Systems biological analyses
of intracellular signal transduction

D i s s e r t a t i o n
zur Erlangung des akademischen Grades
d o c t o r r e r u m n a t u r a l i u m
( Dr. rer. nat.)
im Fach Biophysik
ingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät I
der Humboldt-Universität zu Berlin

von

Herrn Diplom-Biochemiker Stefan Legewie
geboren am 31.10.1977 in Aachen

Präsident/Präsidentin der Humboldt-Universität zu Berlin
Prof. Dr. Christoph Markschies

Dekan/Dekanin der Mathematisch-Naturwissenschaftlichen Fakultät I
Prof. Dr. Christian Limberg

Gutachter: Prof. Dr. Hanspeter Herzel
Prof. Dr. Jens Timmer
Prof. Dr. Olaf Wolkenhauer

Tag der mündlichen Prüfung: 31. 10. 2008
für meine Eltern
Summary

Intracellular regulatory networks involved in the sensing of extracellular cues are crucial to all living organisms. Signal transduction networks allow unicellular organisms sensing nutrient availability, finding mating partners and responding to stress. Moreover, intercellular communication is the fundamental basis for the functioning and homeostasis of multicellular organisms. Accordingly, many diseases including cancer are caused by deregulation of signal transduction networks. Extracellular signals are typically transmitted rapidly from the cell membrane to the nucleus by activation of multi-level enzymatic cascades which ultimately elicit slow changes in gene expression, and thereby affect the cell fate. These signalling cascades are highly interconnected, thus giving rise to complex networks, that are hard to understand intuitively. In this thesis, a combination of kinetic modeling and analysis of quantitative experiments is applied to get insights into the principles of intracellular signalling.

In the first part, the dynamics of enzymatic signalling cascades involved in transducing signals from the cell membrane to the nucleus are investigated. The Ras-MAPK cascade plays a central role in various physiological processes such as cell cycle progression, cell differentiation and cell death, and mutational cascade activation appears to be crucial for cancer development. Overexpression of wildtype Ras is also frequently observed in tumours, but its functional relevance remains unclear. By analysis of an experimentally validated kinetic model of Ras signalling, it is shown in Chapter 2 that the basal state MAPK signalling can be completely insensitive towards overexpression of the uppermost cascade member Ras. Thus, the simulations reveal a “kinetic tumour suppression effect” inherent to the Ras (de)activation cycle, and also explain experimental studies showing that overexpression events within the MAPK cascade, though phenotypically silent in isolation, frequently cooperate to bring about strong cellular deregulation (“oncogene cooperation”). In Chapter 3, it is analysed how an experimentally validated MAPK cascade model responds to more physiological, transient inputs and converts them into an all-or-none, irreversible cell fate decision. More specifically, it is shown that bistability arises in the core MAPK cascade by a previously unrecognised enzyme sequestration effect that establishes a hidden positive feedback loop. Chapter 4 is focussed on the proteolytic caspase cascades controlling apoptosis, a form of cell suicide activated in response to extracellular stress. The simulations suggest an unanticipated role for inhibitors of apoptosis proteins (IAPs): Simultaneous inhibition of multiple caspases by IAPs can result in strong positive feedback regulation, and may thus be essential to establish all-or-none and irreversible initiation of cell death.

Cellular commitment to a new fate typically requires ongoing extracellular stimulation and/or intracellular signalling for several hours, so that the long-term dynamics of signalling cascades are important for cellular responses. In the second part of the thesis, it is investigated how slow signal-induced changes in gene expression feed back into the signalling network and modulate its dynamical activation pattern. In Chapter 5, the general design principles underlying transcriptional feedback regulation of mammalian signalling pathways are investigated by analysing the stimulus-induced gene expression profiles of 134 intracellular signalling proteins. It turns out that transcriptional feedback regulation occurs in each of the five signalling cascades considered, and that negative feedback strongly dominates over positive feedback. Moreover, negative feedback exclusively occurs by transcriptional induction of a subgroup of signal inhibitors, termed rapid feedback inhibitors (RFIs), while downregulation of signal transducers plays no role. Systematic analysis of mRNA and protein half-lives reveals a remarkable separation of the signalling network into flexible and static parts: transcriptionally regulated RFIs are unstable at the mRNA and protein level, while other signalling proteins are generally stable. Kinetic modelling, also presented in Chapter 5, is employed to get insights into the functional implications of RFI-mediated transcriptional feedback regulation. In Chapter 6, transcriptional feedback regulation of TGFβ signalling via Smad transcription factors is analysed in more detail in primary hepatocytes to confirm the physiological relevance of transcriptional feedback.
regulation at the protein level. The TGFβ family of cytokines constitute major inhibitors of cell growth, and accordingly they play important roles in various physiological and pathological processes including development, tissue homeostasis, tissue regeneration, and cancer. Genome-wide microarray analyses and protein measurements in response to TGFβ stimulation (presented in Chapter 6) suggest that the SnoN oncoprotein is the central transcriptional feedback regulator in primary mouse hepatocytes. A mathematical model including TGFβ-induced Smad signalling and SnoN-mediated feedback is fitted to experimental data obtained under various stimulation conditions, and predictions derived from the model are then quantitatively confirmed in primary hepatocytes isolated from SnoN knock-out mice. The modelling results in Chapter 6 mechanistically explain how a small pool of SnoN proteins can efficiently regulate a much larger pool of Smad proteins, and further support the relevance of transcriptional negative feedback regulation in signal transduction.

Cells face a specificity problem as different extracellular stimuli frequently engage the same set of intracellular signalling pathways even though they elicit completely different biological responses. Experimental evidence suggests that stimulus-specific biological information is frequently encoded in the quantitative aspects of stimulus-specific activation kinetics (e.g., signal amplitude and/or duration). If biological information is encoded in the quantitative characteristics of intracellular signals, proper cell fate decisions require that the downstream gene expression machinery is able to accurately decode signal amplitude and/or duration. Part III of this thesis deals with such decoding of upstream signals by the gene expression machinery, and thus represents a first step towards more integrated systems biological models that include both, upstream signal transduction and downstream phenotypic responses such as cell growth. The results presented in Section 7 identify competitive inhibition and regulated degradation as mechanisms that allow intracellular regulatory networks to efficiently discriminate transient vs. sustained signals. More specifically, a combination of mathematical modelling and quantitative experimental analyses reveals that a recently discovered small non-coding RNA, IsrR, establishes a pronounced delay and duration decoding in the cyanobacterial gene expression response towards iron stress. In other words, it is shown that the small non-coding RNA, IsrR, restricts the potentially harmful and costly expression of late-phase stress proteins to severe, prolonged and ongoing stress conditions. Many of the downstream target genes induced by signalling pathways are transcription factors, thus giving rise to a complex transcriptional regulatory network. Therefore, signal decoding at the level of gene expression cannot be fully understood by insights into the functioning of small transcriptional regulatory motifs, but additionally requires integrated analyses of multiple transcription factors. In Chapter 8, a recently proposed reverse engineering approach, called modular response analysis (MRA), is applied to derive the topology of an oncogenic transcription factor network from high-throughput and knock-down data. Statistical analyses of the MRA results are used to derive predictions that can be verified experimentally, and also identify a key transcription factor cascade whose existence is supported by the published literature.

In conclusion, this thesis shows how systems biological analyses can enhance our understanding of intracellular signalling networks. The approaches presented here include quantitative analyses of small regulatory motifs (Chapters 2, 3, 4 and 7), systematic investigation of high-throughput data (Chapters 5 and 6), kinetic modelling and thus integration of multiple time course experiments (Chapter 6), as well as reverse engineering of regulatory network topologies (Chapter 8).
Zusammenfassung


In den meisten Fällen muss eine Zelle über mehrere Stunden kontinuierlich stimuliert werden, um auf eine entsprechende phänotypische Antwort festgelegt zu sein. Also ist die Langzeit-Dynamik intrazellulärer Signalnetzwerke über mehrere Stunden entscheidend für zelluläre Antworten. Im zweiten Teil der Arbeit wird untersucht, wie langsame signalinduzierte Änderungen der Genexpression in das Signalnetzwerk rückkoppeln und somit die Langzeit-Dynamik beeinflussen. In Kapitel 5 werden durch eine Analyse signalinduzierter Genexpressionsmuster generelle Design-Prinzipien der transkriptionellen Rückkopplung in Säugetier-Signalkaskaden identifiziert. Die Untersuchung zeigt, dass transkriptionelle Rückkopplung in allen fünf untersuchten Signalwegen auftritt, und dass transkriptionelle Regulation nahezu ausschließlich negative Rückkopplung bewirkt. Des Weiteren weist die Analyse darauf hin, dass negative Rückkopplung einzig und allein durch die Induktion von hemmenden Faktoren ("Inhibitoren") geschieht, während die Repression


Zusammenfassend zeigt diese Arbeit, wie systembiologische Ansätze das Verständnis intrazellulärer Signalnetzwerke vertiefen können. Die in dieser Arbeit angewendeten Ansätze beinhalten quantitative Analysen kleiner regulatorischer Motive (Kapitel 2, 3, 4 und 7), die systematische Untersuchung von Hochdurchsatzdaten (Kapitel 5 und 6), die integrative kinetische Modellierung vieler Zeitreihen (Kapitel 6) und das Ableiten von Netzwerk-Topologien aus experimentellen Daten ("Reverse Engineering"; Kapitel 8).
# CONTENTS

Summary ............................................................................................................................................. v

Zusammenfassung ............................................................................................................................... vii

1 General introduction ......................................................................................................................... 1
   1.1 Signalling from the cell membrane to the nucleus ................................................................. 1
   1.2 Time scales of signalling and transcriptional feedback regulation ...................................... 5
   1.3 The quantitative model of signal specificity and decoding at the level of gene expression .... 8

I Signalling via fast post-translational regulatory mechanisms

2 Kinetic mechanisms for tumour suppression and oncogene cooperation ................................ 11
   2.1 Introduction ............................................................................................................................ 11
   2.2 Basal Ras signalling activity can be insensitive towards Ras overexpression ................... 13
   2.3 Requirements for Kinetic Tumour Suppression ................................................................ 14
   2.4 Implications for oncogene cooperation ............................................................................... 16
   2.5 Conclusions .......................................................................................................................... 17

3 Competing docking interactions can bring about bistability in the MAPK cascade ................. 21
   3.1 Introduction ............................................................................................................................ 21
   3.2 Rationale and model implementation .................................................................................. 23
   3.3 Structural and kinetic requirements for bistability ............................................................... 25
   3.4 Synergism of Bistability Mechanisms .................................................................................. 30
   3.5 Conclusions .......................................................................................................................... 31

4 Simultaneous inhibition of initiator and effector caspases by XIAP establishes implicit bistability in caspase activation ................................................................. 33
   4.1 Introduction ............................................................................................................................ 33
   4.2 Model derivation ..................................................................................................................... 34
   4.3 Dynamic and steady state characteristics of caspase activation ........................................ 38
   4.4 Mechanism of bistability and kinetic requirements ............................................................. 39
   4.5 The mitochondrial pathway acts as an efficient integrator of multiple regulatory inputs ....... 43
   4.6 Conclusions .......................................................................................................................... 45
II Transcriptional feedback regulation of signalling cascades

5 Recurrent design patterns in the feedback regulation of the mammalian signalling network

5.1 Introduction
5.2 Expression and half-life data reveal design principles of transcriptional feedback
5.3 Kinetic modelling of rapid feedback inhibitor (RFI) action
5.4 Conclusions

6 Transcriptional feedback regulation of TGFβ signalling

6.1 Introduction
6.2 mRNA expression profiling reveals transcriptional feedback
6.3 Experimental verification of transcriptional feedback at the protein level
6.4 Implementation of a mathematical model
6.4.1 Rationale
6.4.2 Modelling receptor trafficking
6.4.3 Modelling the Smad activation cycle
6.4.4 Modelling negative feedback via SnoN induction
6.5 Model calibration
6.5.1 Experimental data for model calibration
6.5.2 Scaling factors and error estimation
6.5.3 Parameter estimation
6.5.4 Input functions
6.5.5 Results of model calibration
6.6 Testing the model predictions
6.6.1 Prediction of Smad signalling dynamics in SnoN-depleted cells
6.6.2 Analysis of SnoN-mediated regulation of gene expression
6.7 Conclusions and outlook

III Regulation at the level of gene expression

7 Small RNAs Establish Delays and Temporal Thresholds in Gene Expression

7.1 Introduction
7.2 Model Implementation
7.3 Theoretical Analysis of Steady State and Dynamical Behaviour
7.4 Experimental Verification of the Simulated Dynamic Behaviour
7.5 Pulse Filtering Properties of sRNA Circuits
7.6 Conclusions
1 General introduction

Parts of this chapter will be published as a review in a forthcoming issue of Essays in Biochemistry. Parts of Section 1.2 are published as a supplement in [1].

1.1 Signalling from the cell membrane to the nucleus

The intracellular signalling network transduces extracellular signals (e.g., growth factor stimulation) from the cell membrane to the nucleus, and ultimately mediates changes in gene expression to induce cellular responses. Information is typically transmitted through multi-layered enzymatic cascades, and various signalling mechanisms including protein-protein interactions, covalent modification and nucleo-cytoplasmic shuttling are employed. The three-tiered mitogen-activated protein kinase (MAPK) signal transduction cascades are probably the best-characterised signalling pathways in eukaryotic cells. Active MAP kinases control various physiological processes and cell fate decisions in all eukaryotic cells, mainly by regulating gene expression.

The classical, most studied Ras-MAPK cascade in higher mammalian cells is the pathway comprising the small G-protein Ras, and the serine/threonine kinases Raf, MEK and ERK (Fig. 1.1A; reviewed in [2]). The activity of Ras is controlled by guanine nucleotide binding, while downstream kinase activity is regulated by (de)phosphorylation. Cascade activation is typically initiated by hormone binding to cognate transmembrane receptors. The resulting active receptor-ligand complexes employ various mechanisms to convert the small G-protein Ras from its inactive GDP-bound form into the active GTP-bound form. RasGTP then recruits the uppermost MAPK cascade member, Raf, to the plasma membrane, where it is phosphorylated and thereby activated. Phospho-Raf activates MEK by phosphorylating it at two residues, and phospho-MEK then in turn acts as an ERK kinase (Fig. 3.1A). Through double-phosphorylation, ERK is activated and translocates to the nucleus, where it phosphorylates several nuclear kinases and transcription factors, which change the expression of about one hundred genes. The Raf-Mek-Erk cascade received much scientific attention since it plays a central role in various physiological processes such as cell cycle progression, cell differentiation and cell death. Additionally, cascade activation is dysregulated in many diseases. Moreover, the system serves as the best-understood paradigm for biological regulation by protein phosphorylation/dephosphorylation cycles. The human genome encodes for about 500 kinases and 200 phosphatases [3,4], so that insights into the design principles of MAPK cascades are expected to be applicable to cellular regulation in general.

Many diseases are characterised by deregulated signalling. For example, growth-promoting signalling is constitutively elevated in many cancers, so that cells divide at a high rate, and are no longer dependent on external stimulation. The Ras-MAPK signalling cascade is central to tumour development, as it is deregulated in more than 30% of all cancers, and even more frequently in some specific types of cancer, such as pancreatic cancer (90%) [5]. Activating mutations in Ras and B-Raf are common in cancer, and were shown to be crucial for tumour progression [6]. Overexpression of Ras and downstream cascade members due to promoter deregulation and/or genomic amplification has also been observed in tumours, but the functional significance of such wildtype protein overexpression remains unclear (e.g., [7,8,9]). In Section 2, the role of Ras overexpression is studied theoretically by numerical analysis of an experimentally validated model of the RasGDP/GTP cycle [10]. It is shown that basal Ras signalling activity can be completely insensitive to the total Ras protein expression level, which might explain why overexpression of wildtype Ras is often insufficient to enhance downstream signalling and to transform cells. Interestingly, this "tumour suppression effect" is only observed in the weak stimulus regime, and disappears upon strong stimulation, so that Ras overexpression, though phenotypically silent on its own, can
prime cells for strong transformation by mutational activation of signalling pathways upstream of Ras. Cancer development is a multistep process, and it is well established that multiple oncogenic mutations often cooperate to bring about cellular transformation. The simulations presented in Section 2 explain experimental studies showing that overexpression events within the MAPK cascade, though weakly transforming in isolation, frequently cooperate to bring about strong cellular transformation ("oncogene cooperation"). The results hold for cellular signal transduction via activation-deactivation cycles in general, and are thus likely to be relevant for many intracellular signalling pathways.

![Figure 1.1: Signal transduction via MAPK and caspase cascades.](image)

(A) Schematic representation of signalling via the Ras-MAPK cascade. Solid arrows indicate mass transfer, while dotted lines denote regulatory control (typically by enzymatic catalysis) (B) Schematic representation of apoptosis signalling via caspase cascades (adapted from [11]). Solid arrows indicate mass transfer, dotted lines indicate regulatory control, and blunted solid arrows indicate reversible inhibition by protein-protein interaction.

Under physiological conditions, cells are typically exposed to transient extracellular stimulation. Provided that stimulation is sufficiently strong and/or sustained, cells become irreversibly committed to a new fate. Any biochemical signalling network involved in cell fate decisions must be able to reliably discriminate between physiological relevant signals and background fluctuations in order to avoid improper responses. Such filtering is typically realised by all-or-none, switch-like stimulus-response behaviour. These stimulus-response relationships are sigmoid, meaning that sub-threshold stimuli fail to elicit significant responses, while strong signalling is observed in response to supra-threshold stimulation. As these sigmoid curves are highly sensitive around the threshold, this phenomenon has been often referred to as ultrasensitivity. Three basic ultrasensitivity mechanisms have been described for biochemical signalling networks:
Multistep ultrasensitivity: Ultrasensitivity arises if an input regulates its downstream effectors at multiple points [12]. In the MAPK cascade, the kinases Raf and Mek phosphorylate their downstream effectors, Mek and Erk, at two sites, and thus establish multistep regulation (Fig. 1.1A). Such double phosphorylation can give rise to quadratic (instead of linear) kinase control over substrate phosphorylation [13].

Zero-order ultrasensitivity: Early theoretical work by Goldbeter and Koshland [14] showed that phosphorylation cycles can exhibit very strong ultrasensitivity if the catalyzing enzymes (i.e., kinase and phosphatase) operate near saturation, i.e., if the Michaelis-Menten constants (K_M-values) of the enzymes are much smaller than the substrate concentration. This phenomenon has been termed zero-order ultrasensitivity to reflect that the (de)phosphorylation velocities are independent of the substrate concentration in the saturated regime. Recent studies however casted doubt on the physiological relevance of zero-order ultrasensitivity in mammalian MAPK signalling, as ultrasensitivity is strongly weakened if the enzyme and substrate concentrations are comparable, i.e., if the substrate is significantly sequestered on the catalyzing enzymes [13,15].

Inhibitor ultrasensitivity: A third way to generate ultrasensitivity is through competitive inhibition [16]. In the extreme case of strong inhibition, low-level pathway activation is completely suppressed by binding of the inhibitor. Once the concentration of the target molecule, however, exceeds that of the inhibitor, inhibition no longer occurs, so that the system suddenly switches on.

Switch-like stimulus-response behaviour can be further enhanced if the basic ultrasensitivity mechanisms described above are embedded into a larger biochemical network. In particular, multi-layered signalling cascades such as the MAPK pathway typically show much stronger and more robust ultrasensitivity than each level in isolation [12,17,18]. Another way to amplify ultrasensitivity is positive feedback, as feedback signalling only starts to become significant once the threshold has been exceeded, thus making the near-threshold response even steeper [19]. An extreme manifestation of feedback amplification is bistability, where the system switches between two discrete states (on and off) in a true all-or-none manner. Bistable systems also display hysteresis, meaning that different stimulus-response curves are obtained depending upon whether the system began in its off or its on state. Thus, the systems' behaviour is history-dependent, as the positive feedback can act as a memory device, which maintains high activity even if stimulation strength decreases. In some cases, the on state is maintained indefinitely after the stimulus is completely removed, so that the system shows irreversible activation [20].

Bistability is an important mechanism for noise-resistant cell fate decisions, and it has been proposed that bistability arises in the mammalian MAPK cascade, because Erk activates Raf in a positive feedback loop via PKC [21]. Intriguingly, recent theoretical work indicated that implicit positive feedbacks and bistability can emerge even in the core Raf-Mek-Erk module due to enzyme sequestration effects [22,23]. Single-cell measurements strongly suggest that the Raf-Mek-Erk module is bistable in mammalian neuronal precursors (PC12 cells), because all-or-none and irreversible Erk activation was observed in response to extracellular stimulation [24]. However, this phenomenon appears to be cell-type-specific, as Erk activation at the single-cell level was gradual but ultrasensitive in fibroblasts [25]. Taken together, these data suggest that filtering due to ultrasensitivity is a general property of the mammalian Raf-Mek-Erk module, while bistability appears to be a context-dependent phenomenon.

The molecular mechanisms underlying ultrasensitivity and bistability in MAPK activation remain incompletely understood, but recent single-cell measurements suggest that bistability can arise within the core MAPK cascade [26]. Experimental studies revealed that unphosphorylated Mek and Erk form a stable complex in unstimulated cells, which
dissociates upon stimulation with growth factors (see, e.g., [27]). In Section 3, the impact of such basal Mek~Erk association is analysed numerically using a mathematical model consisting of experimentally validated parameters. The simulations reveal that an implicit positive feedback and bistability can arise in the MAPK cascade from stimulus-induced dissociation of the Mek-Erk complex. Further simulations demonstrate that bistability due to Erk-Mek complex formation strongly cooperates with other potential bistability mechanisms in the MAPK cascade. In conclusion, the results in Section 3 help to explain how MAPK signalling can establish all-or-none and irreversible cell fate decisions in response to transient extracellular stimulation.

Apoptosis, an evolutionary conserved form of cell suicide, allows multicellular organisms to eliminate damaged or excess cells in order to maintain tissue homeostasis (reviewed in [11]). Aspartate-specific cysteine proteases, also known as caspases, are the central executioners of apoptosis. In most cases, apoptotic stimuli activate initiator caspases, whose substrates, the effector caspases, ultimately cause cellular demise by cleaving various cellular substrates (e.g., nuclear DNA processing enzymes). Figure 1.1B schematically depicts the so-called intrinsic and extrinsic apoptotic pathways that elicit apoptosis by cleaving and thereby activating caspase-3, the major cellular effector caspase. The extrinsic pathway is initiated by ligand binding to death receptors (e.g., CD95), which then oligomerise and recruit various proteins, including pro-caspase8, into the so-called death-inducing signalling complex. Formation of the death-inducing signalling complex leads to autoprocessing of pro-caspase8 into active (initiator) caspase8, which then cleaves (effector) caspase-3. Cytotoxic stress (e.g., chemotherapeutic treatment) or death-receptor–stimulated caspase-8 engage the intrinsic, or mitochondrial, apoptosis pathway by inducing the translocation of pro-apoptotic Bcl-2 family members such as Bid to mitochondria. Pro-apoptotic Bcl-2 family members permeabilise the mitochondrial membrane and thereby elicit the release of pro-apoptotic proteins (cytochrome c and Smac) into the cytosol. Cytosolic cytochrome c then induces the oligomerisation of Apaf-1 into an active high molecular-weight complex, the apoptosome, which recruits and stimulates (initiator) caspase-9, and thereby allows activation of effector caspases such as caspase-3. Smac and inhibitors of apoptosis (IAPs) such as X-linked IAP (XIAP) establish an additional layer of regulation in the intrinsic pathway: XIAP inhibits the catalytic activities of caspase-9 and caspase-3 through reversible binding, and cytosolic Smac relieves this inhibition by sequestering XIAP away from caspases.

Experimental analyses revealed that apoptosis is an all-or-none process at the single-cell level [28,29,30], and that cells are irreversibly committed to death once the apoptosis program has been initiated [31]. Previous theoretical analyses indicated that bistability might underlie rapid all-or-none and irreversible caspase activation in the extrinsic apoptosis pathway [32], but the kinetic characteristics of the intrinsic caspase pathway have been less well studied. In Section 4, a mathematical model of the core intrinsic apoptosis pathway is derived based on published biochemical and kinetic measurements. Subsequent numerical analyses reveal an unanticipated role for the inhibitor of apoptosis (IAP) proteins, which simultaneously inhibit caspase-9 and its downstream effector caspase-3 (Fig. 1.1B): Caspase-3, once activated, sequesters XIAP away from caspase-9, and thus allows for further caspase-3 and -9 activation in a positive feedback loop. In other words, IAP proteins, though inhibitors of caspase enzymatic activity, might establish an implicit positive feedback loop similar to that discussed for MAPK signalling in Section 3.

Taken together, the results in Part I provide insights into the steady state dose-response behaviour of signal transduction by rapidly acting post-translational regulatory mechanisms. The following parts of the thesis deal with much slower transcriptional feedback regulation of signal transduction (Part II), and with decoding of intracellular signals by the downstream gene expression machinery (Part III).
1.2 Time scales of signalling and transcriptional feedback regulation

Signalling networks employ post-translational regulatory mechanisms such as phosphorylation, ubiquitination and nucleo-cytoplasmic shuttling to transduce inputs from the cell membrane to the nucleus. These events typically occur on a time scale of seconds to minutes, and are thus much faster than transcriptional regulation. Thus, it might be expected that the dynamics of signalling networks are mainly governed by post-translational feedback mechanisms, while slow transcriptional feedback loops play no major role. Experimental evidence summarised in the following, however, suggests that transcriptional feedbacks are central to the regulation of decision making in mammalian cells.

The c-fos proto-oncogene belongs to the group of immediate-early genes directly regulated by phospho-Erk, and c-fos protein starts to accumulate 20 – 40 minutes after growth factor stimulation [33,34,35]. Thus, the time scale of stimulus-induced gene expression and transcriptional feedback is beyond ~20 minutes. Depending on the cellular context and on the stimulus, signals are shorter than 20 minutes [36,37], which suggests that transcriptional feedback does not play a role in these systems. Accordingly, it was explicitly shown in PC12 cells that short-term Erk activation in response to EGF is not affected if de novo transcription is blocked by the general transcription inhibitor Actinomycin D [36].

Table A.1 summarises the experimentally observed signal duration ranges of several signalling pathways, and reveals that they frequently remain active over several hours under physiological conditions. Most importantly, Table A.1 also lists measurements in vivo, and reveals that signalling cascades are frequently active over several hours in living animals. These data strongly suggest that transcriptional feedback regulates the temporal dynamics of many signalling pathways under physiological conditions.

The impact of transcriptional feedback cannot simply be shown by specific knock-down or inhibition of signalling proteins, as even constitutively expressed signalling proteins modulate the signal duration [38]. Three lines of evidence nevertheless strongly support that transcriptional feedback loops regulate the dynamics of signalling cascades. First, specific inhibition of signalling proteins that are transcriptionally induced upon stimulation does not affect early phase signalling and the peak amplitude of the signal, but selectively affects signal termination at later times [39,40,41]. Second, a knock-down of transcription factors mediating gene expression and thus transcriptional feedback regulation downstream of the JNK cascade was shown to affect JNK signal magnitude and duration in response to oxidative stress [39]. Third, the several signalling pathways show altered dynamics if downstream gene expression was blocked by pre-incubation with inhibitors of translation or transcription (cycloheximide, actinomycin D). Table A.2 summarises literature studies where protein biosynthesis inhibitors sustained signalling activity, and thus strongly supports the relevance of transcriptional negative feedback regulation in the mammalian signalling network.

Late-acting transcriptional feedback loops can only be physiologically relevant if late phase signalling is required for phenotypic responses towards extracellular stimulation. Tables A.3 and A.4 summarise published experimental studies where signalling was terminated by delayed addition of small-molecule kinase inhibitors or by delayed removal of extracellular ligands. The threshold times indicate the signal duration required to irreversibly commit the cell population to the given phenotypic response (stimulation occurred at t = 0 h). The data can be summarised as follows: The decision whether any (immediate-early) gene expression occurs or not is typically made within minutes [36,42,43], while late-phase gene expression requires ongoing signalling for 30’ – 4 h (Tables A.3 and A.4). Importantly, the commitment to new cell fates such as apoptosis, differentiation or S-phase entry often requires ongoing signalling or stimulation for more than 5 h (Tables A.3 and A.4), even though a few seconds might already be sufficient in neuronal signalling [44]. It should be noted that Tables A.3 and A.4 summarise threshold times, i.e., the signalling time required to irreversibly commit to an event. In other cases, commitment is not observed and ongoing signalling over long periods...
is required to maintain a phenotype. For example, differentiation of PC12 occurs after ~3 days stimulation with NGF, but cells still de-differentiate if NGF is removed after 7 days [45].

In conclusion, different experimental approaches (i.e., pharmacological inhibition, ligand removal and stimulation with different ligand doses) strongly suggest that long-term, supra-basal signalling over several hours is often required for cells to commit to phenotypic responses. This indicates that decision making in mammalian cells is profoundly affected by slowly acting transcriptional feedback mechanisms. In Section 5, the general design principles underlying transcriptional feedback regulation of mammalian signalling pathways are therefore investigated. Published microarray studies in response to extracellular stimulation were collected, and the stimulus-induced expression of 134 intracellular signalling proteins was investigated. The analysis revealed that transcriptional feedback regulation occurs in each of the five mammalian signalling pathways considered, and that negative feedback strongly dominates over positive feedback. Moreover, negative feedback exclusively occurs by transcriptional induction of a subgroup of signal inhibitors, termed rapid feedback inhibitors (RFIs), while downregulation of signal transducers plays no role. Systematic analysis of mRNA and protein half-lives reveals a remarkable separation of the signalling network into flexible and static parts: transcriptionally regulated RFIs are unstable at the mRNA and protein levels, respectively, while other signalling proteins are generally stable. Kinetic modelling, also presented in Section 5, suggests that this design principle allows for swift feedback regulation and establishes latency phases after signalling, and that it might be an optimal design due to a trade-off between energy efficiency and flexibility.

The analyses presented in Section 5 confirm that rapid feedback inhibitors (RFIs) are induced at the mRNA level, but the physiological relevance of these feedbacks remains to be verified at the protein level. In Section 6, transcriptional feedback regulation was therefore analysed in more detail using TGF\(\beta\) signalling in primary hepatocytes as an example. The TGF\(\beta\) family of cytokines constitute major inhibitors of cell growth, and accordingly they play an important role in various physiological processes such as development, tissue homeostasis, and tissue regeneration. Moreover, TGF\(\beta\) signalling is dysregulated under pathological conditions including cancer, organ fibrosis, and Marfan syndrome (reviewed in [46,47,48]). TGF\(\beta\) signalling is initiated by binding of extracellular TGF\(\beta\) to transmembrane serine/threonine kinase type I and type II receptors. Ligand binding triggers receptor-mediated phosphorylation of Smad2/3 transcription factors, which then homotrimerise or heterotrimerise (with Smad4), and subsequently translocate into the nucleus (Fig. 6.1). Nuclear Smad trimers control the expression of several hundred target genes, many of which are involved in cell cycle control (e.g., p21, c-myc, cdc25A).
Figure 1.2: The quantitative model of signalling specificity.

(A) - (C) Signalling motifs mediating signal adaption. (A) Sequestration: Signal termination occurs when the activated molecule ($X^*$) is sequestered into an inactive form ($X_1^*$), e.g., through binding, internalisation or degradation. (B) Incoherent feed-forward loop: The stimulus first activates the molecule through a fast process, and then terminates signalling through a slow, delayed mechanism. (C) Negative feedback: The activated molecule inhibits its own production. (D) The signalling network from NGF, Glutamate and EGF stimulation to Erk activation. Various adaptation mechanisms control the duration of Erk activation in a stimulus-specific manner (see main text).

