n, \quad u \in \mathbb{R}^m, \; t \geq 0 \in \mathbb{R}. \tag{4.35}$$

The necessary assumptions concerning f (piecewise continuous in t and locally Lipschitz in g and u) and u (piecewise continuous in t and bounded) are usually fulfilled for models of cellular signaling systems. In particular, they are fulfilled for Equation 4.29.

Theorem (input-to-state stability): *Assume that there exists a continuously differentiable function V(t,g) that satisfies*

$$\alpha_1(||g||) \leq V(t, g) \leq \alpha_2(||g||) \tag{4.36a}$$

$$\dot{V} = \frac{\partial V}{\partial t} + \frac{\partial V}{\partial g} \cdot \frac{\partial g}{\partial t} \leq -W(g) \qquad \forall \quad ||g|| \geq \rho(||u||) > 0 \tag{4.36b}$$

$\forall\ (t, g, u)$ *where* α_1 *and* α_2 *are class* $\mathcal{K}_\infty$ *functions,* $W(g)$ *is a continuous positive definite function on* $\mathbb{R}^n$ *and* ρ *is a class* $\mathcal{K}$ *function. Then, the system (Equation 4.35) is input-to-state stable and the solution* $g(t)$ *is ultimately bounded by* $\gamma = \alpha_1^{-1} \circ \alpha_2 \circ \rho(||u||)$.

Remarks:

- A continuous function $\alpha(||g||)$ belongs to class $\mathcal{K}$ if it is zero for $||g|| = 0$ and strictly increasing [113, Definition 4.2].
- A class $\mathcal{K}_\infty$ function $\alpha(||g||)$ is of class $\mathcal{K}$ and has the additional property $\alpha(||g||) \to \infty$ for $||g|| \to \infty$ [113, Definition 4.2].
- A solution $g(t)$ is ultimately bounded if it is bounded after a transient period has passed.
- A function $V(g)$ that is not a function of t satisfies Equation 4.36a if it is positive definite ($V(0) = 0$, $V(g) > 0\ \forall\ g \neq 0$) and radially unbounded ($V(g) \to \infty$ for $||g|| \to \infty$) [113, Lemma 4.3]. If $\alpha_1(||g||) = V(g) = \alpha_2(||g||)$, the solution $g(t)$ is ultimately bounded by $\gamma = \rho(||u||)$.

In the following paragraph, we show input-to-state stability of the approximation error $g(t)$ (Equation 4.29). A Lyapunov function candidate for the approximation error is $V(g) = \frac{1}{2}g^2$ which satisfies $\alpha_1(||g||) = V(g) = \alpha_2(||g||)$. The derivative of V along the trajectory of the system (Equation 4.29) is given as

$$\begin{aligned} \dot{V} = \frac{\partial V}{\partial g} \cdot \frac{\partial g}{\partial t} &= -a(t)g^2 + gu \\ &\leq -a_{min}|g|^2 + |g|\ |u| \end{aligned} \tag{4.37}$$

where $a_{min} = \min\limits_{0 \leq \tau \leq t}(a(\tau)) > 0$. Introducing θ with $0 < \theta < 1$ leads to

$$\begin{aligned} \dot{V} &\leq -a_{min}(1-\theta)|g|^2 - a_{min}\theta|g|^2 + |g|\ |u| \\ &\leq \underbrace{-a_{min}(1-\theta)|g|^2}_{-W(g)} \qquad \forall \quad |g| \geq \underbrace{\frac{|u|}{a_{min}\theta}}_{\rho(||u||)} \end{aligned} \tag{4.38}$$

which shows input-to-state stability of the approximation error g (Equation 4.29). Furthermore, the approximation error is ultimately bounded by $\gamma = \frac{|u_{ss}|}{a_{min}\theta}$ where u_{ss} is the stationary value of $u(t)$ and θ can be chosen close to one. Input-to-state stability of each approximation error $g_i(t)$ can be shown by analogous calculations.

Equation 4.33 provides the ultimate bound $\frac{u_{max}}{a_{min}}$ for the approximation error g in the case of $u(t) \geq 0\ \forall\ t$. The ultimate bound γ is less conservative than $\frac{u_{max}}{a_{min}}$ since $u_{ss} \leq u_{max}$. It follows from input-to-state stability [113] that the approximation error g finally goes to zero if the input u equals zero or goes to zero.

4.5.4.3 Evaluation of the approximation quality

Formal aspects: All $u_i(t)$ vanish in thermodynamic equilibrium where all fluxes vanish [22]. Note that a system is able to reach thermodynamic equilibrium if the Wegscheider condition [22] is fulfilled for the entire reaction network and if no species or fluxes are set to a constant value (or follow an externally imposed trajectory). All $u_i(t)$ also vanish if there are no graded interactions. This follows from the work of Borisov et al. [99, 112] because in this case the processes forming the reaction cycles are independent.

Therefore, layer-based modeling provides an *exactly reduced model* if there are no graded interactions (i.e. if all processes are connected by all-or-none interactions or do not interact) or if the layers containing graded interactions are not coupled to the other layers. Both scenarios lead to $u_i(t) \equiv 0 \; \forall \; (i,t)$. Note that an isolated layer containing graded interactions corresponds to a conventional model of a subsystem and causes no approximation error. In both cases, the initial conditions of the corresponding conventional model have to fulfill the equilibrium conditions for the diagonal reactions $\Psi(\hat{x}) = 0$ (e.g. Equation 4.20). If they do not, the approximation is only *stationarily exact*.

If all domains are completely independent, the initial conditions may not fulfill $\Psi(\hat{x}) = 0$, but the reduction is nevertheless exact.

The approximation is also stationarily exact if all fluxes vanish in the stationary case, i.e. if the system reaches thermodynamic equilibrium which for $t \to \infty$ leads to $u_i(t) \to 0 \; \forall \; i$. Another possibility for stationary exactness is the presence of a stationary flux distribution, characterized by $u_i(t) \equiv 0 \; \forall \; i$ even if the system does not reach equilibrium.

Altogether, the approximation is stationarily exact if $u_i(t) \to 0 \; \forall \; i$ for $t \to \infty$. This is for example the case if the system is able to reach thermodynamic equilibrium where all fluxes vanish. Layer-based models provide dynamically exact descriptions if in addition the layers containing graded interactions (if there are any) are not coupled to the other layers and if the initial conditions fulfill $\Psi(\hat{x}) = 0$, the equilibrium conditions for the diagonal reactions.

In all other cases, layer-based modeling is an *approximative* method. However, for typical scenarios the approximation error is very small. This is demonstrated for an extended example system (Section 4.6.1) where the approximation quality is high within wide ranges around parameter values from literature [105].

Note that there are cases in which some macroscopic variables are described exactly whereas others are approximations. Assume that a subsystem S fulfills the conditions for stationary or dynamic exactness. This subsystem may unidirectionally influence processes of other subsystems but is not influenced by processes of other subsystems. In this case, the macroscopic variables describing the state of the processes of S are exactly described. As an example, the process of ligand binding in the small example system (Section 4.3.6) is not influenced by any other process since the graded interaction with the process of receptor phosphorylation is unidirectional. Therefore, the corresponding degree of occupancy F_L is exactly described in the layer-based model.

Unfortunately, there is no easy method to quantify the approximation error of large layer-based models. A very laborious method (which is not applicable for very large systems)

is to build the conventional model and compare typical simulation results for macroscopic variables with those of the layer-based model. Another method which also needs to simulate the corresponding conventional model is to check all errors g_i. If all g_i decline rapidly, the approximation quality is high. However, even for large errors g_i a layer-based model may provide very good approximations of the states z_1 (Equation 4.12). This is the case if the erroneous states z_2 that are approximated have only little influence on z_1, which is closely related to the concept of observability [114]. However, since in most cases even the state variables of the conventional model $\hat{x} = \phi^{-1}(z_1)$ are approximated quite well, this shall not be discussed in detail. More qualitative checks, described below, can be performed in order to assess the approximation quality.

Qualitative checks: In the typical case, effector binding and the corresponding binding site phosphorylation are described in different layers. If the process of effector binding is relatively fast (as the binding of most effectors is) and in particular equilibrates faster than binding site phosphorylation, the approximation quality is usually high. If there are fast processes in the layer describing the binding site phosphorylation that connect all or most of the different phosphorylated species, the approximation quality is usually high as well. High parameter values for these fast processes result in high values of the exponents a_i and therefore in a high approximation quality (Equation 4.33). The process of ligand binding is a candidate for such a fast process in the small example system (Figures 4.3 and 4.4).

If the binding process of an effector is slow, it should be considered to merge the layers describing this effector binding and the corresponding binding site phosphorylation to a single one. In this case, the potentially erroneous rapid equilibrium assumptions are not introduced and do not lead to an approximation error (but do also not contribute to the model reduction).

If one is not sure about the approximation quality resulting from certain parameter combinations, one could build a conventional model of a subsystem comprising the processes belonging to the considered parameters and all processes that directly or indirectly interact with them via graded interactions. The approximation error can be analyzed quantitatively for the comparison of simulation results for macroscopic quantities of the conventional subsystem with those of a corresponding layer-based model. This can be done for different constant input signals x_i and all $x_j b$ $(i \neq j)$ entering the model from the environment set to zero. The initial conditions of the conventional model have to be set such that the effect of the external (constant) x_i entering the system is mirrored adequately. If the approximation quality is high, the analysis can be extended to time-variant input signals

$$x_i(t) = \int_0^{t_{end}} f_{x_i}(t)dt \tag{4.39}$$

guaranteeing $x_i \geq x_i b \ \forall \ i$. External fluxes f_{x_i} that reflect the changes in x_i have to be included in the conventional model. Each flux f_{x_i} adds or removes the corresponding species with phosphorylated and unoccupied binding site from the conventional model. Note that

this species is defined by $x_i - x_i b$ in the layer-based model. The velocity with which a certain x_i may change depends on the kinetic parameters for phosphorylation and dephosphorylation of the corresponding binding site.

4.6 Modeling and analysis of larger systems

Applying the step by step procedure given in Section 4.3.8, layer-based modeling of large systems is easily possible. This section discusses layer-based modeling of two systems. The extended example system discussed in Section 4.6.1 mainly serves to show that the approximation quality is also high for larger systems. It also serves to demonstrate the approximation of selected variables of a corresponding conventional model. The models of insulin signaling with 51 ODEs and 214 ODEs discussed in Section 4.6.2 replace a conventional model of insulin signaling with $1.5 \cdot 10^8$ ODEs. This demonstrates that layer-based modeling of extremely large systems usually results in models that can be used for efficient simulation or parameter estimation.

4.6.1 An extended example system

Assume that a receptor R can bind L (e.g. insulin) and perform autophosphorylation on two sites, one being a binding site for the scaffold E (e.g. IRS) and one being a regulatory phosphorylation site whose phosphorylation negatively affects autophosphorylation of the binding site. Unphosphorylated scaffold proteins that are bound to the receptor can be phosphorylated. An effector F (e.g. PI3K) can bind to the phosphorylated scaffold E. The concentration of the ligand L is considered as the input of the system.

A conventional model of this system consists of 24 ODEs. If there is no protein synthesis or degradation, three of these 24 differential equations can be replaced by conservation relations for the receptor, E and F, leading to 21 differential equations for the conventional model which are given in [105, Additional file 4].

The interaction graph of this system is given in Figure 4.2 B on page 75. A layer-based model of this system consists of three layers, the reaction network is shown in Figure 4.8. The receptor layer contains the processes of L binding and receptor phosphorylation, the E layer describes binding and phosphorylation of E. The binding of F is described in the F layer. The receptor layer consists of 8 differential equations, the E layer of four and the F layer of two. Three of these 14 differential equations can be replaced by conservation relations, leading to 11 ODEs for the layer-based model (see Appendix A.6). The signal flows between the layers are: x (sum of all receptor species phosphorylated on the binding site for E), xb (sum of all species of E bound to the receptor), $x2$ (sum of all species of E phosphorylated on the binding site for F), $x2b$ (sum of all species of F bound to E), and R_{active} (sum of catalytically active receptor species).

Simulation with parameter values from literature (Appendix A.6, Table A.1) was performed and showed a very high approximation quality of the layer-based model (Figure 4.9).

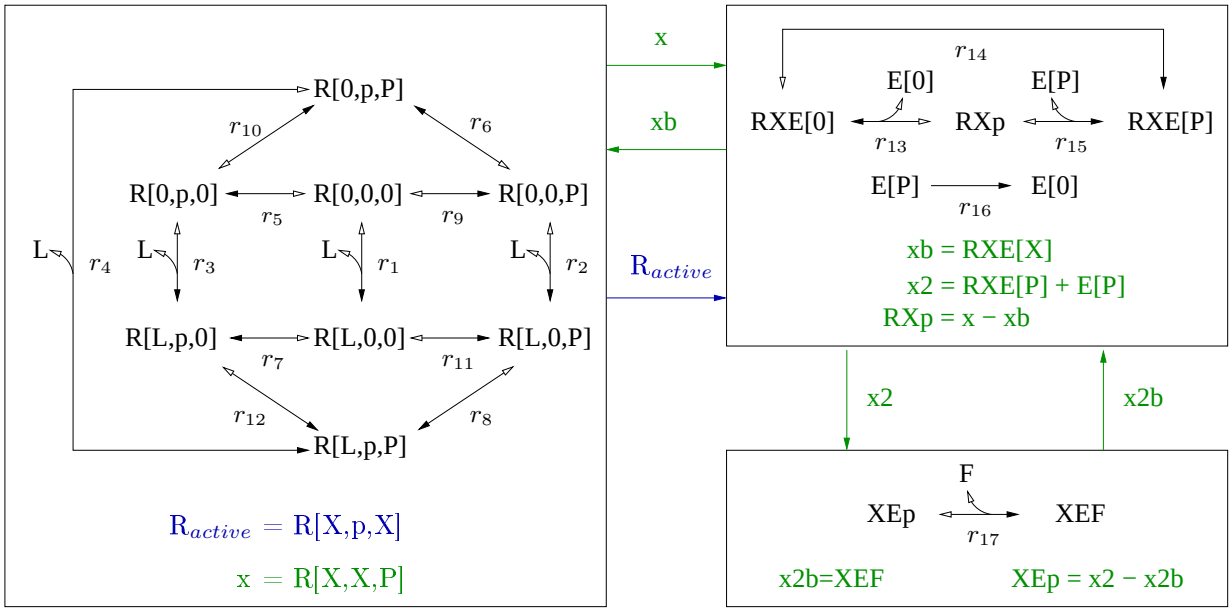

Figure 4.8: Layer-based reaction network of the extended example system
The left box shows all reactions of the receptor layer, the upper right box shows all reactions of the E layer, and the lower right box shows all reactions of the F layer. The molecule definition of the receptor is `R{0,L}{0,p}{0,P}`, the molecule definition for receptor-bound E is `RXE{0,P}` and that of E not bound to the receptor is `E{0,P}`. RXp and XEp represent sums of species with phosphorylated but unoccupied binding sites. Arrows with two heads indicate reversible reactions, whereas arrows with one head indicate irreversible reactions. Filled arrowheads indicate the positive direction of rates.

Simulation and optimization studies showed that the approximation quality is high within wide parameter ranges [105].

4.6.1.1 Approximation of variables

The approximation of selected variables of the corresponding conventional model is demonstrated for the variables $D[X, X, F]$ and $D[L, p, F]$ which both have no direct counterpart in the layer-based model. $D[X, X, F]$ corresponds to the sum of all microscopic species of F that are bound to the receptor via phosphorylated E and may be of physiological relevance. $D[L, p, F]$ is used to demonstrate the approximation of microscopic state variables of the conventional model. It represents liganded receptors that are phosphorylated on both sites. F is bound to the receptor via phosphorylated E.