Experimental evidence suggests that transcriptional feedback loops play an important role for the dynamics of TGFβ signalling via Smad transcription factors, and thus for phenotypic responses triggered by TGFβ: Smad activity remains elevated over several hours after TGFβ stimulation (Appendix A), thus being vulnerable for transcriptional feedback regulation. Accordingly, various transcriptional feedback regulators acting at all levels of the signalling cascade have been reported of TGFβ signalling (Appendix E). Moreover, Smad activation was shown to be prolonged upon incubation with protein synthesis inhibitors which block de novo gene expression, and thus transcriptional feedback loops (Appendix A). In support for the relevance of transcriptional feedback regulation, it is known that sufficiently sustained Smad signalling over several hours is required for TGFβ-mediated inhibition of cell growth [49]. Genome-wide microarray analyses presented in Section 6 suggest that transcriptional feedback regulation exists in primary mouse hepatocytes stimulated with TGFβ, and measurements at the protein level indicate that the SnoN oncoprotein is the central feedback regulator. Theoretical predictions regarding the mechanisms of transcriptional feedback regulation are then confirmed in primary hepatocytes isolated from SnoN knock-out mice, thus further supporting the relevance of transcriptional negative feedback regulation in signal transduction.
Taken together, the results in Part II reveal how transcriptional feedback regulation shapes the dynamics of signal transduction. Part III is mainly focussed on the decoding of intracellular signals by the downstream gene expression machinery, as further outlined in Section 1.3.

1.3 The quantitative model of signal specificity and decoding at the level of gene expression

The quantitative model of signal specificity: Cells face a specificity problem as different extracellular stimuli frequently engage the same set of intracellular signalling pathways even though they elicit completely different biological responses. Experimental evidence suggests that stimulus-specific biological information is frequently encoded in the quantitative aspects of stimulus-specific activation kinetics. The Raf-Mek-Erk pathway serves as a paradigm for quantitative encoding of signalling specificity, as both the amplitude and the duration of Erk activation are critical determinants for the cell fate. The amplitude of the Erk signal seems to be important in fibroblasts, where low level Erk signalling induces proliferation, while cell cycle arrest is observed in response to strong Erk activation [50]. In neuronal precursors (PC12 cells), it seems to be mainly the Erk signal duration that matters, as these proliferate upon short-term Erk activation, but differentiate if Erk phosphorylation is sustained [51]. Yet, the fate of PC12 cells can be induced in a stimulus-specific manner, as EGF stimulation elicits transient Erk activation, while prolonged Erk signalling is observed in response to NGF (see Fig. 1.2B) [51].

Various mechanisms are thought to contribute to such stimulus-specific modulation of Erk signal duration. In simple terms, transient Erk activation is observed if the MAPK cascade or upstream pathways are subject to delayed negative regulation, which ensures efficient signal termination, also referred to as signal adaptation. In contrast, Erk exhibits sustained activation if such adaptation mechanisms are absent. Three basic regulatory mechanisms depicted in Fig. 1.2A – C are known to mediate signal adaptation in biochemical regulatory networks (sequestration/internalisation, incoherent feed-forward regulation and negative feedback regulation). All of them seem to be involved in the modulation of MAPK signal duration as discussed in the following.

Growth factor receptors are taken up into the cell upon ligand binding, and are deactivated and/or degraded within intracellular vesicles and therefore the signal is terminated (arrow 1 in Fig. 1.2D). It is known that EGF receptors internalise more rapidly than NGF receptors, and it has been suggested that rapid receptor sequestration accounts for transient Erk signal duration in EGF-treated PC12 cells [52].

Glutamate induces transient Erk activation in neurons, whereas sustained signalling is observed in response to potassium chloride [53]. Strikingly, the rapid decline in Erk phosphorylation depends on glutamate’s ability to specifically activate PTP-STEP, an Erk phosphatase (arrow 2 in Fig. 1.2D). This kind of regulation is known as incoherent feed-forward loop, because a common upstream regulator (i.e., glutamate) both activates and inhibits a downstream effector (i.e., Erk) via two independent branches. Incoherent feed-forward loops generate transient signals if positive regulation (Erk phosphorylation by Mek) proceeds faster than negative regulation (PTP-STEP activation), as observed in glutamate-induced neuronal MAPK signalling [53].

Early theoretical work indicated that differential feedback control in EGF vs. NGF signalling may underlie transient vs. sustained Erk signalling in PC12 cells [54]. A recent experimental and theoretical study confirmed that stimulus-specific feedback regulation indeed occurs in PC12 cells [24]. It was shown that EGF specifically induces negative feedback within the MAPK cascade and thereby induces early termination of Erk activity (e.g., arrow 3 in Fig 1.2D). NGF signalling appeared to escape signal adaptation because strong initial Erk
activation (compared to EGF stimulation) specifically triggers a positive feedback which outweighs negative feedback regulation (arrow 4 in Fig. 1.2D). In particular, this positive feedback was shown to establish bistable and irreversible Erk activation, thus ensuring prolonged signalling [24].

Both theoretical and experimental studies demonstrated that a pathway consisting of the adaptor protein FRS and the small G-protein Rap plays a key role for sustained Erk signalling in NGF-treated PC12 cells [55,56,57]. Importantly, this pathway is specifically engaged by NGF for receptor-mediated Raf activation, but not by EGF [55,57]. Moreover, sustained activation of Erk in response to NGF is abolished if Rap signalling is blocked [57]. Accordingly, the FRS-Rap pathway shows sustained activation kinetics, while the functionally redundant Shc-Ras pathway that mediates EGF-induced Raf activation is only transiently active [55,57]. Various explanations have been proposed why the duration of Ras and Rap signalling differs: (i) incoherent feed-forward regulation via RasGAP limits signal duration of the Shc-Ras pathway (arrow 5 in Fig. 1.2D), but not that of the FRS-Rap module [55]. (ii) the Shc-Ras branch is subject to negative feedback control by Erk (arrow 3 in Fig. 1.2D), while no such feedback has been described for FRS-Rap signalling [54,57]. (iii) cytosolic adaptor proteins such as Shc are only functional when recruited to active transmembrane receptors, whereas signalling via membrane-anchored adaptors like FRS can continue even without such recruitment. Thus, the FRS-Rap pathway is expected to be less susceptible to signal termination by receptor downregulation, and should therefore exhibit more sustained activation [56]. A summary of the network shaping Erk activation upon stimulation with NGF, glutamate and EGF is shown in Fig. 1.2D.

In conclusion, it appears that stimulus-specific upstream regulatory pathways (e.g., receptors, adaptors or small G-proteins) are major determinants of Erk signal duration, because they exhibit differential susceptibility to negative regulation and thus signal termination. Additionally, the shared MAPK cascade integrates incoming inputs and generates transient or sustained output depending on their amplitude.

**Decoding at the level of gene expression:** If biological information is encoded in the quantitative aspects of intracellular signals, proper biological responses require that the downstream gene expression machinery is able to accurately decode their amplitude and duration. Amplitude decoding most likely occurs by the ultrasensitivity mechanisms discussed in Section 1.1 (i.e., inhibitor ultrasensitivity, zero-order ultrasensitivity and multistep ultrasensitivity). The mechanisms of duration decoding have been less well studied, even though the duration of intracellular signals is known to determine the characteristics of gene expression downstream of TGFβ signalling [49], NF-κB signalling [43], cAMP signalling [42], glucose signalling [58] and MAPK signalling (see above).

Previous theoretical work indicated that multistep regulatory motifs such as the coherent feed-forward loop [59] or multisite phosphorylation [15,60] are able to discriminate transient and sustained stimuli. Moreover, it has recently been shown that bistable systems decode the input signal duration in an all-or-none manner [61]. A multistep regulation mechanism that has been implicated in the decoding of Erk signal duration is the simultaneous induction and stabilisation of the transcription factor c-Fos [5,34]. The transcription of c-Fos is induced by phosphorylated Erk. After translation, the protein is very unstable (half-life $t_{1/2} \approx 15$ min) and cannot accumulate, unless it is stabilised by Erk-mediated phosphorylation ($t_{1/2} \approx 4$ hours). Thus, sufficiently sustained Erk signalling is required for accumulation of the c-fos protein. Such a mechanism, where Erk positively regulates c-Fos at two levels (transcription and protein stability), is known as coherent feed-forward loop [62].

The results presented in Section 7 identify competitive inhibition and regulated degradation as alternative mechanisms that allow intracellular regulatory networks discriminating transient and sustained inputs in an ultrasensitive manner. The temporal characteristics of the cyanobacterial iron stress response are analysed by mathematical modelling and
quantitative experimental analyses. A recently discovered small non-coding RNA, IsrR, is shown to be responsible for a pronounced delay in the accumulation of \textit{isiA} mRNA encoding the late-phase stress protein, IsiA. Moreover, it is demonstrated that sRNA-mediated regulation ensures a rapid decline in \textit{isiA} levels once external stress triggers are removed. These kinetic properties allow the system responding selectively to sustained (as opposed to transient) stimuli. Thus, sRNA-mediated regulation establishes a temporal threshold which ensures that costly \textit{isiA} accumulation is restricted to severe, prolonged and ongoing stress conditions. Non-protein-coding RNA regulators control diverse processes in metazoans including development, cell differentiation, and cell proliferation \cite{63,64}, and also play important roles in bacterial stress responses \cite{65,66}, so that the results presented in Section 7 are likely to be of broader physiological relevance.

Many of the downstream target genes induced by signalling pathways are transcription factors, thus giving rise to a complex transcriptional regulatory network \cite{67}. Therefore, signal decoding at the level of gene expression cannot be fully understood by insights into the functioning of small transcriptional regulatory motifs, but additionally requires integrated analyses of multiple transcription factors. In Section 8, it is described how systems biological approaches allow to derive the topology of complex transcriptional regulatory networks from high-throughput and knock-down data. More specifically, modular response analysis (MRA), a recently proposed reverse engineering method \cite{68}, is applied to get insights into a transcription factor network that is regulated by oncogenic Ras and involved in the regulation of cell division. Statistical analyses of the MRA results allow proposing new experiments that can be used to verify the predicted regulatory interactions. The results in Section 8 thus represent a first step towards an integrated model including both, upstream signal transduction and downstream phenotypic responses such as cell growth.
Part I

Signalling via fast post-translational regulatory mechanisms
2 Kinetic mechanisms for tumour suppression and oncogene cooperation

SYNOPSIS

The basal state activity of signalling pathways is deregulated in many diseases. For example, cancer cells show permanent activation of mitogenic signalling protein kinase cascades and thus proliferate constitutively. In many cases, however, even strong overexpression of signalling intermediates does not appear to initiate tumourogenesis although relatively minor pathway activation (5 – 10 fold) is typically sufficient to induce phenotypic responses such as cell division. This suggests that tumour suppression mechanisms exist which prevent permanent deregulation of signal transduction by protein overexpression. By theoretical analysis of an experimentally validated kinetic model of the Ras activation cycle, it is shown in this chapter that basal Ras signalling activity can be completely insensitive to the Ras expression level, which might explain why overexpression of wildtype Ras is insufficient to enhance downstream signalling and to transform cells. Interestingly, this “tumour suppression effect” is only observed in the weak stimulus regime, and disappears upon strong stimulation, so that Ras overexpression, though phenotypically silent on its own, can prime cells for strong transformation by mutational activation of signalling pathways upstream of Ras. Cancer development is a multistep process, and it is well established that multiple oncogenic mutations often cooperate to bring about cellular transformation. The simulations presented in this chapter explain experimental studies showing that overexpression events within a single signalling cascade, though weakly transforming in isolation, frequently cooperate to bring about strong cellular transformation.

2.1 Introduction

Normal cells require extracellular mitogenic signals before they can move from a quiescent state into an active proliferative state. These signals are transmitted into the cell by transmembrane receptors that bind distinctive classes of signalling molecules: diffusible growth factors, extracellular matrix components, and cell-to-cell adhesion/interaction molecules. Cancer cells escape the requirement for extracellular stimuli by constitutive activation of intracellular mitogenic signalling pathways. Such aberrant signalling is typically induced by mutational activation or overexpression of signal transducers, or by deactivation as well as downregulation of signal inhibitors [69]. Thus, cancer cells are often characterised by deregulated steady state signal transduction in the basal state.

The small GTPase Ras transduces signals from extracellular growth factors, and controls various cellular responses including proliferation, apoptosis and migration. The Ras protein binds to guanine nucleotides, and cycles between inactive and active states by switching between GDP- and GTP-bound forms. Guanine nucleotides bound to Ras are exchanged even in the absence of other proteins by slow association-dissociation reactions, and Ras-bound GTP is hydrolysed to GDP by the weak intrinsic GTPase activity of Ras. Within the cell, these reactions are strongly enhanced by guanine nucleotide exchange factors (GEFs) and GTPase-activating proteins (GAPs), respectively. Growth factors enhance GEF activity and/or inhibit GAPs, and thereby trigger the accumulation of active GTP-bound Ras [6].
Figure 2.1: Mathematical Modelling of Ras overexpression.

(A) Schematic representation of the mathematical model. Ras cycles between the inactive GDP-bound state and the active GTP-bound state. Cycling occurs either by intrinsic nucleotide exchange and intrinsic Ras GTPase (grey arrows), or is catalysed by guanine nucleotide exchange factors (GEF) and GTPase-activating proteins (GAP), respectively (black arrows). GEF- and GAP-catalysed reactions were modelled with Michaelis-Menten kinetics, and the intrinsic steps with first-order kinetics. GTP and GDP were assumed to be present in excess, and were thus treated as external variables, as described in more detail in the Supplement. (B) Effects of Ras overexpression on signal transduction. The steady state RasGTP concentration is plotted as a function of total Ras expression (i.e., the sum of RasGTP, RasGDP and Ras). A simple linear relationship is observed if only the first-order intrinsic reaction steps (grey arrows in A) are taken into account (grey line). The black lines show simulations using a model where GEF- and GAP-catalysed reactions were taken into account, while intrinsic rates were neglected, and thus reflect the situation in living cells, where Ras cycling typically seems to be dominated by GEF- and GAP-catalysed reactions (see text). The default kinetic parameters were taken from [10], and the experimentally measured Ras expression level in living cells is indicated by the vertical dashed line. The KM-values for both GEF-catalysed reactions (i.e., RasGTP → RasGDP and RasGDP → RasGTP) were simultaneously varied as indicated by the arrow (line 1: KM,GDP = 3.86 · 10^-6 mol/l; KM,GTP = 3 · 10^-6 mol/l; line 2: KM,GDP = 3.86 · 10^-5 mol/l; KM,GTP = 3 · 10^-5 mol/l; line 3: KM,GDP = 3.86 · 10^-4 mol/l; KM,GTP = 3 · 10^-4 mol/l; line 4: KM,GDP = 3.86 · 10^-3 mol/l; KM,GTP = 3 · 10^-3 mol/l).

Aberrant Ras signalling is a common feature of tumours, and is, for example, caused by enhanced GEF activity or by decreased GAP expression. Moreover, ~30% of all tumours are characterised by Ras mutations which impair GTPase activity, and thus lock Ras in the active GTP-bound state [6]. The oncogenic potential of Ras mutants is usually analysed in overexpression studies, and the cells are assayed for their ability to induce tumours in mice, to proliferate in starvation media, or to induce a tumour-like (i.e., a transformed) phenotype in cell culture. Hallmarks of transformed cells are anchorage-independent growth, focus formation and a lack of contact-inhibition. Based on these criteria, it seems that even a relatively low cellular Ras-GTP content activation is sufficient to promote normal cell proliferation and tumourigenesis: (i) Even saturating growth factor concentrations typically induce no more than 5-10-fold Ras activation in normal cells when compared to the starved basal state (e.g., [70,71]). (ii) Heterozygous knock-in mice expressing mutant Ras under the control of the endogenous promoter develop tumours [72,73]. (iii) Overexpression of GEFs allows for transformation [74,75,76,77] although Ras-GTP is only increased by 3 – 4-fold in these cells [74,77]; (iv) Neurofibromatosis patients do not express the NF1 GTPase-activating protein, and it is thought that enhanced Ras signalling contributes to tumour formation in neurofibromatosis [78]. Yet, these tumours exhibit only a 3 – 10-fold increase in RasGTP content when compared to normal cells [78,79,80].

One would thus expect that even relatively minor (5–10 fold) overexpression of wildtype Ras (‘c-Ras’) should be sufficient to induce tumourigenesis in vivo and cellular transformation in culture, as a significant fraction of c-Ras is in the GTP-bound active form even under starving conditions (e.g., [57,81]). In fact, even very strong overexpression of c-Ras does not induce cellular transformation or does so only very weakly in most cell lines and experimental settings [76,82,83,84,85,86,87]. Yet, the wildtype protein has, depending on the context, the
potential to induce transformation in culture [88,89] or to induce the very same tumours as its mutant counterparts [90].

Biochemical studies of Ras-transfected cell lines revealed that the fractional activation of c-Ras (i.e., the amount of Ras-GTP divided by the total Ras concentration) in the basal state decreases with increasing c-Ras expression in several cell lines [80,91]. This suggests that kinetic features of the Ras signalling network can suppress aberrant basal signalling in cells strongly overexpressing c-Ras. Accordingly, it was shown experimentally that c-Ras overexpression alone does not enhance downstream signalling in the unstimulated basal state [87]. In this chapter, an experimentally validated model of the core Ras activation cycle [10] is analysed, and it is shown by kinetic modelling that the signal activity can be completely insensitive to the total Ras expression levels. The requirements for this ‘kinetic tumour suppression’ effect are derived analytically and the implications for cooperation among oncogenes are discussed. The results also account for experimental observations in the downstream MAPK cascade, and might explain why oncogene overexpression is generally inefficient for cellular transformation.

### 2.2 Basal Ras signalling activity can be insensitive towards Ras overexpression

The core Ras cycle (Fig. 2.1A) comprising the intrinsic and GAP-catalysed GTPase reactions, and spontaneous as well as GEF-catalysed nucleotide exchange has been intensively characterised by kinetic analyses of recombinant proteins in vitro. Stites et al. implemented a mathematical model for the core Ras activation using the kinetic parameters measured in vitro, and showed that the model predictions agreed well with experiments in living cells [10]. In this chapter, a slightly simplified version of their model, termed the “default model”, is analysed (Appendix B and Fig. 2.1A). The degree of DNA synthesis and of cellular transformation are known to be titrable functions of the intracellular Ras-GTP concentration [88,92,93]. Thus, the concentration of RasGTP is used as the physiologically relevant output of the model, and conclusions regarding cellular transformation are generally based on the assumption that transformation is proportional to the simulated RasGTP levels.

The fraction of Ras protein in the GTP-bound state in starved cells ranges from less than 1% to 30% (e.g., [57,81]). Some background stimulation most likely exist in vivo, so that even higher basal Ras activation levels are expected in the living animal. Intuitively, one would assume that the basal RasGTP concentration (i.e., basal Ras signalling) is proportional to the absolute Ras expression level. In other words, enhanced basal RasGTP signalling and thus cellular transformation might be inducible by Ras overexpression, especially because relatively minor (5 – 10 fold) increases in RasGTP already elicit phenotypic responses (see Section 2.1).

To investigate the effect of Ras overexpression in the model, the basal steady state RasGTP level was simulated as a function of total Ras concentration. For some cells, it was shown that the intrinsic GTPase and exchange reactions predominate over the GEF- and GAP-catalysed steps in the unstimulated basal state [71]. A corresponding simplified model solely consisting of the intrinsic grey reaction steps in Fig. 2.1A showed a simple linear relationship between the amount of GTP-bound and total Ras (Fig. 2.1B, grey line). Such direct proportionality is expected in this simplified model, because the intrinsic GTPase and exchange reactions occur with first-order kinetics (note that GTP and GDP are assumed to be in excess). Assuming that the amount of RasGTP governs cellular behaviour, one can conclude that sufficiently strong Ras overexpression induces phenotypic responses and cellular transformation, as long as the intrinsic reaction steps predominate basal Ras cycling.

In most cells, however, basal Ras cycling in starvation media was shown to be dominated by the GEF- and GAP-catalysed reactions and/or is strongly affected by GEF and GAP knock-
downs [78,79,80,94,95,96,97], respectively. In the “GEF-GAP model” it was therefore assumed that the GEF- and GAP-catalysed steps are much faster than the spontaneous exchange and GTPase reactions, so that the model comprises only the black arrows in Fig. 2.1A. Weak basal state signalling was modelled by assuming that maximal velocity of the GEF-catalysed reaction significantly exceeds that of the GEF-catalysed exchange reaction. The simulation result using the default kinetic parameters from the Stites-model (black line in Fig. 2.1B marked with default) reveals that basal state RasGTP level responds in a nonlinear fashion towards changes in Ras expression. More specifically, the RasGTP concentration no longer increases with increasing Ras expression, once the intracellular Ras expression is sufficiently high. This suggests that the kinetic features of the core Ras cycle can explain why in some cells the fraction of RasGTP decreases with increasing Ras expression [80,91], and why Ras overexpression alone does not enhance downstream signalling in the basal state [87].

This ‘kinetic tumour suppression’ effect, however, requires strong Ras overexpression in the default model proposed by Stites et al. [10], as the RasGTP level starts to become completely insensitive only for more than 10-fold Ras expression over the measured intracellular concentration (which is indicated by vertical dashed line in Fig. 2.1B). The kinetic requirements for saturation were therefore analysed in more detail to get insights into the physiological relevance of the proposed mechanism.

2.3 Requirements for Kinetic Tumour Suppression

In the Stites-model, the GEF- and GAP-catalysed reactions are modelled by reversible and irreversible Michaelis-Menten equations, respectively. Detailed numerical simulations (not shown) revealed that the behaviour of the Ras activation cycles does not appreciably change if the GEF-catalysed back-reaction (RasGTP $\rightarrow$ RasGDP) is eliminated from the model. A simplified GEF-GAP model comprising only irreversible GEF and GAP reactions was therefore analysed in order to derive an analytical expression for kinetic tumour suppression effect. For the half-saturation Ras expression level, where RasGTP reaches half of its saturation value, one obtains:

$$ Ras_{50} = \frac{1}{2} \cdot \frac{K_{M,GAP} + 2 \cdot \frac{V_{\text{max,GAP}}}{V_{\text{max,GEF}}} \cdot K_{M,GEF}}{V_{\text{max,GAP}}/V_{\text{max,GEF}} - 1} $$

(2.1)

Assuming weak basal state signalling ($V_{\text{max,GAP}} >> V_{\text{max,GEF}}$), this simplifies to:

$$ Ras_{50} = \frac{1}{2} \cdot \frac{V_{\text{max,GAP}}}{V_{\text{max,GEF}}} \cdot K_{M,GAP} + K_{M,GEF} $$

(2.2)

In the Ras activation cycle, the Michaelis-Menten constant of the GAP-catalysed reaction ($K_{M,GAP}$) is typically lower than that of the GEF-catalysed reaction (and $V_{\text{max,GAP}} >> V_{\text{max,GEF}}$), so that the half-maximal Ras expression level is approximately given by $Ras_{50} \approx K_{M,GEF}$. One therefore expects that the Michaelis-Menten constant of the GAP-catalysed reaction should be a major determinant for the RasGTP insensitivity towards Ras expression. The predicted linear relationship between $Ras_{50}$ and $K_{M,GEF}$ was confirmed by numerical simulations of the full GEF-GAP model with a reversible GEF reaction (black lines in Fig. 2.1B).

The simulations in Fig. 2.1B strongly suggest that the proposed insensitivity mechanism is physiologically relevant: The Michaelis-Menten constant of the GEF-catalysed reaction used in the Stites model was measured for Cdc25Mm on unprocessed Ras [98]. However, it is well established that post-translational Ras farnesylation drastically enhances the Ras-GEF interaction [99,100,101]. Moreover, it is known that Michaelis-Menten constants of other
GEFs than Cdc25Mm are lower, especially if Ras is farnesylated [99,101,102,103]. Kinetic measurements taken from the literature (see Appendix B) indicate that the $K_{M, GEF}$ will often be two or more orders of magnitude lower than that used in the default Stites model. In other words, the $K_{M, GEF}$ will be close to or even below the intracellular Ras concentration, which suggests that the proposed kinetic tumour suppression effect operates in vivo.

Figure 2.2: Ras overexpression primes cells for transformation.

(A) Effects of Ras overexpression on signal transduction for varying upstream stimulus levels in a model comprising GEF- and GAP-catalysed reactions only (thus neglecting the intrinsic reaction steps, i.e., the grey arrows in Fig. 2.1 A). The RasGTP concentration is plotted as a function of total Ras expression (i.e., the sum of RasGTP, RasGDP), and the GEF concentration (i.e., the $V_{\text{max}}$ of GEF-catalysis) is varied. The GEF concentrations are $10^{-10}$, $2 \cdot 10^{-9}$, $7 \cdot 10^{-9}$ and $5 \cdot 10^{-8}$ mol/l (this corresponds to $V_{\text{max,GEF}} / V_{\text{max,GAP}} = 0.02, 0.4, 1.4$ and 10). The Michaelis-Menten constants of the GEF reactions were set to $K_{M,GDP} = 3.86 \cdot 10^{-6}$ mol/l and $K_{M,GTP} = 3 \cdot 10^{-6}$ mol/l. All other parameters were as in [10] (see Appendix B), and the experimentally measured Ras expression level in living cells is indicated by the vertical dashed line. Points A – D are described in B. (B) Schematic representation of phenotypically silent priming for strong transformation by overexpression of wildtype Ras. Normal untransformed cells are typically characterised by low upstream GEF activity, low Ras expression and accordingly low downstream MAPK activity (point A in panel A). Signalling can be enhanced to some extent by upstream pathway (GEF) activation (point B in panel A), resulting in a weakly transformed cell, which can be converted into a highly transformed cell by Ras overexpression (point D in panel A). The simulations in this chapter reveal that, in many cases, Ras overexpression alone is insufficient to enhance downstream signalling and to induce transformation (‘kinetic tumour suppression’; point C in panel A). However, these Ras expressing cells are primed for transformation in the sense that they can be directly converted from a silent state into a highly transformed state by a single mutational step resulting in minor upstream pathway activation (‘oncogene cooperation’; point C $\rightarrow$ point D in panel A).

Taken together, the above simulations might explain why some cell lines can be transformed by overexpression of wildtype Ras, while others cannot (see Section 2.1). In some cells, basal Ras signalling increases linearly with the Ras expression level [104,105,106]. The results presented here suggest that the intrinsic GTPase and exchange rates dominate basal Ras cycling in these cells or that the Michaelis-Menten constant of the GEF-catalysed reaction significantly exceeds the intracellular Ras concentration. On the other hand, cells with Michaelis-Menten constants close to or below the Ras concentration are predicted to be insensitive towards Ras overexpression, and such insensitivity has also been reported experimentally [80,87,91].
2.4 Implications for oncogene cooperation

Equation 2.1 suggests that insensitivity of signalling towards Ras overexpression is only observed under basal state conditions, as $R_{\text{basal}}$ becomes negative for $V_{\text{max,GEF}} > V_{\text{max,GAP}}$. Accordingly, numerical simulations of the full GEF-GAP model with a reversible GEF reaction (Fig. 2.2A) confirm that the insensitivity disappears for strong signalling ($V_{\text{max,GEF}} > V_{\text{max,GAP}}$). Moreover, it seems that Ras overexpression, though phenotypically silent in the basal state (point A $\rightarrow$ point C in Fig. 2.2A), primes cells for strong transformation by a mutational pathway activation event upstream of or at the level of GEFs. Here, priming means that the fold change in Ras activation by GEF activation is greater in Ras overexpressing cells (point C $\rightarrow$ point D in Fig. 2.2A) than in wildtype cells (point A $\rightarrow$ point B in Fig. 2.2A). Thus, the simulation results is in accordance with previous work by Goldbeter and Koshland [14] who showed that the sensitivity in covalent modification cycles increases with increasing enzyme saturation. Experimental evidence supports the existence of priming by Ras overexpression, as cellular transformation by transfection with constitutively active Sos GEF constructs is much more efficient in Ras-overexpressing cells than in wildtype cells, even though Ras overexpression alone was without effect [75,76]. Likewise, it was shown that basal MAPK signalling downstream of Ras is not affected by Ras overexpression in LNCAP cells, and that Ras overexpression drastically enhances EGF-induced MAPK activity, thus priming the signalling network for stimulus-induced activation [87]. Figure 2.2B summarises how Ras overexpression and GEF activation interact to bring about strong cellular transformation according to the kinetic simulations in Fig. 2.2A. A normal untransformed cell can either be weakly transformed by GEF activation, and transformation can be further enhanced by Ras overexpression (Fig. 2.2B, right), or phenotypically silent Ras overexpression ('kinetic tumour suppression') primes cells for strong transformation by subsequent GEF activation (Fig. 2.2B, left).

The simulation results in Fig. 2.2, though generally applicable to activation-deactivation cycles, do not fully explain all experimental observations for the Ras activation cycle, at least in some cells. Experimental studies revealed that transfection with constitutively active GEF constructs elicits only negligible cellular transformation (as does overexpression of wildtype Ras), while co-transfection of GEF constructs with Ras induces strong transformation [75,76]. Hence, it seems that both ‘mutational’ events strongly cooperate similar to a logical AND-gate for cellular transformation, a property that is not fully described Fig. 2.2A, where GEF activation has an effect (point A $\rightarrow$ point B in Fig. 2.2A). We have previously reported that sequestration effects can efficiently suppress signalling even upon strong stimulation, and that suppression is relieved in an ultrasensitive manner once the signalling intermediate concentration exceeds that of the sequestration species ('ultrasensitization') [107]. Such signal suppression effects due to Ras sequestration might explain why transfection with constitutively active GEF constructs alone fails to transform cells. In Figs. 2.1 and 2.2, the GEF- and GAP-catalysed reactions were modelled using the Michaelis-Menten approximation, which implies that sequestration effects due to Ras sequestration were neglected. In contrast, the simulations in Fig. 2.3 were obtained using the elementary-step description of the GEF- and GAP-catalysed reactions, and thus take sequestration of Ras in enzyme-substrate complexes into account. It can be seen in Fig. 2.3A that neither Ras overexpression (point A $\rightarrow$ point C) nor GEF activation (point A $\rightarrow$ point B) alone has a significant effect on steady state Ras signalling, while both effects strongly cooperate (point D). Thus, the sequestration model can explain the experimentally observed logical AND-gate between GEF activation and Ras overexpression (as also summarised in Fig. 2.3B). This is due to the fact that signal suppression for low Ras expression levels and ultrasensitization for higher Ras expression levels (indicated in Fig. 2.3A) enhance the increase in sensitivity upon Ras overexpression which arises from zero-order ultrasensitivity (compare point A $\rightarrow$ point B with point C $\rightarrow$ point D in Figs. 2.2A and 2.3A). Experimental measurements revealed that RasGAP proteins are highly abundant at least in some cells, and thus corroborate the physiological relevance of sequestration effects [108].
2.5 Conclusions

Oncogenes such as Ras are frequently overexpressed in tumours, e.g., due to genomic amplification or due to promoter deregulation [109]. However, in many cases, including Ras, the relevance of such overexpression remains unclear (e.g., [7,8,9]). This chapter was focussed on the steady state basal activation level in an enzymatic futile cycle. Intuitively, one would expect that the basal activation level is proportional to the total concentration of the cycling substrate. Thus, strong overexpression might be expected to result in strongly enhanced basal signalling, and possibly in tumourigenesis. The modelling analyses revealed that the kinetic properties of enzymatic futile cycles can prevent deregulation of signal transduction by substrate overexpression, and thus identified a possible mechanism for tumour suppression.

The differential equations of the model are based on the assumption that spontaneous Ras (de)activation occurs with first-order kinetics, while GEF- and GAP-catalysed reactions were assumed to proceed with Michaelis-Menten kinetics. Therefore, the results presented above also apply for covalent modification cycles (e.g., phosphorylation / dephosphorylation cycles) which might explain why mutational activation of oncogenic kinases is typically required to transform cells, and/or to activate downstream signalling, while even strong overexpression of wildtype proteins is insufficient [110,111,112,113,114,115].

Experimental studies support that the proposed tumour suppression mechanism is physiologically relevant for protein kinase cascades: (i) Experiments with kinase or phosphatase inhibitors in starved cells [116,117] suggest that the basal state activation level of cascade intermediates is mainly determined by the ratio of kinase and phosphatase activities. In other words, the basal state seems to be controlled by nonlinear Michaelis-Menten kinetics, and not simply by linear auto(de)phosphorylation. (ii) Protein kinase cascade intermediates are typically expressed at concentrations close to (or even above) the \( K_m \)-values of the upstream kinase. This implies that the kinetic requirements for insensitivity (Section 2.3) are often fulfilled in kinase cascades. (iii) Quantitative measurements revealed that more than 5% of total molecules of cascade intermediates are frequently activated even in serum-depleted medium [57,81,117,118,119,120]. Accordingly, even strong stimulation often induces only about 10-fold increases in kinase activation [19,25,57,117,121], and biological responses can be induced by a 10-fold kinase activation in cell culture [57,122] and in vivo [123]. Thus, 10 – 20 fold overexpression of kinase cascade intermediates should be sufficient to induce biological responses unless the suppression effects such as those discussed in this chapter avoid deregulation of signalling by protein overexpression.

Few theoretical studies thoroughly analysed basal state signalling so far [81,124], even though it is the background activity (and not the dynamics) of signalling pathways is deregulated in many diseases. It is becoming increasingly clear that basal signal transduction serves to control important functions such as cell survival [125], gene expression [126], and cell adhesion [127]. The simulations in Sections 2.2 and 2.3 reveal how cells robustly maintain such background activity of signalling pathways, independent of fluctuations in protein expression. Other mechanisms that have been reported for robust signal transduction include assembly into multisubunit protein complexes [128], negative feedback [129], and co-expression of antagonistic enzymes [130]. Moreover, it is known that transcription factors often act as repressors in the inactive state, while being transcriptional activators in the active state (e.g., [131]). Competition of active and inactive states thus gives rise to robustness towards fluctuations in total transcription factor protein expression.
Figure 2.3: Implications for oncogene cooperation.

(A) Effects of Ras overexpression on signal transduction for varying upstream stimulus levels in the "sequestration model" that takes Ras sequestration in enzyme-substrate complexes into account. The GEF- and GAP-catalysed reactions are modelled using an elementary step description of the irreversible Michaelis-Menten mechanism (E + S ↔ ES → E + P; Appendix B), and the intrinsic reaction steps (i.e., the grey arrows in Fig. 2.1 A) were neglected. The RasGTP concentration is plotted as a function of total Ras expression (i.e., the sum of RasGTP, RasGDP, the Ras-GEF complex and the Ras-GAP complex), and the total GEF concentration is varied. The total GEF concentrations are $3 \times 10^{-8}$, $3 \times 10^{-7}$ and $3 \times 10^{-6}$ mol/l (this corresponds to $[\text{GEF}] / [\text{GAP}] = 0.04$, 0.4 and 4). Note that the curve for $[\text{GEF}] / [\text{GAP}] = 4$ also applies for higher stimulus levels ($[\text{GEF}] / [\text{GAP}] > 4$), and thus represents the maximal effect induced by strong stimulation. The Michaelis-Menten constants of the GEF reactions were set to $K_M,\text{GDP} = 3.86 \times 10^{-6}$ mol/l and $K_M,\text{GTP} = 3 \times 10^{-6}$ mol/l, and the $k_{\text{cat}},\text{GEF}$, the $K_M,\text{GAP}$, and the $k_{\text{cat}},\text{GAP}$ were as in [10] . The undetermined off-rates of the GEF- and GAP-Ras complexes were estimated to be $10 \text{ s}^{-1}$ (Appendix B). The experimentally measured Ras expression level in living cells is indicated by the vertical dashed line. Points A – D are described in B.

(B) Schematic representation of the logical AND-gate for oncogene cooperation. Normal cells (point A in panel A) remain completely untransformed by both, GEF activation (point B in panel A) and Ras overexpression (point C in panel A), but they become primed for strong transformation by a secondary mutational event (point D in panel A).

The robustness mechanism (“kinetic tumour suppression”) presented here is similar to that discussed for bacterial two-component systems [132,133]. However, two component systems are topologically different from mammalian signalling cascades, as a single phosphate group is transferred from upstream proteins to downstream effectors, while no such mass transfer is observed in mammalian systems. Using an experimentally validated model of Ras signalling, it is shown in this chapter that robustness is observed for parameter values that are physiologically relevant for Ras signalling in mammalian cells. Moreover, the simulations reveal that the robustness in mammalian cells is restricted to weak stimulation, while it is observed for the full dose-response in two-component systems [133]. Figures 2.2 and 2.3 indicate that loss of robustness upon strong stimulation might give rise to priming and oncogene cooperation effects in mammalian cancer development. The physiological relevance of priming and cooperation is further supported by a study focussing on cellular transformation by Ras and its downstream effector Raf: Transfection with either wildtype protein did not induce any significant phenotypic response, while strong transformation upon co-transfection of Ras and Raf [115]. The easiest way to test the model predictions regarding tumour suppression and oncogene cooperation is to mix recombinant Ras, GAP and GEF proteins in vitro, and to measure the steady state concentration of RasGTP as a readout.

This chapter was focussed on the pathological deregulation of basal state signal transduction. More specifically, it was analysed how long-term alterations (e.g., at the transcriptional or the genomic level) affect information transfer via the basic motif of intracellular signalling networks, the activation-deactivation cycle. The following two chapters
(Section 3 and 4) deal with the question of how transient physiological signals are converted into persistent cell-fate decisions by the intracellular signalling network in the absence of slow transcriptional regulation. More specifically, the analysis is focussed on two signal transduction pathways (MAPK signalling and apoptosis signalling) that have been shown experimentally to remain irreversibly activated even after input stimuli were removed. In Sections 5 and 6, it is then investigated how slow transcriptional feedback regulation affects the signals generated by the signalling network. Therefore, Sections 5 and 6 extend the analyses presented here, as slow transcriptional regulation is no longer assumed to be an 'external' event, but is rather a part of the signalling network.
3 Competing docking interactions can bring about bistability in the MAPK cascade

Parts of this chapter are published in [22].

SYNOPSIS

MAP kinases are crucial regulators of various cell fate decisions including proliferation, differentiation and apoptosis. Depending on the cellular context, the Raf-Mek-Erk MAPK cascade responds to extracellular stimuli in an all-or-none manner, most likely due to bistable behaviour. Here, a previously unrecognised positive feedback mechanism is described that emerges from experimentally observed sequestration effects in the core Raf-Mek-Erk cascade. Un-/monophosphorylated Erk sequesters Mek into Raf-inaccessible complexes upon weak stimulation, and thereby inhibits cascade activation. Mek, once phosphorylated by Raf, triggers Erk phosphorylation, which in turn induces dissociation of Raf-inaccessible Mek–Erk heterodimers, and thus further amplifies Mek phosphorylation. It is shown that this positive circuit can bring about bistability for parameter values that were experimentally measured in living cells. Previous studies revealed that bistability can also arise from enzyme depletion effects in the Erk double (de)phosphorylation cycle. It is demonstrate that the feedback mechanism proposed in this chapter synergises with such enzyme depletion effects to bring about a much larger bistable range than either mechanism alone. The results show that stable docking interactions and competition effects, which are common in protein kinase cascades, can result in sequestration-based feedback, and thus can have profound effects on the qualitative behaviour of signalling pathways.

3.1 Introduction

The three-tiered mitogen-activated protein kinase (MAPK) pathways are known to be crucial regulators of various physiological processes such as proliferation, differentiation, senescence, and apoptosis [2]. Cell-fate decisions such as differentiation are thought to occur in an all-or-none fashion and, once initiated, should be stably maintained in an irreversible manner. Theoretical and experimental work [134] suggested that such switch-like and irreversible signal transduction could arise due to bistability at the level of MAPK activation.

Single-cell measurements confirm that both the Raf-Mek-Erk pathway and the JNK cascade are indeed activated in all-or-none manner in Xenopus oocytes [135,136]. Additionally, switch-like activation was recently shown to occur in the yeast mating MAPK signalling module [137,138]. In mammalian systems, all-or-none activation of the Raf-Mek-Erk pathway was observed in T cells [139], in BHK cells [26], in PC12-D2R cells [140], in dopaminergic SN4741 neurons [140], and in Hek 293 cells (Boris Kholodenko, pers. communication). In contrast, gradual MAPK activation at the single-cell level was seen in growth-factor-stimulated Swiss 3T3 fibroblasts [25], in HeLa cells [141], and in human foreskin fibroblasts [141]. The qualitative behaviour of Erk activation can also depend on the stimulus strength, being all-or-none at weak stimulation, but gradual upon strong stimulation in LpT2 gonadotrope cells [142]. Finally, the opposite, i.e., a gradual response upon weak stimulation and all-or-none activation at high stimulus levels, was reported to occur in PC12 cells [24].

These single cell measurements reveal that MAPK cascades frequently exhibit all-or-none behaviour over a broad range of stimulus concentrations, which suggests that these pathways can indeed be bistable under physiological conditions. Bistable systems display hysteresis, meaning that different stimulus-response curves are obtained depending upon
whether the system began in its off or its on state [20,143]. Experimental studies confirmed hysteresis for the JNK cascade in Xenopus oocytes [135] and for the Raf-Mek-Erk pathway in PC12 neuronal precursor cells [24]. Bistability is thought to require a positive signalling circuit, which may be established either by feedback activation of upstream pathway intermediates or by relief from upstream pathway inhibition [20,143]. All-or-none activation in the Mos-Mek-Erk MAPK cascade was indeed shown to depend on feedback pathway activation in Xenopus oocytes: Active phosphi-Erk stimulates transcription and thereby upregulation of the constitutively active Raf homolog Mos, which is the uppermost member of the cascade [136]. Similar positive feedback loops, which rely on Erk-dependent Raf activation, have been proposed to exist in mammalian cells [41,134].

However, direct experimental evidence for functionally significant positive feedback loops is scarce. Transfection with constitutively active Mek seemed to activate Raf-1 in NIH3T3 cells [144] and in Hek 293 cells [145]. In contrast, constitutively active mutants of Raf, Mek or Erk failed to activate their endogenous counterparts when exogenously expressed in C7 3T3 cells (Raf, [146]), in BHK cells (Mek, [26]), in Hek 293 cells (Erk, [113]), in PC12 cells (Erk, [113]) and in Cos7 cells (Erk, [147]). These data suggest that functionally relevant positive feedback activation is the exception rather than the rule in the mammalian Raf-Mek-Erk pathway. Additionally, positive feedback activation does not seem to correlate with all-or-none Erk activation at the single cell level, and is therefore unlikely to account for bistable behaviour in the mammalian MAPK cascade.

Instead, these overexpression data support a model, where bistability arises from a positive feedback circuit that relies on elimination of upstream cascade inhibition, and not on upstream cascade activation. Such relief-from-inhibition is expected to be insufficient for full pathway activation in the absence of upstream input signals, and would thereby explain why overexpressed constitutively active mutants of Raf, Mek or Erk failed to activate their endogenous counterparts. A recent study suggests that such relief from upstream inhibition occurs downstream of Raf kinase, i.e., within the core MAPK cascade: All-or-none activation of the MAPK cascade was observed even if cascade activation was triggered by an exogenously expressed Raf construct, which would most likely overcome endogenous feedback mechanisms acting upstream of Raf [26].

Recent theoretical studies indicated that implicit feedback and bistability can indeed arise in the core MAPK signalling module: Markevich et al. [23] described how relief from inhibition and hysteresis emerge in the basic motif of MAPK cascades, the double phosphorylation cycle, if realistic kinetic parameters are assumed. Additionally, it is shown in Section 4 that bistability due to relief from inhibition can be observed if two consecutive cascade members (e.g., Mek and Erk) are deactivated by the same phosphatase. This ‘shared phosphatase motif’ applies for the mammalian Erk-MAPK signalling module, as PP2A was reported to dephosphorylate both Mek and Erk [148]. However, these implicit mechanisms exhibit a relatively narrow range of bistability, and might thereby require amplification in order to bring about robust hysteresis in vivo.

In this chapter a previously unrecognised relief-from-inhibition feedback mechanism in the core Raf-Mek-Erk cascade is described. Un-/monophosphorylated Erk is proposed to sequester Mek into Raf-inaccessible complexes upon weak stimulation, and thereby inhibits the cascade. Mek, once phosphorylated by Raf, triggers Erk phosphorylation, which in turn induces dissociation of Raf-inaccessible Mek–Erk heterodimers (relief from inhibition), and thus further stimulates Mek phosphorylation. The suggested mechanism is in accord with experimental studies, which showed that Mek and Erk form a stable complex under resting conditions, and dissociate upon Erk phosphorylation (see, e.g., [27]). The simulations reveal that this positive circuit can bring about bistability for parameter values that were experimentally measured in living cells [149]. Additionally, it is demonstrated that the feedback mechanism proposed in this paper synergises with that implicit in double
phosphorylation [23] to bring about a much larger bistable range than either mechanism alone.

3.2 Rationale and model implementation

Rationale: Experimental studies revealed that unphosphorylated Mek and Erk form a stable complex in unstimulated cells, which dissociates upon stimulation with growth factors (see, e.g., [27]). Such basal Mek~Erk association has been neglected in most mathematical models of the MAPK cascade. It was the aim of this work to study the impact of Mek-Erk complex formation in more detail.

Figure 3.1: Proposed Bistability Mechanism and Model Structure.

(A)-(C) Schematic representation of the proposed bistability mechanism. Upon weak stimulation (i.e., at low pRaf levels), Erk and Mek are mostly un-/monophosphorylated (indicated by white boxes), and pathway activation is suppressed by Mek sequestration into Raf-inaccessible (p)Mek-(p)Erk heterodimers (A). Stronger stimulation increases the amount of double phosphorylated ppMek, which then triggers Erk phosphorylation (B). Erk phosphorylation in turn induces dissociation of Raf-inaccessible Mek~Erk heterodimers (relief from inhibition), and thus amplifies Mek phosphorylation in a positive feedback circuit (B), so that finally the pathway is completely activated (C). (D) Schematic representation of model topology (see Appendix C for differential equations). The black arrows indicate the previously described 'basic model' that includes Raf-mediated Mek phosphorylation, Mek-mediated Erk phosphorylation and the antagonizing phosphatase reactions [150]. The 'sequestration model' analysed in this paper additionally includes association of various Mek species with un-/monophosphorylated Erk (grey arrows), and the resulting Mek sequestration complexes (i.e., Mek~Erk, Mek~pErk, pMek~Erk, and pMek~pErk).

Figures 3.1A-C show how a positive feedback circuit can arise in the core Raf-Mek-Erk cascade due to basal Mek-Erk association. At low levels of the phospho-Raf stimulus, Mek and Erk are mostly unphosphorylated or monophosphorylated (i.e., inactive) as indicated by the white (p)Mek and (p)Erk boxes in Fig. 3.1A. Stable heterodimer formation between inactive Mek and Erk molecules efficiently inhibits Raf-mediated Mek phosphorylation, and thereby suppresses pathway activation. Stronger stimulation results in the formation of some fully phosphorylated Mek and Erk as indicated by the black ppMek and ppErk boxes in Fig. 3.1B. This depletes non-/monophosphorylated Erk pools, and thereby promotes the
dissociation of the Raf-inaccessible (p)Mek~(p)Erk complexes (relief from inhibition). In other words, Mek, once fully phosphorylated, triggers the release of monomeric, Raf-accessible Mek and hence further Mek phosphorylation in a positive feedback circuit (Fig. 3.1B). Upon sufficiently strong stimulation, all (p)Mek~(p)Erk are dissociated and almost all Mek and Erk molecules are fully phosphorylated (Fig. 3.1C). Previous studies [23,151] suggested that such a relief-from-inhibition feedback mechanism could result in bistability. It will therefore be addressed in the following whether hysteresis can be observed in physiologically relevant parameter ranges.

Model implementation: A mathematical model of the core MAPK signalling module, schematically depicted in Fig. 3.1D, was implemented (see Appendix C for differential equations). Here, the black arrows indicate the previously described MAPK model (referred to as 'basic model' hereafter) that includes Raf-mediated Mek phosphorylation, Mek-mediated Erk phosphorylation and the antagonizing phosphatase reactions [150]. For simplicity, it was assumed that the phosphatases for Mek and Erk are less abundant than their substrates, and the corresponding phosphatase reactions were thus modelled using the Michaelis-Menten approximation. In contrast, kinase catalysis was modelled using the elementary step description in order to take possible sequestration and back-propagation effects into account [15,107,152]. Experimental studies revealed that Mek and Erk form stable heterodimers under basal conditions, and that these heterodimers dissociate upon stimulus-induced Erk phosphorylation [27]. To account for such complex formation, association of un-/monophosphorylated Erk (but not of bisphosphorylated Erk) with various Mek species was also considered in the ‘sequestration model’ (black and grey arrows in Fig. 3.1D).

In order to demonstrate that the proposed feedback mechanism is in fact responsible for bistability in the ‘sequestration model’, other possible sources of hysteresis were excluded (except for the results shown in Fig. 3.5). First, implicit positive feedback which could arise if both the Mek and the Erk cycles were deactivated by the same phosphatase [151] was eliminated. This was accomplished by assuming that Mek and Erk proteins are dephosphorylated by different phosphatases. Second, it was excluded that bistability implicitly arises in double phosphorylation [23,153] by assuming that similar reaction steps are characterised by the same kinetics. For example, the same kinetic parameters were assumed for the phosphorylation of the first and the second modification site in Mek (‘non-cooperative phosphorylation’). Similar non-cooperativity was also assumed for Mek dephosphorylation, Erk phosphorylation and Erk dephosphorylation. Finally, the same association and dissociation rate constants were assumed for all Mek~Erk complexes (i.e., Mek~Erk, pMek~Erk, ppMek~Erk, Mek~pErk, pMek~pErk, ppMek~pErk). The resulting model comprised 10 kinetic parameters and three total protein concentrations (pRaf, Mek and Erk), all of which could be taken from a recent quantitative experimental study by Fujioka et al. [149] (see Table 3.1).
3.3 Structural and kinetic requirements for bistability

*Bistability Due to Mek Sequestration*: Single-cell analyses revealed that the MAPK cascade can respond to extracellular stimuli in an all-or-none manner, most likely due to bistable behaviour of the system (see Section 3.1). Extracellular stimulation was simulated by varying the total concentration of active Raf protein (pRaf), and the amount of doubly phosphorylated, active Erk (ppErk) was analysed as the response. The system showed a simple, monostable stimulus-response (not shown) when simulations were run using the kinetic parameters and the protein concentrations, which were measured by Fujioka et al. [149] in HeLa cells (Table 3.1; column ‘Fujioka et al.’). This result is in accordance with experimental measurements, as gradual Erk activation at the single-cell level was demonstrated in EGF-treated HeLa cells [141].

**Table 3.1: Kinetic Parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Notes</th>
<th>Fujioka et al.</th>
<th>This Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mek_tot</td>
<td>Total Cellular Mek Concentration</td>
<td>1.4 µM</td>
<td>1 µM</td>
</tr>
<tr>
<td>Erk_tot</td>
<td>Total Cellular Erk Concentration</td>
<td>0.96 µM</td>
<td>10 µM</td>
</tr>
<tr>
<td>kon,Raf–Mek</td>
<td>Association Rate Constant of Raf–Mek complex</td>
<td>0.65 µM(^{-1}) s(^{-1})</td>
<td>0.65 µM(^{-1}) s(^{-1})</td>
</tr>
<tr>
<td>koff,Raf–Mek</td>
<td>Dissociation Rate Constant of Raf–Mek complex</td>
<td>0.065 s(^{-1})</td>
<td>0.065 s(^{-1})</td>
</tr>
<tr>
<td>kcat,Raf–Mek</td>
<td>Catalytic Turnover Constant of Raf–Mek complex</td>
<td>0.18 s(^{-1})</td>
<td>0.18 s(^{-1})</td>
</tr>
<tr>
<td>(V(<em>{\text{max}})/K(</em>{\text{M}}))Mek–PPase</td>
<td>First-order Rate Constant of Mek-Phosphatase</td>
<td>0.01 s(^{-1})</td>
<td>(0.01 s(^{-1}))</td>
</tr>
<tr>
<td>V(_{\text{max}}),Mek–PPase</td>
<td>Maximal Velocity of Mek-Phosphatase</td>
<td>-</td>
<td>0.001 µM s(^{-1})</td>
</tr>
<tr>
<td>K(_{\text{M}}),Mek–PPase</td>
<td>Michaelis-Menten Constant of Mek-Phosphatase</td>
<td>-</td>
<td>0.1 µM</td>
</tr>
<tr>
<td>kon,Mek–Erk</td>
<td>Association Rate Constant of Mek–Erk complex</td>
<td>0.88 µM(^{-1}) s(^{-1})</td>
<td>0.88 µM(^{-1}) s(^{-1})</td>
</tr>
<tr>
<td>koff,Mek–Erk</td>
<td>Dissociation Rate Constant of Mek–Erk complex</td>
<td>0.088 s(^{-1})</td>
<td>0.088 s(^{-1})</td>
</tr>
<tr>
<td>kcat,Mek–Erk</td>
<td>Catalytic Turnover Constant of Mek–Erk complex</td>
<td>0.22 s(^{-1})</td>
<td>0.22 s(^{-1})</td>
</tr>
<tr>
<td>(V(<em>{\text{max}})/K(</em>{\text{M}}))Erk–PPase</td>
<td>First-order Rate Constant of Erk-Phosphatase</td>
<td>0.014 s(^{-1})</td>
<td>(0.08 s(^{-1}))</td>
</tr>
<tr>
<td>V(_{\text{max}}),Erk–PPase</td>
<td>Maximal Velocity of Erk-Phosphatase</td>
<td>-</td>
<td>0.04 µM s(^{-1})</td>
</tr>
<tr>
<td>K(_{\text{M}}),Erk–PPase</td>
<td>Michaelis-Menten Constant of Erk-Phosphatase</td>
<td>-</td>
<td>0.5 µM</td>
</tr>
</tbody>
</table>

The total protein concentrations and the kinetic parameters of the model depicted in Fig. 3.1D are listed (column ‘This Study’), and compared to the values that were measured by Fujioka et al. [149]. Fujioka et al. estimated the apparent first-order rate constant (V\(_{\text{max}}\)/K\(_{\text{M}}\)) of Mek and Erk dephosphorylation only. Saturated Michaelis-Menten kinetics were assumed in the model, as: (i) the time course data in [149] indicates saturation in the dephosphorylation reactions; (ii) the K\(_{\text{M}}\)-values of phosphatases towards full-length substrates are frequently in the sub-micromolar range [154,155]. See Supplement for differential equations of the model.

The intracellular concentrations of Mek and Erk depend on the cellular context, and have been reported to be 0.6–40 µM [16,139] and 0.8–30 µM [150,156] in mammalian cells. Thus, bistability might still be observed in other cell types than HeLa cells, especially because phosphatase activity in the MAPK signalling module is known to be intensely regulated as well [157]. Importantly, bistability occurred within the physiologically relevant kinase/phosphatase concentration ranges. Figure 3.2A (grey line) shows a representative stimulus-response curve. Here, the total Erk concentration (10 µM) and the Mek-phosphatase activity were increased relative to the default values measured in HeLa cells, while the total Mek concentration (1 µM) and the Mek-phosphatase activity were kept essentially unchanged (Table 3.1, column ‘This study’).

Markevich et al. [23] reported that hysteresis can implicitly arise in double (de)phosphorylation cycles even in the absence of allosteric feedback. In order to exclude that their mechanism is responsible for the observed bistability, Mek sequestration was
eliminated from the model. The resulting basic model (only black arrows in Fig. 3.1D) exhibits a simple, monostable response, which demonstrates that Mek sequestration into Raf-inaccessible complexes (grey arrows in Fig. 3.1D) is indeed responsible for bistability. This conclusion also holds in general, i.e. regardless of the parameters chosen, because bistability in double phosphorylation cannot arise with non-cooperative protein (de)phosphorylation reactions [153], which was assumed in the default sequestration model (see Section 3.2).

Figure 3.2: Bistability Due to Mek Sequestration.

(A) Bistable stimulus-response of the core MAPK cascade. Extracellular stimulation was simulated by varying the total concentration of active Raf \((pRaf_{tot} = pRaf + pRaf\_Mek + pRaf\_pMek)\), and bisphosphorylated Erk \((ppErk)\) was taken as the response. The black curve corresponds to the previously analysed basic model (black solid arrows in Fig. 3.1D), while the grey stimulus-response was obtained for the sequestration model, which additionally takes Mek sequestration by Erk into account (black solid and grey arrows in Fig. 3.1D). Kinetic parameters are given in Table 3.1 (column 'This study'). (B) Mek release from inactive sequestration complexes upon cascade activation. The amount of sequestered Mek \((i.e., Mek\_Erk + Mek\_pErk + pMek\_Erk + pMek\_pErk)\) and the total amount of bisphosphorylated Mek \((i.e., ppMek + ppMek\_Erk + ppMek\_pErk)\) is shown as function of total active Raf for the kinetic parameters given in Table 3.1 (column ‘This study’).

The amount of sequestered Mek \((i.e., Mek\_Erk + Mek\_pErk + pMek\_Erk + pMek\_pErk)\) was analysed in more detail to further corroborate that the positive circuit described in ‘Rationale’ is responsible for the hysteresis. The corresponding simulations confirm a pronounced Mek release from Raf-inaccessible sequestration complexes upon switching from the lower to the upper steady state branch (Fig. 3.2B, grey line). This Mek release relieves the cascade from strong inhibition (compare black and grey lines in Fig. 3.2A), and allows for coordinated activation of Mek and Erk (grey lines in Fig. 3.2A and B). Taken together, these data reveal that Mek sequestration into Raf-inaccessible complexes and its subsequent ppMek-dependent release are responsible for hysteresis in Fig. 3.2.

**Kinetic Requirements for Bistability:** MAPK activation was shown to proceed in an all-or-none manner in some, but not in all cells (see Section 3.1). In order to get insights into such cell-type specific behaviour, the requirements for bistability in terms of protein expression and kinetic parameters were investigated.

The impact of alterations in kinase expression was analysed by classifying the stimulus-response curves (similar to those in Fig. 3.2A) for varying total Mek and Erk concentrations into monostable and bistable. Figure 3.3A (grey area) shows that the stimulus-response is bistable over a relatively broad range of Mek and Erk expression levels, which match those previously measured experimentally (see above). Hysteresis seems to require that the Erk concentration exceeds that of Mek, as the bistable range is bounded by the dashed line in
Fig. 3.3A, which corresponds to equal Mek and Erk expression. Excess of Erk ensures efficient Mek sequestration into Raf-inaccessible complexes upon weak stimulation, and thereby strengthens the relief-from-inhibition feedback mechanism discussed in Section 3.2. Experimental studies confirmed that Erk is indeed more abundant than Mek in a variety of mammalian cell lines including CHO cells (Erk/Mek = 2.15 [16]), Cos-1 cells (Erk/Mek = 2.03 [158]), Cos7 cells (Erk/Mek >> 1 [147]), 208F cells (Erk/Mek = 2.5 [158]), NIH3T3 cells (Erk/Mek = 12.86 [158]), and Rat1 cells (Erk/Mek = 1.5 [158]). Additionally, the yeast homologues Kss1p and Fus3p were reported to significantly exceed their shared upstream activator, the Mek homologue Ste7p ((Kss1p + Fus3p)/Ste7p > 5.71 [16]).

![Figure 3.3: Kinetic Requirements for Bistability.](image)

(A) Bifurcation diagram for alterations in kinase expression. The stimulus-response curves of the default model were calculated for varying total Mek and Erk concentrations, and were then classified into monostable (white area) and bistable (grey area). The dashed line corresponds to equal Mek and Erk expression. Default indicates the parameter set given in Table 3.1 (column ‘This Study’). Point I indicates the situation, where the Mek concentration is low relative to that of the Erk phosphatase, so that Erk activation is completely abolished. Point II corresponds to a cell, which expresses high levels of Mek relative to Erk phosphatase. This provokes strong Erk activation before the Mek cycle is switched on, and therefore excludes coordinated activation of both kinases in a positive feedback circuit. (B) Bifurcation diagram for alterations in phosphatase expression. Similar to (A), but bistable behaviour was analysed for varying maximal velocities (i.e., varying expression) of the phosphatases that dephosphorylate Mek and Erk. See (A) for explanation of points I and II. Point III indicates the situation, where strong Mek phosphatase expression necessitates high levels of active Raf to elicit Mek phosphorylation. Under these conditions Mek is strongly sequestered by active Raf, and this abolishes hysteresis.

The group of Gertraud Müller from the University of Stuttgart quantified the intracellular Erk concentration in Rat1 cells. They found $2.3 \cdot 10^6$ molecules per Rat1 cell using Western blotting and a calibration curve of recombinant GST-Erk fusion proteins as described previously [150]. Assuming a cell volume of 1 pl [16] and a Erk:Mek ratio of 1.5 in these cells [158], one arrives at $\text{Mek}_{\text{tot}} = 2.56 \mu\text{M}$ and $\text{Erk}_{\text{tot}} = 3.83 \mu\text{M}$. These values lie within the range of bistability (Figure 3.3A), and therefore further corroborate the physiological relevance of the implicit feedback mechanism discussed in this paper.
Figure 3.4: Structural Requirements for Bistability.

The sequestration model depicted in Fig. 3.1D was extended in order to study the topological constraints for bistability. More specifically, Raf-mediated phosphorylation of Erk-bound Mek (i.e., non-competitive binding of Raf and Erk to Mek) was taken into account. Additionally, ppErk binding to Mek, pMek and ppMek (i.e., by product inhibition in Erk phosphorylation) was considered. The stimulus-response curves of the resulting extended sequestration model (see Supplement for differential equations) were classified into monostable and bistable for varying degrees of competition and product inhibition. The competition factor, \( c \), equals the fold-change in Raf’s affinity for Mek brought about by Erk binding to Mek (and vice versa). Likewise, the product inhibition factor, \( p \), quantifies how the affinity between Erk and Mek is altered by Erk double phosphorylation (relative to un-/monophosphorylated Erk). The grey bistability range was calculated using the default parameters given in Table 3.1 (column ‘This study’). The dashed black line indicates the bistable-monostable transition for a 10-fold lower Michaelis-Menten constant of the Mek phosphatase (\( K_{M,Mek-PPase} = 0.01 \mu M \)).

It was also analysed how altered Mek- and Erk-phosphatase expression (i.e., changes in the corresponding \( V_{max} \)-values) affect the qualitative behaviour of the stimulus-response curve, and it turned out that bistability is retained over a relatively broad range of phosphatase concentrations (Fig. 3.3B). The bifurcation analysis with respect to kinase and phosphatase expression (Figs. 3.3A and B) reveal several kinetic constraints for the existence of a bistable stimulus-response: (i) Too low Mek concentrations or too high Erk-phosphatase concentrations abolish any significant Erk activation and thereby also hysteresis (points I in Figs. 3.3A and B). (ii) Too high Mek levels or too low Erk-phosphatase levels provoke strong Erk activation before the Mek cycle is switched on, and therefore exclude coordinated activation of both kinases in a positive circuit (points II in Figs. 3.3A and B). (iii) Strong Mek-phosphatase expression necessitates high levels of active Raf to elicit Mek phosphorylation. Under these conditions Mek sequestration by active Raf becomes significant and this abolishes bistability, because Mek activation is both subsensitive [15] and submaximal [107] (point III in Fig. 3.3B). (iv) Bistability requires that the Mek and the Erk concentration exceed the dissociation constant \( K_d \) of the Raf-inaccessible \((p)Mek-(p)Erk\) sequestration complexes, since otherwise Mek sequestration is relatively inefficient. Experimental evidence suggests that this requirement holds in living cells, as the measured dissociation constant \( K_d = 30-300 \text{ nM} \) [149,159,160] is indeed lower than typical intracellular Mek and Erk levels.