As discussed in Section 4.5.2, a variable of the conventional model is approximated by multiplying a species of the layer-based model containing the variable to be approximated by factors specifying the characteristics of the desired variable. The approximation of $D[X, X, F]$ from variables of the layer-based model is as follows. $D[X, X, F]$ is contained in XEF, a state variable of the F layer (Figure 4.8) which corresponds to the sum of all microscopic species of F bound to E. The fraction of phosphorylated E that is bound to the receptor is

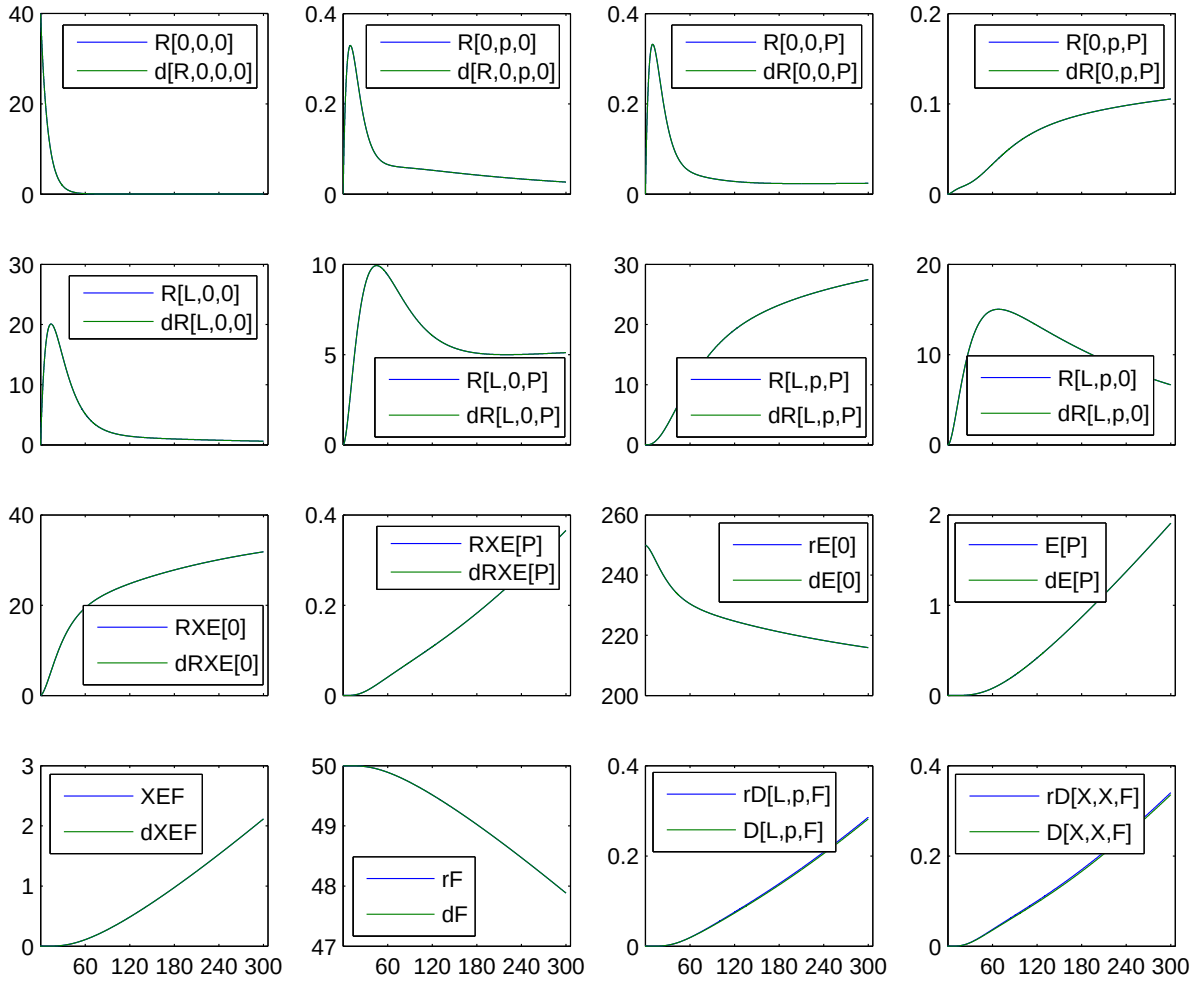

Figure 4.9: Simulation results for the extended example system

For parameter values from literature (Appendix A.6, Table A.1), the deviations of the simulation results of the layer-based model from those of the corresponding conventional model are negligible. As demonstrated for $D[L,p,F]$, the approximation of state variables of the conventional model is possible with a high accuracy. A prefix r or d of the species notation indicates that this species is computed from the states of the layer-based or conventional model, respectively, according to the transformation equations [105, Additional file 4]. The axis of abscissae indicates time in s, the axis of ordinates indicates concentration in nM.

$\frac{RXE[P]}{x2}$, which is multiplied by XEF to approximate $D[X,X,F]$.

$$D[X,X,F] = XEF \cdot \frac{RXE[P]}{x2} \tag{4.40}$$

$D[L,p,F]$ is contained in $R[L,p,P]$. The fraction of $R[L,p,P]$ that has an occupied binding site is $\frac{xb}{x}$. The fraction of receptor-bound E that is phosphorylated is $\frac{RXE[P]}{xb}$. The fraction of phosphorylated E that is bound to F is $\frac{x2b}{x2}$.

$$D[L,p,F] = R[L,p,P] \cdot \frac{xb}{x} \cdot \frac{RXE[P]}{xb} \cdot \frac{x2b}{x2} \tag{4.41}$$

Note that the approximation can be performed starting from each species of the layer-based model containing the variable to be approximated. In the example of $D[L, p, F]$, approximation is also possible starting from $RXE[P]$ or XEF. All possibilities give identical results which can be verified by considering $x2b = XEF$.

4.6.2 Layer-based modeling of insulin signaling

A conventional model of insulin signaling including all processes mentioned in the introduction would consist of $1.5 \cdot 10^8$ ODEs (see Section 1.3). In contrast to that, a layer-based model consists of only $64+128+4+11+5+2 = 214$ ODEs. An interaction graph of insulin signaling that is the basis of the layer-based model is shown in Figure 4.10.

The 214 ODEs of the layer-based model form part of the different layers as follows. $2^6 = 64$ equations in the receptor layer describe the bindings of two insulin molecules and the phosphorylation of four binding sites (two for Shc and two for IRS). The insulin concentration is considered as the input of the system. The $2^7 = 128$ equations of the IRS layer result from the presence of six binding sites, each of which can be phosphorylated and unphosphorylated. In addition, IRS can be bound to the receptor and unbound. The Shc layer is described by 4 ODEs (Shc binding to the receptor and Shc phosphorylation). All binding and modification processes of SOS and Grb2 are described in a common layer which consists of 11 ODEs (six ODEs describing the binding of complexes of Grb2 and SOS to IRS and Shc, five ODEs for free species). The PI3K layer consists of 5 ODEs (binding to four binding sites) and the SHP2 layer of 2 (binding to IRS). Free species are considered in all effector layers.

The two binding sites on the receptor for Shc and IRS are in each case assumed to be equivalent with respect to effector binding. This leads to a reduction of the model size as described in Section 4.4.2. Without exploiting this equivalence, the IRS layer would consist of $3 \cdot 2^6 = 192$ ODEs and the Shc layer of 6. A layer-based model of the system is given in [105, Additional files 3 and 6].

The number of ODEs can be strongly reduced if it is assumed that the four binding sites for PI3K on IRS are equivalent and that the two binding sites each for IRS and Shc on the receptor are also equivalent. As described in Section 4.4.1, it is only necessary to model each qualitatively different binding site once. The IRS layer can be described by $2^4 = 16$ ODEs instead of $2^7 = 128$, the receptor layer can be described by $2^4 = 16$ ODEs instead of $2^6 = 64$, and the PI3K layer can be described by 2 instead of 5 ODEs. This results in 51 ODEs for the final model. A layer-based model definition for the 51 ODE model is given in Appendix A.8. Evaluating this model definition with ALC (presented in Chapter 5) results in documentation files and ready-to-run simulation files in different formats.

4.7 Reduction of the model size

A typical situation in cellular signaling is that there are multiple sites within the same protein. In this case, the size of the conventional model is a product of the number of states of each site, where each possible state of a binding partner must be considered as a separate

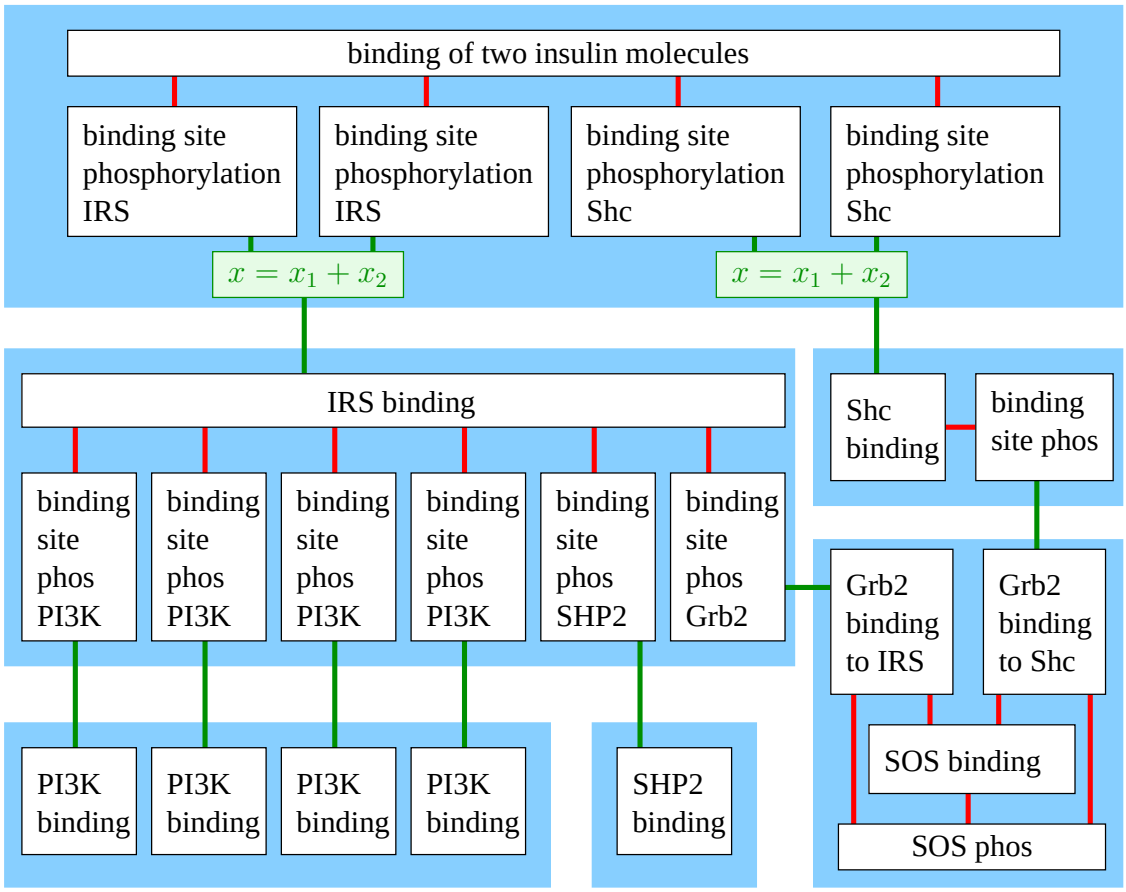

Figure 4.10: Interaction graph of insulin signaling
All processes (white boxes) that are coupled by graded interactions (red lines) are described in the same layer (blue boxes). Processes of different layers are only connected by all-or-none interactions (green lines) or do not interact. The green boxes $x = x_1 + x_2$ indicate that the corresponding binding sites are equivalent with respect to effector binding (see Section 4.4.2). According to the definition of layers, each process of PI3K binding to a binding site on IRS should be described in a separate layer. However, merging the four PI3K layers (done for graphical reasons) has no impact on the model equations. This figure shows a simplified scenario where the different phosphorylation sites of the same molecules do not interact. Such interactions would neither change the modular structure of the model nor the number of necessary ODEs.

state of the site. In the layer-based model however, each binding site that is phosphorylated has usually two states (unphosphorylated and phosphorylated). The states of the binding partner are modeled in a separate module (layer) and do not increase the number of states of the original protein. Another typical case is that binding partners are phosphorylated on binding sites and bind other effectors. In this case, the same argumentation can be applied to each additional effector in such an association chain. Both scenarios lead to implicitly assumed equilibrium conditions for diagonal reactions (or independence assumptions) that are the basis of the approximation in layer-based models (see Section 4.5).

Table 4.2: Number of ODEs necessary for models of the insulin signaling system
Layer-based modeling, domain-oriented model reduction and their combination are compared in three scenarios. The insulin concentration is considered as the input of the system, and the two binding sites on the receptor for Shc and IRS each are assumed to be equivalent with respect to effector binding. All binding site phosphorylations and the corresponding effector bindings undergo all-or-none interactions. **1)** All binding sites on the same molecule may perform graded interactions. **2)** The phosphorylation state of the binding sites does not influence the phosphorylation of other binding sites on the same molecule. **3)** As scenario 2), but all binding sites for the same effectors are assumed to be equivalent.

Scenario	Conventional modeling	Domain-oriented model reduction	Layer-based modeling	Combination
1	145 156 468	145 156 468	214	214
2	145 156 468	212	214	56
3	145 156 468	60	51	39

The exponential growth of the number of state variables with the number of binding sites and binding partners in conventional models is partly replaced by linear growth with the number of layers. Within the layers there is still exponential growth with respect to the number of molecular sites included in the layer. Applying the advanced strategies of layer-based modeling (Section 4.4) may strongly reduce the exponential growth within the layers.

Combinatorial complexity increases if there are additional modification and binding events (see Section 1.3). However, the reduction potential of layer-based modeling strongly increases with increasing combinatorial complexity. For the small example system (Section 4.3.6), where a layer-based model with 4 ODEs approximates a conventional model with 5 ODEs, the number of ODEs is decreased by 20%. For the extended example system (Section 4.6.1), where 11 ODEs approximate 21 ODEs, the number of ODEs is decreased by 48%.

Table 4.2 shows how many ODEs are necessary to model the insulin signaling system as described in Section 1.3. Conventional modeling, domain-oriented model reduction [93,103], layer-based modeling and a combination of the reduction techniques are compared in three scenarios. Scenario 1 is a very general setting where all sites of a protein may interact. In scenario 2 it is assumed that all phosphorylation sites of the same protein do not interact, and in scenario 3 it is additionally assumed that all binding sites on a protein for the same effector are equivalent. In scenarios 1 and 2, a layer-based model consists of 214 ODEs, whereas only 51 ODEs are necessary in scenario 3 (see Section 4.6.2). Considering the model with 214 ODEs, this means that the number of necessary equations is decreased by 99.99985 %. Scenario 3 allows the modeler to exploit the equivalence of binding sites (Section 4.4.1) which reduces the number of ODEs to 51. This corresponds to a total decrease in the number of necessary ODEs by 99.99996 %.

Domain-oriented model reduction [93, 103] (Sections 3.7 and 4.4.4) allows for a further reduction of the model size in scenarios 2 and 3. In scenario 2, the 214 ODEs can be reduced to 56 ODEs [105, Additional files 7 and 8], whereas in scenario 3 the 51 ODEs can be reduced to 39 ODEs (not shown). Note that the reduction is macroscopically exact in both cases. The 214 ODE model and its domain-oriented reduction are given in [105, Additional files 6-8]. A model definition for the 51 ODE model is given in Appendix A.8. Evaluating this model definition with ALC (see Chapter 5) results in documentation files and executable simulation files.

Altogether, layer-based modeling allows for a strong reduction of the model size compared to conventional modeling. Even large systems can be described by a model of acceptable size.

4.8 Conclusions

The layer-based approach allows for a macroscopic description of cellular signaling systems with inherent combinatorial complexity. It is an approximative modeling technique that results in a reduced order model with a pronounced modular structure. Layer-based models are characterized by a high approximation quality. They can be directly built, the preceding generation of a conventional model is not necessary.

An interaction graph representing the interactions between the considered processes defines the modularity of the model. The resulting modules, called layers, are connected by signal flows and can be modeled separately from each other once their connections are defined. The signal flows between the layers are macroscopic quantities representing phosphorylation and occupancy of binding sites.

In the derivation of the modules from the interaction graph one has to distinguish between graded interactions and all-or-none interactions. The layers are chosen such that processes of different layers only interact via all-or-none interactions. All processes of the same layer are directly or indirectly connected via graded interactions. A detailed reaction network comprising all processes of the current layer but neglecting all other processes is built for each layer. The reaction rates and the ODEs are derived from these reactions like in conventional modeling. Subsequently, dephosphorylation rates of binding sites to which effector binding is described in different layers are modified by correction terms guaranteeing that only unoccupied binding sites are dephosphorylated.

The basis of the approximation are rapid equilibrium assumptions for virtually introduced reactions, called diagonal reactions. These virtual reactions connect processes of different layers that do not interact directly. A different interpretation of the same equations leads to independence assumptions for the corresponding processes. The approximation errors follow linear first order ODEs and are characterized by input-to-state stability. Dynamic and ultimate bounds for the approximation errors are given.

The approximation is stationarily exact under certain conditions, e.g. if the system is able to reach thermodynamic equilibrium. Layer-based models are dynamically exact representations of the corresponding systems if in addition graded interactions do not occur

in layers that are coupled to other layers and if the initial conditions fulfill the equilibrium conditions for the diagonal reactions. If a subsystem fulfills the conditions for stationary or dynamic exactness and is not influenced by processes of other subsystems, the macroscopic variables describing the state of the processes of this subsystem are exactly described. In all other cases, layer-based modeling is a purely approximative method characterized by a small approximation error for physiologically relevant parameter values.

Layer-based modeling allows for an enormous reduction of the number of necessary ODEs compared to a conventional model. As an example, it is possible to model the insulin signaling system with only 214 ODEs compared to the $1.5 \cdot 10^8$ ODEs needed for a conventional model. If certain equivalence conditions for binding sites are fulfilled, 51 ODEs are sufficient.