Structural Requirements for Bistability: Two key topological assumptions were made when deriving the model depicted in Fig. 3.1D. First, it was assumed that Mek and Erk no longer associate once Erk has been fully phosphorylated by Mek, that is, product inhibition of ppMek-mediated Erk phosphorylation was neglected. This seems justified, since it has been shown that Mek and Erk form stable heterodimers under basal conditions, and that these heterodimers dissociate almost completely upon stimulus-induced Erk phosphorylation [27]. Second, it was assumed that Erk and Raf bind to Mek in a mutually exclusive manner (i.e.,
Experimental studies revealed that Raf associates with a C-terminal domain in Mek [161], while Erk is recruited to the N-terminus of Mek [162]. Importantly, the C- and the N-terminal domains adjoin to each other in the Mek crystal structure [163], which suggests competitive binding, especially because both Raf (74 kD) and Erk (44 kD) are relatively bulky and are known to homodimerise. Competition of Raf and Erk for Mek is further suggested by the fact that the Mek proline-rich domain, which adjoins to the Mek C-terminus [163], has been implicated in both Raf and Erk recruitment [164,165]. Finally, mutually exclusive binding is also supported by biochemical analyses of the JNK pathway which showed that the Raf homolog, Mekk-1, competes with JNK for binding to the Mek homolog, JNKK1 [166].

Figure 3.5: Synergism of Bistability Mechanisms.

A broad range of bistability is observed in the stimulus-response (curve 2) of the sequestration model (black and grey arrows in Fig. 3.1D) if the second step of Mek-mediated Erk phosphorylation proceeds faster than the first ('positive cooperativity'). Such pronounced hysteresis can be explained by synergism of the feedback mechanism discussed in this paper with that described by Markevich et al. [23], which arises from enzyme depletion effects in the Erk cycle. The grey lines correspond to the stimulus-response curves of reduced models, where one of the two feedback mechanisms was eliminated, and thereby directly demonstrate such synergism. Curve 1 depicts the stimulus-response of the basic model (black arrows in Fig. 3.1D), which is devoid of Mek sequestration into Raf-inaccessible complexes. Curve 3 corresponds to a sequestration model (black and grey arrows in Fig. 3.1D), where positive cooperativity and enzyme depletion effects in the Erk cycle are eliminated. See Appendix C for differential equations and kinetic parameters.

Taken together, these data suggest that the scheme depicted in Fig. 3.1D applies for the core MAPK cascade. However, scaffold proteins, which bring kinases and their substrates into close proximity, allow for cascade activation even if otherwise essential docking interactions are absent [162], and might thus alleviate competition effects. To characterise the topological requirements of the proposed bistability mechanism more generally, an ‘extended sequestration model’ was implemented. The extended sequestration model includes Raf-mediated phosphorylation of Erk-bound Mek (i.e., non-competitive Raf and Erk binding to Mek) as well as ppErk binding to Mek, pMek and ppMek (see Supplement). Nine additional molecular species (compared to Fig. 3.1D) are considered in the extended model: six ternary Raf-Mek-Erk complexes (i.e., pRaf~Mek~Erk, pRaf~pMek~Erk, pRaf~Mek~pErk, pRaf~pMek~pErk, pRaf~Mek~ppErk, and pRaf~pMek~ppErk) arising from non-competitive Raf and Erk binding to Mek, and three Mek~ppErk heterodimers (i.e., Mek~ppErk, pMek~ppErk, and ppMek~ppErk), which can be considered as product inhibition complexes in Erk phosphorylation. Raf is assumed to catalyse Mek phosphorylation within the ternary
Raf-Mek-Erk complexes (e.g., pRaf~Mek~Erk → pRaf + pMek~Erk), so that Mek sequestration by inactive Erk no longer prevents cascade activation.

Figure 3.4 shows the qualitative behaviour of the extended sequestration model for varying degrees of competition and product inhibition. It can be seen that bistability is abolished in the extended sequestration model if competition between Raf and Erk for Mek is too weak, and if product inhibition in the Erk phosphorylation becomes significant. The competition factor, c, equals the fold-change in Raf’s affinity for Mek brought about by Erk binding to Mek (and vice versa). Thus, hysteresis requires that Erk binding to Mek decreases the affinity between Raf and Mek (and vice versa) at least by a factor of 25 (c < 0.04). Likewise, the product inhibition factor, p, quantifies how the affinity between Erk and Mek is altered by Erk double phosphorylation (relative to un-/monophosphorylated Erk). According to Fig. 3.4, ppErk must have a 40-fold (p < 0.025) lower affinity for Mek than its un-/monophosphorylated precursors.

These above constraints for p and c can be relaxed if one assumes kinetic parameters which differ from those given in Table 1 (column ‘This study’). The dotted line in Fig. 3.4 shows for example how the bistable-monostable border of the extended sequestration model is shifted if the Michaelis-Menten constant of the Mek-phosphatase is assumed to be 10-fold less than in Table 1 (i.e., K_{M,Mek-PPase} = 0.01 µM). Such strong phosphatase saturation, which increases zero-order ultrasensitivity in the Mek cycle [14], allows the system to filter out leakage from Raf-inaccessible sequestration complexes, and thereby broadens the range of bistability.

In conclusion, it was shown in this section that significant competition between Raf and Erk for Mek, and pronounced release of doubly phosphorylated Erk from Mek is required for bistability to be observed.

### 3.4 Synergism of Bistability Mechanisms

Markevich et al. [23] reported that hysteresis can implicitly arise in double (de)phosphorylation cycles if the kinetic parameters for the first and the second phosphorylation site differ (‘kinetic asymmetry’). More specifically, hysteresis is favoured if the kinase (ppMek) has significantly higher affinity for the unphosphorylated substrate (Erk) than for the monophosphorylated substrate (tyrosine-phosphorylated Erk) [153]. Experimental studies revealed that such kinetic asymmetry may occur in the Erk (de)phosphorylation cycle, as monophosphorylated Erk seems to have weaker affinity for ppMek than unphosphorylated Erk [27].

It was therefore analysed whether the bistability mechanism proposed in this chapter and that described by Markevich et al. [23] synergise to yield a larger bistable region than either mechanism alone. Simulations were done using the model structure depicted in Fig. 3.1D. In contrast to the default model (see ‘Model Implementation’), it is now assumed that ppMek-mediated Erk phosphorylation occurs with positive cooperativity. More specifically, the first phosphorylation step (Erk → pErk) is modelled with a low Michaelis-Menten constant, but with a slow catalytic rate constant. On the contrary, a higher Michaelis-Menten constant and a much faster catalytic rate constant were assume for the second phosphorylation step (pErk → ppErk) as this should favour bistability in the Erk cycle [153]. These parameters yield a narrow bistable range for the basic MAPK model, which neglects Mek sequestration (Fig. 3.5, curve 3). Hysteresis of curve 3 in Fig. 3.5 can be attributed to the mechanism described by Markevich et al. [23]. Figure 3.5 (curve 2) shows that this narrow bistable region is strongly enlarged if Mek sequestration into Raf-inaccessible complexes is additionally taken into account. Thus, both feedback mechanisms in combination bring about much more pronounced bistability than enzyme depletion effects in double phosphorylation alone. It was also analysed how the bistable range of the sequestration model is affected if kinetic asymmetry in Erk phosphorylation is eliminated from
the model. This was done by assuming equal catalytic rate constants for the first and second phosphorylation steps. As expected, the bistable range got significantly narrower once kinetic asymmetry was removed from the model (Fig. 3.5, curve 1).

Taken together, it was shown that the bistability mechanism proposed in this paper and that described by Markevich et al. [23] synergise to yield a much larger bistable region than either mechanism alone. Generally, the bistable range due to Mek sequestration, which was analysed for the non-cooperative system in Figs. 3.2-3.4, can be enlarged if kinetic asymmetries in the Mek and/or Erk (de)phosphorylation cycles are taken into account. In this context, it has recently been discussed that strong positive cooperativity occurs in Raf-mediated Mek phosphorylation [153]. Finally, Figure 3.5 demonstrates that positive circuits, which are bistable in isolation, cooperate to bring about even more pronounced bistability when combined in a network of nested positive circuits. Thus, bistability due to Mek sequestration might be even further enhanced by outer positive feedback circuits, which act at or upstream of Raf.

3.5 Conclusions

In this paper, it was shown that bistability is caused by an implicit positive feedback circuit that emerges from the network structure of the core MAPK cascade: Un-/monophosphorylated Erk sequesters Mek into Raf-inaccessible complexes upon weak stimulation, and thereby inhibits the cascade (see Fig. 3.1A). Mek, once phosphorylated by Raf, triggers Erk phosphorylation, which in turn induces dissociation of Raf-inaccessible Mek-Erk heterodimers (relief from inhibition) and thus further Mek phosphorylation (Fig. 3.1B-C). The suggested mechanism is in accord with experimental studies, which showed that Mek and Erk form a stable complex under resting conditions, and dissociate upon Erk phosphorylation (see, e.g., [27]). Positive feedback due to Mek sequestration can bring about bistability for experimentally measured parameters [149], and is expected to enhance ultrasensitive behaviour of the MAPK signalling module outside the bistable range [19]. Experimental studies suggest that the MAPK cascade is frequently bistable even though overexpression of constitutively active Raf, Mek or Erk mutants does not result in positive feedback activation of their endogenous counterparts (see Section 3.1). The relief-from-inhibition mechanism discussed in this chapter resolves this apparent contradiction (Section 3.1). Markevich et al. [23] demonstrated that relief from inhibition and bistability can arise in double phosphorylation cycles, but hysteresis was restricted to a relatively narrow parameter range. It is shown here that implicit feedback in double phosphorylation and feedback due to Mek sequestration synergise to yield a significantly larger bistable region than either mechanism alone.

Feedback due to Mek sequestration can be verified experimentally by initial velocity analysis of Raf-mediated Mek phosphorylation in vitro. The proposed feedback mechanism requires that un-/monophosphorylated Erk, but not bisphosphorylated Erk acts as a competitive inhibitor of Raf-mediated Mek phosphorylation. Competitive inhibition can be shown by analyzing Lineweaver-Burk plots for varying Erk concentrations (see Biophysical Textbooks). Kinase-defective Mek and Erk mutants should be used in these assays in order to prevent Mek-mediated Erk phosphorylation and Erk-mediated feedback phosphorylation of Mek. Significant competitive inhibition of Raf-mediated Mek phosphorylation should be seen with unphosphorylated Erk as a competitor, but not with bisphosphorylated Erk. Bistability due to Mek sequestration can be directly proven in an in vitro reconstitution system. Mek, a Mek-phosphatase (e.g., PP2A), Erk, and an Erk-phosphatase (e.g., MKP3) should be incubated with varying amounts of active Raf, and the stimulus-response is expected to exhibit true all-or-none behaviour (Fig. 3.2A). Hysteresis can be shown by varying the time of phosphatase addition to the system: (i) a weak response is expected within the bistable range if the Mek/Erk-phosphatases are added simultaneously with Raf to Mek and Erk; (ii) a strong
response will be observed if the Mek/Erk-phosphatases are added after Mek and Erk have been fully activated by Raf.

Recent theoretical studies revealed that sequestration-based feedback (i.e., feedback without explicit allosteric regulation) might be a common principle in signal transduction, and that it allows for bistable [23,151,167] or oscillatory behaviour [138,168]. Feedback emerges in these systems due to high affinity protein-protein interactions, which appear to be an ubiquitous and robust means to achieve nonlinear behaviour in biochemical networks [32,151,169]. Sequestration-based feedback also requires that protein-protein interactions are competitive at least to some extent, because otherwise the bound protein can still participate in other cellular reactions (i.e., it cannot be sequestered). It seems likely that sequestration-based feedback is a common feature of protein kinase cascades: Enzyme-substrate binding in these cascades is generally mediated by relatively stable docking/domain interactions in addition to transient recognition by the enzyme's active site [170]. Additionally, cascade intermediates frequently engage a single binding site to recruit upstream kinases, phosphatases and downstream substrates in a competitive manner [170,171]. The results presented in this paper demonstrate that such competition effects profoundly affect the qualitative behaviour of protein kinase cascades. Scaffold proteins, which bring kinases and their substrates into close proximity, allow for cascade activation even if otherwise essential docking/domain interactions are absent [162]. It is tempting to speculate that scaffold proteins might alleviate competition effects within the MAPK kinase cascade, and thereby regulate the qualitative behaviour of the stimulus-response (monostable vs. bistable).

The simulations in this chapter revealed how irreversibility can arise in the MAPK cascade, and the results generally apply for protein kinase cascades that operate via reversible (de)phosphorylation. Some signalling cascades such as the Notch pathway or the caspase pathway involved in apoptosis are activated by irreversible proteolysis of signalling intermediates, and their regulatory logic is thus fundamentally different from protein kinase cascades. Apoptosis, a regulated form of cell suicide, is activated in an all-or-none and irreversible manner, which means that cells, once committed to apoptosis, fully proceed through the death program. Commitment is thought to arise at the level of caspase activation. The requirements for bistability to be observed in a mathematical model of caspase regulation is analysed in Section 4, and it is shown that an implicit positive feedback similar to that discussed in the present chapter arises in the apoptosis cascade operating by irreversible proteolysis.
4 Simultaneous inhibition of initiator and effector caspases by XIAP establishes implicit bistability in caspase activation

Parts of this chapter are published in [151].

SYNOPSIS

The intrinsic, or mitochondrial, pathway of caspase activation is essential for apoptosis induction by various stimuli including cytotoxic stress. It depends on the cellular context, whether cytochrome c (cyto c) released from mitochondria induces caspase activation gradually or in an all-or-none fashion, and whether caspase activation irreversibly commits cells to apoptosis. By theoretical derivation and analysis of a quantitative kinetic model, it is shown that simultaneous inhibition of caspase-3 (Casp3) and caspase-9 (Casp9) by inhibitors of apoptosis (IAPs) results in an implicit positive feedback, since cleaved Casp3 augments its own activation by sequestering IAPs away from Casp9. It is demonstrated that this positive feedback brings about bistability (i.e., all-or-none behaviour), and that it cooperates with Casp3-mediated feedback cleavage of Casp9 to generate irreversibility in caspase activation. The calculations also unravel how cell-specific protein expression brings about the observed qualitative differences in caspase activation (gradual vs. all-or-none and reversible vs. irreversible). Finally, known regulators of the pathway are shown to efficiently shift the apoptotic threshold stimulus, suggesting that the bistable caspase cascade computes multiple inputs into an all-or-none caspase output. As cellular inhibitory proteins (e.g., IAPs) frequently inhibit consecutive intermediates in cellular signalling cascades (e.g., Casp3 and Casp9), the feedback mechanism described in this paper is likely to be a widespread principle how cells achieve ultrasensitivity, bistability and irreversibility.

4.1 Introduction

Apoptosis, an evolutionary conserved form of cell suicide, allows multicellular organisms to eliminate damaged or excess cells in order to maintain tissue homeostasis. Deregulation of apoptosis is associated with various pathological conditions including cancer and neurodegenerative disorders. Aspartate-specific cysteine proteases, also known as caspases, are the central executioners of apoptosis. In most cases, apoptotic stimuli activate initiator caspases, whose substrates, the effector caspases, ultimately cause cellular demise by cleaving various cellular substrates [172]. Figure 4.1A schematically depicts the so-called extrinsic and intrinsic apoptotic pathways that elicit apoptosis by cleaving and thereby activating Casp3, the major cellular effector caspase. The extrinsic pathway is initiated by ligand-binding to death receptors (e.g., CD95), which then oligomerise and recruit various proteins including pro-Casp8 into the so-called death-inducing signalling complex (DISC). Formation of the DISC complex leads to autoprocessing of pro-Casp8 into active (initiator) Casp8, which then cleaves (effector) Casp3. Cytotoxic stress or death-receptor-stimulated Casp8 engage the intrinsic, or mitochondrial, apoptosis pathway by inducing the translocation of pro-apoptotic Bcl-2 family members such as Bax and Bid to mitochondria. This event, which is negatively regulated by anti-apoptotic Bcl-2 family members (e.g., Bcl-2), results in the release of pro-apoptotic proteins (cyto c and SMAC) from mitochondria into the cytosol. Cytosolic cyto c then elicits the oligomerisation of Apaf-1 into an active high-molecular weight complex, the apoptosome, which recruits and stimulates (initiator) Casp9, and thereby allows activation of effector caspases such as Casp3. SMAC and inhibitors of apoptosis such as XIAP establish an additional layer of regulation in the intrinsic pathway: XIAP inhibits the catalytic activities of Casp9 and Casp3 through reversible binding, and cytosolic SMAC relieves this inhibition by sequestering XIAP away from caspases [11].
Experimental studies revealed that the qualitative behaviour of caspase activation in the intrinsic pathway depends on the cellular context. Cyto c added to cytosolic extracts activates Casp3 in an all-or-none fashion in some cells [173,174,175,176,177], while gradual activation was observed in other systems [178,179,180]. As cyto c release from mitochondria can be a reversible event [181], which does not affect mitochondrial function [182,183,184], it has been suggested that downstream caspase activation irreversibly commits cells to apoptosis [31,185]. Accordingly, cyto c-induced Casp3 activation remained elevated even after a strong decline in cytosolic cyto c [186] or after apo-cyto c, an inhibitor of apoptosome formation, was added [187]. Furthermore, the time course of caspase activation via the intrinsic pathway equals that for irreversible commitment to apoptosis [31,185], and caspase-inhibition allows for long-term cellular recovery and/or proliferation after removal of apoptotic stimuli [31,185,188,189,190,191]. Finally, Fas-treated Jurkat T cells, which enter apoptosis by the intrinsic pathway, escaped commitment to death as judged by maintenance of clonogenic potential if Casp3 was inhibited [192]. On the contrary, Casp3 activation was found to be a reversible event in glycochenodeoxycholate-treated hepatocytes [193].

These qualitative differences in caspase activation suggest that the intrinsic pathway is bistable in some cells, but monostable in others. While simple monostable systems respond in a gradual and reversible manner, complex bistable systems exhibit true all-or-none responses and in some cases irreversibility. Bistability is thought to require a positive circuit, which may be established either by positive feedback or by double negative feedback. Once a threshold stimulus is exceeded such positive circuits allow bistable systems to switch from low activation levels (off state) to high activation levels (on state) in an all-or-none fashion. Bistable systems display hysteresis, meaning that different stimulus-response curves are obtained depending upon whether the system began in its off or its on state. In some cases, the on state is maintained indefinitely after the stimulus is removed, so that the system shows irreversible activation [20]. Experimental studies confirmed that bistability indeed occurs in natural and artificial biological networks [20,194,195,196,197].

Recent mathematical modelling demonstrated that bistability can arise from ‘hidden’, or implicit, feedback loops that are usually not explicitly drawn in biochemical reaction schemes [23,198]. Similarly, it is shown here that inhibition of Casp3 and Casp9 by IAPs results in an implicit positive feedback and in bistability. As cellular inhibitory proteins (e.g., IAPs) frequently inhibit consecutive intermediates in cellular signalling cascades (e.g., Casp3 and Casp9), the mechanism described in this paper is likely to be a widespread principle how cells achieve ultrasensitivity, bistability and irreversibility (Appendix D).

### 4.2 Model derivation

Based on the published literature, a core model of the intrinsic apoptosis pathway was derived, which includes general regulatory mechanisms, while cell-type-specific events were not taken into account. The grey-shaded area in Fig. 4.1A indicates the regulatory interactions considered in the model: Active Apaf-1, which was taken as the input in most simulations, recruits and thereby stimulates (initiator) Casp9. Casp9 then in turn activates the output species, (effector) Casp3, by proteolytic processing. Additionally, Casp3-mediated cleavage of Casp9 results in positive feedback amplification. Finally, both Casp3 and Casp9 are subject to stoichiometric inhibition by IAPs. For simplicity, only the most potent caspase inhibitor among the IAP family of proteins, the X-linked inhibitor of apoptosis (XIAP), was considered. The corresponding kinetic scheme is depicted in Fig. 4.1B.
Fig. 4.1: Mathematical Model of the Intrinsic Apoptosis Pathway.

(A) Schematic representation of intrinsic and extrinsic apoptosis pathways. Dotted lines indicate positive regulation (arrows) or negative regulation (blunted arrows), and the solid lines refer to release of SMAC and cyto c from mitochondria. The regulatory interactions considered in the model are highlighted in grey. The numbers 1-4 refer to additional feedbacks described in Section 4.6 (B) Kinetic scheme of the model, where the following abbreviations were used: A* = activated Apaf-1, C9 = Casp9, C3 = Casp3, and X = XIAP. The reactions depicted in grey, which are involved in Casp3-mediated feedback cleavage of Casp9, were eliminated in the Casp9-mutant model in order to dissect the role of XIAP-mediated feedback.

Cyto c released from mitochondria is known to elicit heptamerisation of Apaf-1 into active apoptosomes. As detailed kinetic measurements of apoptosome formation are currently lacking, apoptotic stimulation was modelled by altering the total concentration of activated Apaf-1 molecules assembled in apoptosomes ($A_{tot}^*$ = $A^* + A^*C9 + A^*C9X + A^*C9^* + A^*C9^*X$). Each active Apaf-1 monomer assembled in apoptosomes was shown to reversibly bind to a single Casp9 molecule [199], and Casp9 is then auto-proteolytically processed at amino acid Asp-315 [200]. Importantly, Casp9 auto-proteolysis neither affects enzymatic activity of Casp9 [201], nor its recruitment to apoptosomes [202,203]. Owing to these data, the model did not distinguish between auto-proteolytically processed and unprocessed...
Casp9. The enzymatic activity of Casp9 is thought to be mainly determined by apoptosome recruitment, as apoptosome-bound Casp9 was shown to be much more active than free Casp9 [204,205]. It was therefore assumed in the model that reversible association of Casp9 (C9) and Apaf-1 (A*) (reaction 1) yields a highly active Apaf1-Casp9 complex (A*-C9), which cleaves pro-Casp3 (C3) much more efficiently (reaction 3) than free Casp9 (reaction 2; see Table 4.1). The latter reaction was nevertheless taken into account, since free Casp9 was shown to have significant basal activity towards pro-Casp3 [199].

Processing of pro-Casp3 into mature Casp3 by upstream initiator caspases such as Casp9 was reported to occur by a sequential two-step mechanism: Pro-Casp3, which has negligible enzymatic activity [206], is initially processed by Casp9 into active p12-p20-Casp3, and this intermediate is subsequently auto-catalytically cleaved into active p12-p17-Casp3 [207]. As shown in Fig. 4.1B, Casp3 activation was modelled by a single-step mechanism (C3 → C3*). This seems justified, as the p12-p20-Casp3 intermediate and mature p12-p17-Casp3 exhibit similar catalytic activities [208], and as they are both subject to inhibition by XIAP (see below).

Casp3 is known to cleave its own activator, Casp9, at amino acid Asp-330 in vitro [201,209], and in cytosolic extracts treated with cyt c [175,200]. As Casp9 processing by Casp3 was shown to significantly enhance Casp9 activity [201], feedback cleavage by Casp3 results in auto-amplification of the apoptotic signal. The physiological relevance of this positive feedback loop was confirmed in several studies, which showed that inhibition of Casp3-mediated cleavage of Casp9 prevented full activation of both Casp3 and Casp9 in response to cyt c [201,210,211]. Casp3-mediated feedback processing of Casp9 was modelled by assuming that active Casp3 (C3*) cleaves both free and Apaf1-associated Casp9 (reactions 4 and 5), thereby generating the Asp330-cleaved Casp9 species, C9* and C9*-A*. These feedback-cleaved Casp9 species in turn cleave pro-Casp3 more efficiently (reactions 6 and 7; see Table 4.1) when compared to their precursors, C9 and C9-A*, thus establishing a feedback amplification loop. Feedback-processed Casp9 (cleaved at Asp330) was shown to be associated with apoptosomes [201,203], much like its precursors that are not cleaved at Asp330 (see above). It was therefore assumed in the model that the kinetics of Casp9-binding to Apaf-1 (reactions 1 and 8) are unaffected by Casp3-mediated feedback cleavage (see Table 4.1).

Inhibitors of apoptosis such as XIAP act as stoichiometric inhibitors of Casp3 and Casp9 [11], and accordingly caspase inhibition can be described by simple reversible binding [212,213]. Experimental evidence suggests that XIAP can bind to and inhibit Casp9, even if the latter is associated with apoptosomes [201]. Accordingly, it was assumed in the model that active Apaf-1 (A*) and XIAP (X) bind to Casp9 in a non-competitive manner, so that Apaf1-bound Casp9 intermediates (A*C9 and A*C9*) recruit XIAP with the same kinetics as free Casp9 (C9 and C9*). Additionally, XIAP binding to Casp9 was modelled such that it is neither affected by Casp9 autocleavage (at Asp-315), nor by Casp3-mediated feedback cleavage (at Asp-330). As contradictory experimental results were obtained on how Casp9 cleavage modulates inhibition by XIAP, the impact of the latter assumption is stressed in Section 4.6.
Table 4.1: Kinetic Parameters.

<table>
<thead>
<tr>
<th>#</th>
<th>Reaction</th>
<th>( k_+ )</th>
<th>( k_- )</th>
<th>( k_+ / k_- )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( C_9 + A^* \leftrightarrow A^*C_9 )</td>
<td>( 2 \times 10^{-3} ) nM s(^{-1} )</td>
<td>0.1 s(^{-1} )</td>
<td>50 nM ([199] )</td>
</tr>
<tr>
<td>2</td>
<td>( C_9 + C_9 \rightarrow C_9^* + C_9 )</td>
<td>( 5 \times 10^{-7} ) nM s(^{-1} ) ([199] )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>( C_3 + A^<em>C_9 \rightarrow C_3^</em> + A^*C_9 )</td>
<td>( 3.5 \times 10^{-5} ) nM s(^{-1} ) (70-times #2 ([204] ) )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>( C_9 + C_3 \rightarrow C_9^* + C_3^* )</td>
<td>( 2 \times 10^{-7} ) nM s(^{-1} ) ([214] )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>( A^<em>C_9 + C_3 \rightarrow A^<em>C_9^</em> + C_3^</em> )</td>
<td>( 2 \times 10^{-7} ) nM s(^{-1} ) (same as #4)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>( C_3 + C_9 \rightarrow C_9^* )</td>
<td>( 5 \times 10^{-7} ) nM s(^{-1} ) (10-times #2 ([201] ) )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>( A^<em>C_9 + C_3^</em> \rightarrow A^<em>C_9^</em> + C_3 )</td>
<td>( 3.5 \times 10^{-7} ) nM s(^{-1} ) (10-times #3 ([201] ) )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>( C_9^* + A^* \leftrightarrow A^<em>C_9^</em> )</td>
<td>( 5 \times 10^{-7} ) nM s(^{-1} ) ([201] )</td>
<td>0.1 s(^{-1} )</td>
<td>50 nM [199]</td>
</tr>
<tr>
<td>9</td>
<td>( C_9 + X \leftrightarrow C_9X )</td>
<td>( 10^{-3} ) s(^{-1} )</td>
<td>( 10^{-3} ) s(^{-1} )</td>
<td>0.02 nM s(^{-1} ) (adjusted)</td>
</tr>
<tr>
<td>10</td>
<td>( A^*C_9 + X \leftrightarrow A^*C_9X )</td>
<td>( 10^{-3} ) s(^{-1} ) ( )</td>
<td>( 10^{-3} ) s(^{-1} ) ( )</td>
<td>10^{-3} s(^{-1} ) ( )</td>
</tr>
<tr>
<td>11</td>
<td>( C_9^* + X \leftrightarrow C_9^*X )</td>
<td>( 10^{-3} ) s(^{-1} ) ( )</td>
<td>( 10^{-3} ) s(^{-1} ) ( )</td>
<td>10^{-3} s(^{-1} ) ( )</td>
</tr>
<tr>
<td>12</td>
<td>( A^<em>C_9^</em> + X \leftrightarrow A^*C_9^*X )</td>
<td>( 10^{-3} ) s(^{-1} ) ( )</td>
<td>( 10^{-3} ) s(^{-1} ) ( )</td>
<td>10^{-3} s(^{-1} ) ( )</td>
</tr>
<tr>
<td>13</td>
<td>( C_3^* + X \leftrightarrow C_3^*X )</td>
<td>( 3 \times 10^{-3} ) s(^{-1} ) ([212] )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>( C_3^* + X \leftrightarrow C_3^*X )</td>
<td>( 3 \times 10^{-3} ) s(^{-1} ) ([212] )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>( A^* \leftrightarrow 10^{-3} ) s(^{-1} )</td>
<td>0.02 nM s(^{-1} ) (adjusted)</td>
<td>20 nM [218]</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>( C_3^* \leftrightarrow 10^{-3} ) s(^{-1} )</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>( C_9 \leftrightarrow 10^{-3} ) s(^{-1} )</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>( C_9X \leftrightarrow 10^{-3} ) s(^{-1} )</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>( A^*C_9X \leftrightarrow A^*C_9X )</td>
<td>( 10^{-3} ) s(^{-1} ) ( )</td>
<td>( 10^{-3} ) s(^{-1} ) ( )</td>
<td>( 10^{-3} ) s(^{-1} ) ( )</td>
</tr>
</tbody>
</table>

The reactions numbered according to Fig. 4.1B (Column ‘#’) are listed, and the corresponding reactants and products are indicated (Column ‘Reaction’). The column ‘\( k_+ \)’ contains the rate constants of the reactions from left to right, and the column ‘\( k_- \)’ those for the opposite direction. The ratio of \( k_+ / k_- \) equals the dissociation constant for reversible bimolecular reactions, while it refers to the steady state protein concentration for synthesis and degradation reactions (#16-18 and #22). Similar reactions were assumed to proceed with same kinetics as indicated below the parameter values. References for parameters chosen are indicated in parentheses (e.g., \[1\]).

Owing to the assumptions made in the previous paragraph, there is reversible recruitment of XIAP to all Casp9 species in the model (reactions 9-12), and also free exchange of Apaf1 between the resulting Casp9-XIAP complexes (reactions 13-14). All Casp9-XIAP complexes were assumed to be catalytically inactive, which is in accordance with experimental studies \[221,222\] . Furthermore, Casp3-mediated feedback processing of XIAP-bound Casp9 was neglected in the model, as the Casp9-XIAP binding interface is nearby the corresponding cleavage site (Asp-330) \[222\] . It is well established that XIAP binds to both partially processed Casp3 (p12-p20) and to mature Casp3 (p12-p17), but not to its inactive precursor pro-Casp3 \[206\] . In accordance with experimental data \[212,213\] , reversible association between Casp3 and XIAP (reaction 15) was modelled to result in a catalytically inactive complex (C3*X). Due to the enzymatic inactivity of pro-Casp3 (C3) \[206\] and of the Casp3-XIAP complex (C3*X), free active Casp3 (C3*) was taken as the response in the simulations.

Finally, protein synthesis and degradation was included in the model (reactions 16 – 28). More specifically, the unmodified proteins A*, C9, X, and C3 are produced with a constant rate, and all molecular species in Fig. 4.1B are subject to first-order degradation. While the total cellular concentrations of Apaf-1, Casp9, Casp3 and XIAP (i.e., the ratio of protein synthesis and degradation rates) have been measured \[32,202,218,219,220\] , the kinetics of synthesis and degradation are not known. For simplicity, the same degradation rate was assumed for all molecular species in the model, and adjusted the synthesis rates in order to obtain previously measured protein concentrations (Table 4.1). This implies that the total concentrations of Apaf-1, Casp9, Casp3 and XIAP remain constant throughout the simulations.
From the model described above (Fig. 4.1B), which will be referred to as the ‘wildtype model’ in the following, molecular balances could be derived for each considered molecular species resulting in a system of 13 ordinary differential equations (Appendix D). In general, protein-protein association (reactions 1, 4 - 6, 7, and 10 – 13 in Fig. 4.1B) was modelled as a reversible second-order process, and caspase-mediated cleavage (reactions 2, 3, 8, 9, 14, and 15 in Fig. 4.1B) as an irreversible second-order process. As many similar reactions (e.g., 1 and 13 in Fig. 4.1B) were assumed to proceed with the same kinetics (see Table 4.1), the model comprises 16 kinetic parameters. The unknown kinetic parameters were set to reasonable values (Table 4.1) in order to reproduce the previously reported time courses of caspase activation (see Section ‘Time Course of Casp3 Activation’).

Besides the wildtype model, two modified models were also analysed in order to get insights into the mechanisms that are responsible for bistability in caspase activation: (i) In the ‘Casp9-mutant model’, which comprises only the black reactions in Fig. 4.1B, Casp3-mediated feedback cleavage of Casp9 (reactions 8 and 9 in Fig. 4.1B) was eliminated from the wildtype model. (ii) Based on available experimental data (see Section 4.6), competitive (i.e., mutually exclusive) binding of Casp3 and Casp9 to XIAP was assumed in the wildtype model. By contrast, Casp3 and Casp9 were allowed to bind XIAP simultaneously in the ‘non-competitive model’, that is, the wildtype model was extended by four ternary Casp9-XIAP-Casp3 complexes (Appendix D).

4.3 Dynamic and steady state characteristics of caspase activation

Time Course of Casp3 Activation: Experiments in cytosolic extracts revealed that exogenously added cyto c induces maximal Casp3 cleavage within ~15 minutes in some cells [203], while completion takes longer (up to ~60 minutes) in other systems [175,176,200,224]. More specifically, the Casp3 cleavage seems to be fast upon strong stimulation, but slower if stimulation is weak [176,220,225]. It was investigated whether the model was able reproduce these observations if previously measured protein concentrations of Apaf-1 (20 nM), Casp9 (20 nM), Casp3 (200 nM), and XIAP (40nM) were assumed [32,202,218,219,220]. Exogenous addition of cyto c was simulated by a step-like increase in the total amount of active Apaf-1 monomers, A* tot, as cyto-c-induced apoptosome formation was reported to be a very rapid process [203,205]. Such a step-input is also expected to reflect input characteristics within living cells reasonably well, since cyto c release from mitochondria was shown to complete within 5 minutes [226,227].

The results shown in Fig. 4.2A reveal that the simulated time courses of caspase activation agree well with those measured experimentally, and that simulated response time is indeed inversely related to the stimulus strength (Fig. 4.2A). Full activation of all cellular Apaf-1 molecules (A* tot = 20 nM) elicits fast Casp3 activation, while a critical slowing down is observed near the threshold (A* tot ~ 3 nM; see Fig. 4.2B), as expected for a bistable system [32,196]. Notably, the slope of the time courses shown in Fig. 4.2A is only marginally affected by the onset time of caspase activation (i.e., by the stimulus level). This is in accordance with experimental results obtained in cyto-c-treated cytosolic extracts [175,203,224] and in single living cells [29], which showed that, once initiated, Casp3 activation is rapidly completed within less than 15 minutes.
Fig. 4.2: Dynamic and Steady State Behaviour of the Caspase Cascade.

(A) Time course of Casp3 activation (‘wildtype model’) upon a step-like increase in the amount of active Apaf-1 (\(A^*_\text{tot}\)) at \(t = 0\) from zero to the concentration indicated. (B) Steady state stimulus-response curves of the wildtype model (black line) and of the Casp9-mutant model (grey line), where Casp3-mediated feedback cleavage of Casp9 does not occur. Stable and unstable steady states are indicated by solid and dashed lines, respectively.

**Bistability in Caspase Activation:** Experimental evidence suggests that cyto c-induced caspase activation can be bistable and irreversible (see Section 4.1). The simulated steady state Casp3 activity (\(C^*_3\)) was indeed bistable and irreversible (Fig. 4.2B, black line). The system exhibits three steady states, two stable (solid black lines) and one unstable (dashed black line), for \(A^*_\text{tot}\) between 0 and \( \sim 3 \) nM, and shows hysteresis: Starting from the resting state (point 1) the system retains low Casp3 activity even for increasing stimuli, \(A^*_\text{tot}\), until a threshold (point 2) is reached, where Casp3 activity switches to the higher steady state (point 3) in an all-or-none fashion. The system remains at this higher steady state even if the stimulus is removed (point 4), so that caspase activation is irreversible, and thus represents the point of no return for apoptosis.

The mechanism of bistability was addressed next, and it was hypothesised that Casp3-mediated feedback cleavage of Casp9 was responsible, since bistability is thought to require a positive circuit [20]. Therefore, reactions 8 and 9 in Fig. 4.1B were blocked to simulate a mutant Casp9 (D330A), which is refractory to cleavage by Casp3 (‘Casp9-mutant model’). Unexpectedly, bistability was retained (Fig. 4.2, grey line), which suggests that a hidden positive feedback loop operates in the Casp9-mutant model.

### 4.4 Mechanism of bistability and kinetic requirements

**XIAP Establishes an Implicit Positive Feedback in Caspase Activation:** More detailed simulations revealed that XIAP establishes an implicit positive feedback in the Casp9-mutant model, and Fig. 4.3 schematically depicts how this mechanism contributes to irreversibility in the wildtype model: Upon weak stimulation (point 1 in Fig. 4.2B) the vast majority of Apaf1-associated, highly active Casp9 molecules is inhibited by excess XIAP, so that cleavage of pro-Casp3 is negligible (top left in Fig. 4.3). As the stimulus strength is increased above the threshold (point 2 in Fig. 4.2B), active Apaf-1 also recruits some free Casp9 that is not subject to inhibition by XIAP, so that Casp3 activation is initiated (top right in Fig. 4.3). Active Casp3 then further promotes its own activation by sequestering XIAP away from Apaf1-associated Casp9 (‘redistribution’), so that finally the vast majority of XIAP is bound to Casp3 (bottom right in Fig. 4.3). This XIAP redistribution results a positive feedback loop, which,
together with Casp3-mediated Casp9 feedback cleavage, suddenly switches the system from low to high Casp3 activity (transition from point 2 to point 3 in Fig. 4.2B). Caspase activity is maintained even if the stimulus is removed, as Casp3, once activated, retains XIAP, and thereby prevents full Casp9 deactivation (bottom left in Fig. 4.3). Additional simulations, which corroborate XIAP-mediated feedback can be found in Appendix D.

Fig. 4.3: Schematic Representation of XIAP-mediated Feedback.

At resting state (top left) Casp9 is efficiently inhibited by XIAP, so that Casp3 is inactive. Upon stronger stimulation (top right) some Casp9 escapes XIAP-mediated inhibition and activates Casp3, which then sequesters XIAP away from Casp9 (‘redistribution’). This XIAP redistribution finally results in strong activation of both Casp9 and Casp3 (bottom right), and retains the system in an active state even if the stimulus is reduced (bottom left). The numbers on the top of each scheme correspond to those indicated next to the stimulus-response in Fig. 4.2B (black line).

In order to determine how the protein concentrations in the caspase cascade affect bistability, the stimulus-response curves (similar to those in Fig. 4.2) were analysed for varying total Casp3 and Casp9 concentrations. Five types of qualitative behaviour in caspase activation could be distinguished in the physiological range of stimulus concentrations ($A^*_\text{tot} = 0-200$ nM): (i) the system is essentially devoid of any Casp3 activation (MN; Fig. 4.4A); (ii) Casp3 activation occurs in a gradual manner (MG; Fig. 4.4B); (iii) the caspase cascade is bistable and reversible (BR; Fig. 4.4C); (iv) Casp3 activation is bistable and irreversible (BI; Fig. 4.4D); (v) constitutive Casp3 activity is observed (MB; Fig. 4.4E). The corresponding bifurcation diagram (Fig. 4.4F) reveals that bistability in the Casp-9 mutant model can only be observed if the total Casp9 concentration is below that of XIAP (40 nM), which ensures that the system is in the off-state as long as Casp3 is inactive. Additionally, Casp3 must be
significantly more abundant than XIAP, in order to sequester it away from Casp9, i.e., to establish positive feedback.

**Fig. 4.4: Determinants for Bistability and Irreversibility**

The dose-response curves of the Casp9-mutant model (F), those of the wildtype model (G), and those obtained for non-competitive caspase binding to XIAP (H) were analysed for varying Casp3 and Casp9 expression levels. Five types of qualitative behaviour, which are schematically depicted in A-E, could be distinguished in the physiological range of Apaf-1 expression levels (0-200 nM). The light and dark grey areas in F-H correspond to the bistable regions of the model (BR, BI), and the abbreviations MN, MG and MB indicate the qualitative behaviour outside the bistable region. Experimentally measured caspase concentrations (see Table 4.1) are highlighted by dashed lines in F-H.

**Determinants for Bistability and Irreversibility:** The relative contribution of XIAP-mediated feedback and that of Casp3-mediated feedback cleavage (of Casp9) to bistability and irreversibility in caspase activation remained to be determined. To this end, the bifurcation plot of the wildtype model (Fig. 4.4G) was compared with those of mutant models, where either XIAP-mediated feedback (Fig. 4.4H; ‘Non-competitive model’) or Casp3-mediated feedback cleavage (Fig. 4.4F; ‘Casp9-mutant model’) was selectively blocked. XIAP-mediated feedback is abolished in the non-competitive model (Appendix D), since XIAP was assumed to be capable of simultaneous binding to Casp3 and Casp9 in these simulations. As schematically depicted above Fig. 4.4H, this corresponds to a caspase cascade, which is controlled by the XIAP fragments, BIR1-BIR2 (specific for Casp3) and BIR3-RING (specific for Casp9), rather than by full-length XIAP. Figures 4.4F and 4.4H demonstrate that each feedback mechanism alone can bring about bistability for experimentally measured caspase expression levels (interception of dashed lines in Figs. 4.4 F-H). By contrast, irreversibility is restricted to a narrow range of caspase concentrations in both mutant models, and is never observed in the vicinity of experimentally measured caspase expression levels. Importantly, the wildtype model exhibits robust irreversibility in the physiological range of caspase expression levels, which suggests that irreversibility in caspase activation requires coordinated action of both XIAP- and cleavage-mediated feedbacks.
Fig. 4.5: Determinants for Bistability and Irreversibility II.

The qualitative behaviour of caspase activation according to Figs. 4.4 A-E is shown as a function of the XIAP level, and of the competition ratio $\alpha$. The competition ratio $\alpha$ equals the fold-change in XIAP’s affinity for Casp9 brought about by Casp3 binding to XIAP (and vice versa), and thereby quantifies the degree of competitive caspase binding to XIAP as indicated on the top (see also Appendix D).

The computational results shown in Fig. 4.4G also explain why various cell types show qualitatively different patterns of caspase activation and unravel the underlying mechanisms: Casp3 activation is efficiently inhibited in cells, where the total XIAP concentration exceeds those of Casp3 and Casp9 (MN; Fig. 4.4A). Gradual Casp3 activation is predicted to occur in cells, where Casp9 expression is high compared to XIAP and Casp3 expression (MG; Fig. 4.4B). In this situation XIAP is effectively sequestered by excess Casp9, and the remaining free Casp9 molecules efficiently cleave Casp3 as if XIAP was not present. In case that both caspases are expressed at intermediate levels, the feedback loops discussed above cooperate to reversibly switch on the system in an all-or-none fashion (BR; Fig. 4.4D). Even higher caspase expression levels relieve the cascade from XIAP-mediated inhibition, so that Casp3 can be highly active even in the absence of stimulation. Such constitutive activation either arises spontaneously (MB; Fig. 4.4C) or it requires previous suprathreshold Casp3 activation (BI; Fig. 4.4E).

The preceding conclusions could be confirmed by analyzing the qualitative behaviour of caspase activation as a function of the competition ratio $\alpha$, and of XIAP expression (Fig. 4.5). The competition ratio $\alpha$ equals the fold-change in XIAP’s affinity for Casp9 brought about by Casp3 binding to XIAP (and vice versa), and thereby quantifies the degree of competitive caspase binding to XIAP (Appendix D). Figure 4.5 demonstrates that the range of bistability is significantly broadened even if the Casp3-binding to XIAP reduces XIAP’s affinity for Casp9 (and vice versa) less than 5-fold ($\alpha > 0.2$). By contrast, reliable irreversibility requires significant competition of caspases for XIAP at least with the default protein concentrations (Table 4.1) assumed here. As shown in Fig. 4.5, high XIAP levels completely abolish caspase activation (MN), bistability is observed for intermediate XIAP concentrations (BR,
BI), and low XIAP levels fail to prevent caspase activation even in the absence of external stimulation (MB).

The model simulations regarding the qualitative behaviour of caspase activation are supported by experimental data: (i) Overexpression of XIAP abolishes apoptosis and Casp3 activation in response to microinjection of cyto c (type MN) [228]. (ii) Overexpression of Casp3 [229,230] or Casp9 [200,231,232] results in caspase activation and/or apoptosis (type MB). In contrast, Casp3 overexpression failed to elicit its own activation in another study [233], and the model suggests that this may be due to low Casp9 expression (see Fig. 4.4G). (iii) High levels of IAP antagonists such as SMAC were shown to activate the Casp9 → Casp3 pathway [234,235] and to elicit spontaneous apoptosis [236] even in cell types that are devoid of basal cyto c release or Casp8 activation (type MB). The inability of others to reproduce Casp3 activation by XIAP depletion or SMAC addition [237] is probably due to the fact that the threshold BI → MB (Fig. 4.5) was not exceeded in these studies, e.g., due to the expression of SMAC-resistant IAP proteins such as NAIP [238]. (iv) Gradual Casp3 activation (type MG) was observed in cyto-c-treated cytosolic extracts [178,179,180], and also in flow cytometric analyses of living cells [239,240]. (v) The existence of bistable states (types BI and BR) is supported by all-or-none Casp3 activation in response to cyto c, and by the fact that Casp3 activation can irreversibly commit cells to death (see Section 4.1), although definitive proof for these types of behaviour is lacking (see Section 4.6).

4.5 The mitochondrial pathway acts as an efficient integrator of multiple regulatory inputs

In Section 4.4, it was demonstrated that excess of XIAP over Casp3 and Casp9 abolishes cyto-c-induced caspase activation even if high concentrations (200 nM) of the stimulus, active Apaf-1, were assumed (type MN). However, various experimental studies in cells, where Casp3 activation was inhibited downstream of cyto c release, have shown that caspase activation can be rescued by Apaf-1 overexpression (see [241] and references therein). This suggests that Casp3 activation does not occur if the concentration of the bottleneck, active Apaf-1, is below the threshold stimulus concentration, where the bistable system switches from the lower to the upper steady state (point 2 in Fig. 4.2B). In support for such a threshold model, it was recently shown that a minor (~ 2-fold) decrease in Apaf-1 expression dramatically decreases caspase activation in response to cyto c microinjection [241]. These studies also suggest that the apoptotic threshold can be regulated downstream of Apaf-1, as SMAC, an inhibitor of XIAP action, rescued cyto-c induced caspase activation in Apaf1-knock-down cells [241]. It was therefore investigated how the threshold of the bistable cascade is affected by transcriptional and post-transcriptional regulation of Casp3, Casp9 and/or XIAP. The corresponding simulation results are shown in Fig. 4.6: Starting from the default model (point of intersection), the predicted threshold stimuli, $A^{*}_{\text{tot,T}}$, of the bistable system were plotted as a function of Casp3 (grey dotted line), Casp9 (grey solid line), and XIAP (black solid line) expression. Additionally, simultaneous alterations of Casp3 and Casp9 to the same relative extent was also considered (black solid line) in order to understand how the apoptotic threshold is affected by nitric oxide (NO), a covalent inhibitor of Casp3 and Casp9 active sites [172]. These simulations demonstrate that decreasing levels of Casp3 moderately increase the threshold, $A^{*}_{\text{tot,T}}$, while alterations in Casp9 shift the threshold more efficiently. Regulation of XIAP levels is predicted to allow even more effective control over the apoptotic threshold, and similar arguments also hold for NO-mediated inhibition of both Casp3 and Casp9 [172].
The threshold stimulus, $A_{\text{tot,T}}^*$, where the bistable system switches from the lower to the higher steady state (point 2 in Fig. 4.2B), is plotted as a function of Casp3 (grey dotted line), Casp9 (grey solid line), or XIAP (black solid line) expression. Additionally, the impact of simultaneous alterations of Casp3 and Casp9 (black solid line) or of Casp3, Casp9 and XIAP (black dash-dotted line) to the same relative extent is shown. The intersection of the graphs corresponds to the default protein concentrations (see Table 4.1). The terms ‘linear’, ‘quadratic’ and ‘quartic’ indicate the relationship between protein expression and the apoptotic threshold, $A_{\text{tot,T}}^*$. The results regarding XIAP as an efficient modulator of an all-or-none threshold are in accordance with experimental studies, as a two-fold drop in XIAP expression was sufficient to allow cyto-c-induced Casp3 activation [242,243]. Moreover, increasing amounts of SMAC, a high affinity inhibitor of XIAP (Fig. 4.1B), elicited all-or-none Casp3 activation in cyto-c-treated HeLa cell cytosol [237]. Finally, the threshold cyto c concentration that is required to achieve switch-like Casp3 activation was shown to be cell-type-dependent, and low thresholds correlated with low IAP expression levels [177]. The simulation results are also corroborated by the fact that PKB/Akt-mediated inhibitory phosphorylation of Casp9 completely abolished cyto-c-induced Casp3 activation, even though Casp9 enzymatic activity was only partially suppressed [232]. Additionally, Casp3 overexpression sensitises cells to apoptosis in response to cytotoxic stress [233], which is also in accordance with the simulations shown in Fig. 4.6.

Thus, one can conclude that bistable behaviour in the mitochondrial caspase cascade serves to compute multiple regulatory inputs into a binary decision whether caspase activation occurs or not (‘binary integrator’). Further calculations, where relative changes in protein expression were related to relative changes in the threshold stimulus, $A_{\text{tot,T}}^*$, suggest that the following order of input potency holds in general: Regulation of active Apaf-1 < Casp3 regulation < Casp9 regulation < XIAP regulation ≈ Simultaneous regulation of Casp3 and Casp9. The simulations also predict that the apoptotic threshold is essentially constant if all components (i.e., Casp3, Casp9 and XIAP) are simultaneously changed to the same relative extent (Fig. 4.6; black dash-dotted line). Hence, the life-or-death decision appears to be remarkably insensitive towards random fluctuations in gene expression, which are thought to result in correlated changes in cellular protein levels [244]. Additionally, these simulations...
suggest that general inhibitors of protein synthesis or degradation, which are known to be inducers of apoptosis [226,233], do not affect the threshold of the Casp9 → Casp3 cascade.

4.6 Conclusions

In this Chapter, it was shown that inhibition of Casp3 and Casp9 by IAPs results in an implicit positive feedback, since cleaved Casp3 augments its own activation by sequestering IAPs away from Casp9 (Fig. 4.3). Additionally, the simulations revealed that XIAP-mediated feedback cooperates with Casp9 cleavage by Casp3 to bring about bistable and irreversible Casp3 activation in the range of experimentally measured kinetic parameters and protein concentrations (Figs. 4.2, 4.4 and 4.5).

Model Assumptions: XIAP-mediated feedback can only be observed if Casp3 and Casp9 compete for binding to XIAP at least to some extent (Figs. 4.3 and 4.5). Such competition is supported by the fact that Casp3 and Casp9 cannot be co-immunoprecipitated in cells [202]. Casp3 (and not only Casp9) is recruited to the apoptosome at least in some cells [202], and it is conceivable that this occurs by means of a sequential Apaf1-Casp9-XIAP-Casp3 complex. Even if such a complex exists, it seems to be rather instable, as Casp3 can be eluted from the apoptosome (i.e., from Apaf-1) by low ionic strength [206], while much higher ionic strength is required to elute Casp9 [199]. Recent co-immunoprecipitation experiments revealed the existence of a ternary Casp9-XIAP-Casp3 complex \textit{in vitro} [245]. However, only minor amounts of Casp3 were found in the complex even if XIAP was incubated with excess Casp3 and Casp9. Taken together, these data suggest that Casp3 and Casp9 significantly compete for binding to XIAP. Co-immunoprecipitation studies with Casp3, Casp9 and XIAP might underestimate the degree of competition of caspases for a single XIAP molecule (i.e., XIAP-mediated feedback), as IAP family members are often homodimers. In case that each XIAP molecule in a dimer independently couples to caspases, a ternary Casp9-XIAP-Casp3 complex will be seen, even if Casp3 and Casp9 compete for a single XIAP molecule. Therefore, it is most reasonable to directly test for XIAP-mediated feedback \textit{in vitro}. As further outlined in Appendix D, a Casp9 mutant (D330A), which is refractory to Casp3-mediated feedback cleavage, should be incubated with active apoptosomes and XIAP either in the presence or in the absence of pro-Casp3. Co-incubation with XIAP alone is expected to result in low Casp9 activity [201], but excess pro-Casp3 should reverse this inhibition by sequestering XIAP away from Casp9.

In the model it was also assumed in the model that XIAP inhibits all forms of Casp9, i.e., that the affinity between Casp9 and XIAP is neither affected by Casp9 autocleavage (at Asp-315) nor by Casp3-mediated feedback cleavage of Casp9 (at Asp-330). While it is clear that autoprocessed Casp9 (cleaved at Asp-315 only) is efficiently inhibited by XIAP [201,218,221,223,246], some authors reported that XIAP also binds to and inhibits uncleaved pro-Casp9 [201,221,247], at least partially [246,248], but others could not reproduce these results [218,223]. As explained in the context of Fig. 4.3, bistability requires that XIAP binds to and inhibits Apaf1-activated Casp9 upon weak stimulation, so that low Casp3 activity can be maintained. Importantly, such XIAP-mediated control over Casp9 activity will be ensured even if XIAP does not associate with uncleaved pro-Casp9, since pro-Casp9 recruitment to the apoptosome was shown to result in its fast and complete autoprocessing (at Asp-315) [205,206]. Casp3-mediated feedback cleavage (at Asp-330) was reported to relieve Casp9 from inhibition by XIAP [218], and might thereby establish an additional positive feedback, which would further broaden the ranges of bistability and irreversibility. As other experimental studies do not support the existence of this additional feedback [201,223], the conservative assumption was made that XIAP inhibits feedback-cleaved Casp9 as well.

The core model of the intrinsic pathway presented in this Chapter considers only Casp3 and XIAP, but not functionally redundant molecules. For example, Casp7, which is activated by
Casp9 [175] also mediates XIAP-mediated feedback, since it efficiently binds to IAPs [213]. Likewise, molecules such as c-IAP1, c-IAP2 and NAIP are functionally redundant to XIAP, as they inhibit both Casp3 and Casp9 [11,238]. In case that such functionally redundant proteins are expressed, the protein concentrations varied in the simulations (e.g., C3\text{tot} in Fig. 4.3F-H) represent combinations (e.g., sums) of functionally redundant protein concentrations (e.g., C3\text{tot} and C7\text{tot}), so that the results given in the paper continue to hold.

**Input Signals:** The concentration of active Apaf-1 assembled into apoptosomes was used as the varying input signal in the simulations, rather than the amount of cyto c released from mitochondria. This seems justified, as available experimental evidence suggests that apoptosome formation increases gradually with increasing cyto c concentration [180,204], and that signal amplification occurs in the caspase cascade considered in this Chapter [180]. The present model explains how cells reject erroneous cyto c release from single mitochondria, and also predicts that reversible cyto c release can elicit irreversible caspase activation. It should be noted that cyto c release upon apoptotic stimulation was reported to be all-or-none under many [226,227], but not all [249,250] circumstances. Importantly, dose-response curves using active Apaf-1 as the input (e.g., Fig. 4.2B) are physiologically relevant even if cyto c release is all-or-none, as they help to explain why caspase activation is completely abolished for limiting Apaf-1 expression (see [241] and references therein). More in general, the model provides insights into how the intrinsic pathway integrates multiple regulatory inputs including cyto c release, cyto c sequestration [206], transcriptional regulation of Apaf-1 [241], Apaf-1 sequestration [206], transcriptional regulation of IAPs [11], SMAC-mediated IAP sequestration [11], Casp9 phosphorylation [232], and caspase S-Nitrosylation [172]. As shown in Fig. 4.6, the caspase cascade acts as a binary integrator in the range of bistability (BI and BR in Figs. 4.4 and 4.5). In contrast, gradual integration will be seen if the system resides in the ‘monostable-gradual’ (MG) range, and this is particularly relevant for apoptotic stimuli that directly regulate caspase cascade members (e.g., Apaf-1) in addition to releasing cyto c (‘feed-forward regulation’). For example, p53 is known to induce Apaf-1 expression [251], and thereby can elicit gradual Casp3 activation even if cyto c release is all-or-none. Alternatively, gradual Casp3 activation, which was seen in flow cytometric analyses of living cells [239,240], may be due to cell-to-cell variability in the intrinsic pathway. Such cellular heterogeneity seems to be significant, as cyto c injection alone or in combination with SMAC does not elicit Casp3 activation [252] or cell death [228,253] in all cells of a population. The present model provides a reasonable basis for further studies that focus on cell-to-cell variability in the intrinsic pathway.

In Section 4.2 – 4.5, experimental studies were discussed, where SMAC, a competitive, high affinity inhibitor of IAP-binding to caspases [254], was either added to cytosolic extracts or microinjected into living cells. In living cells, SMAC is eventually released simultaneously with cyto c from mitochondria [11] (see Fig. 4.1A). Importantly, such physiological release of SMAC simply corresponds to decreasing XIAP levels in the model, as most experiments with caspase inhibitors have shown that SMAC release does not require caspase-mediated feedback [227,255,256,257]. Thus, the results shown in Fig. 4.6 explain why simultaneous release of cyto c and SMAC is required to elicit Casp3 activation in many cell types (e.g., [252]), and predict that these two stimuli are integrated in an all-or-none manner.

**Upstream, Downstream and Feedback Signalling:** In accordance with previous experimental studies (see Section 4.1), it was shown that, depending on the protein expression levels in the intrinsic pathway, caspase activation irreversibly commits cells to apoptosis (BI-regions in Figs. 4.4 and 4.5). However, some cells die by a delayed and morphologically distinct form of cell death, so-called caspase-independent cell death (CICD), even if caspases are inhibited [258]. As CICD is thought to be initiated at the level of mitochondria, the present simulations do not unravel the determinants for commitment to death in these cells, but only those for commitment to the fastest death pathway (i.e., apoptosis). As the precise kinetics of cell death may, for example, be important in development [258], the present results are likely to be relevant even in cells subject to CICD. The physiological importance of the caspase
cascade considered in the model is further supported by the fact that Apaf-1, Casp9 and Casp3 knockout mice show morphological defects and die early in development [258]. Additionally, caspase inhibition, e.g. due to IAP overexpression, allowed for long-term cellular survival and mitochondrial recovery in response to cytotoxic stress [188,189,190,191,192] and/or after cyto c was released [31,181,183].

Other positive feedbacks than those included in the model have been described in the literature. For example, Casp3 was shown to induce processing of Casp6, which in turn cleaves Casp8, an activator of Casp3 [175] (feedback 1 in Fig. 4.1A). This feedback is unlikely to account for bistable Casp3 activation via the intrinsic pathway, since Casp3 activation in response to cyto c is unaffected when the delayed Casp6 → Casp8 pathway is abrogated [175]. This conclusion is likely to hold in general, as Casp8 cleavage alone is not sufficient to stimulate its catalytic activity, but recruitment to the DISC complex (i.e., ligand-binding to death receptors) is required [231].

It has been suggested that active Casp3 amplifies cyto c release from mitochondria by directly cleaving upstream regulators such as Bid and Bcl-2 (feedbacks 2 and 3 in Fig. 4.1A), or by cleaving modulators of these Bcl2-family members such as Mekk1 [172]. However, the relevance of this feedback for the intrinsic pathway remains unclear, as experiments with caspase inhibitors revealed that cyto c release is caspase-independent in most cell types (e.g., [174,181,223,226,227,246]). Furthermore, the concept of Casp3-induced cyto c release is inconsistent with the fact that Casp3 activation fails in various cell types even though large amounts of cyto c were released from mitochondria (see [241] and references therein).

XIAP was shown to be cleaved by Casp3 and/or Casp8 in response to apoptotic stimulation, and such XIAP processing may result in auto-amplification of Casp3 activity (feedback 4 in Fig. 4.1A) [259,260]. In line with a predominant role of Casp8, cleavage of XIAP seems to be especially pronounced when cells are subjected to death-receptor stimulation [259,260]. By contrast, moderate [260], minor [202,203] or even no XIAP processing [261,262] was seen in response to apoptotic stimuli that initiate apoptosis via the intrinsic pathway. Additionally, Casp3 may also establish a positive feedback loop by cleaving inhibitors of XIAP auto-ubiquitination and proteasomal degradation such as PKB/Akt (feedback 4 in Fig. 4.1A) [263,264]. Accordingly, the total XIAP abundance was shown to decrease during apoptosis (e.g., [264]), but this seems to be a cell-type specific phenomenon, as the total amount of full-length XIAP remains essentially unchanged [260,261,262] or even increases [265] in other models of apoptosis.

Owing to these data and due to the fact that most molecular species of the caspase cascade were shown to be continuously synthesised during apoptosis [265,266], constant total protein concentrations were assumed in the model. In order to get insight into how Casp3-mediated XIAP degradation affects the behaviour of the model, an extended model was also implemented, which takes such regulation into account (Appendix D). Importantly, Casp3-mediated feedback cleavage of XIAP did not result in physiologically relevant bistability in a system devoid of other feedback amplification loops (Appendix D). Additionally, the qualitative conclusions drawn from Figs. 4.2, 4.4 and 4.5 were still valid when XIAP-mediated feedback was included in the wildtype model (Fig. 4.1B). However, these calculations also indicated that Casp3-mediated XIAP degradation may cooperate with the feedback loops discussed above, as it lowered the apoptotic threshold, A*_{tot,T}, and significantly broadened the range of XIAP concentrations, where caspase activation is irreversible (Bl in Fig. 4.5).

Active Casp3 cleaves a variety of cellular substrates, and thereby initiates the execution phase of apoptosis [172]. Experimental evidence suggests that Casp3 activates multiple execution pathways in parallel and not in a sequential, cascade-like manner, since mutational inactivation of Casp3 cleavage sites abrogates specific features of apoptosis depending on the target mutated [172]. Some Casp3 substrates (e.g., PARP) are cleaved...
almost simultaneously with Casp3, while the processing of others (e.g., Topo I) is delayed by several hours [267,268]. Taken together, these data suggest that transient activation of the branch point molecule, Casp3, elicits a partial apoptotic program, which might lead to potentially harmful cellular deregulation or tissue inflammation. Active Casp3 is known to be a rather unstable protein [269], which suggests that irreversible behaviour of the caspase cascade is required to maintain Casp3 activation if upstream stimuli are removed. Experimental evidence indeed suggests that such transient stimulation occurs in living cells: (i) Cyto c release from mitochondria is thought to be a reversible as long as mitochondrial membrane potential (MMP) is maintained. As the MMP can remain unchanged long after caspases have been activated [173,174], cytosolic cyto c (i.e., the stimulus) will decline as soon as the apoptotic trigger is removed. (ii) Experiments with antibodies towards the caspase-activating form of cyto c, holo-cyto c, revealed that holo-cyto c is rapidly degraded after its release into the cytosol [186]. The irreversibility mechanisms described in this chapter ensure that apoptosis will fully proceed even after a decline in cyto c, and render apoptotic execution program insensitive towards survival signalling once apoptosis has been initiated. Such insensitivity is then further enhanced by delayed Casp3-mediated cleavage and thereby inactivation of various anti-apoptotic signalling proteins [263].