The strong reduction and the high approximation quality clearly demonstrate the suitability of layer-based modeling for description and analysis of cellular signaling networks. However, building large models is laborious and sometimes also error-prone. To facilitate layer-based modeling, the next chapter introduces ALC (Automated Layer Construction), a tool for automated layer-based modeling that performs the laborious and error-prone parts of the model generation automatically.

5 ALC: Rule-based automation of layer-based modeling

ALC (Automated Layer Construction) is a computer program that converts model definitions for layer-based models into ready-to-run simulation files in different formats as well as into documentation files. This highly simplifies layer-based modeling since the laborious and error-prone steps of the model generation are performed automatically.

Section 5.1 provides a short introduction to the main features and the architecture of ALC. The syntax and the preparation of the model definitions are described in Section 5.2. One of the error-prone steps in manual layer-based modeling is the assignment of correction terms guaranteeing that only unoccupied binding sites are dephosphorylated. ALC automatically performs this assignment following an algorithm presented in Section 5.3. ALC can be used offline or online. Both possibilities are described in Section 5.4 which also provides a step by step procedure and a description of the output files.

5.1 Introduction to ALC

ALC (Automated Layer Construction) is a computer program highly simplifying the rule-based generation of layer-based models. The user has to prepare a model definition specifying the reaction network of each layer and the signal flow between the layers. Such a model definition only contains the minimal information uniquely defining the model equations. The intuitive syntax of model definitions is simple, but powerful and supports the concepts of rules, macrostates and modularity. Generating a model definition is much simpler and faster than generating the corresponding model equations manually. After running ALC, the resulting model files can be directly used for simulations and model analysis.

ALC performs detailed syntax checks and consistency checks on the model definition. It is checked if all molecule definitions, rules, reactions, assignments and outputs match the corresponding general format. In addition, it is verified that each used variable is uniquely defined. There are numerous more specific checks, described in the ALC user guide [111, Additional file 1], which result in informative error messages and warnings ensuring that most errors are easy to find and easy to correct. Critical inconsistencies result in error messages indicating the nature and the location of the errors. In this case, the model generation is not continued. Minor inconsistencies, e.g. undefined initial conditions or parameter values, result in detailed warnings. The model generation is continued and the model files will be executable.

The output files of ALC are ready-to-run simulation files in the formats C MEX, MATLAB (The MathWorks), *Mathematica* (Wolfram Research) and SBML [115]. The models are also provided in LaTeX [116, 117] and plain text format to simplify their publication or presentation. ALC can be used offline or via a form on the ALC website [118]. The program is freely available and published under the GNU Lesser General Public License.

5.1.1 Architecture and flow diagram

ALC is written in the programming language Perl [119, 120]. When used offline, the model definition is read in by the script file `ALC.pl` which also starts the modeling procedure. The source code necessary for the generation of output files is organized in three separate Perl modules. The Perl module `Procedures.pm` contains procedures that guide the application through the general steps of the model generation process. The Perl module `Functions.pm` contains functions that are called by the procedures from the module `Procedures.pm` and functions that are called by other functions from the same module. The Perl module `Output.pm` contains procedures that generate all output files after the model generation process is finished.

The form on the ALC website [118] is linked to a Perl CGI script that is running on an Apache web server [121]. This CGI script extracts the model definition from the uploaded model definition file or from the text field in which the model definition was pasted and starts the model generation. It also displays the results and provides the download links for the output files. The model generation and the generation of the output files is performed by the same Perl modules that are used offline.

ALC processes the following tasks, which constitute the most important steps in the generation of the output files from the model definition.

- Syntax checks on the model definition
- Consistency checks on the model definition
- Generation of reactions from rules
- Generation of reaction rates
- Generation of ODEs from reactions and reaction rates
- Output of the simulation files and documentation files in different formats

The sequence of steps and the assignment of these steps to the most important procedures of the Perl module `Procedures.pm` is shown in more detail as a flow diagram in Figure 5.1.

5.2 Preparing the model definition

The model definition uniquely defines the resulting model equations. It is prepared in plain text and consists of distinct sections that are encapsulated by `#name` and `#end name`, where

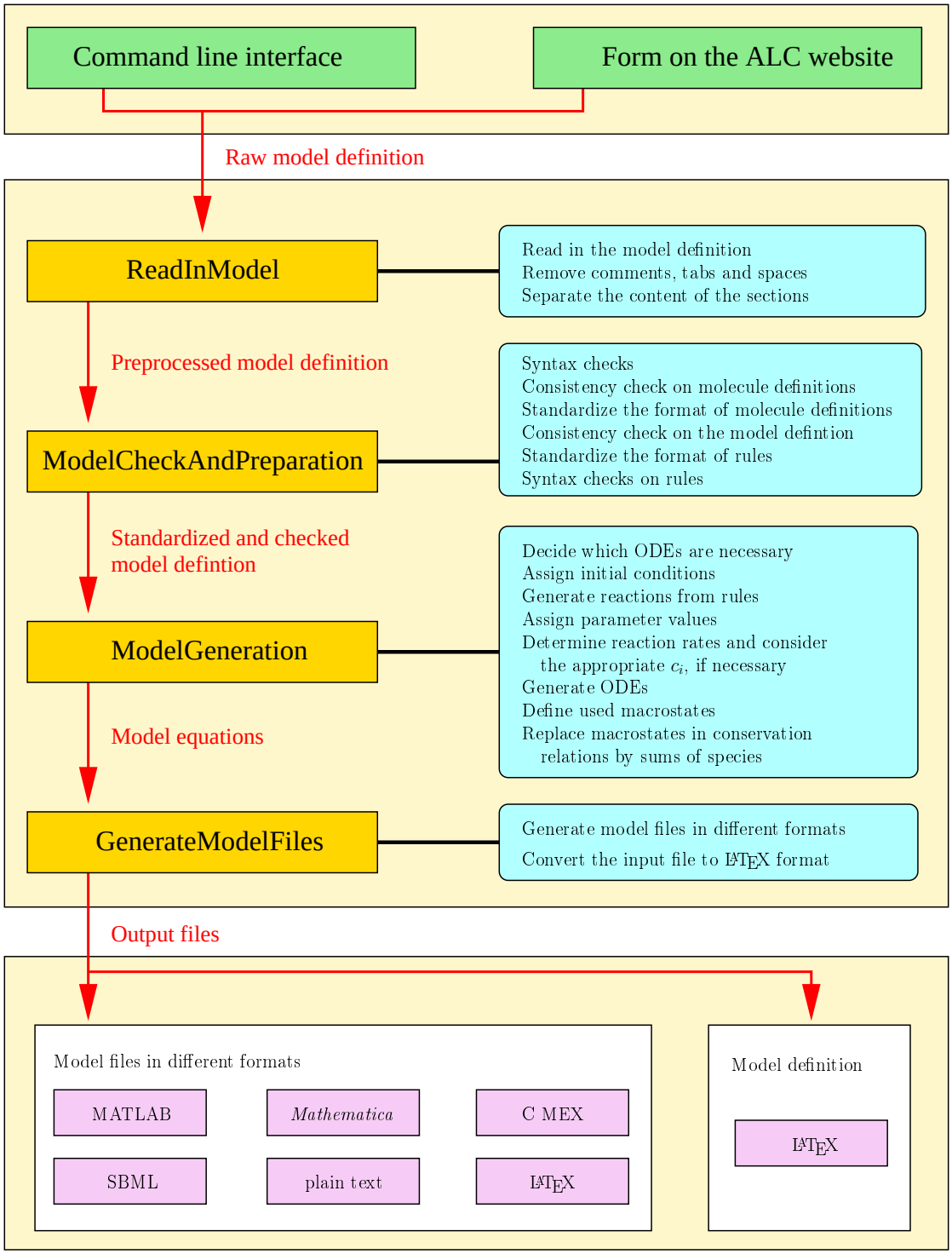

Figure 5.1: Flow diagram of ALC

The green boxes represent the different possibilities when starting ALC. The gold boxes represent procedures of ALC (stored in the Perl module `Procedures.pm`) that initiate the general processes from reading in the model definition to generating the finished model files. The red arrows represent the input and the output of the procedures. A short description of the processes that are initiated by the procedures is given in the cyan boxes. The white boxes define the principle class of the output files, while the magenta boxes define the format of the output files.

'name' is the name of the section. All sections may occur as often as desired in the model definition. Therefore, each layer can be defined separately, supporting the modular structure of layer-based models (see e.g. Equations 4.3 and 4.6).

The layer-based model definition for the small example system (Figure 4.4) is shown in Figure 5.2. The model definition for the 51 ODE model of insulin signaling is given in Appendix A.8. Model definitions for other example systems, including the model definition for the layer-based model of insulin signaling with 214 ODEs [105], can be found in [111, Additional file 3] and on the ALC website [118].

Note that ALC can also be used for conventional (Section 1.2) and rule-based modeling (Section 1.5). Using the site modification 'P' is not allowed in these cases since it is a special feature of layer-based modeling associated with signal flows between the layers.

The syntax of ALC and the distinct sections are introduced below. More detailed descriptions can be found in the ALC user guide. A tutorial demonstrates the generation of the model definition for an example system in detail. The first releases of the user guide and the tutorial are given in [111, Additional files 1 and 4], recent releases are available on the ALC website [118] and on the ALC project page on SourceForge.net [122].

5.2.1 Molecule definitions: the section '#molecules'

All molecule definitions are given in the section `#molecules` (Figure 5.2). A molecule definition consists of the molecule name that is optionally followed by the successive definition of all sites (see Section 4.2.1). The molecule name has to start with a capital letter which is optionally followed by an arbitrary sequence of alphanumeric symbols (letters, digits and underscores are allowed). Each site definition consists of a comma-separated sequence of all possible modifications which is encapsulated by curly brackets. A site modification is defined by a sequence of alphanumeric symbols. The definition of a site modification 'X' is forbidden, since 'X' is reserved for indicating macrostates and patterns. A site modification 'P' has a predefined meaning. It marks the site as a binding site to which effector binding is described in another layer. In this case, 'P' indicates the phosphorylated state.

As an example, a molecule definition `R{0,L}{0,P}` defines the molecule R that can have the modifications '0' and 'L' at the first site and the modifications '0' and 'P' at the second site. If a molecule is defined, its instances (e.g. species like $R[L, 0]$ or macrostates like $R[X, P]$) can be used in other sections of the model definition.

Complexes: ALC supports the simple notation of complexes introduced in Section 4.2.1. Complexes of species can be treated as species of one of the corresponding molecule definitions, or as species of new molecules that are defined for complexes of the molecules. As an example, the species $R[L, 0]$ is the complex of the species L and $R[0, 0]$ as this is already contained in the molecule definition `R{0,L}{0,P}`, whereas the dimer of `R{0,L}{0,P}` has to be explicitly defined, e.g. as `R2{0,L}{0,P}{0,L}{0,P}`.

```
############################
## receptor layer

#molecules
R{0,L}{0,P}
L
#end molecules

#parameters
totR=40
k1=0.001
k1d=4*10^(-4)
k2=0
k2d=0.00385
k3=0.0231
k3d=0.00385
#end parameters

#initial conditions
R[L,0]=0.1
#end initial conditions

#clamped concentrations
L=10
#end clamped concentrations

#reactions
L+R[0,X]<->R[L,X]   k1   k1d
R[0,0]<->R[0,P]   k2   k2d
R[L,0]<->R[L,P]   k3   k3d
#end reactions

#layer connections
x=R[X,P]
#end layer connections

#molecular balances
R[0,0]=totR-R[X,X]+R[0,0]
#end molecular balances
```

```
############################
## effector layer

#molecules
E
RXE
#end molecules

#initial conditions
RXE=0
#end initial conditions

#parameters
totE=250
kE=0.033
kEd=0.113
#end parameters

#reactions
RXp+E<->RXE   k4   k4d
#end reactions

#layer connections
xb=RXE
RXp=x-xb
#end layer connections

#molecular balances
E=totE-RXE
#end molecular balances


########################


#output
R[L,X]
RXE
#end output
```

Figure 5.2: Model definition for the small example system

Default values which can be freely chosen are assigned to all initial conditions and parameters that are not given (see Section 5.4.2 and Appendix A.7). Note that for all non-zero default values this model definition guarantees $x \geq xb$ and $x > 0$ for the initial conditions. When running the C MEX, MATLAB and *Mathematica* models, the simulation results for $R[L, X]$ and RXE are visualized since they are defined as outputs. The simulation results for RXp, x and xb are also visualized as this is automatically done for all variables defined in the section `#layer connections`.

5.2.2 Defining rules and reactions: the section '#reactions'

All rules and reactions are defined in the section `#reactions` (Figure 5.2). The standard format of a rule or reaction is: chemical reaction equation, tab, parameter of the forward reaction, tab, parameter of the backward reaction.

$$\underbrace{\texttt{aA+bB <-> cC+dD}}_{\textit{reaction equation}} \overbrace{\quad}^{\textit{tab}} \underbrace{\texttt{parameter}}_{\textit{forward}} \overbrace{\quad}^{\textit{tab}} \underbrace{\texttt{parameter}}_{\textit{backward}} \tag{5.1}$$

The reaction symbol is '<->' for reversible reactions and '->' or '<-' for irreversible reactions. For irreversible reactions only one parameter has to be given. ALC generates the corresponding rates following the generalized law of mass action where the parameters can be general nonlinear functions (Section 1.2.3).

If a reaction describes the transition of a site modification 'P' to anything but 'P' (i.e. the dephosphorylation of a binding site to which effector binding is described in another layer), the reaction rate has to be modified using a correction term c_i to guarantee that only the unoccupied fraction of binding sites with a site modification 'P' becomes dephosphorylated. ALC automatically searches the corresponding x_i and $x_i b$ in the model definition and considers $c_i = (x_i - x_i b)/x_i$ in the rate laws for dephosphorylation. As the correction terms c_i in the reaction rates are automatically assigned by ALC, one need not include them in the reaction parameters. The algorithm for this assignment is described in Section 5.3.

5.2.2.1 Defining rules

A rule represents a class of reactions with common properties (see Section 1.5). In particular, all reactions of this class are parameterized by the same kinetic constants. The hallmark of rules is that they must contain at least one pattern in the reaction equation. Note that the sole occurrence of macrostates (which have the same notation as patterns) in the parameter part of a reaction does not result in a rule, as at this position macrostates represent the sums of all corresponding species.

As an example, for `R{0,L}{0,P}{0,p}` the rule

```
R[0,X,X]+L <-> R[L,X,X]     k0     k1
```
(5.2)

represents four reactions.

$$\begin{aligned} R[0,0,0] + L &\leftrightharpoons R[L,0,0] \quad k0 \quad k1 \\ R[0,0,p] + L &\leftrightharpoons R[L,0,p] \quad k0 \quad k1 \\ R[0,P,0] + L &\leftrightharpoons R[L,P,0] \quad k0 \quad k1 \\ R[0,P,p] + L &\leftrightharpoons R[L,P,p] \quad k0 \quad k1 \end{aligned} \tag{5.3}$$

Rules may also be used for association reactions, as exemplified for dimerization in the rule

```
Rmon[X]+Rmon[X] <-> Rdim[X,X]      k0     k1
```
(5.4)

which for `Rmon{0,L}` and `Rdim{0,L}{0,L}` is evaluated to

$$\begin{array}{lcllll} Rmon[0] & + & Rmon[0] & \leftrightharpoons & Rdim[0,0] & k0 \quad k1 \\ Rmon[0] & + & Rmon[L] & \leftrightharpoons & Rdim[0,L] & k0 \quad k1 \\ Rmon[L] & + & Rmon[0] & \leftrightharpoons & Rdim[L,0] & k0 \quad k1 \\ Rmon[L] & + & Rmon[L] & \leftrightharpoons & Rdim[L,L] & k0 \quad k1. \end{array} \tag{5.5}$$

Stoichiometric coefficients can be used in rules. As an example, the evaluation of the rule

```
2 Rmon[X] <-> Rdim[X,X]    k0    k1
```
(5.6)

leads to the same results as the evaluation of the rule given by Equation 5.4.

Note that compared to the binding of L, binding sites for the aggregation of two larger molecules (e.g. the dimerization of *Rmon*) do not have to be defined. In this case, the complex (e.g. *Rdim*) is defined as a new molecule whose site definitions are the sequence of the site definitions of the reactants.

5.2.2.2 Symmetric reaction rules

ALC deals with the problem of symmetric reaction rules correctly. In Equation 5.5, the second and the third reaction are symmetric to each other. If the two sites on *Rdim* are really equivalent, $Rdim[0,L]$ and $Rdim[L,0]$ are indistinguishable species. *Rdim* with one bound L is represented by $Rdim[0,L] + Rdim[L,0]$. Therefore, the rate constant for the association of $Rmon[0]$ and $Rmon[L]$ is twice the nominal rate constant $k0$.

This parameterization differs from the one used by BioNetGen 2 [47], where homodimerization is parameterized by 0.5 times the nominal rate constant. Both solutions are equivalent but differ by a constant factor of two in the parameterization.

5.2.2.3 Generalized mass action kinetics for rules

Following the generalized law of mass action (Section 1.2.3), rules can be parameterized by complex kinetic laws in the same way as reactions. This is demonstrated for the degradation of a molecule M that has the molecule definition `M{0,1,...,q}` representing the species $M[i]$, $i \in [0,q]$. The degradation follows macroscopic Hill kinetics for $M[X]$.

```
M[X] ->        kmax*M[X]^(n-1)/(M[X]^n+km^n)
```
(5.7)

Applying the generalized law of mass action, this results in the rates

$$r_i = \frac{kmax \cdot M[i] \cdot M[X]^{n-1}}{M[X]^n + km^n} \tag{5.8}$$

which together are

$$r_{tot} = \sum_{i=0}^{q} r_i = \frac{kmax \cdot M[X]^n}{M[X]^n + km^n}. \tag{5.9}$$

It can be seen that the macroscopic rate r_{tot} is correctly allocated to the species $M[i]$.

$$r_i = \frac{M[i]}{M[X]} \cdot r_{tot} \tag{5.10}$$

Michaelis-Menten kinetics can be realized using $n = 1$ in Equation 5.7.

5.2.2.4 Advanced association rules

A very strong restriction for rule definitions is that corresponding sites with a modification 'X' must have the same site definitions (Section 1.5). The evaluation of rules to reactions is not uniquely defined if this constraint (checked by ALC) is violated. Defining rules which describe the association of two molecules can be problematic if site modifications of a molecule include other molecules which can be modified themselves. An important example is the association of a ligand $L1$ to a receptor R which is defined by the rule

```
R[X,0]+L1[X] <-> R[X,X]     k1     k2.
```
(5.11)

This rule can only be evaluated to reactions if the second site of the receptor has the same site definition as the sole site of the ligand $L1$. In addition, the option `StrictRS=0` slightly relaxing the syntax checks for rules (see ALC user guide) has to be set in the file `Config_ALC.txt` or via the command line when executing ALC (see Section 5.4.2 and Appendix A.7). The required identity of site definitions implies that a possible site modification of the ligand $L1$ has to be '0' and that the evaluation of the rule also leads to reactions defined by

$$R[X,0] + L1[0] \leftrightharpoons R[X,0] \qquad k1 \qquad k2. \tag{5.12}$$

This is in most cases not the desired result since such reactions are degradation reactions for $L1[0]$. A possible solution of this problem is to define $L1$ in a special way and to remove the species $L1[0]$ completely. Using the section `#remove` allows for the complete removal of defined species from the model. This includes the removal of all reactions where these species occur and the omission of these species in all macrostates. We illustrate this on an example. The model definition

```
#molecules
R{0,1}{0,L1,L12,L123}
L1{0,L1,L12,L123}
#end molecules
```

```
#reactions
R[X,0]+L1[X]<->R[X,X]    k1    k2
#end reactions
```

defines the association of R and $L1$ but also leads to reactions defined by Equation 5.12. Let us consider the species $L1[L1]$ as the molecule $L1$ having the configuration 'L1' which indicates that it is not modified. In this case, the species $L1[0]$ (usually representing what in this example is represented by $L1[L1]$) is not necessary any more. Adding

```
#remove
L1[0]
#end remove
```

to the model definition completely removes $L1[0]$ from the system and results in the reactions

$$\begin{array}{lcllll}
R[0,0] & + & L1[L1] & \leftrightharpoons & R[0,L1] & k1 \quad k2 \\
R[0,0] & + & L1[L12] & \leftrightharpoons & R[0,L12] & k1 \quad k2 \\
R[0,0] & + & L1[L123] & \leftrightharpoons & R[0,L123] & k1 \quad k2 \\
R[1,0] & + & L1[L1] & \leftrightharpoons & R[1,L1] & k1 \quad k2 \\
R[1,0] & + & L1[L12] & \leftrightharpoons & R[1,L12] & k1 \quad k2 \\
R[1,0] & + & L1[L123] & \leftrightharpoons & R[1,L123] & k1 \quad k2.
\end{array} \tag{5.13}$$

In addition, the macrostate $L1[X]$ does also not contain $L1[0]$ and is defined as

$$L1[X] = L1[L1] + L1[L12] + L1[L123]. \tag{5.14}$$

If the section `#remove` contains macrostates, all corresponding species are removed. Obviously, this makes only sense for molecules with more than one site.

Modeling the association of complexes is very error-prone if there are no strict consistency checks and in particular the condition that corresponding sites with a site modification 'X' need to have the same site definition. The solution proposed here, which is defining a species in the corresponding molecule definition, subsequent removal of the same species and specific relaxation of syntax checks, may seem intricate. However, it is an efficient possibility allowing for a rule-based description of advanced association processes. The used syntax allows for consistency checks avoiding erroneous model definitions.

5.2.3 Defining signal flows: the section '#layer connections'

The signal flow between the layers is defined in the section `#layer connections` (Figure 5.2). Notations of the sums of species with a phosphorylated binding site x_i have to start with a lowercase 'x' and are optionally followed by a sequence of letters and digits. Notations of the sums of species with occupied binding sites x_ib start with the notation of the corresponding x_i and end with a lowercase 'b'. Note that notations of x_i are not allowed to end with a lowercase 'b'.

The phosphorylation of binding sites (site modification 'P') and the corresponding effector bindings are described in different layers. In this case, the binding partners for the effectors are defined by algebraic equations, typically as differences of x_i and x_ib. The definition of these binding partners, which represent sums of microscopic species with phosphorylated but unoccupied binding sites, is also done in the section `#layer connections`. Their notation is analogous to that for species, however, the binding partners do not have to be defined in the section `#molecules`.

5.2.4 Defining parameters: the section '#parameters'

The section `#parameters` is used to define parameter values (Figure 5.2). In contrast to species notation, parameter notations have to start with a lowercase letter (which must not be 'r' or 'x', as these letters are reserved for the notation of reaction rates and layer connections, respectively). Parameters can also be functions of other parameters (e.g. `k1=k2/k3`). Parameters for which no value is assigned, are set to the default value (the value of the

option `Param`) that can be defined in the file `Config_ALC.txt` or via the command line when running ALC (see Section 5.4.2 and Appendix A.7). A warning is given for each omitted parameter definition.

5.2.5 Defining initial conditions: the section '#initial conditions'

The section `#initial conditions` is used to define initial conditions (Figure 5.2). Each species that is not defined by an algebraic equation requires an initial condition. An initial condition for a species is defined by an equation that assigns a value to the species. Note that initial conditions can be functions of parameters (e.g. `R[0,P]=a/b+1`).

Initial conditions for which no value is assigned, are set to the default value (the value of the option `InCond`) that can be defined in the file `Config_ALC.txt` or via the command line when running ALC (see Section 5.4.2 and Appendix A.7). A warning is given for each undefined initial condition. Setting the default value for initial conditions to zero may result in division by zero in the correction terms c_i. To avoid this, initial conditions for some phosphorylated species have to be set manually to a non-zero value. Alternatively, one can use a default value very close to zero (e.g. 10^{-20}). During the assignment of initial conditions, one has to assure that $x_i \geq x_i b$ and $x_i > 0$ hold for all i (see Section 4.3.5).

5.2.6 Constant concentrations: '#clamped concentrations'

Concentrations of species can be set to a constant value (which can be a function of parameters) in the section `#clamped concentrations` (Figure 5.2). In this case, no ODE is generated for these species. Note that clamping the concentration of a protein species results in a non-constant overall concentration of this protein if there is no conservation relation. In the small example system, clamping L leads to a non-constant sum $L + R[L, X]$.

5.2.7 Defining conservation relations: '#molecular balances'

Conserved moieties are chemical entities that participate in a reaction system without loss of integrity and always remain in the system [22]. The total concentration of conserved moieties is constant, if the volume is constant. As an example, the sum of all species of a protein that is not degraded or synthesized is a conserved moiety. Conservation relations (molecular balances) allow the modeler to replace the ODE of one species for each conserved moiety by an algebraic equation, and can be defined in the section `#molecular balances` (Figure 5.2). ALC does not generate ODEs for these species. Considering `R{0,L}{0,P}`, an example for a conservation relation is

```
R[0,0]=totR-R[0,P]-R[L,0]-R[L,P]
```
(5.15)

where $totR$ is a parameter defining the total (constant) concentration of R. A conservation relation can be defined using macrostates and the species whose ODE is to be replaced by this conservation relation. Therefore,

```
R[0,0]=totR-R[X,X]+R[0,0]
```
(5.16)

is also allowed. ALC replaces the used macrostates by the corresponding sums of species. Subsequently, the species whose concentration is defined by this conservation relation is removed from the right hand side of the assignment. From this it follows that using

```
R[0,0]=totR-R[X,X]
```
(5.17)

leads to the same result as the assignments above.

The automatic removal of the species defined by the conservation relation from the right hand side of the assignment allows for the usage of macrostates, which highly simplifies the definition of conservation relations.

5.2.8 Defining algebraic assignments: '#algebraic relations'

The section `#algebraic relations` allows for the assignment of arbitrary algebraic expressions (that may contain species, macrostates, numerical values and parameters) to variables whose notation is analogous to that of a species. Variables defined in this section may simplify the usage of recurrent expressions in the parameterization of reaction rates. They can also be used to provide the output of lengthy algebraic expressions with a short and meaningful plot legend (see Section 5.2.9). The notations of variables that are defined in this section have to match the restrictions on species notation (Section 4.2.1). However, variables that are defined in this section (by assigning an algebraic expression to them) do not have to be defined in the section `#molecules`.

5.2.9 Defining model outputs: the section '#output'

The section `#output` allows for the declaration of species, macrostates, variables from the section `#algebraic relations`, or arbitrary algebraic expressions as outputs. The simulation results of all outputs and all variables defined in the section `#layer connections` are visualized in the resulting *Mathematica*, C MEX and MATLAB models. In the SBML model, the outputs are defined as $Output_i$. Note that the lines of this section do not contain equations, but algebraic expressions (see Figure 5.2).

The lines of the section `#output` are the legends of the plots when executing the simulation files. If more descriptive legends are preferred, new variables can be defined in the section `#algebraic relations` (e.g. `Out1=...`) and subsequently declared as outputs in the section `#output` (e.g. only `Out1` within one line).

The automatic addition of all variables defined in the section `#layer connections` to the output list is performed as they correspond to macroscopic quantities that are often of interest. It can be disabled by setting `OutLC=0` in the file `Config_ALC.txt` or in the command line when running ALC (see Section 5.4.2 and Appendix A.7).

5.2.10 Order of assignments

The order of the sections in the model definition is irrelevant. The sections where algebraic assignments can be defined are evaluated in the following order: `#parameters`, `#clamped`

`concentrations`, `#molecular balances`, `#layer connections`, `#algebraic relations`. This means for example that all variables whose values are defined in the other sections can be used in the section `#algebraic relations`. Variables defined in the section `#algebraic relations`, however, cannot be used in all the other sections (but of course in the sections `#output` and `#reactions`).

Variables which are defined and used within the same section have to be defined in a row that is above their first usage. The sole exception is the section `#parameters` where this strict rule is slightly relaxed. All parameter definitions where other parameters occur on the right hand side of the assignment are moved to the end of the parameter list. This is done in the order they are defined. All other parameter assignments are sorted alphanumerically and placed above.

The order of assignments is irrelevant when using the *Mathematica* model, which results from the repeated insertion of all assignments into the ODEs.

5.3 Automated assignment of correction terms

This section describes the conditions that are checked in order to decide whether a correction term c_i is necessary and if so, the algorithm to find the appropriate correction term. ALC detects the need for a correction term in the rate law of a reaction if the following two conditions are simultaneously fulfilled:

- A molecule M occurs on both sides of the reaction equation.
- The site modification of the molecule M is 'P' at the jth site on one side of the reaction equation and not 'P' on the other side of the reaction equation.

As an example, these conditions are fulfilled for the second and for the third reaction in Figure 5.2, whereas they are not fulfilled simultaneously in the remaining reactions.

ALC uses the following algorithm to find the appropriate correction term:

1. Replace macrostates by the corresponding sums in each definition of x_i.
2. Count for all x_i how often the molecule M occurs with a site modification 'P' at the jth site (factors preceding species or macrostates are not considered).
3. Take the x_i with the highest score and use $c_i = (x_i - x_i b)/x_i$ as the correction term.

This algorithm is implemented much faster than described above. The first step is only performed once. The second step is performed simultaneously for all molecules and positions with the site modification 'P' that occur in at least one x_i. This is done by building a hash (an associative array, a major data type in Perl) [119] which has the keys: molecule name and position. The values are the corresponding x_i having the highest score. Using this hash, the third step is trivial.

Consider a certain $x_{M,j}$ representing the sum of all species of the molecule M with phosphorylated jth site. Assume that effector binding depends on binding site phosphorylation but not on other modifications of M (which is the normal situation). In this case, the algorithm always leads to the desired correction term $(x_{M,j} - x_{M,j}b)/x_{M,j}$ for the dephosphorylation of the jth site of M. The reason is that the sum $x_{M,j}$ contains more species of the molecule M with phosphorylated jth site than the other x_i describing the phosphorylation of other sites. As an example consider the molecule `M{0,P}{0,P}...`, where the first two sites are binding sites to which effector binding is described in different layers. The sums of species with a phosphorylated binding site are $x_1 = M[P, X, ...]$ and $x_2 = M[X, P, ...]$ where the dots represent a sequence of site modifications 'X'. Applying the algorithm above leads to the correction term $(x_1 - x_1 b)/x_1$ in all dephosphorylation rates of the first site and $(x_2 - x_2 b)/x_2$ for those of the second site. Note that all species with phosphorylated first or second site are included in x_1 or x_2, respectively, but not all of them occur in the other x_i.

Ignoring the factors preceding species or macrostates in Step 2 allows for the application of the advanced strategy of equivalent binding sites (Section 4.4.1). Otherwise, such factors could dominate the analysis and lead to wrong correction terms for the dephosphorylation of other sites.

5.4 Using ALC

5.4.1 Step by step procedure: using ALC

This step by step procedure is a modification of the step by step procedure for layer-based modeling (Section 4.3.8) that can be applied when using ALC.

1. Identify all processes and their interactions.
2. Draw the interaction graph.
3. Deduce from the interaction graph how many layers are in the model and which processes form part of which layer. All processes that are coupled by graded interactions are part of the same layer. Processes of different layers are coupled by all-or-none interactions or do not interact.
4. Model each layer separately as if the processes of all other layers do not exist.
   a) Define all molecules in the section `#molecules` (e.g. `R{0,L}{0,P}`). The site modification (uppercase) 'P' is reserved for indicating the phosphorylation of binding sites to which the effector binds in another layer. A different notation (e.g. lowercase 'p') has to be used to indicate the phosphorylation of other sites.
   b) Define rules and reactions, as in conventional modeling, in the section `#reactions` (e.g. `R[0,X]+L<->R[L,X] k1 k2`). Variables representing frequently used expressions in the parameter part of the reactions can be defined in the section `#algebraic relations`.

c) Define all sums of species with phosphorylated binding sites x_i (e.g. `x=R[X,P]`) and all sums of species with occupied binding sites x_ib (e.g. `xb=RXE`) in the section `#layer connections`. Note that x_ib has the same notation as the corresponding x_i, followed by (lowercase) 'b'.

d) Define the binding partners for effectors whose binding sites are phosphorylated in other layers in the section `#layer connections` (e.g. `RXp=x-xb`).

e) Assign values to parameters in the section `#parameters` (e.g. `k1=3`). Undefined parameters are set to the default value (the value of the option `Param`) which can be defined in the file `Config_ALC.txt` or via the command line.

f) Assign initial conditions in the section `#initial conditions` (e.g. `R[0,0]=2`). Undefined initial conditions are set to the default value (the value of the option `InCond`) which can be defined in the file `Config_ALC.txt` or via the command line. The initial conditions have to guarantee $x_i > 0$ and $x_i \geq x_ib \; \forall \; i$.

g) Concentrations of species can be set to a constant value in the section `#clamped concentrations` (e.g. `L=10`).

h) Conservation relations (e.g. `R[0,0]=totR-R[X,X]+R[0,0]`) can be defined in the section `#molecular balances`. Using macrostates which may also contain the species whose ODE is to be replaced highly simplifies this step.

5. Define outputs (observables) to be visualized in the C MEX, MATLAB and *Mathematica* models in the section `#output`. The simulation results of all variables defined in the section `#layer connections` are visualized in the C MEX, MATLAB and *Mathematica* models without declaring them as outputs. This can be disabled by changing the value of the option `OutLC` which can be defined in the file `Config_ALC.txt` or via the command line.

6. Store the model definition in a file, e.g. 'layer.alc', which is the default file name for the model definition file.

7. Run ALC offline or use the form on the ALC website.

5.4.2 Running ALC

There are two possible ways of running ALC. The first is to use the form on the ALC website [118], where ALC can be accessed using a browser without any additional software.