Proposed Experimental Verification of Bistability: The predictions regarding all-or-none and binary integration of multiple inputs behaviour in caspase activation (Figs. 4.2-4.5) can be addressed experimentally by analyzing Casp3 activation in cytosolic extracts or on a single-cell level. In cytosolic extracts, depletion and re-addition experiments with various Apaf-1, Casp3, Casp9 and/or XIAP concentrations should result in all-or-none caspase activation in the BR and BI ranges in Fig. 4.4F, but the amount of fluorescent Casp3 substrates must be chosen carefully if enzymatic activity is used as a readout. Alternatively, such multivariate analyses can be performed by microinjecting these proteins together cyto c and/or SMAC into living cells. Caspase activation can then be determined using antibodies against active Casp3 either in flow cytometric measurements or in immunofluorescence microscopy. Bistability should be confirmed by adding cyto c in combination with appropriate antagonists such as anti-cyto c antibodies, apo-cyto c or Diarylureas, which are known to inhibit apoptosome activity [187,270]. In the range of bistability, simultaneous addition of suprathreshold cyto c levels and sufficient amounts of antagonist should yield low Casp3 activity, while strong caspase activation should be observed if the antagonist is added after cyto c. Subsequent addition of a Casp9 inhibitor would break the feedback loops discussed in the paper, and is therefore expected to reverse Casp3 activation. The bistability measurements described above can be done on a population level (i.e., by Western Blotting) if caspase activation is irreversible, but require single cell tracking methods, e.g., real-time Casp3 assays or flow-cytometric cell sorting, in the bistable-reversible range.

Concluding Remarks: In conclusion, a theoretical framework for quantitative experimental analyses of the intrinsic apoptosis pathway was presented. Previous mathematical models differ from the present study in (i) the choice of apoptotic pathways, (ii) the network properties focused on, (iii) the cell types analysed. Bentele et al. [271] and Eissing et al. [32] concentrated on the extrinsic apoptosis pathway (see Fig. 4.4.1A), and analysed how switch-like behaviour arises due to stoichiometric inhibition [271] or due to positive feedback [32]. Fussenegger et al. [272] have implemented a large-scale model of both intrinsic and extrinsic pathways, and analysed time course behaviour rather than bistability and apoptotic thresholds. Bagci et al. [273] focussed on how Casp3-mediated feedback cleavage of Bcl2-family members (feedbacks 2 and 3 in Fig. 4.1A) contributes to bistability in the intrinsic apoptosis pathway. As discussed above, these feedbacks appear to be restricted to particular cell types, where they might cooperate with those discussed here. Finally, Stucki and Simon [274] concentrated on the regulation of Casp3 degradation. The mechanisms proposed in this chapter may be combined with those discussed by Bagci et al. [273] and by Stucki and Simon [274] in order to implement more realistic models of the intrinsic apoptosis pathway.
As summarised in Appendix D, cellular inhibitory proteins such as stoichiometric inhibitors, phosphatases and GTPase-activating proteins frequently inhibit consecutive intermediates in cellular signalling cascades. In general, positive feedback and bistability can arise in this ‘shared inhibitor motif’ if: (i) the signalling intermediates compete for binding to the inhibitor at least to some extent; (ii) only the active form of the downstream intermediate (e.g., Casp3), but not its inactive precursor (pro-Casp3), binds to the inhibitor; (iii) the downstream intermediate (e.g., Casp3) is more abundant than the inhibitor (e.g., XIAP), which in turn needs to exceed the upstream intermediate (e.g., Casp9). As available experimental data is in accordance with these requirements, the feedback mechanism described in this paper is likely to be a widespread principle how cells achieve ultrasensitivity, bistability and irreversibility (Appendix D).

Sections 2, 3 and 4 mainly dealt with the regulation of signalling networks by fast post-translational mechanisms such as phosphorylation and dephosphorylation which operate on a time scale of minutes. In contrast, slow regulation by transcriptional induction or repression (operating on a time scale of hours) was excluded from the analysis or assumed to be so slow that it can be considered to be an external event. Many signalling pathways, however, remain activated over several hours (see Section 1.1 and Appendix A), so that slow transcriptional feedback regulation affect their dynamics. In Section 5, it is shown that transcriptional negative feedback regulation is a general design principle of mammalian signalling pathways. The transcriptional feedback regulation of TGFβ signalling is investigated in more detail in Section 6, and the functional relevance of transcriptional negative feedback is confirmed by siRNA-mediated knock-down of the SnoN oncoprotein.
Part II

Regulation of signal transduction by slow transcriptional feedback loops
5 Recurrent design patterns in the feedback regulation of the mammalian signalling network

This chapter is the result of a collaboration with Nils Blüthgen from the University of Manchester. Parts of this chapter are published in [1].

SYNOPSIS

Biochemical networks are characterised by recurrent patterns and motifs, but the design principles underlying the dynamics of the mammalian intracellular signalling network remain unclear. Decay rates of 134 signalling proteins were systematically analysed and their gene expression profiles in response to stimulation were investigated to get insights into transcriptional feedback regulation of signalling pathways. The analysis revealed a clear separation of the signalling pathways into flexible and static parts: for each pathway a subgroup of unstable signal inhibitors is transcriptionally induced upon stimulation, while the other constitutively expressed signalling proteins are long-lived. Kinetic modelling suggests that this design principle allows for swift feedback regulation and establishes latency phases after signalling, and that it might be an optimal design due to a trade-off between energy efficiency and flexibility.

5.1 Introduction

Signalling from the cell membrane to the nucleus typically occurs by post-translational regulatory mechanisms (such as phosphorylation), and thus occurs on a time scale of seconds to minutes. Within the nucleus, transcription factors are activated which induce changes in target gene expression on a time scale beyond 30 minutes. Some of the target genes are again involved in signalling, and thus slowly feed back into the signalling network. Given that such transcriptional regulation occurs on a time scale of hours, one might expect that transcriptional feedback regulation does not influence the dynamics of signalling networks. Experimental studies summarised in Appendix A reveal, however, that intracellular signalling pathways remain of active over several hours after stimulation. In other words, fast post-translational desensitisation mechanisms are often not strong enough to rapidly terminate the signal, so that transcriptional feedbacks are likely to play a role. Experiments with inhibitors of transcription and in cells harbouring reduced levels of signalling proteins due siRNA-mediated knock-down further support that transcriptional feedbacks affect the dynamics of intracellular signalling (Appendix A). Cellular decision making frequently requires ongoing signalling activity over several hours before the cell decides about commitment to a certain fate (Appendix A). This suggests that decision making in mammalian cells is profoundly affected by the slowly acting transcriptional feedback mechanisms.

Evolution gave rise to recurring patterns and motifs in biological networks, e.g., in signal-processing networks in bacteria, in metabolic networks, in neuronal networks and in ecological food-webs [62,130,275,276]. The recognition of these patterns helps to understand network optimisation principles, and to interpret the network structure. Despite its importance, no such general evolutionary footprints have been identified in the mammalian intracellular signalling network. In this chapter, design-principles in early transcriptional feedback regulation of mammalian signalling pathways is investigated.
Figure 5.1: Correlation between half-life and degree of induction after signal activation for signal transducers and signal inhibitors of five important signal transduction networks.

Shown are stimulus-induced expression changes of 134 signalling proteins divided by the standard deviation of all log2-fold-change values in the corresponding array dataset (z-value; horizontal dimension) in relation to the mRNA and protein half-lives, respectively (vertical dimension). Rapid Feedback Inhibitors (RFIs), defined as signal inhibitors whose the mRNAs are induced (z>1) within 4 hours after signal administration (grey box), are generally characterised by mRNAs (squares) and proteins (circles) with very short half life (<2h). By contrast, signal transducers (filled symbols) and most long-lived signal inhibitors (open symbols) are not induced within 4h and have significantly longer half-lives than the RFIs.

5.2 Expression and half-life data reveal design principles of transcriptional feedback

The analysis presented in this chapter was focused on transcriptional feedback regulation of five major mammalian signalling pathways: growth factor signalling via MAPK cascades, growth factor signalling via PI3K, β-adrenergic signalling via cAMP, TGFβ signalling via Smads, and cytokine signalling via JAKs and STATs (Fig. 5.2). The analysis was restricted to early transcriptional feedback, i.e., to gene expression changes within less than four hours after extracellular stimulation. Published gene expression profiles from rat, mouse and human cell lines in response to stimulation were collected from the Gene Expression Omnibus (GEO) database. The expression values were log2-transformed, and for each gene and each stimulus duration in each dataset, the median value of replicas was calculated and the value for unstimulated cells was subtracted. The values were further normalised to account for different fold-changes in the different experiments. This was done by dividing by the standard deviation of all log2-fold-changes values in the array dataset, resulting in “z-values”. A gene was marked as induced if the median z-value over all datasets, conditions and time-points was larger than 1. Expression data where the majority of replicas were reported as not expressed (absent call) were left out.

Key regulators in transcription factor networks are characterised by unstable mRNAs [277]. The same might be true for rapid feedback regulators of signalling pathways, so that mRNA and protein half-life measurements were also included in the analysis. mRNA half-lives were taken from the genome-wide data-sets reported in [278] and [279]. All three different treatment conditions in [279] were considered, and the mRNA half-lives used are the median over all half-life measurements for each gene (including all conditions in [279]). Protein half-lives (listed in Appendix E) were collected from literature studies, which measured protein
decay after administration of the translation inhibitor cycloheximide or by pulse-chase assays. In some cases, the amount of protein decreased less than 50% within the measurement time, $T$. If the protein decayed to almost 50% within the measurement time, the half-life $t_{1/2} = T + 2h$ was used for further analyses, while the value $t_{1/2} = T + 4h$ was taken if the protein hardly decayed within the measurement time. For those proteins whose half-lives were measured in multiple literature studies, the median of all measured half-lives was used in order to reduce the influence of outliers.

Only proteins for which mRNA or protein half-life and microarray measurements were available were taken into account in the analysis. Figure 5.1 relates transcriptional induction in response to extracellular stimulation with half-life data, and thus summarises the results of the analysis. The expression of at least one of the mRNAs changes in each of the five pathways analysed, confirming that transcriptional feedback is a general design principle in biochemical signalling networks. 15 out of 134 signal proteins are significantly upregulated throughout all analysed pathways, with repression (negative values at the horizontal axis in Fig. 5.1) being rare. The half-life analysis revealed that the mRNAs and proteins of transcriptionally induced signalling species are all unstable (vertical axis in Fig. 5.1). On the other hand, mRNAs and proteins of non-induced signalling species tend to be stable.

Signalling proteins were classified into two functional categories (‘signal transducers’ and ‘signal inhibitors’) in order to get further insights into the design principles of transcriptional feedback. The term signal transducer refers to proteins that are required to sense and to transmit the signal (e.g., kinases), while signal inhibitors attenuate information transfer, e.g., by catalyzing kinase dephosphorylation. 14 out of 51 signal inhibitors turned out to be induced in response to extracellular stimulation of their signal transduction chain (open symbols in Fig. 5.1), while only 1 out of 83 signal transducers was marginally upregulated. This revealed two design principles at once: (i) transcriptional negative feedback is the dominant general design principle in intracellular signalling in mammals, whereas positive feedback seems to play no major role and (ii) the negative feedback regulation is completely asymmetric in these biochemical signalling networks, i.e., negative feedback does not occur by downregulation of signal transducers, but relies on the induction of a subset of signal inhibitors ($p<0.0001$). In the following, these induced proteins will be called ‘Rapid Feedback Inhibitors’ (RFIs).

The vertical dimension of Figure 5.1 shows that the induced RFIs have very short average mRNA (with one exception) and protein half-lives (median 0.5 h and 1 h, respectively). The RFI mRNAs and proteins are significantly less stable than all other molecules analysed (both $p<0.0005$, two sided Mann-Whitney test). The average RFI half-lives are an order of magnitude shorter than those of signal transducers (median mRNA and protein half-life of 6 h and 12 h, respectively; both $p<0.0005$, two-sided Mann-Whitney test). The constitutively expressed, non-inducible signal inhibitors have a median mRNA and protein half-life of 2.6 h and 9 h, respectively, and are thus more stable than the inducible RFIs (both $p<0.01$, two-sided Mann-Whitney test).

It was analysed next whether RFIs differ from constitutively expressed signal inhibitors in terms of position or function in the mammalian signal transduction network. Figure 5.2 depicts the signalling pathways considered in this study, with functional groups containing RFIs highlighted in grey. MAPK phosphatases (DUSPs) are induced in response to MAPK signalling [67]. Similarly, in all other analysed pathways RFIs were induced, such as protein phosphatase 1 (PPR15A), phosphodiesterases (PDE4B/D), Smad7, SnoN, and SOCS proteins. In all pathways RFIs attenuate signal transduction, often at multiple levels, ranging from cell-surface receptors to transcription factors. Moreover, there is no common mode of inhibition, as some RFIs inhibit catalytically (PDEs, DUSPs, PPR15A), others act by binding to their targets (Sprouty, SnoN, BAMBI, TGIF) and yet others may combine these mechanisms (SOCS, Smad7). In terms of their mode of action, RFI’s do not differ appreciably from that of constitutively expressed signal inhibitors.
Taken together, the analysis revealed that early transcriptional feedback regulation in the mammalian signalling network is mediated exclusively by induction of a subgroup of signal inhibitors (RFIs). These RFIs are highly unstable in terms of mRNA and protein. This is consistent with the idea that rapid transcriptional regulation of steady-state protein expression requires short mRNA and protein half-lives (Section 5.3). It was surprising to see that negative feedback exclusively relies on rapid RFI induction, down regulation of signal transducers playing no role. Moreover, it was unexpected that the mRNAs and proteins of constitutively expressed signalling proteins are generally stable. Thus, the analysis reveals a clear subdivision of the mammalian signalling network into two parts: (i) a constitutively expressed static part comprising stable signalling proteins required to receive and transmit the signal; (ii) a flexible part that is transcriptionally induced upon stimulation, and mediates negative feedback regulation.

5.3 Kinetic modelling of rapid feedback inhibitor (RFI) action

The separation of the signalling network into flexible and static parts suggests that, as compared to signal transducer downregulation, RFI induction could be advantageous in achieving signal attenuation. This will be further investigated in the following by kinetic modelling. First, the kinetics of transcriptional regulation will be discussed, as gene expression is the rate-limiting step in RFI action. Then, the regulation of signal transduction by signal inhibitor induction will be compared to signal transducer downregulation to get insights into the kinetic implications of RFI action. Finally, the energetic aspects of protein turnover are investigated, and it will be discussed that feedback via RFI induction might be an optimal design due to a trade-off between energy efficiency and flexibility.

Kinetics of transcriptional regulation: The expression of a gene into its protein is determined by four processes: transcription, translation, mRNA degradation and protein degradation (schematically depicted in Fig. 5.3A). The dynamics of gene expression may be described by two differential equations incorporating these four reactions (see Fig. 5.3A). The steady-state protein concentration, $P_{ss}$, of a gene product is given by:

$$P_{ss} = \frac{k_1 \cdot k_2}{d_1 \cdot d_2} \quad (5.1)$$

If a gene is regulated at the transcriptional level (i.e., if the transcription rate is changed to $k_1$ at $t=0$) the time course of protein expression is given by:

$$P(t) = P_{ss} \left(1 - \frac{P_{ss} - P(0) \cdot d_1 \cdot e^{-d_1 t} - d_2 \cdot e^{-d_2 t}}{P_{ss} - d_1 - d_2} \right) \quad (5.2)$$
Signal transducers and signal inhibitors were classified into groups of functionally similar proteins (e.g., Receptor Tyrosine Kinases). Post-translational regulatory interactions among these groups are indicated by arrows. Individual proteins are indicated by small boxes which are divided in three parts. The left part indicates whether the factor is transcriptionally regulated upon external activation of the pathway (red), while the centre and the right parts depict the protein and mRNA half-lives, respectively (see legend). Empty parts indicate missing data. Functional groups containing Rapid Feedback Inhibitors (RFIs) are highlighted in dark grey. RFIs possess short protein and mRNA half-lives and they regulate all levels of the signalling network, from receptors to terminal kinases.

Thus, the response time, defined as the time required to reach the new steady state, is solely determined by the decay rates. The response time depends on both, $d_1$ and $d_2$, if the protein and the mRNA half-lives are of similar magnitude, while it is mainly set by the slowest decay in case mRNA and protein stability differ significantly from each other. This implies that for transcriptional regulation, both the mRNA and the protein have to be unstable in order to attain a new steady state rapidly.

The activation of signalling networks can be modulated by transcriptional regulation of the concentrations of their components. The time required to attain a new signalling steady state by transcriptional regulation of a signalling protein is essentially determined by the stability of mRNA and protein (Eq. 5.2), as subsequent post-translational regulatory mechanisms are much faster than transcriptional induction. The behaviour is slightly more complex if feedback is involved: a negative feedback system subjected to activation reaches a steady state faster than expected from the decay rates of the feedback regulator, while no such acceleration is observed upon deactivation [62]. Rapid transcriptional feedback regulation of the signalling network requires that both the mRNA and the protein of the transcriptional feedback regulator need to be unstable, since otherwise: (i) feedback induction upon stimulus addition
implies continuously increasing feedback strength over many hours; (ii) long latency will be observed upon stimulus removal.

**Signal inhibitor induction vs. signal transducer downregulation:** Transcriptional negative regulation of the signalling network can, in principle, occur by upregulation of signal inhibitors or by downregulation of signal transducers. The analysis presented in Section 5.2 revealed that RFI induction dominates the feedback regulation of the mammalian signalling network which suggests that upregulation of signal inhibitors might be more effective in achieving signal attenuation. Accordingly, it has previously been shown for the MAPK cascade that phosphatases exert stronger control on signal duration than kinases [38]. In the following, the dynamic behaviour of a generic protein kinase cascade will be compared for three different transcriptional regulatory designs to get insights into kinetic implications of RFI action: (i) repression of a kinase acting as a signal transducer (Fig. 5.3B, left); (ii) induction of a phosphatase acting as a catalytic RFI (Fig. 5.3B, middle); (iii) induction of a stoichiometric inhibitor acting as a non-catalytic RFI (Fig. 5.3B, right).

In a weakly activated phosphorylation/dephosphorylation cycle (modelled with linear kinetics), the amount of active phosphorylated protein at steady state is proportional to the ratio of kinase to phosphatase concentrations [280]. Thus, the signal can be reduced to 10% of its original value, either by reducing kinase expression to 10% or by a 10-fold phosphatase upregulation. Figure 5.3B shows how the signal cascade activation level (i.e., the ratio of kinase and phosphatase activities) follows a slow change in kinase or phosphatase expression, when modelled according to Eq. 5.2 (with $d_1 = 2/h$ and $d_2 = 1/h$). A 10-fold phosphatase upregulation allows to switch off the signal much more quickly (Fig. 5.3B, middle, solid line) when compared to 10-fold kinase downregulation (Fig. 5.3B, left, solid line). The recovery times if the kinase and phosphatase expression are regulated in the opposite direction were also analysed. In this case, kinase upregulation (Fig. 5.3B, left, dashed line) allows for faster disappearance of the signalling than phosphatase downregulation (Fig. 5.3B, middle, dashed line). Thus, the signalling activity immediately follows transcriptional regulation of kinase expression (due to direct proportionality), while phosphatases regulate signalling pathways asymmetrically, with a long latency for recovery (this is due to the inverse proportionality). Similar conclusions also hold for strongly activated kinase cascades, although the difference between phosphatase and kinase regulation becomes less pronounced.

Several RFIs act as stoichiometric inhibitors, that is, they inhibit signal transduction non-catalytically by binding reversibly to their targets (as depicted schematically in Fig. 5.3B, right). In the limiting case of strong stoichiometric inhibition the inhibitor binds to a kinase with very high affinity. Then, all available inhibitor I will be bound, unless the inhibitor is in present in excess over its target. Thus, the free, active concentration $K$ of the targeted kinase with the total concentration $K_T$ is given by:

$$K = \max(0, [K_T] - [I]) \quad (5.3)$$

The cascade activity was assumed to be proportional to the free kinase concentration $K$ (see above), and was analysed for slow inhibitor up- and downregulation according to Eq. 5.2 (Figure, right; $d_1 = 2/h$ and $d_2 = 1/h$). The change in the signal level (again 10-fold ultimately) immediately follows alterations in inhibitor protein expression. This statement holds true for as long as the inhibitor is not induced too strongly. Otherwise, the concentration $I$ exceeds $K_T$, so that the system shows some latency before it recovers. In any case, the signalling dynamics in response to inhibitor regulation do not differ from those observed upon kinase regulation (compare Fig. 5.3B, left and right).
Figure 5.3: Kinetic modelling of RFI action.

(A) Schematic representation of a minimal model of gene expression (left), and corresponding differential equations (right). (B) Comparison of three different transcriptional regulatory designs for a phosphorylation-dephosphorylation cycle: (i) repression of a kinase acting as a signal transducer (left); (ii) induction of a phosphatase acting as a catalytic RFI (middle); (iii) induction of a stoichiometric inhibitor acting as a non-catalytic RFI (right). The (de)phosphorylation steps were assumed to be fast relative to transcriptional regulation, so that the system response was taken to be proportional to the ratio of kinase to phosphatase concentrations [280]. The simulations show how the signal cascade activation level (i.e., the ratio of kinase and phosphatase activities) follows a slow change in kinase or phosphatase expression, which was modelled according to Eq. 5.2 (with $d_1 = 2/h$ and $d_2 = 1/h$). The solid lines show the kinetics of cascade deactivation by kinase downregulation (left), phosphatase upregulation, and inhibitor upregulation (right). The recovery times if the kinase, phosphatase or inhibitor expression are regulated in the opposite direction were also analysed (dashed lines). The analysis reveals that signalling activity immediately follows transcriptional regulation of kinase expression (due to direct proportionality), while phosphatases regulate signalling pathways asymmetrically, with fast signal downregulation, and with a long latency for recovery (this is due to the inverse proportionality).

Taken together, these simulations suggest that the design pattern of negative feedback regulation by a small group of RFIs might have been evolved to speed up the adaptation upon activation, and to introduce a lag phase upon deactivation in some (but not all) signalling pathways.

Energetic aspects of RFI action: An alternative but not mutually exclusive explanation for the separation of the signalling network into flexible and static parts might be improved energy efficiency. Unstable proteins and mRNAs need higher translation or transcription rates, respectively, to reach the same steady state protein concentration (Eq. 5.1). Therefore their production consumes more energy, as the energy expenditure is proportional to the transcription and translation rates ($k_1$ and $k_2\cdot[mRNA]$). Thus, a trade-off exists between making the protein network flexible (by increasing $d_1$ and $d_2$, and simultaneously increasing $k_1$ and $k_2$ to maintain the expression level), and making it energy efficient (by decreasing $k_1$ or $k_2$ and, to compensate, simultaneously decreasing $d_1$ or $d_2$).
In other words, the nodes in the network can either be designed in a flexible, rapidly responding manner (unstable mRNAs and proteins) or such that energy expenditure is minimal (stable mRNAs/proteins), but not both. The mammalian signalling network seems to circumvent this trade-off in an especially elegant manner; the network consumes only low amounts of free energy in the unstimulated state because constitutively expressed signal transducers required to receive the signal have stable mRNAs and proteins (static part). In contrast, the mRNAs and proteins of stimulus-induced RFIs are highly unstable, and thus allow for rapid stimulus-dependent negative feedback. This part of the network is flexible, and consumes energy only when the pathway is stimulated.

Experimental evidence supports the physiological relevance of the proposed energy minimisation principle: Depending on the cellular context, protein turnover requires between 30 and 70 % of the total cellular energy budget [281] . From quantitative experimental measurements one can estimate that signalling proteins make up ~5 % of the total cellular protein mass (Appendix E). The most abundant cellular proteins, i.e., house-keeping and structural proteins, are very stable with half-lives of up to 60 days [282] and thus do not contribute strongly to the cellular energy budget. It therefore seems likely that signalling protein turnover consumes much more than 5% of the total energy spent for protein synthesis, and thus constitutes one of the dominant energy sinks in mammalian cells. The present analysis and literature data summarised in Appendix E showed that signalling pathways are transcriptionally regulated at multiple points by the induction of different (and possibly cell-type specific) inhibitors. If these flexible parts of the signalling pathway would be highly turned over constitutively, they would likely represent a strong energy burden. However, the analyses presented in this chapter show that the constitutively expressed signalling proteins are generally stable. Thus, an energy minimisation principle might have contributed to an evolutionary selection pressure favouring this strategy of regulation.

Kinetic modelling therefore reveals that the criteria of rapid feedback regulation and energy efficiency favour the same wiring of the regulatory network. Consequently, two explanations for the observed separation of the signalling network into flexible and static parts are conceivable, but the present data does not allow selecting between them.

5.4 Conclusions

The simultaneous analysis of signal protein expression profiles and half-life data presented in this chapter revealed design principles of the mammalian network: (i) negative feedback dominates over positive feedback in the transcriptional feedback regulation of signalling. (ii) transcriptional feedback regulation is completely asymmetric, i.e., negative feedback does not occur by downregulation of signal transducers, but exclusively relies on the induction of a subset of signal inhibitors. (iii) signalling pathways are separated into flexible and static parts, as transcriptional feedback regulators are unstable, while constitutively expressed (i.e., unregulated) signalling proteins are generally long-lived.

The modelling analyses presented in Section 5.3 suggest that swift feedback regulation and energy minimisation may underlie design principles ii and iii. Rapid transcriptional negative feedback (design principle i) might allow signal transduction cascades to tone down signalling (to ‘adapt’), as suggested by experimental studies where signalling persisted if transcriptional feedback was blocked by protein biosynthesis inhibitors ([67] ; see also Appendix A). Efficient signal adaptation might enable signalling cascades to induce controlled pulses of gene expression in a robust manner, independent of environmental variations and transcriptional noise [62,283,284,285] . Consistent with a role in signal termination, many of the RFIs given in Fig. 5.2 are identified as tumour suppressors or, in the case of cytostatic TGFβ signalling, as oncogenes. Cellular decision-making frequently requires ongoing signalling activity over several hours (see also Appendix E). This suggests that RFIs are key regulators of the cell fate, while rapidly acting post-translational feedbacks might often be more important for initial signal processing and specificity [24,139] .
A circuitry involving RFIs could also be beneficial for simpler eukaryotic organisms like yeast. Accordingly, it is known that all three yeast MAPK signalling cascades induce their phosphatases PTP2, PTP3 and MSG5 (reviewed by [286]). Additionally, transcriptional feedback in response to cAMP signalling was analysed and specific upregulation of the signal inhibitor RGS2 was observed (see Appendix E). If yeast harbours RFIs, one expects these feedback-regulators to be unstable as well. Therefore a genome-wide dataset of yeast protein half-lives was analysed [287]. For the cell integrity, high osmolarity and cAMP pathways, all 20 analysed signal transducers had a protein half life of more than 15 min, while 4 out of 17 inhibitors were short-lived proteins with half lives of 15 min or less. Therefore, unstable proteins in these pathways are inhibitors (p<0.05, two-sided Fisher’s exact test). Moreover, transcriptional feedback occurs by induction of these unstable signal inhibitors, particularly via PTP2 with a half-life of only 3 min. This suggests that the yeast signalling network shows the same design pattern as mammalian cells.

Analysis of the yeast pheromone signalling pathway revealed an exception to the finding that transcriptional regulation of signalling generally occurs through negative feedback: the transducers FUS3, STE12, STE2 in the MAPK pathway required to receive the pheromone signal are transcriptionally upregulated in response to pheromone stimulation, with rapid kinetics (<15 min) well below the cell-cycle time [288]. This is positive feedback therefore. Interestingly, for positive feedback the two possible selection criteria discussed (energy efficiency and quick feedback regulation) both predict that in the case of positive feedback induction of the transducer should be a better strategy than repression of the inhibitor, as constitutively expressed flexible inhibitors will require high amount of energy, while upregulated transducers require only high turnover when the cells are stimulated. In pheromone signalling, some of the signal transducers are indeed short-lived proteins, much like the signal inhibitors involved in transcriptional negative feedback, which ensures that the pathway can be upregulated quickly. The pheromone pathway of yeast is different when compared to many other signalling pathways, as it is only required in certain phases of the cellular life cycle and therefore might require a positive feedback. However, the limited amount of available data does not allow us to substantiate this explanation.

Taken together, the results presented in this chapter strongly support that transcriptional negative feedback regulation plays an important role for intracellular signal transduction. However, the physiological relevance of these feedbacks remains to be verified at the protein level. TGFβ-induced signalling via Smad transcription factors typically remains elevated over several hours after stimulation (see Appendix A), thus suggesting that transcriptional negative feedbacks are important determinants for the dynamics of this signalling pathway. Using a combination of theoretical and experimental approaches, TGFβ/Smad signalling in primary mouse hepatocytes is therefore analysed in Section 6. Genome-wide microarray analyses suggest that transcriptional feedback regulation exists in primary mouse hepatocytes, and measurements at the protein level indicate that the SnoN oncoprotein is the central feedback regulator. Theoretical predictions regarding the mechanisms of transcriptional feedback regulation are then confirmed in SnoN knock-out mice, thus further supporting the relevance of transcriptional negative feedback regulation in signal transduction.
6 Transcriptional feedback regulation of TGFβ signalling

This chapter is the result of a collaboration with Peter Nickel (DKFZ Heidelberg), Thomas Maiwald (University of Freiburg), Jens Timmer (University of Freiburg) and Ursula Klingmüller (DKFZ Heidelberg). All experiments mentioned were performed by Peter Nickel in the lab of Ursula Klingmüller. Thomas Maiwald from the group of Jens Timmer performed the microarray raw data analysis, provided the software environment for data-based modelling (PottersWheel) and did part of the modelling work.

SYNOPSIS

Transforming growth factor β (TGFβ) acts as an important mitosis inhibitor and differentiation factor on hepatocytes both in the resting liver and during liver regeneration. Dysregulated TGFβ signalling can lead to liver fibrosis, cirrhosis and ultimately cancer. Extracellular TGFβ stimulation induces specific changes in gene expression by activating intracellular Smad transcription factors, and thereby affects the cellular phenotype. In this chapter, genome-wide expression profiling at the mRNA level in response to stimulation was employed to identify transcriptional feedback loops of TGFβ-Smad signalling in primary mouse hepatocytes. Subsequent measurements at the protein level further narrowed down the number of potential feedback regulators, and suggested a key role for the SnoN oncoprotein, which inhibits signalling by reversibly binding to Smad proteins. A mathematical model of TGFβ-Smad signalling that was calibrated using multiple quantitative immunoblotting data sets suggested that SnoN-mediated feedback efficiently controls Smad transcription factor complexes in the nucleus and thus Smad-induced gene expression. Subsequent experiments in SnoN-depleted hepatocytes quantitatively confirmed these model predictions and thereby suggested that SnoN is the major negative feedback regulator during the early phase of TGFβ-Smad signalling in primary mouse hepatocytes. The mathematical model mechanistically explains how low levels of SnoN are able to efficiently regulate a much larger pool of Smad proteins, and provides a basis for further quantitative experimental analyses of TGFβ-Smad signalling.

6.1 Introduction

Tight regulation of growth and differentiation processes requires the precisely coordinated interplay of growth stimulatory and inhibitory factors. The TGFβ family of cytokines constitute major inhibitors of cell growth, and accordingly they play an important role in various physiological processes such as development, tissue homeostasis, and tissue regeneration. Moreover, TGFβ signalling is dysregulated under pathological conditions including cancer, organ fibrosis, and Marfan syndrome. TGFβ signalling is initiated by binding of extracellular TGFβ to transmembrane serine/threonine kinase type I and type II receptors. Ligand binding triggers receptor-mediated phosphorylation of Smad2/3 transcription factors, which then homotrimerise or heterotrimerise (with Smad4), and subsequently translocate into the nucleus (Fig. 6.1). Nuclear Smad trimers control the expression of several hundred target genes, many of which are involved in cell cycle control (e.g., p21, c-myc, cdc25A).