The second possibility is to download the latest release of ALC from the same website [118], from SourceForge.net [122], or to use the first release that is given in [111, Additional file 5]. Perl (freely available, e.g. [120]) is required for the offline use. For the installation of ALC it is only necessary to unpack the file 'download_ALC.zip' (or [111, Additional file 5]) in the directory where the model definition files will be stored. If ALC is installed, store the model definition in the file 'layer.alc', open a command line interface, go to the ALC directory and

type

```
ALC.pl
```

followed by the return key. The equivalent command

```
perl ALC.pl
```

can also be used. These commands execute ALC with default values, e.g. for undefined parameters or initial conditions and for the filenames of the model definition file and the resulting model files. Without changing the default values, the model definition has to be stored in the file 'layer.alc', and the names of the resulting output files start with 'model'. All default values can be changed in the file `Config_ALC.txt` or via the command line for each call of ALC separately. A detailed description of default values and command line parameters is given in Appendix A.7. Note that all options have their equivalent as text fields or select elements in the form on the ALC website [118].

5.4.3 Resulting model files

The resulting model is given in MATLAB (default: modelM_call.m and modelM.m), C MEX (default: modelC_call.m, modelC_mdl.mdl and modelC.c), *Mathematica* notebook (default: model.nb) and *Mathematica* input file (default: model.mma.m) formats as directly executable files. They include the visualization of the simulation results for the user-defined output variables and (unless `OutLC=0`) all variables defined in the section `#layer connections`.

The core of the C MEX model is a MATLAB S-function written in C, which results in a much faster simulation of the model compared to the standard MATLAB format. A recent release of MATLAB (earliest version successfully tested: 7.0.4.352) is required for using the C MEX models. The *Mathematica* input file contains the same code as the notebook file, but simplifies the integration of layer-based models into existing *Mathematica* code.

The model is also given in SBML (default: model.xml), LaTeX (default: model.tex) and plain text (default: model.txt) formats, which allows for the direct usage of model equations for presentations or publications. Additionally, the model definition is converted to the LaTeX format (default: model.input.tex).

Note that the SBML standard [115] allows for explicit algebraic assignments, which are used in layer-based models. However, some SBML-supporting tools will not work with layer-based SBML models as these tools do not support algebraic assignments. An important tool that works with layer-based SBML models is MathSBML [123] (release 2.7.0.3 or higher).

5.4.4 Computational aspects

The output files for small layer-based models are generated in less than one second on a desktop PC. The generation of larger layer-based models that can correspond to extremely large conventional models is also very rapid. As an example, the output files for the layer-based model of insulin signaling (214 ODEs) that replaces a conventional model with $1.5 \cdot 10^8$

ODEs are generated in about one second on a desktop PC. A few thousand ODEs are usually generated within seconds.

The model size in the offline version of ALC is only restricted by CPU time, disk space and memory capacity. The online version of ALC which is accessible via a form on the ALC website [118] provides full functionality, but is restricted to model definitions that define no more than 500 species. This is done to keep the traffic and the processor load on the server at a reasonable size.

5.5 Conclusions

ALC converts model definitions given in a simple, but powerful rule-based syntax to computational models in different formats, as well as documentation files. The model definition is divided into distinct sections and can be structured to mirror the modularity of the model. A main benefit of ALC is that it dramatically simplifies layer-based modeling and reduces the risk of creating erroneous model equations.

The assignment of the correction terms c_i to dephosphorylation rates of binding sites to which effector binding is described in different layers is one of the more difficult and error-prone steps in manual layer-based modeling. ALC performs these assignments automatically, such that errors in this step are avoided. This strengthens the analogy of layer-based modeling to conventional modeling and rule-based modeling as the reaction network within each layer can now be defined using rules without considering correction terms.

ALC also supports the usage of macrostates. This highly simplifies the definition of the signals between the layers (x_i and $x_i b$) as well as the definition of conservation relations and the use of enzyme kinetics for rules and reactions.

In many cases, it is comfortable to have the model in different formats. As an example, SBML [115] is becoming the *de facto* standard for model representation in systems biology. Though by far not all modeling and simulation projects use SBML, it is often desired to provide SBML models for publications. In addition, the model equations in plain text are often part of the manuscript or provided as supplementary material. The manual format conversion of models however is probably the major reason for errors in published model descriptions. ALC prevents this problem since it automatically exports the models to several formats (C MEX, MATLAB, *Mathematica* and SBML) and provides the model equations in LaTeX format as well as in plain text format, which simplifies the presentation and publication of the models.

ALC is optimized for building layer-based models. It supports features of layer-based modeling that are not present in conventional or rule-based modeling, and that are not supported by other tools. Though it is not its main application, ALC is also well suited for rule-based or conventional modeling of reaction networks.

Altogether, ALC highly simplifies the generation of layer-based models. ALC is freely available and can be used offline or via a form on the ALC website [118].

6 Conclusions and outlook

Mathematical modeling of cellular signaling systems and subsequent model analysis is a promising way to a deeper understanding of physiological processes and to improved therapies for severe disorders such as cancer and diabetes mellitus. However, combinatorial complexity complicates the modeling of many signaling systems because the association of a few proteins can result in an enormous amount of feasible complexes and reactions. As one ODE is necessary for the balance of each feasible complex, conventional modeling of these systems is rarely possible. A potential solution is to simplify the description of complex formation at the receptors by focussing on the processes assumed to dominate the systems behavior.

This strategy was applied to the modeling of insulin dynamics and hepatic insulin receptor activation *in vivo*. Model analysis showed that insulin clearance and the relative contributions of the liver and the kidney to insulin degradation are highly dependent upon the insulin concentration. At low insulin concentrations, insulin is mainly degraded by the liver, whereas at high insulin concentrations insulin degradation is mainly performed by the kidney. Experimentally determined values of insulin clearance and relative contributions of the different tissues to insulin degradation are therefore only meaningful when the corresponding insulin concentrations are taken into account. This however is not considered in many experimental studies resulting in a wide variation of the reported values. As a consequence, insulin clearance and the relative contributions of the different tissues to insulin degradation should always be measured at a constant insulin concentration given in the resulting publications.

Due to combinatorial complexity, comprehensive modeling of many signaling systems is only possible if a reduced order modeling technique is applied. This thesis introduces layer-based modeling, a reduced order modeling technique for cellular signaling systems which is characterized by a high approximation quality. Layer-based models of signaling systems show a pronounced modular structure and provide a macroscopic description, with state variables that have a direct biochemical interpretation. Due to the macroscopic description, the number of necessary ODEs is highly decreased compared to conventional models. In contrast to most other model reduction techniques, a layer-based model can be directly generated by a procedure quite similar to conventional modeling. The preceding generation of a potentially very large conventional model is not necessary. Nevertheless, it is also possible to derive layer-based models from conventional models applying a formal procedure.

At the beginning of the direct model generation, the modularity of the model is deduced from a graph of all processes and their interactions. Then, the resulting modules, called layers, are modeled separately from each other and their connections are defined. The connections are given by signal flows representing phosphorylation and occupancy of binding

sites which are important physiological characteristics of signaling systems.

The approximation in layer-based modeling is based upon the implicit systematic assumption of rapid equilibrium for additionally introduced virtual reactions, called diagonal reactions. These assumptions are equivalent to independence assumptions for processes of different layers that do not interact directly. The approximation errors in layer-based modeling represent the distances of the diagonal reactions from equilibrium. They can be described by first order ODEs, characterized by input-to-state stability, whose solutions are dynamically and ultimately bounded. For physiologically relevant parameter values, the approximation error is very small. Qualitative checks allow for evaluating the approximation quality of layer-based models without comparing simulation results to those of a conventional model. In well-defined special cases, layer-based models describe the dynamics of the corresponding signaling systems in a stationarily or even dynamically exact way.

Altogether, layer-based modeling is a powerful technique for the modeling of signaling systems with inherent combinatorial complexity. It is the only applicable method if both a conventional model (which may be the result of rule-base modeling) and a domain-oriented model are too large for efficient simulation or parameter estimation. Layer-based modeling is especially suited for comparing many model variants, since in most cases only a few equations in a single layer have to be changed.

Though providing a macroscopic description, layer-based modeling of large systems may be time-consuming. To simplify the generation of layer-based models and to prevent errors in the modeling procedure, this thesis presents the computer program Automated Layer Construction (ALC). Using this tool, models can be defined in a simple, but powerful syntax supporting the concepts of rules and macrostates. The structure of a model definition mirrors the modularity of the model, as deduced from the interaction graph. Model definitions for ALC only contain the minimal information necessary for uniquely defining the models and simultaneously allow for numerous automated consistency checks. Layer-based models of very large systems are in most cases defined by relatively short and simple model definitions.

ALC is freely available and can be used offline or via a form on the ALC website. Running ALC results in directly usable simulation files in the formats C MEX, MATLAB, *Mathematica* and SBML. The model equations are also given in LaTeX and plain text format which simplifies their publication and presentation.

Layer-based modeling and automated model generation, as provided by ALC, highly simplify the modeling of signaling systems with inherent combinatorial complexity. Even the modeling of highly branched signaling networks requires in most cases only a moderate effort.

Potential future applications of layer-based modeling and ALC involve the modeling of mammalian signal transduction systems whose malfunctions may give rise to severe diseases. A promising possibility is to extend the *in vivo* model for insulin dynamics and hepatic insulin receptor activation presented here in close collaboration with experimental biologists. Such a model could cover the signaling cascades emerging from the hepatic insulin receptor and the effect of feedback phosphorylations. Model analysis can be expected to significantly contribute to understanding the effect of insulin on hepatocytes and the entire organism, and finally contribute to improved therapies for diabetes mellitus.

Bibliography

[1] Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P: *Molecular Biology of the Cell: Reference Edition*. Garland Science 2007.

[2] Hancock JT: *Cell Signalling*. Oxford University Press 2005.

[3] Kitano H: **Systems biology: a brief overview.** *Science* 2002, **295**(5560):1662–4.

[4] Kitano H: **Computational systems biology.** *Nature* 2002, **420**(6912):206–10.

[5] Saltiel A, Kahn C: **Insulin signalling and the regulation of glucose and lipid metabolism.** *Nature* 2001, **414**(6865):799–806.

[6] Turnheim K, Waldhäusl W: **Essentials of insulin pharmacokinetics.** *Wien Klin Wochenschr* 1988, **100**(3):65–72.

[7] Duckworth W, Bennett R, Hamel F: **Insulin degradation: progress and potential.** *Endocr Rev* 1998, **19**(5):608–24.

[8] Chang L, Chiang SH, Saltiel AR: **Insulin signaling and the regulation of glucose transport.** *Mol Med* 2004, **10**(7-12):65–71.

[9] Khan A, Pessin J: **Insulin regulation of glucose uptake: a complex interplay of intracellular signalling pathways.** *Diabetologia* 2002, **45**(11):1475–83.

[10] Plum L, Belgardt BF, Brüning JC: **Central insulin action in energy and glucose homeostasis.** *J Clin Invest* 2006, **116**(7):1761–6.

[11] Mounier C, Posner BI: **Transcriptional regulation by insulin: from the receptor to the gene.** *Can J Physiol Pharmacol* 2006, **84**(7):713–24.

[12] Saltiel AR, Pessin JE: **Insulin signaling pathways in time and space.** *Trends Cell Biol* 2002, **12**(2):65–71.

[13] Taniguchi CM, Emanuelli B, Kahn CR: **Critical nodes in signalling pathways: insights into insulin action.** *Nat Rev Mol Cell Biol* 2006, **7**(2):85–96.

[14] Leng Y, Karlsson HKR, Zierath JR: **Insulin signaling defects in type 2 diabetes.** *Rev Endocr Metab Disord* 2004, **5**(2):111–7.

[15] Chakraborty C: **Biochemical and molecular basis of insulin resistance.** *Curr Protein Pept Sci* 2006, **7**(2):113–21.

[16] Ludvigsson J: **Why diabetes incidence increases–a unifying theory.** *Ann N Y Acad Sci* 2006, **1079**:374–82.

[17] Musi N, Goodyear LJ: **Insulin resistance and improvements in signal transduction.** *Endocrine* 2006, **29**:73–80.

[18] Salsali A, Nathan M: **A review of types 1 and 2 diabetes mellitus and their treatment with insulin.** *Am J Ther* 2006, **13**(4):349–61.

[19] Stumvoll M, Goldstein BJ, van Haeften TW: **Type 2 diabetes: principles of pathogenesis and therapy.** *Lancet* 2005, **365**(9467):1333–46.

[20] Hirsch IB: **Insulin analogues.** *N Engl J Med* 2005, **352**(2):174–83.

[21] Atkins P, Paula Jd: *Physical Chemistry*. Oxford University Press 2006.

[22] Heinrich R, Schuster S: *The Regulation of Cellular Systems*. Chapman & Hall 1996.

[23] Kholodenko BN: **Cell-signalling dynamics in time and space.** *Nat Rev Mol Cell Biol* 2006, **7**(3):165–76.

[24] Gillespie DT: **A rigorous derivation of the chemical master equation**. *Physica A* 1992, **188**:404–425.

[25] Brown TE, LeMay HE, Bursten BE, Murphy C, Woodward P: *Chemistry: The Central Science*. Prentice Hall 2008.

[26] Hlavacek WS, Faeder JR, Blinov ML, Perelson AS, Goldstein B: **The complexity of complexes in signal transduction.** *Biotechnol Bioeng* 2003, **84**(7):783–94.

[27] Hlavacek WS, Faeder JR, Blinov ML, Posner RG, Hucka M, Fontana W: **Rules for modeling signal-transduction systems.** *Sci STKE* 2006, **2006**(344):re6.

[28] McKern NM, Lawrence MC, Streltsov VA, Lou MZ, Adams TE, Lovrecz GO, Elleman TC, Richards KM, Bentley JD, Pilling PA, Hoyne PA, Cartledge KA, Pham TM, Lewis JL, Sankovich SE, Stoichevska V, Silva ED, Robinson CP, Frenkel MJ, Sparrow LG, Fernley RT, Epa VC, Ward CW: **Structure of the insulin receptor ectodomain reveals a folded-over conformation.** *Nature* 2006, **443**(7108):218–21.

[29] Ward C, Lawrence M, Streltsov V, Garrett T, McKern N, Lou M, Lovrecz G, Adams T: **Structural insights into ligand-induced activation of the insulin receptor.** *Acta Physiol (Oxf)* 2008, **192**:3–9.

[30] Meyts PD: **The insulin receptor: a prototype for dimeric, allosteric membrane receptors?** *Trends Biochem Sci* 2008, **33**(8):376–84.

[31] White M: **The IRS-signalling system: a network of docking proteins that mediate insulin action.** *Mol Cell Biochem* 1998, **182**(1-2):3–11.

[32] Gual P, Marchand-Brustel YL, Tanti JF: **Positive and negative regulation of insulin signaling through IRS-1 phosphorylation.** *Biochimie* 2005, **87**:99–109.

[33] Pirola L, Johnston A, Obberghen EV: **Modulation of insulin action.** *Diabetologia* 2004, **47**(2):170–84.

[34] Asthagiri A, Lauffenburger D: **A computational study of feedback effects on signal dynamics in a mitogen-activated protein kinase (MAPK) pathway model**. *Biotechnol Prog* 2001, **17**:227–239.

[35] Sedaghat AR, Sherman A, Quon MJ: **A mathematical model of metabolic insulin signaling pathways.** *Am J Physiol Endocrinol Metab* 2002, **283**(5):E1084–101.

[36] Hatakeyama M, Kimura S, Naka T, Kawasaki T, Yumoto N, Ichikawa M, Kim JH, Saito K, Saeki M, Shirouzu M, Yokoyama S, Konagaya A: **A computational model on the modulation of mitogen-activated protein kinase (MAPK) and Akt pathways in heregulin-induced ErbB signalling**. *Biochem J* 2003, **373**(Pt 2):451–463.

[37] Haugh J, Schooler K, Well A, Wiley H, Lauffenburger D: **Effect of epidermal growth factor receptor internalization on regulation of the phosphoslipase C-gamma1 signaling pathway**. *J Biol Chem* 1999, **274**:8958–8965.

[38] Haugh J, Well A, Lauffenburger D: **Mathematical modeling of epidermal growth factor receptor signaling through the phospholipase C pathway: mechanistic insights and predictions for molecular interventions**. *Biotechnol Bioeng* 2000, **70**:225–238.

[39] Kholodenko BN, Demin OV, Moehren G, Hoek JB: **Quantification of Short Term Signaling by the Epidermal Growth Factor Receptor**. *J. Biol. Chem.* 1999, **274**(42):30169–30181.

[40] Moehren G, Markevich N, Demin O, Kiyatkin A, Goryanin I, Hoek JB, Kholodenko BN: **Temperature dependence of the epidermal growth factor receptor signaling network can be accounted for by a kinetic model.** *Biochemistry* 2002, **41**:306–20.

[41] Schoeberl B, Eichler-Jonsson C, Gilles ED, Müller G: **Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors**. *Nat. Biotechnol.* 2002, **20**(4):370–375.

[42] Liu G, Swihart MT, Neelamegham S: **Sensitivity, principal component and flux analysis applied to signal transduction: the case of epidermal growth factor mediated**

signaling. *Bioinformatics* 2005, **21**(7):1194–202.

[43] Hori SS, Kurland IJ, DiStefano JJ: **Role of endosomal trafficking dynamics on the regulation of hepatic insulin receptor activity: models for Fao cells.** *Ann Biomed Eng* 2006, **34**(5):879–92.

[44] Faeder JR, Blinov ML, Goldstein B, Hlavacek WS: **Rule-based modeling of biochemical networks.** *Complexity* 2005, **10**:22–41.

[45] Blinov ML, Faeder JR, Goldstein B, Hlavacek WS: **BioNetGen: software for rule-based modeling of signal transduction based on the interactions of molecular domains.** *Bioinformatics* 2004, **20**(17):3289–91.

[46] Faeder JR, Blinov ML, Hlavacek WS: **Graphical rule-based representation of signal transduction networks**. *Proc. ACM Symp. Appl. Computing* 2005, 133–140.

[47] Blinov ML, Yang J, Faeder JR, Hlavacek WS: **Graph theory for rule-based modeling of biochemical networks**. *Lect. Notes Comput. Sci.* 2006, **4230**:89–106.

[48] Novère NL, Shimizu T: **STOCHSIM: modelling of stochastic biomolecular processes.** *Bioinformatics* 2001, **17**(6):575–6.

[49] Lok L, Brent R: **Automatic generation of cellular reaction networks with Moleculizer 1.0.** *Nat Biotechnol* 2005, **23**:131–6.

[50] Calzone L, Fages F, Soliman S: **BIOCHAM: an environment for modeling biological systems and formalizing experimental knowledge.** *Bioinformatics* 2006, **22**(14):1805–7.

[51] Sedwards S, Mazza T: **Cyto-Sim: a formal language model and stochastic simulator of membrane-enclosed biochemical processes.** *Bioinformatics* 2007, **23**(20):2800–2.

[52] Castillo M, Scheen A, Letiexhe M, Lefèbvre P: **How to measure insulin clearance.** *Diabetes Metab Rev* 1994, **10**(2):119–50.

[53] Hammond B, Christensen J, Smith G: **A model of insulin distribution and uptake in the perfused rat liver.** *J Theor Biol* 1991, **149**:121–39.

[54] Corin R, Donner D: **Insulin receptors convert to a higher affinity state subsequent to hormone binding. A two-state model for the insulin receptor.** *J Biol Chem* 1982, **257**:104–10.

[55] Standaert M, Pollet R: **Equilibrium model for insulin-induced receptor down-regulation. Regulation of insulin receptors in differentiated BC3H-1 myocytes.** *J Biol Chem* 1984, **259**(4):2346–54.

[56] Backer J, Kahn C, White M: **Tyrosine phosphorylation of the insulin receptor during insulin-stimulated internalization in rat hepatoma cells.** *J Biol Chem* 1989, **264**(3):1694–701.

[57] Quon M, Campfield L: **A mathematical model and computer simulation study of insulin receptor regulation.** *J Theor Biol* 1991, **150**:59–72.

[58] Wanant S, Quon M: **Insulin receptor binding kinetics: modeling and simulation studies.** *J Theor Biol* 2000, **205**(3):355–64.

[59] Thorsteinsson B: **Kinetic models for insulin disappearance from plasma in man.** *Dan Med Bull* 1990, **37**(2):143–53.

[60] Hovorka R, Powrie J, Smith G, Sönksen P, Carson E, Jones R: **Five-compartment model of insulin kinetics and its use to investigate action of chloroquine in NIDDM.** *Am J Physiol* 1993, **265**(1 Pt 1):E162–75.

[61] Mosekilde E, Jensen K, Binder C, Pramming S, Thorsteinsson B: **Modeling absorption kinetics of subcutaneous injected soluble insulin.** *J Pharmacokinet Biopharm* 1989, **17**:67–87.

[62] Puckett W, Lightfoot E: **A model for multiple subcutaneous insulin injections developed from individual diabetic patient data.** *Am J Physiol* 1995, **269**(6 Pt 1):E1115–24.

[63] Wilinska ME, Chassin LJ, Schaller HC, Schaupp L, Pieber TR, Hovorka R: **Insulin kinetics in type-I diabetes: continuous and bolus delivery of rapid acting insulin.** *IEEE Trans Biomed Eng* 2005, **52**:3–12.

[64] Berger M, Rodbard D: **Computer simulation of plasma insulin and glucose dynamics after subcutaneous insulin injection.** *Diabetes Care* 1989, **12**(10):725–36.

[65] Trajanoski Z, Wach P, Kotanko P, Ott A, Skraba F: **Pharmacokinetic model for the absorption of subcutaneously injected soluble insulin and monomeric insulin analogues.** *Biomed Tech (Berl)* 1993, **38**(9):224–31.

[66] Cobelli C, Caumo A, Omenetto M: **Minimal model SG overestimation and SI underestimation: improved accuracy by a Bayesian two-compartment model.** *Am J Physiol* 1999, **277**(3 Pt 1):E481–8.

[67] Natali A, Gastaldelli A, Camastra S, Sironi A, Toschi E, Masoni A, Ferrannini E, Mari A: **Dose-response characteristics of insulin action on glucose metabolism: a non-steady-state approach.** *Am J Physiol Endocrinol Metab* 2000, **278**(5):E794–801.

[68] Hann CE, Chase JG, Lin J, Lotz T, Doran CV, Shaw GM: **Integral-based parameter identification for long-term dynamic verification of a glucose-insulin system model.** *Comput Methods Programs Biomed* 2005, **77**(3):259–70.

[69] Toffolo G, Campioni M, Basu R, Rizza RA, Cobelli C: **A minimal model of insulin secretion and kinetics to assess hepatic insulin extraction.** *Am J Physiol Endocrinol Metab* 2006, **290**:E169–E176.

[70] Thorsteinsson B, Fugleberg S, Binder C: **Non-linearity of insulin kinetics.** *Diabetologia* 1986, **29**(12):898.

[71] Gherzi R, Andraghetti G, Versari G, Cordera R: **Effect of insulin receptor autophosphorylation on insulin receptor binding.** *Mol Cell Endocrinol* 1986, **45**(2-3):247–52.

[72] Koschorreck M, Gilles ED: **Mathematical modeling and analysis of insulin clearance in vivo**. *BMC Systems Biology* 2008, **2**:43.

[73] Schlatter E, Schurek H, Zick R: **Renal handling of homologous and heterologous insulin in the isolated perfused rat kidney.** *Pflugers Arch* 1982, **393**(3):227–31.

[74] Doherty J, Kay D, Lai W, Posner B, Bergeron J: **Selective degradation of insulin within rat liver endosomes.** *J Cell Biol* 1990, **110**:35–42.

[75] Backer J, Kahn C, White M: **The dissociation and degradation of internalized insulin occur in the endosomes of rat hepatoma cells.** *J Biol Chem* 1990, **265**(25):14828–35.

[76] Drake P, Bevan A, Burgess J, Bergeron J, Posner B: **A role for tyrosine phosphorylation in both activation and inhibition of the insulin receptor tyrosine kinase in vivo.** *Endocrinology* 1996, **137**(11):4960–8.

[77] Faure R, Baquiran G, Bergeron J, Posner B: **The dephosphorylation of insulin and epidermal growth factor receptors. Role of endosome-associated phosphotyrosine phosphatase(s).** *J Biol Chem* 1992, **267**(16):11215–21.

[78] Overmoyer B, McLaren C, Brittenham G: **Uniformity of liver density and nonheme (storage) iron distribution.** *Arch Pathol Lab Med* 1987, **111**(6):549–54.

[79] Pass D, Freeth G: **The rat**. *ANZCCART News* 1993, **6(4)**:1–4.

[80] Rasch R, Dørup J: **Quantitative morphology of the rat kidney during diabetes mellitus and insulin treatment.** *Diabetologia* 1997, **40**(7):802–9.

[81] Klein HH, Ullmann S, Drenckhan M, Grimmsmann T, Unthan-Fechner K, Probst I: **Differential modulation of insulin actions by dexamethasone: studies in primary cultures of adult rat hepatocytes.** *J Hepatol* 2002, **37**(4):432–40.

[82] Desbuquois B, Lopez S, Burlet H: **Ligand-induced translocation of insulin receptors in intact rat liver.** *J Biol Chem* 1982, **257**(18):10852–60.

[83] Kruse V, Jensen I, Permin L, Heding A: **Fate of insulin analogs in intact and nephrec-**

tomized rats determined by their receptor binding constants. *Am J Physiol* 1997, **272**(6 Pt 1):E1089–98.

[84] Reed BC, Ronnett GV, Lane MD: **Role of glycosylation and protein synthesis in insulin receptor metabolism by 3T3-L1 mouse adipocytes.** *Proc Natl Acad Sci U S A* 1981, **78**(5):2908–2912.

[85] Pringault E, Plas C: **Differences in degradation processes for insulin and its receptor in cultured foetal hepatocytes.** *Biochem J* 1983, **212**(3):529–37.

[86] Blouin A, Bolender R, Weibel E: **Distribution of organelles and membranes between hepatocytes and nonhepatocytes in the rat liver parenchyma. A stereological study.** *J Cell Biol* 1977, **72**(2):441–55.

[87] Harada S, Loten E, Smith R, Jarett L: **Nonreceptor mediated nuclear accumulation of insulin in H35 rat hepatoma cells.** *J Cell Physiol* 1992, **153**(3):607–13.

[88] Rother K, Imai Y, Caruso M, Beguinot F, Formisano P, Accili D: **Evidence that IRS-2 phosphorylation is required for insulin action in hepatocytes.** *J Biol Chem* 1998, **273**(28):17491–7.

[89] Eaton R, Allen R, Schade D: **Hepatic removal of insulin in normal man: dose response to endogenous insulin secretion.** *J Clin Endocrinol Metab* 1983, **56**(6):1294–300.

[90] Bray G, Tartaglia L: **Medicinal strategies in the treatment of obesity.** *Nature* 2000, **404**(6778):672–7.

[91] Stagsted J, Hansen T, Roth R, Goldstein A, Olsson L: **Correlation between insulin receptor occupancy and tyrosine kinase activity at low insulin concentrations and effect of major histocompatibility complex class I-derived peptide.** *J Pharmacol Exp Ther* 1993, **267**(2):997–1001.

[92] Ristic S, Bates PC: **Effects of rapid-acting insulin analogs on overall glycemic control in type 1 and type 2 diabetes mellitus.** *Diabetes Technol Ther* 2003, **5**:57–66.

[93] Conzelmann H, Saez-Rodriguez J, Sauter T, Kholodenko BN, Gilles ED: **A domain-oriented approach to the reduction of combinatorial complexity in signal transduction networks**. *BMC Bioinformatics* 2006, **7**:34.

[94] Blinov ML, Faeder JR, Goldstein B, Hlavacek WS: **A network model of early events in epidermal growth factor receptor signaling that accounts for combinatorial complexity**. *Biosystems* 2006, **83**(2–3):136–151.

[95] Faeder JR, Hlavacek WS, Reischl I, Blinov ML, Metzger H, Redondo A, Wofsy C, Goldstein B: **Investigation of early events in Fc epsilon RI-mediated signaling using a detailed mathematical model.** *J Immunol* 2003, **170**(7):3769–81.

[96] Faeder J, Blinov M, Goldstein B, Hlavacek W: **Combinatorial complexity and dynamical restriction of network flows in signal transduction.** *Syst Biol (Stevenage)* 2005, **2**:5–15.

[97] Pawson T, Nash P: **Assembly of cell regulatory systems through protein interaction domains.** *Science* 2003, **300**(5618):445–452.

[98] Antoulas AC: *Approximation of Large-scale Dynamical Systems*. Cambridge University Press 2006.

[99] Borisov NM, Markevich NI, Hoek JB, Kholodenko BN: **Signaling through receptors and scaffolds: independent interactions reduce combinatorial complexity.** *Biophys J* 2005, **89**(2):951–966.

[100] Horn RA, Johnson CR: *Matrix Analysis*. Cambridge University Press 1990.

[101] Antoulas A, Sorensen D: **Approximation of large-scale dynamical systems: an overview**. *Int. J. Appl. Math. Comput. Sci.* 2001, **11(5)**:1093–1121.

[102] Conzelmann H, Gilles ED: *Functional Proteomics: Methods and Protocols*, Humana Press 2008 chap. Dynamic pathway modeling of signal transduction networks - A domain-oriented approach, 557 – 576.

[103] Conzelmann H, Fey D, Gilles ED: **Exact model reduction of combinatorial reaction networks.** *BMC Syst Biol* 2008, **2**:78.

[104] Borisov N, Chistopolsky A, Faeder J, Kholodenko B: **Domain-oriented reduction of rule-based network models**. *IET Systems Biology* 2008, **2**(5):342–351.

[105] Koschorreck M, Conzelmann H, Ebert S, Ederer M, Gilles ED: **Reduced modeling of signal transduction - a modular approach.** *BMC Bioinformatics* 2007, **8**:336.

[106] Hartwell L, Hopfield J, Leibler S, Murray A: **From molecular to modular cell biology.** *Nature* 1999, **402**(6761 Suppl):C47–52.

[107] Saez-Rodriguez J, Kremling A, Conzelmann H, Bettenbrock K, Gilles ED: **Modular Analysis of Signal Transduction Networks**. *IEEE Contr. Syst. Mag.* 2004, **24(4)**:35–52.

[108] Saez-Rodriguez J, Kremling A, Gilles ED: **Dissecting the puzzle of life: Modularization of signal transduction networks.** *Comput. Chem. Eng.* 2005, **29(3)**:619–629.

[109] Ederer M, Sauter T, Bullinger E, Gilles ED, Allgöwer F: **An Approach for Dividing Models of Biological Reaction Networks into Functional Units**. *Simulation* 2003, **79 (12)**:703–716.

[110] Felder S, Zhou M, Hu P, Ureña J, Ullrich A, Chaudhuri M, White M, Shoelson S, Schlessinger J: **SH2 domains exhibit high-affinity binding to tyrosine-phosphorylated peptides yet also exhibit rapid dissociation and exchange.** *Mol Cell Biol* 1993, **13**(3):1449–55.

[111] Koschorreck M, Gilles ED: **ALC: automated reduction of rule-based models**. *BMC Systems Biology* 2008, **2**:91.

[112] Borisov NM, Markevich NI, Hoek JB, Kholodenko BN: **Trading the micro-world of combinatorial complexity for the macro-world of protein interaction domains.** *Biosystems* 2006, **83**(2-3):152–66.

[113] Khalil HK: *Nonlinear Systems*. Prentice Hall 2001.

[114] Isidori A: *Nonlinear control systems*, Springer 1995 chap. Global decomposition of control systems, 77–105.

[115] Hucka M, Finney A, Sauro H, Bolouri H, Doyle J, Kitano H, Arkin A, Bornstein B, Bray D, Cornish-Bowden A, Cuellar A, Dronov S, Gilles E, Ginkel M, Gor V, Goryanin I, Hedley W, Hodgman T, Hofmeyr JH, Hunter P, Juty N, Kasberger J, Kremling A, Kummer U, Novère NL, Loew L, Lucio D, Mendes P, Minch E, Mjolsness E, Nakayama Y, Nelson M, Nielsen P, Sakurada T, Schaff J, Shapiro B, Shimizu T, Spence H, Stelling J, Takahashi K, Tomita M, Wagner J, Wang J, Forum S: **The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models.** *Bioinformatics* 2003, **19**(4):524–31.

[116] Lamport L: *LaTeX: A Document Preparation System: User's Guide and Reference Manual*. Addison-Wesley Longman, Amsterdam 1994.

[117] **Comprehensive TeX Archive Network**. [http://www.ctan.org].

[118] **ALC website**. [http://layer.mpi-magdeburg.mpg.de].

[119] Wall L, Christiansen T, Orwant J: *Programming Perl. There's More Than One Way To Do It*. O'Reilly Media 2000.

[120] **Comprehensive Perl Archive Network**. [http://www.cpan.org].

[121] **Apache HTTP Server Project**. [http://httpd.apache.org].

[122] **ALC on SourceForge.net**. [http://sourceforge.net/projects/alc/].

[123] Shapiro BE, Hucka M, Finney A, Doyle J: **MathSBML: a package for manipulating SBML-based biological models.** *Bioinformatics* 2004, **20**(16):2829–31.

[124] Shoelson S, Chatterjee S, Chaudhuri M, White M: **YMXM motifs of IRS-1 define substrate specificity of the insulin receptor kinase.** *Proc Natl Acad Sci U S A* 1992, **89**(6):2027–31.

A Appendix

A.1 Scientific contribution of this thesis

The main results presented in Chapters 2, 4 and 5 were published in peer-reviewed journals. The article [72] describes the results presented in Chapter 2. The main results of Chapter 4 were published in [105] and [111], and the results of Chapter 5 were published in [111]. The results were presented at international conferences (SBMC 2006, FOSBE 2007, SBMC 2008).

A.2 Software

The software package *Mathematica* (Wolfram Research) was used for all stationary and analytic computations. Dynamic simulations were performed with the software package MATLAB (The MathWorks). ALC was written in the programming language Perl [119,120]. This thesis and all publications were drafted using LaTeX [116, 117] on a Linux System.