TGFβ signalling plays an important role for homeostasis of the normal liver, as injection of follistatin, an extracellular inhibitor of TGFβ family ligands, increases organ size in vivo [289]. Accordingly, cell culture experiments demonstrated that TGFβ inhibits hepatocyte proliferation in vitro [290]. The liver maintains its function after acute (i.e., short-term) intoxication by counterbalancing cell death with compensatory cell proliferation (‘liver regeneration’). Growth factors are released early during liver regeneration, and this triggers quiescent hepatocytes to re-enter the cell division cycle [291]. The proliferative response is
terminated once liver mass is restored. TGFβ seems to play a major role in the regulation of liver regeneration, as perturbations in TGFβ family signalling affect the initiation kinetics, the extent and/or the termination of hepatocyte DNA synthesis during liver regeneration [292,293,294,295,296,297,298] . However, hepatocytes are only sensitive towards TGFβ stimulation within a limited time period during liver regeneration. More specifically, it has been shown that hepatocytes no longer respond to TGFβ treatment during the intermediate proliferative phase of liver regeneration [299] . Yet, TGFβ expression and Smad phosphorylation are permanently elevated throughout liver regeneration [300,301] . This suggests that unidentified negative regulatory mechanisms downstream of receptor-mediated Smad phosphorylation exist in the regenerating liver.

Figure 6.1 Negative Regulation of TGFβ-Smad signalling.
Schematic representation of negative regulatory mechanisms occurring at different levels of TGFβ-Smad signal transduction. General regulatory mechanisms are written in white letters, specific modulators in yellow letters. Figure taken from [47] .

In this chapter, the negative regulatory mechanisms controlling TGFβ signalling in primary mouse hepatocytes are investigated. Various negative regulators have been described for TGFβ signalling, and they appear to act at all levels of the signalling cascade as schematically depicted in Fig. 6.1. Several of these signal inhibitors including BAMBI, GRK2, Smad7, Smurf2, SnoN, and TGIF are known to be induced upon TGFβ stimulation (see Table E.7 in Appendix E), so that this chapter was focused on transcriptional negative feedback regulation. The quantitative aspects of TGFβ signalling such as amplitude and duration are thought to be crucial determinants of the cellular fate. More specifically, it is known that pro-proliferative genes are specifically induced by transient Smad activation, while anti-proliferative genes require sustained Smad signalling [49] . Smad phosphorylation typically remains elevated for several hours after stimulation, and usually peaks late (at ~ 1 h). The transcriptional feedback loops analysed in this chapter are therefore likely to control the Smad signal amplitude and duration, and may thus govern physiological responses towards TGFβ stimulation. Moreover, they are likely to play a role under pathological conditions, as constitutively elevated TGFβ levels are thought to contribute to liver damage (‘fibrosis’) after chronic liver intoxication (e.g., due to long-term alcohol abuse).
6.2 mRNA expression profiling reveals transcriptional feedback

Using microarray technology, mRNA expression in response to stimulation with 1 ng/ml TGFβ was measured by Peter Nickel in order to identify transcriptional feedback loops. Five time points (0, 1h, 4h, 8h and 12h) were considered and three replicates have been done for each time point. For hybridisation of RNA, the GeneChip Mouse Genome 430 2.0 platform (Affymetrix) comprising 45,100 probe sets representing 20,700 genes was used. The analysis and processing of raw data was done by Thomas Maiwald. Briefly, the RMA convolution approach was used for background correction [302]. Normalisation was done by imposing the same empirical distribution of intensities to each microarray based on the averaged sample quantiles over all arrays [303]. For every available gene, i.e., probe set, a t-test is applied to compare the expression level of time point t>0 with the reference time t=0. In order to cope with the occurring multiple-testing problem for 45101 probe sets, the obtained p values are corrected with the false-discovery-rate method. In total, 2,500 genes were significantly up- or downregulated at one or more stimulation time points. A gene was considered to be unregulated at a given time point, if the absolute log2-fold change compared to the unstimulated state was less than 0.5 or the corresponding corrected p-value was larger than 0.05.

A list of 114 genes functionally associated with TGFβ signalling was compiled and analysed for expression changes in order to get insights into transcriptional feedback regulation. The genes comprised the following functional categories (see Fig. 6.2): Ligands activating Smad2/3 signalling (TGFβ, Activins); proteins controlling ligand processing and activity; receptors for TGFβ and Activin; modulators of receptor activity and degradation; Smad transcription factors; and regulators of Smad activity. Some genes are assigned to more than one category in Fig. 6.2 (e.g., Smad7), because their protein products are known to regulate TGFβ signalling at multiple levels.

Figure 6.2 schematically depicts the time course measurements for each of the 114 genes. The analysis indicated that repression plays no important role in the transcriptional feedback regulation of TGFβ signalling, as only 3 out of 114 genes were weakly downregulated (log2 fold change between -0.5 and -1). On the other hand, 14 of the 114 genes were significantly upregulated in response to TGFβ signalling (log2 fold change > 0.5; Fig. II.1). Among these 14 upregulated genes were TGFβ1, the TGFβ receptor I and several factors involved in ligand processing such as Thrombospondin 1 (THBS1), which suggested that TGFβ stimulation might induce an autocrine amplification loop. In line with literature studies [304], gene expression profiling indicated strong induction of TIEG-1, a transcriptional coactivator of Smad transcription factors, which is thought to be required for feed-forward regulation of specific Smad target genes. Finally, the analysis revealed that known feedback regulators TGFβ receptor signalling (Smad7) and of Smad signalling (SnoN, Smad7 and Ski) are also transcriptionally induced in primary mouse hepatocytes. Interestingly, no significant upregulation of GRK-2, Smurf2, BAMBI and TGIF was observed, even though these proteins have been previously described as negative feedback modulators of receptor/Smad signalling (see Table E.7 in Appendix E).
Figure 6.2 mRNA expression profiling of transcriptional negative feedback of TGFβ-Smad signalling in primary mouse hepatocytes.

Alterations in mRNA expression of 114 genes functionally associated with TGFβ signalling after stimulation of hepatocytes with 1 ng/ml TGFβ for 1 h, 4 h, 8 h or 12 h are expressed as log2 fold-changes relative to the unstimulated state.

Taken together, gene expression profiling indicates that TGFβ might induce a positive feedback at the extracellular level (ligand/receptor), while negative feedback seems to predominate at the intracellular level (receptor/Smad). Two lines of evidence suggest that extracellular positive feedback amplification does not play a major role under the experimental conditions chosen: (i) Peter Nickel quantitatively measured the amount of processed TGFβ protein in the extracellular medium in the presence and in the absence of the general feedback inhibitor Actinomycin D, respectively (Fig. 6.5A). These analyses revealed that medium TGFβ monotonically decays over the 10 hour time course and that the decay kinetics were not affected by a blockade of transcriptional feedback. (ii) Previous studies in various cell types revealed ligand-induced TGFβ receptor downregulation at the protein level, while upregulation has never been reported (e.g., [305]). It thus seemed likely that TGFβ receptor I mRNA induction merely serves to compensate ligand-induced receptor protein degradation, but does not result in positive feedback amplification. Owing to these
data, the following analyses (Section 6.3) will be focussed on intracellular negative feedback regulation via Smad7 and SnoN.

6.3 Experimental verification of transcriptional feedback at the protein level

By analyzing genome-wide transcriptional data on TGFβ-induced target genes, putative feedback regulators of the signalling system could be identified. Next, the central components of Smad signalling were quantified at the protein level by Peter Nickel. More specifically, quantitative immunoblotting was applied to estimate absolute cellular levels and to determine the relative abundance of reaction partners. Using serial dilutions of recombinant proteins (Fig. 6.3), it was estimated that each hepatocyte contains approximately 80 000 molecules of each, Smad2 and Smad4, while Smad3 was present in much lower amount (4000 molecules per cell). Time series analysis revealed that these levels were fairly unchanged upon stimulation with TGFβ (see below).

Next, the putative negative regulators of Smad signalling, Smad7 and SnoN, which were both strongly upregulated at the mRNA level upon TGFβ stimulation, were analysed (Fig. 6.3B). Smad 7 was already present in unstimulated hepatocytes at 80,000 molecules per cell. Surprisingly, these numbers slightly declined upon TGFβ stimulation, despite strong Smad7 induction at the RNA level (see Fig. 6.2). SnoN was induced ~4-fold at the protein level, which suggests that it represents a major negative feedback regulator of TGFβ signalling in hepatocytes. However, the absolute expression level was limited to 2300 molecules per cell, even after stimulation (Fig. 6.3B).

This raises the question of whether the much larger pool of Smad proteins can be efficiently regulated by sub-stoichiometric amounts of SnoN, which is thought to act as non-catalytic inhibitor of Smad signalling. A mathematical model of Smad signalling and transcriptional feedback via SnoN was implemented to resolve this issue (Section 6.4).

![Figure 6.3 Absolute quantification of TGFβ-Smad pathway components using recombinant protein standards.](image)

Endogenous protein levels (n=5) of Smad2, Smad 3, Smad4, Smad 7 and SnoN were quantified by means of dilution series (n=3) of the respective recombinant calibrator proteins added to whole cell lysates prior to immunoprecipitation and quantitative immunoblotting. Lysates were either prepared from 2x10^6 unstimulated cells (black bars) or from 2x10^6 cells stimulated with 1 ng/ml TGFβ for 10 h (white bars). Figure kindly provided by Peter Nickel.
6.4 Implementation of a mathematical model

6.4.1 Rationale

An ordinary differential equation (ODE) model of TGF\(\beta\) signalling in primary mouse hepatocytes, as schematically depicted in Fig. 6.4, was implemented (differential equations can be found in Appendix F). All reactions concerning receptor and Smad shuttling between subcellular compartments as well as protein degradation were modelled with 1st-order kinetics, and protein dimerisation and trimerisation were modelled as bi- and trimolecular reactions, respectively. Michaelis-Menten kinetics were assumed for enzyme-catalysed Smad2 phosphorylation and for Smad-mediated transcriptional induction of SnoN, since both reaction types are known to be saturable. Several protein species in the model are continuously synthesised and degraded even in the absence of stimulation. The basal protein concentrations which are determined by various kinetic parameters in the model, were calculated using analytical steady state equations as described in Appendix F.

![Figure 6.4 Dynamic model of TGF\(\beta\)-Smad signalling.](image)

Schematic representation of biochemical reactions considered in the kinetic model of TGF\(\beta\)-mediated signalling. The Receptor Trafficking module includes signal initiation by ligand binding to the cell-surface receptor, and also takes receptor shuttling between endosomal compartments and the plasma membrane into account. The core module ('Smad Activation Cycle') comprises receptor-mediated Smad2 phosphorylation (S2 -> pS2), Smad2 homotrimerisation (3 pS2 <-> (pS2)\(_3\)), Smad heterotrimerisation (S4 + 2 S2 <-> S4~(pS2)\(_2\)), and considers nucleo-cytoplasmic shuttling of all Smad species. Nuclear Smad proteins are indicated by the prefix n. The SnoN feedback module includes Smad-mediated transcriptional induction of SnoN and subsequent feedback inhibition of Smad trimers by binding to SnoN protein.

Three compartments (extracellular medium, cytoplasm and nucleus) are relevant for the chosen experimental setup. Microscopic analyses of primary mouse hepatocytes indicated a total cell volume of 10 pl (10\(^{-12}\) liters) and a ratio of cytoplasmic and nuclear volumes of \(\sim 20\) (Sebastian Bohl, unpublished). In the model, it was thus assumed cytoplasmic and nuclear volumes of 10 and 0.5 pl, respectively. The medium volume was taken to be 1000 pl per cell.
The following sections contain a description of how the model topology was derived from the published literature. The Smad3 protein was generally neglected in the model owing to its low expression level in primary mouse hepatocytes (Fig. 6.3), and only the much more abundant Smad2 protein was taken into account. Experimentally measured kinetic parameters and the parameter ranges used for the fitting procedure are summarised in Tables F.1 – F.3.

6.4.2 Modelling receptor trafficking

TGFβ initiates intracellular signalling by binding to its cognate transmembrane receptor, the TGFβ receptor [48]. Receptor-ligand association was modelled by a bimolecular and reversible reaction (T + R ↔ TR), as supported by Scatchard analyses [306,307,308]. TGFβ receptor trafficking between the endosomal compartment and the plasma membrane was also considered in the model. The TGFβ receptor is known to constitutively shuttle between both compartments [309,310], and the majority of TGFβ receptors resides in the endosome even in unstimulated cells [310]. Experimental studies revealed that the rates of TGFβ internalisation and recycling are not affected by ligand stimulation [309,310]. In the model, the same rate constant was therefore assumed for endocytosis of free and ligand-occupied receptors (R → Re and TR → TR e). In accordance with previous mathematical models [311,312], free and ligand-occupied receptors in the endosome are recycled back to the cell surface with the same recycling rate constant (Re → R and TR e → R). Recycling of ligand-occupied receptors was modelled to be accompanied by intracellular ligand dissociation/degradation (TR e → R), as the pool of internalised TGFβ was shown to be completely degraded [313].

Lysosomal targeting and degradation of endosomal receptors and receptor-ligand complexes were also considered in the model. Experimental studies using inhibitors of lysosomal progression revealed that both, TGFβ [313,314,315] and the TGFβ receptor [316], are subject to degradation within the lysosome. Owing to these data, it was assumed in the model that early endosomal receptor-ligand complexes eventually progress into late endosomes (TR e → TRle), and thereby escape recycling. These complexes are then in turn targeted to the lysosomal compartment, where they are ultimately degraded (TRle → ). A degradation reaction for inactive endosomal TGFβ receptors (R e → ) was also considered in the model, because lysosomal targeting is thought to be an important degradation pathway for TGFβ receptors in unstimulated cells [316]. Other pathways ligand and/or receptor degradation pathways (e.g., raft-mediated internalisation; [310]) were neglected in the model, as they appear to be cell-type-specific [309], and were not required to accurately describe experimental data in hepatocytes. Transmembrane receptors such as the TGFβ receptor are synthesised via the secretion pathway (ER/Golgi). Thus, a constant influx into the pool of inactive endosomal receptors was included to model de novo synthesis of receptors (→ Re).

6.4.3 Modelling the Smad activation cycle

Smad Phosphorylation: Ligand binding induces autophosphorylation of the TGFβ receptor which in turn activates its receptor serine kinase activity towards cytosolic substrates such as Smad2/3 [48]. Receptor (de)phosphorylation occurs much faster than ligand binding and dissociation [48], thus justifying a lumped reaction step in the model (T + R ↔ TR and TR e → T e + R e). Most experimental analyses suggested that endosomal but not cell-surface receptors catalyse Smad2/3 phosphorylation [310,317,318,319,320,321,322,323]. For simplicity, it was thus assumed in the model that only endosomal receptor-ligand complexes catalyse Smad2 phosphorylation. Recent studies revealed that TGFβ receptors can remain active on their way to lysosomal degradation, i.e., in the late endosome [324,325,326].
Owing to these data, the phosphorylation reaction $S_2 \rightarrow pS_2$ is modelled to be additively catalysed by early and late endosomal receptors ($TR_e$ and $TR_{le}$) with the same efficiency. An endosomal accessory protein, SARA, is thought to be required for recruitment of Smad2 to the TGF$\beta$ receptor, and thus for receptor-mediated Smad2 phosphorylation [323,327]. In the model, SARA was assumed to recruit Smad2 to the membrane with high efficiency. This implies that Smad2 binding to the receptor will essentially behave like an association between to membrane proteins. It has been shown theoretically and experimentally that co-localisation of two proteins to the membrane drastically increases the apparent affinity of protein-protein interactions by a factor of $\sim 1000$ [328,329,330,331]. Low $K_M$ values for receptor-mediated Smad2 phosphorylation were thus allowed in the model, as further described in Table F.2.

**Smad Trimerisation:** Receptor-mediated phosphorylation strongly favours Smad2/3 multimerisation, and most experimental studies indicated that trimers are the predominant multimeric R-Smad species [332,333,334,335,336]. Heterotrimerisation of two Smad2/3 molecules with one Smad4 molecule was shown to be more efficient than homotrimerisation of three Smad2/3 molecules [334,335]. More specifically, the heterotrimERIC complex is slightly more stable than the homotrimer *in vitro* [334,335,336], but this difference might be more pronounced *in vivo* [337]. Smad homo- and heterotrimerisation were modelled as reversible, single-step reactions ($2pS_2 + S_4 \leftrightarrow (pS_2)_2S_4$ and $3pS_2 \leftrightarrow (pS_2)_3$), as the concentration of the Smad dimer intermediate was shown to be negligibly small *in vitro* [334]. The model was implemented such that heterotrimers preferentially formed over homotrimers (see Table F.2). Additionally, it was assumed that nuclear trimers are formed with the same kinetics as their cytosolic counterparts, because experimental studies demonstrated efficient Smad heterotrimer formation in the nucleus [338].

**Smad Import:** TGF$\beta$ stimulation induces translocation of Smad2/3 and Smad4 into the nucleus. Yet, the Smad proteins shuttle continuously between the nucleus and the cytoplasm even in unstimulated cells [339]. Recent studies indicated that the import rate of Smad2/3 is not affected by TGF$\beta$ stimulation [336,339,340]. Therefore, it was assumed in the model that all Smad2 monomer and trimer species are imported with the same import rate constant ($S_2 \rightarrow nS_2; pS_2 \rightarrow npS_2; (pS_2)_2S_4 \rightarrow (npS_2)_2nS_4$ and $(pS_2)_3 \rightarrow (npS_2)_3$). In contrast, the import of monomeric Smad4 was modelled to occur with a different rate constant ($S_4 \rightarrow nS_4$). Thus, the model reflects the experimentally established fact that the nuclear import of Smad heterotrimers is mediated by Smad2, and that it occurs independently of Smad4 [336].
Table 6.1 Data sets and error estimation.

<table>
<thead>
<tr>
<th>Definition of model species</th>
<th>Experimental observation</th>
<th>Stimulation conditions</th>
<th>Fitted/ Predicted</th>
<th>Data Points</th>
<th>Error Estimation</th>
<th>Figure</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_6 + x_{13} + 2x_{14} + 3x_{15} + x_{16} + x_{17} + 2x_{18} + 3x_{19} + 2x_{23} + 3x_{24} )</td>
<td>total Smad2 (whole cell)</td>
<td>1 ng/ml TGFβ</td>
<td>F</td>
<td>20</td>
<td>Spline</td>
<td>6.5D</td>
</tr>
<tr>
<td>( x_6 + x_{13} + 2x_{14} + 3x_{15} + x_{16} + x_{17} + 2x_{18} + 3x_{19} + 2x_{23} + 3x_{24} )</td>
<td>pSmad2 (whole cell)</td>
<td>1 ng/ml TGFβ</td>
<td>F</td>
<td>20</td>
<td>Spline</td>
<td>6.5B</td>
</tr>
<tr>
<td>( x_{14} + x_{16} + x_{23} )</td>
<td>Smad4 (CoIP) (whole cell)</td>
<td>1 ng/ml TGFβ</td>
<td>F</td>
<td>20</td>
<td>Spline</td>
<td>6.5C</td>
</tr>
<tr>
<td>( x_6 + x_{13} + 2x_{14} + 3x_{15} )</td>
<td>Smad2 (cytosol)</td>
<td>1 ng/ml TGFβ</td>
<td>F</td>
<td>22</td>
<td>Spline</td>
<td>6.7A</td>
</tr>
<tr>
<td>( x_{15} + x_{17} + 2x_{18} + 3x_{13} + 2x_{23} + 3x_{24} )</td>
<td>total Smad2 (nucleus)</td>
<td>1 ng/ml TGFβ</td>
<td>F</td>
<td>18</td>
<td>Spline</td>
<td>6.7B</td>
</tr>
<tr>
<td>( x_{17} + 2x_{18} + 3x_{19} + 2x_{23} + 3x_{24} )</td>
<td>pSmad2 (nucleus)</td>
<td>1 ng/ml TGFβ</td>
<td>F</td>
<td>18</td>
<td>Spline</td>
<td>6.7B</td>
</tr>
<tr>
<td>( x_6 )</td>
<td>SnoN (medium)</td>
<td>1 ng/ml TGFβ</td>
<td>F</td>
<td>19</td>
<td>Spline</td>
<td>6.5A</td>
</tr>
<tr>
<td>( x_{21} + x_{23} + x_{24} )</td>
<td>total SnoN (whole cell)</td>
<td>1 ng/ml TGFβ</td>
<td>F</td>
<td>20</td>
<td>Spline</td>
<td>6.5A</td>
</tr>
<tr>
<td>( x_{25} + x_{24} )</td>
<td>SnoN (CoIP) (whole cell)</td>
<td>1 ng/ml TGFβ</td>
<td>F</td>
<td>20</td>
<td>Spline</td>
<td>6.5A</td>
</tr>
</tbody>
</table>

Of the 46 listed data sets 27 were used for fitting (F) and 13 for testing of model predictions (P). Six redundant standard stimulation data sets (grey) were used to introduce a common scaling for results from independent experiments. The relative concentrations of total Smad2, phosphorylated Smad2, total SnoN, as well as Smad4 and SnoN co-immunoprecipitated with Smad2/3 were measured by immunoblotting. The amount of processed TGFβ was determined by ELISA. Errors were either calculated from a cubic spline (Spline), based on a spline-derived linear error model (Model), estimated employing typical errors E_{data} and E_{w} (Estimated) or calculated as the standard deviation of three independent measurements (StdDev). The definition of the observables in the model (according to the differential equations given in Appendix F.1) is given in the left column. In order to transform the simulated time courses into relative units, a common scaling factor was introduced for each observable, that is, each box corresponds to one scaling factor (Section 6.5.2).
Smad Export: TGFβ-induced Smad2/3 translocation into the nucleus is thought to be due to a decrease in the overall Smad2/3 export rate upon stimulation [339]. FRAP measurements indicated that this decrease might be due to selective immobilisation and thus retention of nuclear Smad trimers, e.g., by to binding to DNA [339]. Smad2/3 export has been reported to occur via Smad2/3 binding to Exportin 4, which functions as a nuclear export receptor [341]. Exportin 4 preferentially binds to unphosphorylated Smad2/3 [341], and accordingly phosphorylated Smad2/3 is exported with low efficiency [340,341]. Thus, it appears that two independent mechanisms contribute to stimulus-induced nuclear translocation of Smad2: (i) retention of Smad trimers; (ii) inefficient export of phospho-Smad2. In the model, it was therefore assumed that only unphosphorylated Smad2/3 monomers can be exported from the nucleus (nS2 → S2), and an export reaction for monomeric Smad4 was also included (nS4 → S4). Thus, the model is in accordance with experimental studies which indicated that Smad4 can only be exported in monomeric form, but not when assembled in Smad trimers [336].

Smad Dephosphorylation: Smad2/3 dephosphorylation in the nucleus controls the amount of nuclear Smad trimers, and is therefore an important mechanism to control the magnitude and duration of Smad-mediated gene expression [342]. It is thought that the phosphorylation sites in Smad2/3 constitute the binding interface in Smad trimers [333]. In the model that Smad trimers can therefore not be attacked by phosphatases, and dephosphorylation occurs only for monomeric nuclear phospho-Smad2/3 (npS2 → nS2). Initial numerical analyses indicated that cytoplasmic Smad2/3 dephosphorylation is not required to describe experimental data obtained in primary mouse hepatocytes, so that this reaction was skipped from the model.

Protein Synthesis and Degradation: Smad2 and Smad4 protein levels are typically not affected by TGFβ treatment [338,343,344]. In accordance with these data and with the microarray results show in Fig. 6.2, it was therefore assumed in the model that Smad2 and Smad4 are continuously synthesised (→ S2 and → S4). Degradation of all Smad species was also taken into account (S2 → ; S4 → ; nS2 → ; nS4 → ; pS2 → ; npS2 → ; (pS2)2S4 → ; (npS2)2S4 → ; (pS2)3 → ; (npS2)3 → ). For simplicity, it was assumed that the stability of monomeric Smad4 does not differ in the nucleus and in the cytosol, so that the same rate constant for the corresponding reactions (S4 → ; nS4 → ). Experimental studies indicate that Smad2 stability in the nucleus and in the cytosol may be different [345,346,347] and that stimulus-induced proteasomal degradation of Smad2 might occur in the nucleus at least in some cells [345]. Different degradation rate constants were therefore assumed for nuclear and cytosolic Smad2 monomers, respectively (S2 → ; nS2 → ; pS2 → ; npS2 → ). Experimental evidence suggests that Smad2 recruits ubiquitin ligases to Smad2/Smad4 heterotrimers [348]. This suggests that Smad2 plays a dominant role for the degradation of Smad heterotrimers. In the model, it was thus assumed that all Smad trimer species are degraded with the same rate constant as their monomeric Smad2 counterparts ((pS2)2S4 → ; (npS2)2nS4; (pS2)3 → ; (npS2)3 → ).

6.4.4 Modelling negative feedback via SnoN induction

Based on the microarray and the protein quantification analysis (Figs. 6.2 and 6.3), it seemed likely that SnoN is the major transcriptional feedback regulator of TGFβ signalling in primary mouse hepatocytes. SnoN-mediated feedback was therefore included in the model to quantitatively describe TGFβ signalling over a 10 h interval after stimulation.

SnoN Localisation: The SnoN oncoprotein is confined to the nucleus in most cultured cell lines, but is additionally localised to the cytoplasm in some freshly-isolated primary cells [349]. A model comprising nuclear and cytoplasmic SnoN did not lead to significant advantages compared to exclusively nuclear SnoN, whereas a model where SnoN was restricted to the cytoplasm was unable to explain experimental data in primary mouse hepatocytes (not
shown). This suggested that the nuclear SnoN pool is most important in primary mouse hepatocytes, possibly because it is much more abundant than the cytoplasmic pool. For simplicity, it was thus assumed in the model that SnoN is exclusively localised to the nucleus.

**Figure 6.5 Model fits to extracellular TGFβ data and to whole-cellular Smad2/4 data under standard stimulation conditions.**

The kinetic parameters of the model were estimated by fitting the model to experimental data (data points) obtained for various stimulation conditions. Solid lines represent the corresponding simulated time series obtained for the best-fit parameters. (A) The level of active TGFβ in the medium under standard stimulation conditions (1 ng/ml TGFβ) was measured in the absence (black points) and the presence (grey points) of the generic transcription inhibitor Actinomycin D, respectively. (B) Same as A, but the level of phospho-Smad2 was measured in whole-cell lysates. (C) Same as A, but the level of Smad4 co-immunoprecipitated with Smad2 was measured in whole-cell lysates. (D) Same as A, but the total level of Smad2 was measured in whole-cell lysates. Spline-based error estimates were used (Section 6.5.2).

**SnoN Synthesis:** The SnoN gene is regulated by TGFβ-dependent and -independent mechanisms, and accordingly SnoN mRNA is constitutively expressed in most cell lines [300,343,350]. In the model, an unregulated synthesis term was thus included for SnoN, and the corresponding synthesis rate was adjusted such that basal SnoN expression matched the experimentally measured level (~600 molecules per cell; Fig. 6.3 and Table F.3). Experimental studies indicated that TGFβ-dependent SnoN induction is mediated by Smad2-containing transcription factor complexes [350]. For simplicity, Smad homo- and heterotrimers stimulate SnoN synthesis in the model with the same efficiency (→ SnoNprecursor). Transcriptional induction is typically saturable (e.g., [351,352]), so that Smad-mediated SnoN induction was modelled as a Michaelis-Menten process with respect to Smad trimer concentration. During the fitting procedure, sub-nanomolar $K_M$-values for SnoN induction were allowed (Table F.3), as multiple adjacent Smad-binding elements (SBEs) have been reported for the SnoN promoter [350] to which Smad3 and Smad4 proteins are known to bind in a cooperative manner [353,354,355,356]. Due to the multistep nature of protein biosynthesis, a delay ~20 minutes was included in SnoN protein production (SnoNprecursor → SnoN).

**SnoN-Smad2 complex Formation:** The ability of SnoN to inhibit TGFβ signalling depends on its ability to form complexes with Smad2/3 [357]. It is known that SnoN and Smad2/3 complex formation within living cells is strongly enhanced upon TGFβ stimulation [358,359], and in vitro association studies indicate that SnoN selectively binds to phosphorylated R-Smads and to Smad trimers, respectively [358,360]. In the model it was thus assumed that SnoN specifically binds Smad homo- and heterotrimers in a reversible, bimolecular reaction.
(SnoN + (pS2)_2S4 ↔ SnoN−(pS2)_2S4 and SnoN + (pS2)_3 ↔ SnoN−(pS2)_3). Hence, the model reflects the experimental observation that SnoN and the related Ski protein do not disrupt Smad trimers [343,360,361].

Fig. 6.6 Model fits to whole-cellular Smad2/4 data under various stimulation conditions.

The kinetic parameters of the model were estimated by fitting the model to experimental data (data points) obtained for various stimulation conditions. Solid lines represent the corresponding simulated time series obtained for the best-fit parameters. Cells were either stimulated with (A – C) 1 ng/ml TGFβ or (D – F) 0.05 ng/ml TGFβ or (G – I) with 1 ng/ml TGFβ and the TGFβ receptor inhibitor SB431542 was added after 50 min or (J – L) with 1 ng/ml TGFβ which was removed after 50 min by an exchange of cell culture medium. Relative concentrations of phosphorylated Smad2 (A, D, G, J), Smad4 co-immunoprecipitated with Smad2/3 (B, E, H, K) and total Smad2 (C, F, I, L) in whole cell lysates were determined by quantitative immunoblotting and are directly comparable between conditions due to common scaling. The experimental variability was estimated using the error model (Section 6.5.2).

SnoN Degradation: It is well established that TGFβ stimulation strongly enhances SnoN degradation [343,358,362,363]. This seems to be due to the fact that trimers containing Smad2 or Smad3 recruit various ubiquitin ligases to SnoN, and thereby promote its proteasomal degradation [359,362]. Experimental studies indicate that Smad trimers ‘catalyse’ SnoN degradation, as R-Smad stability is only marginally affected by SnoN overexpression [358,362], while SnoN degradation is strongly enhanced by R-Smad overexpression [343,359]. Hence, it was assumed in the model that Smad-SnoN complexes decay into degraded SnoN and intact Smad trimers, while degradation of whole Smad-SnoN complexes was neglected. Experimental studies mapped SnoN-mediated transcriptional repression and Smad-mediated SnoN degradation to the same amino acids in SnoN [357,359], and hence justify the model assumption that the same Smad-SnoN complexes are responsible for both effects.
6.5 Model calibration

6.5.1 Experimental data for model calibration

Various stimulation experiments were performed and several pathway components were monitored by Peter Nickel in primary hepatocytes as a premise for data-based modelling. The following five stimulation conditions were tested experimentally: (i) standard TGFβ stimulation (1 ng/ml), (ii) weak TGFβ stimulation (0.05 ng/ml), (iii) standard TGFβ stimulation in the presence of the generic transcriptional inhibitor actinomycin D, (iv) standard TGFβ stimulation with addition of the specific TGFβ receptor kinase inhibitor SB431542 after 50 min, and (v) standard TGFβ stimulation with the ligand being removed after 50 min by medium exchange. Being central to TGFβ signalling, total SnoN, total Smad2, Smad2 phosphorylation and the amount of Smad4 co-immunoprecipitated with Smad2/3 were measured in whole cell lysates. Total and phosphorylated Smad2 pools were additionally measured in cytosolic and nuclear extracts to get insights into the dynamics of nucleocytoplasmic shuttling. Moreover, the amount of processed TGFβ in the cell culture medium was measured in the presence and absence of actinomycin D. All time course data sets used to calibrate the kinetic model by global parameter estimation are summarised in Table 6.1. For background correction, the lowest value of a time course was generally subtracted from all data points for all observed species except for total Smad2, for total SnoN and for TGFβ measured in the cell culture medium.