A.3 Insulin dynamics and insulin receptor activation

A.3.1 Model parameters and initial conditions

This section provides additional information about the derivation of parameter values and initial conditions given in Table 2.1. *In vitro* insulin receptor autophosphorylation has a half-life of about 0.5 min (Figure 1 in [77]). Assuming linear first order kinetics, this corresponds to a rate constant of $kyp = 0.0231\ s^{-1}$. *In vitro* insulin receptor dephosphorylation at the plasma membrane has a half-life of about 3 min (Figure 2 in [76]). Assuming linear first order kinetics, this corresponds to a rate constant of $kyd = 0.00385\ s^{-1}$. *In vitro* insulin receptor dephosphorylation at endosomal membranes has a half-life of 1.6 min (Figure 2 in [77]). Assuming linear first order kinetics, this corresponds to a rate constant of $kyden = 0.00722\ s^{-1}$.

Weights of livers and bodies given in literature [53] were compared. In average, the liver contributes about 5 % to the body weight of rats. There are 10^5 insulin receptors per hepatocyte [88]. Assuming that the hepatocyte is a sphere with a diameter of 20 μm, this corresponds to a total receptor concentration of $R_{tot} = totR = 40\ nM$. The basal insulin concentration in fasted mice is $0.3 - 0.5\ ng \cdot ml^{-1}$ (Gisela Drews, personal communication). As the molecular weight of insulin is 5.7 kDa (computed from Swiss-Prot entry P01317),

$0.4\ ng \cdot ml^{-1}$ corresponds to $0.07\ nM$. The same basal insulin concentration was assumed for rats. All other parameters were directly taken from the cited references.

The stationary model equations (all derivatives set to zero) were solved for the state variables at the constant basal insulin concentration to get initial conditions that correspond to the basal insulin concentration (Additional files 4 and 5 in [72], Table 2.1).

A.3.2 Dynamic model validation

Kruse et al. [83] used rats with a body weight of $238 \pm 20\ g$. Insulin injection was $100\ \mu l$ of $12 - 21\ pM$ radioactively labeled insulin. A body weight of $m_{body} = 238\ g$ and an insulin injection of $n_{*,in} = 100\ \mu l \cdot 16.5\ pM = 1.65 \cdot 10^{-6}\ nmol$ and $n_{in} = 0\ nmol$ were used for the simulation of this experiment. The experimental data set is given in % dose per ml serum. Multiplied by the amount of injected labeled insulin and divided by 100, these values give the concentrations of labeled insulin in plasma (unit: $nmol/ml$).

Desbuquois et al. [82] used rats with a body weight of $180 - 200\ g$. Insulin injection was $25\ nmol/100\ g$ body weight. A body weight of $m_{body} = 190\ g$ and an insulin injection of $n_{in} = 47.5\ nmol$ and $n_{*,in} = 0\ nmol$ were used for the simulation of this experiment. The experimental data set is already given as concentration values. As the procedure of injection is not exactly described in either study, it was assumed that the injection had been given as a bolus at $t = 0\ s$.

Backer et al. investigated insulin receptor internalization in Fao hepatoma cells [56]. Surface receptors were radioactively labeled at low temperature (on ice), stopping receptor internalization. Incubation at 37 °C initiated receptor internalization in the assay. This experiment was simulated by setting initial conditions such that all receptors are at the surface ($R = 40$). If the insulin concentration is constant, the receptor model is linear and the superposition principle holds. Therefore, the assay can be simulated with this choice of initial conditions.

A.4 Measuring insulin clearance

There are several methods available to measure insulin clearance [52]. This section shortly describes the two most important techniques.

A.4.1 Analysis of plasma insulin concentration after insulin administration

In many studies, exogenous insulin is administered as a bolus. The insulin concentration $Ins(t)$ is measured at as many time points a possible. Subsequently, a continuous curve is fitted to the data points and interpolated to zero (if necessary) and infinity. Insulin clearance c is determined as the quotient of the injected amount (dose) of insulin and the area under

the concentration curve which are referred to as ID and AUC, respectively [52].

$$c = \frac{ID}{AUC}\ , \qquad AUC = \int_0^\infty Ins(t)\ dt \tag{A.1}$$

Only in some special cases however, this method leads to results in accordance with the formal definition of insulin clearance as the quotient of insulin degradation rate and insulin concentration [59]. In addition, AUC is only defined if $Ins(t) \to 0$ for $t \to \infty$, which only holds if there is no endogenous insulin synthesis.

Assume that there is neither endogenous insulin synthesis nor basal insulin, and that insulin is injected as a bolus at $t = 0$. In this case, it holds that $Ins(0) = \frac{ID}{v_p}$ and $Ins(t) \to 0$ for $t \to \infty$. In addition, the insulin degradation rate $r(t)$ is the negative derivative of the insulin concentration. Assume that the time course of the insulin concentration is given by an exponential function

$$Ins(t) = \frac{ID}{v_p} \cdot e^{-k \cdot t} \tag{A.2}$$

implying linearity of insulin degradation.

$$r(t) = k \cdot Ins(t) = k \cdot \frac{ID}{v_p} \cdot e^{-k \cdot t} \tag{A.3}$$

Under these assumptions, the two definitions are identical.

$$c = \frac{r(t) \cdot v_p}{Ins(t)} = \frac{k \cdot ID \cdot e^{-k \cdot t} \cdot v_p}{ID \cdot e^{-k \cdot t}} = \frac{\int_0^\infty r(t)\ dt \cdot v_p}{\int_0^\infty Ins(t)\ dt} = \frac{ID}{AUC} \tag{A.4}$$

However, Equation A.4 does not hold in realistic scenarios. Assuming linearity of insulin degradation results in insulin clearance not depending on the insulin concentration, which is in contrast to a main result of Chapter 2. Note that linearity of insulin degradation is only a good approximation when analyzing a small concentration interval. However, for small amounts of injected insulin (ID), the neglected endogenous insulin becomes important.

Altogether, this method does not give reliable results for insulin clearance according to its formal definition [59].

A.4.2 Continuous insulin infusion: euglycemic insulin clamp

Another technique to measure insulin clearance is called euglycemic insulin clamp [52] and requires a continuous insulin infusion with a constant flux q_{Ins} (unit: $mol \cdot s^{-1}$). After a transient time, the insulin concentration in the plasma Ins (unit: $mol \cdot l^{-1}$) remains constant. In the resulting steady state, the rate of insulin degradation equals the rate of insulin infusion plus the rate of endogenous insulin synthesis. Neglecting endogenous insulin

synthesis, insulin clearance can be calculated according to its formal definition [59] as the quotient of insulin degradation rate (which equals q_{Ins}) and insulin concentration.

$$c = \frac{q_{Ins}}{Ins} \tag{A.5}$$

If large amounts of insulin are infused, the simultaneous administration of glucose is necessary to avoid a strong decrease in plasma glucose concentration. If q_{Ins} is not large compared to endogenously secreted insulin, it is necessary to correct Equation A.5 for endogenous insulin which can be assessed by the additional measurement of the C-peptide[1] concentration Cp [52]. According to [52], insulin clearance in the steady state is given as

$$c = \frac{q_{Ins}}{Ins - Ins_{basal} \cdot \frac{Cp}{Cp_{basal}}} \tag{A.6}$$

where an index *basal* indicates the basal value of the corresponding concentration and endogenous insulin is approximated as $Ins_{basal} \cdot \frac{Cp}{Cp_{basal}}$.

A.5 Approximation error $g(t)$

In this section, the ODE for the error $g(t)$ (Equation 4.29) is derived for the reaction cycle shown in Figure 4.7. The four reactions forming the reaction cycle are given as

$$\begin{array}{llcll} D[0,E] & + L & \leftrightharpoons & D[L,E] & k_1 \quad k_{-1} \\ D[0,p] & + L & \leftrightharpoons & D[L,p] & k_1 \quad k_{-1} \\ D[0,p] & + E & \leftrightharpoons & D[0,E] & k_4 \quad k_{-4} \\ D[L,p] & + E & \leftrightharpoons & D[L,E] & k_4 \quad k_{-4} \end{array} \tag{A.7}$$

and, according to the law of mass action (Section 1.2.3), proceed with the reaction rates

$$\begin{aligned} r_0 &= k_1 \cdot D[0,E] \cdot L - k_{-1} \cdot D[L,E] \\ r_1 &= k_1 \cdot D[0,p] \cdot L - k_{-1} \cdot D[L,p] \\ r_2 &= k_4 \cdot D[0,p] \cdot E - k_{-4} \cdot D[0,E] \\ r_3 &= k_4 \cdot D[L,p] \cdot E - k_{-4} \cdot D[L,E]. \end{aligned} \tag{A.8}$$

The balances of the concentrations of the species forming the reaction cycle are given as

$$\begin{aligned} \frac{d}{dt} D[0,p] &= -r_1 - r_2 + J_1 \\ \frac{d}{dt} D[0,E] &= -r_0 + r_2 + J_2 \\ \frac{d}{dt} D[L,p] &= r_1 - r_3 + J_3 \\ \frac{d}{dt} D[L,E] &= r_0 + r_3 + J_4 \end{aligned} \tag{A.9}$$

[1] Prior to the secretion of insulin, proinsulin is proteolytically processed to insulin (consisting of an A-chain and a B-chain) and C-peptide. The endogenous formation of one insulin molecule is therefore associated with the formation of one C-peptide [1].

where J_i are the in-fluxes into the reaction cycle (Figure 4.7). The diagonal reaction of this cycle is

$$D[0,p] + D[L,E] \leftrightharpoons D[L,p] + D[0,E] \tag{A.10}$$

(see Equation 4.21). In the next step, we assume rapid equilibrium (see Section 3.6.2) for this virtual reaction. As the processes of L binding and E binding do not interact directly and are assumed to be independent, we assume an equilibrium constant of one.

$$D[0,p] \cdot D[L,E] - D[L,p] \cdot D[0,E] = 0 \tag{A.11}$$

The error $g(t)$ is introduced as the distance of the diagonal reaction from equilibrium.

$$g(t) = D[0,p] \cdot D[L,E] - D[L,p] \cdot D[0,E] \tag{A.12}$$

Differentiating Equation A.12 leads to

$$\dot{g}(t) = D[0,p] \cdot \frac{d}{dt} D[L,E] + D[L,E] \cdot \frac{d}{dt} D[0,p] - D[L,p] \cdot \frac{d}{dt} D[0,E] - D[0,E] \cdot \frac{d}{dt} D[L,p] \tag{A.13}$$

which together with Equations A.8 and A.9 results in

$$\begin{aligned} \dot{g}(t) =& J_1 \cdot D[L,E] - J_2 \cdot D[L,p] + D[0,p](J_4 - D[L,E](k_{-1} + k_{-4} + k_4 \cdot E + k_1 \cdot L)) \\ &+ D[0,E](-J_3 + D[L,p](k_{-1} + k_{-4} + k_4 \cdot E + k_1 \cdot L)) \\ =& -(k_{-1} + k_{-4} + k_4 \cdot E + k_1 \cdot L)(D[0,p] \cdot D[L,E] - D[L,p] \cdot D[0,E]) \\ &+ J_1 \cdot D[L,E] - J_2 \cdot D[L,p] - J_3 \cdot D[0,E] + J_4 \cdot D[0,p]. \end{aligned} \tag{A.14}$$

If one defines

$$\begin{aligned} a(t) &= k_{-1} + k_{-4} + k_4 \cdot E + k_1 \cdot L \\ u(t) &= J_1 \cdot D[L,E] - J_2 \cdot D[L,p] - J_3 \cdot D[0,E] + J_4 \cdot D[0,p] \end{aligned} \tag{A.15}$$

Equation A.14 simplifies to Equation 4.29 which is given as

$$\dot{g} = -a(t) \cdot g + u(t). \tag{A.16}$$

A.6 Layer-based model of the extended example system

According to the interaction graph (Figure 4.2 B), the layer-based model of the extended example system (Section 4.6.1) consists of three layers which can be modeled independently from each other once their connections are defined. The reaction network within the layers is shown in Figure 4.8, parameters and initial conditions are given in Table A.1.

A.6.1 The receptor layer

The receptor layer describes the processes of ligand binding and receptor phosphorylation. The receptor `R{0,L}{0,p}{0,P}` has three sites. The first site is a binding site for the ligand L. The second site is a regulatory phosphorylation site whose phosphorylation negatively affects autophosphorylation of the third site, which is a binding site for the effector E.

$$
\begin{aligned}
r_1 &= k_1 \cdot L \cdot R[0,0,0] - k_{-1} \cdot R[L,0,0] \\
r_2 &= k_1 \cdot L \cdot R[0,0,P] - k_{-1} \cdot R[L,0,P] \\
r_3 &= k_1 \cdot L \cdot R[0,p,0] - k_{-1} \cdot R[L,p,0] \\
r_4 &= k_1 \cdot L \cdot R[0,p,P] - k_{-1} \cdot R[L,p,P] \\
r_5 &= k_2 \cdot R[0,0,0] - k_{-2} \cdot R[0,p,0] \\
r_6 &= k_2 \cdot R[0,0,P] - k_{-2} \cdot R[0,p,P] \\
r_7 &= k_3 \cdot R[L,0,0] - k_{-3} \cdot R[L,p,0] \\
r_8 &= k_3 \cdot R[L,0,P] - k_{-3} \cdot R[L,p,P] \\
r_9 &= k_2 \cdot R[0,0,0] - k_{-2} \cdot \frac{x - xb}{x} \cdot R[0,0,P] \\
r_{10} &= k_4 \cdot R[0,p,0] - k_{-4} \cdot \frac{x - xb}{x} \cdot R[0,p,P] \\
r_{11} &= k_3 \cdot R[L,0,0] - k_{-3} \cdot \frac{x - xb}{x} \cdot R[L,0,P] \\
r_{12} &= k_5 \cdot R[L,p,0] - k_{-5} \cdot \frac{x - xb}{x} \cdot R[L,p,P]
\end{aligned}
\tag{A.17}
$$

The signal x represents the sum of all phosphorylated binding sites on the receptor.

$$x = R[X,X,P] \tag{A.18}$$

Note that x represents both occupied and unoccupied phosphorylated binding sites. The correction term $(x - xb)/x$ represents the fraction of phosphorylated binding sites that is not occupied. Using the same correction term in all dephosphorylation rates of the binding site corresponds to the assumption that the same fraction of phosphorylated binding sites is occupied in all species with a site modification (uppercase) 'P'.

$$\begin{aligned}
\frac{d}{dt}R[0,0,0] &= -r_1 - r_5 - r_9 \\
\frac{d}{dt}R[0,0,P] &= -r_2 - r_6 + r_9 \\
\frac{d}{dt}R[0,p,0] &= -r_3 + r_5 - r_{10} \\
\frac{d}{dt}R[0,p,P] &= -r_4 + r_6 + r_{10} \\
\frac{d}{dt}R[L,0,0] &= r_1 - r_7 - r_{11} \\
\frac{d}{dt}R[L,0,P] &= r_2 - r_8 + r_{11} \\
\frac{d}{dt}R[L,p,0] &= r_3 + r_7 - r_{12} \\
\frac{d}{dt}R[L,p,P] &= r_4 + r_8 + r_{12}
\end{aligned} \tag{A.19}$$

R_{active} is an additional signal representing receptor activity transferred to the E layer, where the activated receptor mediates the phosphorylation of E.

$$R_{active} = R[X,p,X] \tag{A.20}$$

The signal R_{active} fulfills the conditions for additional signals between layers given in Section 4.4.3 as it introduces no graded interaction between processes of different layers.

A.6.2 The E layer

The E layer describes the processes of E binding to the receptor and E phosphorylation. Receptor-bound E is defined by `RXE{0,P}`, E not bound to the receptor is defined by `E{0,P}`. The binding partner RXp represents the sum of all microscopic receptor species with phosphorylated and unoccupied binding site for E.

$$\begin{aligned}
r_{13} &= k_6 \cdot E[0] \cdot RXp - k_{-6} \cdot RXE[0] \\
r_{14} &= k_7 \cdot R_{active} \cdot RXE[0] - k_{-7} \cdot \frac{x2 - x2b}{x2} \cdot RXE[P] \\
r_{15} &= k_6 \cdot E[P] \cdot RXp - k_{-6} \cdot RXE[P] \\
r_{16} &= k_{-7} \cdot \frac{x2 - x2b}{x2} \cdot E[P]
\end{aligned} \tag{A.21}$$

The signal $x2$ represents the sum of all species of E phosphorylated on the binding site for F.

$$x2 = RXE[P] + E[P] \tag{A.22}$$

Using the correction term $(x2-x2b)/x2$ in the dephosphorylation rates of $RXE[P]$ and $E[P]$ (Equation A.21) corresponds to the assumption that the same fractions of phosphorylated binding sites are occupied in both receptor-bound E and E not bound to the receptor.

Table A.1: Kinetic parameters and initial conditions for the extended example system Initial conditions were 40 nM for $R[0,0,0]$, 250 nM for $E[0]$ and 50 nM for F. All other initial conditions were about 10^{-20} nM and determined according to the transformation equations [105]. L is set to 100 nM, a typical insulin concentration for *in vitro* experiments. Autophosphorylation of the receptor with bound ligand and without regulatory phosphorylation is parameterized by k_3. Autophosphorylation in the absence of ligand and regulatory phosphorylation is parameterized by $k_2 = k_3 \cdot f_{ins}$. Autophosphorylation without ligand but with regulatory phosphorylation is parameterized by $k_4 = k_3 \cdot f_{ins} \cdot f_{reg}$. Autophosphorylation with bound ligand and regulatory phosphorylation is parameterized by $k_5 = k_3 \cdot f_{reg}$. It is assumed that dephosphorylation of the receptor does not depend on other receptor modifications: $k_{-2} = k_{-3} = k_{-4} = k_{-5}$. It is further assumed that E and F bind with the same kinetic constants to their respective binding sites: $k_6 = k_8$ and $k_{-6} = k_{-8}$.

Parameter	Literature value	Unit	Source
k_1	0.001	$nM^{-1}s^{-1}$	[58]
k_{-1}	$4 \cdot 10^{-4}$	s^{-1}	[58]
k_3	0.0231	s^{-1}	[77]
k_{-3}	0.00385	s^{-1}	[76]
k_6	0.033	$nM^{-1}s^{-1}$	[110]
k_{-6}	0.113	s^{-1}	[110]
k_7	10^{-5}	$nM^{-1}s^{-1}$	[124]
k_{-7}	0.000385	s^{-1}	assumption
f_{ins}	0.1	-	assumption
f_{reg}	0.1	-	assumption

The signal xb represents the sum of all species of E bound to the receptor.

$$\begin{aligned} xb &= RXE[X] \\ RXp &= x - xb \end{aligned} \tag{A.23}$$

Note that the letter 'X' in the molecule names of RXE and RXp represents two sites, namely the ligand binding site and the regulatory phosphorylation site. $RXE[0]$, $RXE[P]$ and RXp are not macrostates. The 'X' in their notations is not a site configuration but a part of the molecule name. $RXE[X]$ however is a macrostate (see Section 4.2.1).

$$\begin{aligned} \frac{d}{dt}RXE[0] &= r_{13} - r_{14} \\ \frac{d}{dt}RXE[P] &= r_{14} + r_{15} \\ \frac{d}{dt}E[0] &= -r_{13} + r_{16} \\ \frac{d}{dt}E[P] &= -r_{15} - r_{16} \end{aligned} \tag{A.24}$$

A.6.3 The F layer

The F layer describes the process of F binding. The binding partner XEp represents the sum of all microscopic species of E with phosphorylated and unoccupied binding site for F. XEF represents the sum of all microscopic species of F bound to E.

$$r_{17} = k_8 \cdot XEp \cdot F - k_{-8} \cdot XEF \tag{A.25}$$

The signal $x2b$ represents the sum of all species of F bound to E.

$$\begin{aligned} x2b &= XEF \\ XEp &= x2 - x2b \end{aligned} \tag{A.26}$$

The letter 'X' in the notations of XEp and XEF indicates that E may be bound to the receptor or not, without considering other sites of the receptor. XEp and XEF are not macrostates as the 'X' in their notations is a part of the molecule name.

$$\begin{aligned} \frac{d}{dt}XEF &= r_{17} \\ \frac{d}{dt}F &= -r_{17} \end{aligned} \tag{A.27}$$

A.6.4 Conservation relations

Three ODEs can be replaced by conservation relations for R, E and F.

$$\begin{aligned} R_{total} &= R[X, X, X] \\ E_{total} &= E[X] + RXE[X] \\ F_{total} &= F + XEF \end{aligned} \tag{A.28}$$

A.7 Command line parameters for ALC

Section 5.4.2 showed how ALC can be run in a shell with default values. As a short repetition, the commands '`ALC.pl`' or '`perl ALC.pl`' execute ALC with default values, e.g. for undefined parameters or initial conditions and for the filenames of the model definition file and the model files.

The default values can be changed via the command line or in the form on the ALC website for each call of ALC separately. Changing the default values via the command line is done by appending a sequence of assignments to the call of `ALC.pl`. This sequence may contain assignments for the parameter default value (`Param`), the initial condition default value (`InCond`), the path and the filename of the model definition file (`Source`), the path and the start of the filenames of the model files (`Target`), and the simulation time for the model files (`SimTime`). The visualization of uncritical warnings (not defined initial conditions or parameters and balanced species without turnover, `Warn`), model equations (`Show`), and the time consumption of the modeling steps (`Time`) in the shell can be enabled or suppressed. Additionally, it can be decided if the simulation results for the variables defined in the section `#layer connections` are automatically visualized in the model files (`OutLC`), and if very strict consistency checks are performed on rules (`StrictRS`). As an example, the command

```
perl ALC.pl InCond="0.5*10^(-3)" Source="input.alc" Target="output"
```

calls ALC and sets all undefined initial conditions to $0.5 \cdot 10^{-3}$. In this example, the model definition is read in from the file 'input.alc', and the notation of the model files starts with 'output'.

The order of the assignments is arbitrary. Default values defined in the file `Config_ALC.txt` are taken for all omitted assignments (Table A.2). If this file does not exist (or is not changed), the values given in Table A.2 are taken. Changing the default values in the file `Config_ALC.txt` is an efficient possibility if the same default values are taken for several models.

All default values can also be easily set in the form on the ALC website.

Table A.2: ALC: command line parameters and default values
The default values defined in the file `Config_ALC.txt` are taken for all options not defined in the command line. The inverted commas encapsulating the values of the options can be omitted if there are no special characters (e.g. '^') inside.

Option	Description	Default	Example
`Source`	Name of the model definition file[1]	layer.alc	`Source=input.alc`
`InCond`	Value for undefined initial conditions[2]	0.001	`InCond="3.1*10^2"`
`Param`	Value for undefined parameters	1	`Param="2.5*10^3"`
`SimTime`	Simulation time for the model files	100	`SimTime=10`
`Target`	Start of the names of the model files[3]	model	`Target=output`
`OutLC`	Add (1) or do not add (0) variables defined in the section `#layer connections` to the output list	1	`OutLC=0`
`Warn`	Show (1) or do not show (0) warnings for undefined initial conditions and parameters and for balanced species without turnover	1	`Warn=0`
`Show`	Show (1) or do not show (0) model equations	0	`Show=1`
`Time`	Show (1) or do not show (0) the time consumption of the modeling steps	0	`Time=1`
`StrictRS`	Perform (1) or do not perform (0) very strict consistency checks on rules[4]	1	`StrictRS=0`

[1] A relative path can also be given. `Source="test/layer.alc"` or `Source="test\layer.alc"` for example direct ALC to use 'layer.alc' in the directory 'test' as the model definition file. Use your system-specific notation for the relative path[5].

[2] Setting `InCond=0` may result in division by zero when simulating the model. Initial conditions have to guarantee $x_i > 0 \; \forall \; i$ to avoid this problem. Initial conditions also have to guarantee $x_i \geq x_i b \; \forall \; i$ to prevent negative concentrations of binding partners.

[3] A relative path can also be given, e.g. `Target="test/model"` or `Target="test\model"` where 'test' is a directory and 'model' is the start of the names of the model files. Use your system-specific notation for the relative path[5]. The directories in the path have to exist, ALC will not create them.

[4] `StrictRS=1` minimizes the probability of errors in the definition of rules. For `StrictRS=0`, only the minimal syntax requirements that are necessary to successfully finish the modeling procedure are checked. `StrictRS=0` is necessary when using functionalities provided by the section `#remove` (Section 5.2.2.4).

[5] Assigning the relative path will not always work if the filename or the names of the directories include special characters. If the actual directory is desired as the relative path, give the filename (`Source`) or the start of the filenames (`Target`) directly.

A.8 The 51 ODE model of insulin signaling

This section provides a model definition for a layer-based model of insulin signaling approximating a conventional model with $1.5 \cdot 10^8$ ODEs (see Section 4.6.2). The advanced strategy of equivalent binding sites (Section 4.4.1) is applied in the receptor layer and the IRS layer. The binding sites for IRS and Shc on the receptor represent two equivalent sites each and the binding site for PI3K on IRS represents four equivalent sites.

Note that the symbol '%' labels the remainder of the line as a comment which is ignored by ALC.

```
%## receptor layer #######################################

#molecules
R{0,I}{0,I}{0,P}{0,P} % 2 binding sites for insulin, one for IRS, one for Shc
Ins
#end molecules

#initial conditions
R[0,0,0,0]=40
#end initial conditions

#clamped concentrations
Ins=100 % Input of the system
#end clamped concentrations

#reactions
% Insulin binding:
R[0,0,X,X]+Ins<->R[I,0,X,X]   ki1   ki1d
R[0,0,X,X]+Ins<->R[0,I,X,X]   ki1   ki1d
R[I,0,X,X]+Ins<->R[I,I,X,X]   ki2   ki2d
R[0,I,X,X]+Ins<->R[I,I,X,X]   ki2   ki2d
% Phosphorylation of the binding site for IRS:
R[0,0,0,X]<->R[0,0,P,X]   kp10   kp10d
R[I,0,0,X]<->R[I,0,P,X]   kp11   kp11d
R[0,I,0,X]<->R[0,I,P,X]   kp11   kp11d
R[I,I,0,X]<->R[I,I,P,X]   kp12   kp12d
% Phosphorylation of the binding site for Shc:
R[0,0,X,0]<->R[0,0,X,P]   kp20   kp20d
R[I,0,X,0]<->R[I,0,X,P]   kp21   kp21d
R[0,I,X,0]<->R[0,I,X,P]   kp21   kp21d
R[I,I,X,0]<->R[I,I,X,P]   kp22   kp22d
#end reactions

#layer connections
% Each phosphorylation site on the receptor represents two equivalent sites
xRIrs=2*R[X,X,P,X]
xRShc=2*R[X,X,X,P]
#end layer connections
```

```
%## IRS layer #######################################

#molecules
IRS{0,R}{0,P}{0,P}{0,P} % Binding sites for the receptor, PI3K, SHP2 and Grb2
#end molecules

#initial conditions
IRS[0,0,0,0]=250
#end initial conditions

#reactions
% Binding of IRS to the receptor:
RXpIrs+IRS[0,X,X,X]<->IRS[R,X,X,X]   kbirs   kbirsd
% Phosphorylation of unbound IRS:
IRS[0,0,X,X]<->IRS[0,P,X,X]   kirsp1      kirsp1d
IRS[0,X,0,X]<->IRS[0,X,P,X]   kirsp2      kirsp2d
IRS[0,X,X,0]<->IRS[0,X,X,P]   kirsp3      kirsp3d
% Phosphorylation of receptor-bound IRS:
IRS[R,0,X,X]<->IRS[R,P,X,X]   kirsbp1      kirsbp1d
IRS[R,X,0,X]<->IRS[R,X,P,X]   kirsbp2      kirsbp2d
IRS[R,X,X,0]<->IRS[R,X,X,P]   kirsbp3      kirsbp3d
#end reactions

#layer connections
xRIrsb=IRS[R,X,X,X]
RXpIrs=xRIrs-xRIrsb
% The binding site for PI3K represents four equivalent sites
xIrsPi3k=4*IRS[X,P,X,X]
xIrsShp2=IRS[X,X,P,X]
xIrsGrb2=IRS[X,X,X,P]
#end layer connections

%## Shc layer ###################################

#molecules
Shc{0,R}{0,P} % Binding sites for IRS and Grb2
#end molecules

#initial conditions
Shc[0,0]=250
#end initial conditions

#reactions
% Binding of Shc to the receptor:
RXpShc+Shc[0,X]<->Shc[R,X]   kbshc   kbshcd
% Phosphorylation of Shc:
Shc[0,0]<->Shc[0,P]      kshcp    kshcpd
Shc[R,0]<->Shc[R,P]      kshcbp   kshcbpd
#end reactions
```

```
#layer connections
xRShcb=Shc[R,X]
RXpShc=xRShc-xRShcb
xShc=Shc[X,P]
#end layer connections

%## Grb2 layer ###################################

#molecules
Grb2{0,Shc,IRS}{0,SOS,SOSp} % Binding site for IRS and Shc, binding site for SOS
SOS{0,p} % Regulatory phosphorylation site
#end molecules

#initial conditions
Grb2[0,0]=40
SOS[0]=40
#end initial conditions

#reactions
% Grb2 binding to IRS:
IrsXpGrb2+Grb2[0,0]   <->Grb2[IRS,0]      kbirsgrb2       kbirsgrb2d
IrsXpGrb2+Grb2[0,SOS] <->Grb2[IRS,SOS]    kbirsgrb2sos    kbirsgrb2sosd
IrsXpGrb2+Grb2[0,SOSp]<->Grb2[IRS,SOSp]   kbirsgrb2sosp   kbirsgrb2sospd
% SOS binding to Grb2:
Grb2[X,0]+SOS[0]<->Grb2[X,SOS]    kbsos    kbsosd
Grb2[X,0]+SOS[p]<->Grb2[X,SOSp]   kbsosp   kbsosdp
% Grb2 binding to Shc:
ShcXp+Grb2[0,0]   <->Grb2[Shc,0]      kbshcgrb2       kbshcgrb2d
ShcXp+Grb2[0,SOS] <->Grb2[Shc,SOS]    kbshcgrb2sos    kbshcgrb2sosd
ShcXp+Grb2[0,SOSp]<->Grb2[Shc,SOSp]   kbshcgrbsosp2   kbshcgrb2sospd
% Phosphorylation of SOS bound to Grb2:
Grb2[0,SOS]  <->Grb2[0,SOSp]     kpgrb2sos      kpgrb2sosd
Grb2[IRS,SOS]<->Grb2[IRS,SOSp]   kpirsgrb2sos   kpirsgrb2sosd
Grb2[Shc,SOS]<->Grb2[Shc,SOSp]   kpshcgrb2sos   kpshcgrb2sosd
% Phosphorylation of free SOS:
SOS[0]<->SOS[p]   kpsos   kpsosd
#end reactions

#layer connections
xIrsGrb2b=Grb2[IRS,X]
IrsXpGrb2=xIrsGrb2-xIrsGrb2b
xShcb=Grb2[Shc,X]
ShcXp=xShc-xShcb
#end layer connections
```

```
%## PI3K layer ###################################

#molecules
PI3K{0,IRS} % Binding site for IRS
#end molecules

#initial conditions
PI3K[0]=40
#end initial conditions

#reactions
% Binding to IRS:
IrsXpPi3k+PI3K[0]<->PI3K[IRS]   kbpi3k   kbpi3kd
#end reactions

#layer connections
xIrsPi3kb=PI3K[IRS]
IrsXpPi3k=xIrsPi3k-xIrsPi3kb
#end layer connections

%## SHP2 layer ###################################

#molecules
SHP2{0,IRS} % Binding site for IRS
#end molecules

#initial conditions
SHP2[0]=40
#end initial conditions

#reactions
% Binding to IRS:
IrsXpShp2+SHP2[0]<->SHP2[IRS]   kbshp2   kbshp2d
#end reactions

#layer connections
xIrsShp2b=SHP2[IRS]
IrsXpShp2=xIrsShp2-xIrsShp2b
#end layer connections
```

List of Tables

2.1 Model parameters and initial conditions 34

4.1 Kinetic parameters and initial conditions for the small example system . . . 83
4.2 Number of ODEs necessary for models of the insulin signaling system 106

A.1 Kinetic parameters and initial conditions for the extended example system . 140
A.2 ALC: command line parameters and default values 143

List of Figures

1.1 Combinatorial complexity in the insulin signaling system 21

2.1 Insulin receptor activation in hepatocytes 32
2.2 Dynamic model validation: physiological insulin concentrations 38
2.3 Dynamic model validation: extremely high insulin concentrations 39
2.4 Dynamic model validation: receptor internalization 40
2.5 Stationary model validation: insulin binding and receptor phosphorylation . 41
2.6 Stationary analysis of renal and hepatic insulin degradation 44
2.7 Stationary analysis of renal and hepatic insulin clearance 45
2.8 Sensitivity of simulation results to changes in $k1ub$ and $k2ub$ 46
2.9 Sensitivity of simulation results to changes in $Kkidney$ 47
2.10 Sensitivity of simulation results to changes in t_{in} 49

3.1 Stationary approximation error of intuitive model reduction 55
3.2 Stationary approximation error for parameter values from literature 56

4.1 Types of interactions between processes 73
4.2 Interaction graphs of two example systems 75
4.3 Conventional reaction network of the small example system 79
4.4 Layer-based reaction network of the small example system 81
4.5 Simulation results for the small example system 84
4.6 Reaction cycle and diagonal reaction 92
4.7 Typical reaction cycle with in-fluxes 96
4.8 Layer-based reaction network of the extended example system 102
4.9 Simulation results for the extended example system 103
4.10 Interaction graph of insulin signaling 105

5.1 Flow diagram of ALC 111
5.2 Model definition for the small example system 113