![Fig. 6.7 Model fits to nuclear and cytoplasmic Smad2/4 data.](Image)

The kinetic parameters of the model were estimated by fitting the model to experimental data (data points). Solid lines represent the corresponding simulated time series obtained for the best-fit parameters. Relative concentrations of phosphorylated and total Smad2 (A) in cytoplasmic extracts after stimulation with 1 ng/ml TGFβ in absence (black) or presence (grey) of actinomycin D and (B) in cytoplasmic and nuclear extracts after stimulation of cells with 1 ng/ml TGFβ were determined by quantitative immunoblotting. Spline-based error estimates were used (Section 6.5.2).
6.5.2 Scaling factors and error estimation

**Scaling Factors:** Western blot measurements are typically semi-quantitative in the sense that the results can only be expressed in arbitrary units, but not in absolute concentrations. Thus, scaling factors were introduced to fit the model (formulated in absolute concentrations) to time course data given in arbitrary units. An individual scaling factor was included for each observable (total Smad2, phosphorylated Smad2, Smad4 (CoIP), total SnoN, SnoN (CoIP), TGFβ) in the respective compartments, as indicated in Table 6.1, where each box corresponds to a single scaling factor. It is important to note that a single scaling factor was used for different stimulation conditions and for SnoN-depleted vs. wildtype cells in order to take biologically relevant scaling into account (Table 6.1).

![Fig. 6.8 Model fits to whole-cellular SnoN data under various stimulation conditions.](image)

The kinetic parameters of the model were estimated by fitting the model to experimental data (data points). Solid lines represent the corresponding simulated time series obtained for the best-fit parameters. Cells were either stimulated with (A) 1 ng/ml TGFβ or (B) with 0.05 ng/ml TGFβ or (C) with 1 ng/ml TGFβ and actinomycin D was added at 0 min or (D) SB431542 was added after 50 min. Relative concentrations of total SnoN in whole cell lysates were determined by quantitative immunoblotting and are directly comparable between conditions (A-D) due to common scaling. Error estimation was based on a spline-derived error model (Section 6.5.2).

The quality of a mathematical model is determined by its deviation from the data relative to the measurement error, so that the variability of the experimental data had to be estimated. All measurements listed were performed at least twice with similar results. As replica measurements frequently differed in the observed time points, errors could not be directly derived by calculating standard deviations. The errors of 14 densely measured time courses (each comprising 18 or more time points) were therefore estimated by taking the distance of each data point from a cubic spline. From these spline-based error estimates linear error models were derived for the observables total Smad2, phosphorylated Smad2, Smad4 (CoIP), SnoN (CoIP) and TGFβ. For each time series y(t), these error models, considering relative errors $E_{\text{rel}}$ and absolute errors $E_{\text{abs}}$, had the form

$$E_{\text{tot},i} = E_{\text{rel}} \cdot |y(t_i)| + E_{\text{abs}} \cdot C.$$
Here, the constant C is given by the 0.9-Quantile of the whole time course, i.e., \( C = Q_{0.9}(y(t_1...t_i)) \). These error models were used to estimate the errors in less densely sampled time courses (Table 6.1). For total SnoN, no spline-based error model could be derived from the measurements, so that typical values (\( E_{\text{rel}} = 10\% \); \( E_{\text{abs}} = 15\% \)) were applied. The relevance of SnoN-mediated feedback was confirmed by analyzing Smad signalling in cells harbouring reduced levels of functional SnoN (see below). The corresponding experiments in SnoN siRNA cells and in knock-in cells expressing mutant SnoN were performed in triplicates with identical time points analysed in each experiment, so that the errors could be estimated by calculating standard deviations. For a comprehensive overview of the error estimation for different data sets see Table 6.1.

### 6.5.3 Parameter estimation

Mathematical modelling was done using the multi-experiment fitting Matlab toolbox PottersWheel (www.potterswheel.de, Maiwald et al., under revision). The reaction scheme (Fig. 6.4) was translated into a set of 19 ordinary differential equations. The final model additionally included 9 auxiliary variables (see Equations F.1 and F.2), which were required to model a delay in SnoN induction and the mRNA decay in response to actinomycin D treatment. Several of the 57 model reactions were assumed to proceed with identical kinetics, so that 36 kinetic parameters were subject to fitting. All parameters were optimised globally, that is, a single parameter set was able to describe all time courses. During the parameter estimation, each parameter was restricted to the range specified in Tables F.1 – F.3. According to Table 6.1, there are 10 scaling parameters, all of which were also fitted, and allowed to vary in a large range \( (10^{-5} – 10^{5}) \). The initial conditions (i.e., the initial protein concentrations) are not subject to separate fitting, as they are determined by the kinetic parameters for basally expressed species (see Section F.3) or, in the case of the extracellular TGF\( \beta \) concentration, by the experimental conditions (Section 6.5.4). The start values for the auxiliary mRNA species were also fixed and set to 1 (see Appendix F). Finally, the experimentally measured volumes of medium, cytosol and nucleus were set to experimentally measured values (see Section 6.4.1), and were also excluded from the fitting procedure. PottersWheel uses the weighted \( \chi^2 \) value for parameter optimisation,

\[
\chi^2 = \sum_{i=1}^{N} \frac{(y_{\text{Model}}(i) - y_{\text{Meas}}(i))^2}{\sigma_{\text{Meas}}(i)}.
\]

\( N \) is the number of all data points and \( \sigma_{\text{Meas}}(i) \) is the error of the i-th measurement, \( y_{\text{Meas}}(i) \). In order to circumvent local minima, a two-step strategy was applied. First, 2800 quasi randomly distributed positions in the space of physiologically reasonable parameter values (see Supplementary Tables F.1 – F.3) were used as starting conditions. Fitting was applied with a deterministic trust region optimiser with a \( \chi^2 \) tolerance of \( 10^{-4} \), a fit parameter tolerance of \( 10^{-4} \), and a maximum of 600 iterations. Then, the best fit was disturbed and refitted 1600 times, i.e., the fitted parameter values were transformed to \( p_{\text{new}} = p_{\text{old}} \times 10^{s \times e} \), with \( s = 0.3 \) and \( e \) being normally distributed with variance 1 and mean 0.

### 6.5.4 Input functions

The model trajectories depend on the characteristics of five external input functions, termed \( u_1 – u_5 \) in the differential equations (Equations F.1 and F.2). The input \( u_1 \) equals the initial TGF\( \beta \) concentration, and thus determines the strength of stimulation. Ligand removal experiments by medium exchange were simulated by changing the input \( u_5 \) from 1 to 0, which abolishes de novo ligand binding to TGF\( \beta \) receptors in the model (see \( v_7 \) in Equations F.2). The input \( u_3 \) controls receptor-mediated Smad2 phosphorylation (\( v_{16} \) in Equations F.2),
and was changed from 0 to 1 to simulate addition of the TGFβ receptor kinase inhibitor SB431542. In some experiments, TGFβ signalling was monitored over 10h in the presence of the general transcription inhibitor, actinomycin D, and this was modelled by setting the input $u_2 = 1$ (with $u_2 = 0$ otherwise). The setting $u_2 = 1$ induces an exponential decay of the auxiliary mRNA species ($v_1$, $v_3$, $v_5$, $v_{43}$ in Equations F.2) and inhibits Smad-induced SnoN synthesis ($v_{45}$ in Equations F.2), thereby mimicking general inhibition of transcription in the model. Finally, experiments in SnoN-depleted hepatocytes were simulated by setting $u_4 = 1$ (with $u_4 = 0$ otherwise), as this blocks both induced and basal SnoN expression ($v_{44}$ and $v_{45}$ in Equations F.2). The parameter $k_1$ controls the effective degree of SnoN depletion, and was set to 1 in experiments with the Smad binding-deficient SnoN mutant (complete depletion), while the value was adjusted to 0.7 to simulate siRNA-mediated SnoN knock-down (partial depletion).

### 6.5.5 Results of model calibration

As shown in Figs. 6.5 – 6.8, the model could simultaneously be fitted convincingly to all data sets. The final best fit had a $\chi^2$ value of 355. Given $N = 506$ data points, the model hypothesis could not be rejected ($p < 0.05$).

The amount of active TGFβ in the medium continuously decreases over the 10 h time course (Fig. 6.5A). This seems to be due to receptor-mediated ligand endocytosis and degradation, as the observed ligand decay can be quantitatively described by the model if experimentally measured kinetic parameters are assumed for the receptor trafficking module (Table F.1). Interestingly, the measured ligand decay time course is not affected by actinomycin D treatment (Fig. 6.5A), thus suggesting autocrine stimulation does not play a significant role in primary mouse hepatocytes, even though induction of TGFβ and of ligand processing enzymes was observed at the mRNA level (Fig. 6.2).

Despite the strong decay in extracellular TGFβ levels, sustained Smad2 phosphorylation and Smad2-Smad4 complex formation occurred in primary mouse hepatocytes (Fig. 6.5B and C), and the model suggests that is due to strong saturation at the receptor level. Incubation of cells with actinomycin D induced some degradation of Smad2 (Fig. 6.5D). Yet, the level of phosphorylated Smad2 was not significantly affected by actinomycin D treatment (Fig. 6.5B), while the amount of Smad4 co-immunoprecipitated to Smad2 dropped to approximately half of the control (Fig. 6.5C). Thus, Smad2 phosphorylation and Smad2-Smad4 complex show distinct responses towards actinomycin D, while they behave very similarly under various other stimulation protocols (Fig. 6.6). This suggests that complex formation between Smad2 and Smad4 (but not Smad2 phosphorylation) might be controlled by transcriptional feedback loops, as analysed in more detail below (Section 6.6).

Immunoblot measurements of Smad signalling in nuclear and cytoplasmic compartments, respectively, revealed weak translocation of Smad proteins into the nucleus in TGFβ-stimulated primary mouse hepatocytes (Fig. 6.6). In the model, the relatively small stimulus-induced change in both, nuclear and cytoplasmic, Smad2 pools could only be explained if Smad2 was roughly equally distributed between both compartments in resting cells. In line, it has recently been reported in the literature that Smad2 is present at similar amounts in the cytoplasm and in the nucleus of unstimulated primary mouse hepatocytes [364].

The SnoN protein expression level was measured under various stimulation conditions (Fig. 6.8). SnoN was induced with a delay of ~ 60 min in primary mouse hepatocytes (Fig. 6.8A), as reported previously for cell lines [343]. Actinomycin D incubation led to a rapid decline in SnoN protein levels (Fig. 6.8C). In the model this was mainly due to TGFβ-induced SnoN degradation mediated by phosphorylated Smad2 (see Section 6.4.4) which cannot be
compensated for by TGFβ-induced SnoN synthesis in actinomycin- treated cells. The alternative

**Fig. 6.9 Experimental testing of model predictions for mutant SnoN form abrogating SnoN-Smad interactions.**

Primary hepatocytes from Smad binding deficient mSnoN knock-in animals or wild-type animals were stimulated with 1 ng/ml TGFβ for the indicated times. Relative concentrations of either SnoN or Smad4 co-immunoprecipitated with Smad2/3 (A and B) as well as phosphorylated and total Smad2 (C and D) in whole cell lysates were determined by quantitative immunoblotting (data points). Trajectories (dashed lines) represent model predictions for mSnoN (grey) or wt cells (black). Standard deviations of experimental data points were calculated from triplicates (see Section 6.5.2).

explanation, i.e., very rapid degradation of constitutively expressed SnoN mRNA and thus swift termination of constitutive SnoN protein synthesis (without stimulus-enhanced degradation), seems unlikely in light of RNA experiments done by Peter Nickel under the same conditions (1 ng/ml TGFβ + actinomycin D). He found that the SnoN mRNA half-life in primary mouse hepatocytes is ~ 100 min, which is inconsistent with a half-maximal protein decay in actinomycin-treated cells within 50 min. SnoN measurements were also performed in cells where Smad signalling was terminated 50 min after TGFβ stimulation by addition of the TGFβ receptor inhibitor SB431542 (Fig. 6.8D). A very slow SnoN protein decay was observed after inhibitor treatment, most likely because Smad2-mediated SnoN degradation was blocked in addition to Smad2-mediated SnoN transcription.

The core components of TGFβ signalling such as the TGFβ receptor, Smad2 and Smad4 are all characterised by relatively stable mRNAs and/or proteins (Tables F.1 – F.3). Thus, the alterations in Smad signalling done in the presence of actinomycin D are likely to reflect at least in part transcriptional feedback regulation, since actinomycin can be assumed to mainly inhibit Smad-induced de novo transcription. Actinomycin D had little impact on Smad2 phosphorylation, but strongly affected the amount of Smad4 co-immunoprecipitated with Smad2. The observation that SnoN, a putative feedback regulator in hepatocytes, rapidly decayed in actinomycin-treated cells suggested that SnoN might be responsible for altered Smad heterotrimerisation. In Section 6.6, the impact of SnoN on Smad signalling was therefore investigated numerically and the model predictions were then confirmed experimentally.
6.6 Testing the model predictions

6.6.1 Prediction of Smad signalling dynamics in SnoN-depleted cells

Experiments in the presence of the generic transcription inhibitor actinomycin D suggested that transcriptional feedback regulation might affect complex formation between Smad2 and Smad4 (Section 6.5.5). Experimental measurements of TGFβ–induced mRNA and protein expression indicated that SnoN is the major transcriptional feedback regulator for Smad signalling in primary mouse hepatocytes (Sections 6.2 and 6.3). Numerical simulations of a ‘SnoN-- model’ devoid of constitutive and induced SnoN expression were performed to investigate the role of SnoN in more detail. The simulated time courses, shown in Fig. 6.9 (dashed lines), reveal that SnoN depletion selectively decreases complex formation between Smad2 and Smad4 (Fig. 6.9B), but does not affect Smad2 phosphorylation (Fig. 6.9C). Thus, the model predicted that: (i) altered Smad2-Smad4 heteromerisation in actinomycin-treated cells is indeed due to abrogation of transcriptional feedback regulation; (ii) SnoN efficiently modulates Smad signalling despite being present at sub-stoichiometric amounts when compared to Smad2 (Fig. 6.3). Interestingly, these model predictions were very robust, as very similar results were obtained for the top 100 fits of the global fit sequence (see Section 6.5.3 and not shown).

Fig. 6.10 Experimental testing of model predictions for siRNA-mediated SnoN knock-down.

After transfection of SnoN siRNA or control siRNA primary hepatocytes were stimulated with 1 ng/ml TGFβ for the indicated times. Relative concentrations of either SnoN or Smad4 co-immunoprecipitated with Smad2/3 (A and B) as well as phosphorylated and total Smad2 (C and D) in whole cell lysates were determined by quantitative immunoblotting (data points). Trajectories (dashed lines) represent model predictions for SnoN siRNA (grey) or control siRNA (black). Standard deviations of experimental data points were calculated from triplicate measurements (see Section 6.5.2).

To confirm the model predictions in Fig. 6.9, experiments were performed by Peter Nickel in a transgenic knock-in mouse model, kindly provided by Kunxin Luo (University of California, Berkeley). In these mice, both wild-type alleles for SnoN were replaced by the mutant form, mSnoN which was described to be deficient in binding to Smad2/3 and Smad4 [357]. Therefore, expression of mSnoN instead of the wild-type protein was expected to fully disrupt SnoN-mediated feedback regulation of TGFβ-Smith signalling. TGFβ time course experiments were performed using hepatocytes isolated from knock-in mice harbouring the mSnoN mutant and compared to the wild-type situation. As expected, no association of mSnoN with Smad2/3 could be detected (Fig. 6.9A), although full-length mSnoN was expressed at levels comparable to SnoN levels in wild-type cells (Peter Nickel, not shown). In line with the model predictions, total and phosphorylated Smad2 did not appreciably differ.
between wild-type and mutant cells upon TGFβ stimulation (Figs. 6.9C and D). More importantly, levels of co-immunoprecipitated Smad4 were significantly lowered in mutant cells when compared to wild-type, in quantitative agreement with the model predictions (Fig. 6.9B).

![Fig. 6.11 Experimental testing of model predictions for SnoN-Smad interactions.](image)

Trajectories (dashed lines) represent model predictions for previously untested experimental conditions (data points). (A) Association of SnoN with Smad2/3 after stimulation with 1 ng/ml TGFβ in absence (black) and presence (grey) of actinomycin D. Association of SnoN with Smad2/3 after stimulation with (B) 1 ng/ml TGFβ or (C) 0.05 ng/ml TGFβ or (D) 1 ng/ml TGFβ and addition of TGFβ receptor inhibitor SB431542 at 50 min or (E) 1 ng/ml TGFβ and removal of ligand by exchanging medium after 50 min. Relative concentrations of SnoN co-immunoprecipitated with Smad2/3 in whole cell lysates were determined by quantitative immunoblotting and are directly comparable between conditions (B-E) due to common scaling. In (A) spline-based error estimates and in (B-E) a spline-derived error model were used (see Section 6.5.2).

An RNAi approach was employed by Peter Nickel to obtain an independent experimental confirmation of the model prediction that SnoN selectively modulates Smad2-Smad4 complex formation. Primary hepatocytes were transfected with siRNA directed against SnoN mRNA prior to stimulation experiments in order to specifically knock-down SnoN expression. As a negative control, siRNA against Renilla luciferase was used. The experimental results again reveal a selective effect of SnoN depletion on Smad2-Smad4 heteromerisation (Fig. 6.10), and thus further support the validity of the model predictions. It was then addressed whether the model was able to quantitatively predict the results in SnoN siRNA cells. The siRNA-mediated SnoN knock-down was simulated by decreasing basal and Smad-induced SnoN transcription to the same extent. The efficiency of knock-down was estimated by fitting the (basal and induced) SnoN transcription rates such that the model optimally matched the measurements shown in Fig. 6.10. The fitting results (dashed lines in Fig. 6.10) are in quantitative agreement with the measured data points. Moreover, the residual SnoN
expression level in the fitted model (~30% relative to wild-type control) was close to the experimentally measured value (~35% relative to wild-type control).

Taken together, the model was able to quantitatively predict the dynamics of Smad signalling in SnoN knock-down and knock-in cells. Interestingly, the selective decrease in Smad2-Smad4 complex formation (but not Smad2 phosphorylation) observed in SnoN-depleted cells (Fig. 6.10) is in quantitative agreement with the dynamical behaviour seen in actinomycin-treated wildtype cells (Fig. 6.5). In other words, complete SnoN depletion pheno-copies generic inhibition of all transcriptional feedback loops, thus strongly suggesting that SnoN is the central transcriptional feedback regulator in primary mouse hepatocytes.

Fig. 6.12 Model simulations of different nuclear Smad populations.

(A) Relative levels of nuclear pSmad2 monomers (dotted line) and nuclear Smad trimers (solid line) as predicted by the best-fit model. The maximum level of pSmad2 monomers was set to 100%. (B) The effect of decreased SnoN expression in cells harbouring siRNA against SnoN on formation of nuclear Smad complexes. Levels of active nuclear Smad trimers (solid lines) and repressing SnoN-bound nuclear Smad trimers (dashed lines) are shown for the unperturbed state (black) and for simulated SnoN knock-down (grey). The parameters used for the knock-down simulations are the same as those in Fig. 6.10. The maximum concentration of active nuclear Smad trimers (i.e., the sum of free homo- and heterotrimers) in the unperturbed state was set to 100%.

6.6.2 Analysis of SnoN-mediated regulation of gene expression

The model successfully predicted Smad signalling in cells harbouring reduced levels of SnoN protein (Section 6.6.1), but it remained to be clarified whether SnoN-mediated feedback operates via complex formation between Smad2 and SnoN, as suggested by the model. Complex formation between SnoN and Smad2 in wild-type cells was simulated for various stimulation conditions, and the amount of SnoN co-immunoprecipitated with Smad was measured by Peter Nickel to test for the model predictions. The simulated time courses were again in quantitative agreement with the experimental measurements (Fig. 6.11), thus suggesting that SnoN indeed controls Smad signalling by directly binding to Smad trimers.
It remained to be determined by which mechanism the heteromerisation of Smad proteins could be efficiently modulated by SnoN, even though SnoN is expressed at much lower levels than the Smad proteins (Fig. 6.3). To this end, the best-fit model was analysed numerically. According to the simulations, the concentration of nuclear Smad homo- and heterotrimers is approximately 10-fold lower than that of nuclear Smad2 monomers (Fig. 6.12A). As nuclear Smad trimers are specifically targeted by SnoN-mediated feedback inhibition (Section 6.4.4), the effective Smad/SnoN ratio shifts in favour of SnoN by approximately 10-fold. The interacting populations of nuclear SnoN and nuclear Smad trimers are thus present at similar levels of some thousand molecules per cell. This explains why low amounts absolute amounts of SnoN can have a substantial impact on Smad complex formation. The results shown in Fig. 6.12A imply that Smad trimerisation is inefficient for the Smad2 and Smad4 concentrations prevailing in hepatocytes. Some support for this hypothesis comes from co-immunoprecipitation studies of endogenous Smad2 and Smad3 proteins in various cell lines. More specifically, it was shown that immunoprecipitates of endogenous Smad3 protein contain negligible amounts of Smad2 when compared to Smad3 [365]. Given that Smad2 and Smad3 are structurally and functionally very similar, one would expect efficient Smad2-Smad3 complex formation (i.e., co-immunoprecipitation) if Smad trimerisation occurred with high affinity.

Fig. 6.13 Differential expression of TGFβ-induced target genes after perturbation of SnoN feedback.

Fold changes in JunB, CTGF, PAI-1, and RhoB mRNA levels after stimulation with 1 ng/ml TGFβ for 8 h or 24 h are expressed as fold-changes relative to the unstimulated state. For the siRNA-mediated SnoN knock-down in wild-type cells (left panel), values for SnoN siRNA (black bars) and control siRNA (white bars) are shown. For the mutant/wild-type comparison (right panel), mRNA levels in cells from knock-in mice harbouring the Smad binding deficient mSnoN mutant (black bars) and wild-type mice (white bars) are shown. Relative mRNA levels were determined by quantitative RT-PCR. Figure kindly provided by Peter Nickel.
Fig. 6.14 Model predictions for perturbed SnoN feedback regulation.

(A – C) The total cellular SnoN expression level in response to stimulation with 1 ng/ml TGFβ is simulated for the best-fit model (‘wildtype cells’), for a model with 10-fold increased basal SnoN expression, and for a mutant model devoid of Smad-induced SnoN degradation (i.e., SnoN in Smad complexes is degraded with the same rate constant as free SnoN). The experimentally measured data points shown in (A) are the same as those in Fig. 6.8A. (D – F) The amount of Smad4 co-immunoprecipitated with Smad2 is plotted for the three model variants mentioned above. The experimentally measured data points shown in (D) are the same as those in Fig. 6.5C. (G – I) The levels of active nuclear Smad trimers (solid lines; sum of homo- and heterotrимерs) and repressing SnoN-bound nuclear Smad trimers (dashed lines) are shown for the three model variants mentioned above. The ratio of active nuclear Smad trimers and repressing SnoN-bound nuclear Smad trimers is likely to reflect Smad-induced gene expression, as both species are expected to compete for the same promoter sites (see Section 6.6.2). Note that the plots in each row have the same y-axis scaling, and are thus directly comparable.

Complex formation between Smad trimers and SnoN is thought to be necessary and sufficient for SnoN-mediated inhibition of gene expression, as SnoN binding interferes with the recruitment of histone acetylases to Smad trimers [366]. The dynamic interactions between nuclear Smad trimers and nuclear SnoN were therefore investigated by numerical simulations (Fig. 6.12B). To get insights into SnoN-mediated regulation of gene expression, the analysis was focused on the ratio of active free Smad trimers to repressing SnoN-bound Smad trimers. Both species are expected to compete for the same promoter sites, as SnoN family proteins do not interfere with DNA binding of Smad trimers [367,368]. Thus, the transcription of Smad target genes will be efficiently blocked in case that the ratio of free to SnoN-bound Smad trimers is low. The simulations, shown in Fig. 6.12B, suggest that SnoN-bound trimers dominate over free Smad trimers in wild-type cells, while the opposite is true in hepatocytes transfected with siRNA directed against SnoN mRNA. Thus, the model predicts that Smad-induced gene expression is significantly enhanced in SnoN siRNA cells, and, in particular, in mSnoN knock-in cells.
It was therefore determined experimentally by Peter Nickel whether depletion of SnoN has a significant effect on TGFβ-induced gene expression. To this end, the two perturbation conditions of siRNA-mediated SnoN knock-down and the Smad binding deficient mSnoN mutant were again employed. Hepatocytes were stimulated with 1 ng/ml TGFβ for 8 h and 24 h, and the induction of known TGFβ target genes (JunB, CTGF, PAI-1 and RhoB) at the mRNA level was measured by RT-PCR. Fold changes in mRNA levels over the unstimulated state are shown in Fig. 6.13. Firstly, siRNA-mediated knock-down of SnoN was performed in wild-type primary hepatocytes and a knock-down using siRNA against Renilla luciferase was used as a negative control (Fig. 6.13, left panel). Secondly, hepatocytes isolated from knock-in animals harbouring the Smad binding deficient mSnoN mutant were compared to hepatocytes from wild-type mice (Fig. 6.13, right panel). Compared to unperturbed control stimulations (Fig. 6.13, white bars) all four genes showed a clear tendency for enhanced upregulation if inhibition of Smad signalling by SnoN was partially or fully disrupted (Fig. 6.13, black bars). This SnoN-dependent effect on mRNA induction tended to be more pronounced after 24 h than after 8 h of TGFβ treatment. In conclusion, the quantitative mRNA measurements in Fig. 6.13 clearly demonstrate that SnoN is involved in the regulation of TGFβ-induced genes at the transcriptional level in primary hepatocytes. This strongly supports the model prediction that SnoN regulates the physiological TGFβ-Smad signalling responses in primary hepatocytes despite being present at sub-stoichiometric concentrations. It should, however, be noticed that SnoN depletion did not affect the expression of all TGFβ target genes analysed (not shown), suggesting that SnoN acts to regulation Smad-induced transcription in a promoter-specific manner.

6.7 Conclusions and outlook

In the present chapter, a systems biological approach was applied to get insights into transcriptional feedback regulation of TGFβ signalling. Putative transcriptional feedback loops were identified by genome-wide expression profiling (Fig. 6.2). Analysis of TGFβ-induced expression changes at the protein level (Fig. 6.3) allowed to further narrow the number of transcriptional feedback loops, and indicated that SnoN is the major transcriptional feedback regulator in primary mouse hepatocytes. Autocrine TGFβ stimulation of primary mouse hepatocytes could be discarded as a feedback mechanism, as the amount of extracellular TGFβ decayed monotonically (Fig. 6.5A), most likely due to receptor-mediated ligand degradation. Moreover, the strong induction of Smad7 at the mRNA level was surprisingly not accompanied by increased Smad7 protein expression (Fig. 6.3), possibly due to regulation at the post-transcriptional level, e.g., by microRNA. For several cell lines, Smad7 has been discussed as a crucial transcriptional feedback regulator leading to the termination of TGFβ-Smad signal transduction by blocking TβRI-mediated R-Smad phosphorylation as well as by triggering receptor dephosphorylation and degradation [47]. In primary mouse hepatocytes, Smad7 can indeed inhibit Smad signalling, as strong adenoviral overexpression of Smad7 by Peter Nickel almost totally abolished TGFβ-dependent activation of Smad2/3 (not shown). On the other hand, transgenic hepatocytes expressing a functionally impaired form of Smad7 [369] did not display a significant increase in Smad2/3 activation (Peter Nickel, not shown). This indicates that Smad7 might not even act as a transcriptionally unregulated negative modulator of TGFβ-Smad signalling in primary mouse hepatocytes, despite being constitutively expressed at high levels (Fig. 6.3). Accordingly, it has recently been reported that physiological levels of Smad7 are insufficient to inhibit Smad2/3 signalling in endothelial cells [370].

A mathematical model of TGFβ-Smad signalling (Fig. 6.4) was implemented in order to get further insights into SnoN-mediated transcriptional feedback regulation. Smad3 was not explicitly considered in the modelling approach for several reasons. First, the absolute concentration of Smad3 is substantially lower than that of Smad2 (Fig. 6.3). Second, the general dynamics of Smad3 phosphorylation closely resembled those of Smad2 in primary
hepatocytes (Peter Nickel, not shown). Finally, Smad2 and Smad3 are known to be highly homologous in their sequence and structure, and are thought to be largely equivalent in their binding affinities for TGFβ receptors, Smad4 and SnoN [366,371]. It should, however, be noted that Smad2 and Smad3 are not redundant and have been described to be differentially involved in the regulation of various physiological processes [372], despite being highly similar. For example, hepatocyte-specific Smad2/3 knock-out models suggest a central role of Smad2 in suppressing hepatocyte growth in the basal state, while Smad3 seems to be the primary mediator of TGFβ-dependent apoptosis [373]. Thus, future model versions taking into account phenotypic responses induced by TGFβ stimulation must distinguish between Smad2 and Smad3.

The kinetic model of TGFβ-Smad signalling was fitted to quantitative measurements of total SnoN, total Smad2, Smad2 phosphorylation, of the amount of Smad4 co-immunoprecipitated with Smad2/3 and of TGFβ in the medium (Figs. 6.5 – 6.9). Based on published experimental measurements, a physiological relevant range was defined for each parameter (Tables F.1 – F.3), and a single set of parameters was able to simultaneously describe all time courses. Fitting analyses during the construction of the final model revealed that (at least some of) the relatively stringent parameter ranges in Tables F.1 – F.3 were required to obtain a physiologically relevant and predictive optimisation result. For example, the selective decrease in Smad4 co-immunoprecipitated with Smad2 in actinomycin-treated cells despite unchanged Smad2 phosphorylation (Fig. 6.5) could also be explained by an alternative model devoid of SnoN feedback if Smad4 is assumed to be very unstable (with a half-life in the order of minutes). Then, actinomycin D reduces Smad2-Smad4 complex formation by inducing a rapid Smad4 decay. Published experimental studies, however, revealed that Smad4 mRNA and protein are relatively stable (Table F.2). Moreover, the experiments in SnoN-depleted cells (Figs. 6.9 and 6.10) strongly favour SnoN-mediated stabilisation of the Smad2-Smad4 complex, and argue against a mechanism with Smad4 decay. In conclusion, the fitting analyses indicate that parameter estimation in systems biology requires a definition of physiologically relevant parameter ranges, and cannot solely be based on unconstrained optimisation of the $\chi^2$-value.

Many of the experimental data sets used for model calibration were actually required to obtain a physiologically relevant and predictive optimisation result. For example, total Smad2, was analysed in cytoplasmic and nuclear compartments, in addition to measurements in whole-cell lysates. It turned out that two of three data sets (whole-cell, nuclear, cytoplasmic) were required for a reasonable fitting result, as: (i) whole-cell data reveals whether the intracellular Smad2 pool is subject to stimulus-induced synthesis or degradation; (ii) nuclear and/or cytoplasmic data provides insights nucleo-cytoplasmic shuttling of Smad proteins. Accordingly, early fitting attempts using cytoplasmic data only failed to predict nuclear and whole-cell data (both in a quantitative and qualitative sense). Thus, a critical number of experiments are necessary to generate predictive data-based models for systems biology. In my opinion, the particular choice of experiments depends on the experience of researchers, and cannot be solely derived from
Fig. 6.15 Model analysis of TGFβ signalling in HaCaT cells.

Top: SnoN-mediated feedback regulation was eliminated from the original model (Fig. 6.4) in order to describe TGFβ signalling in HaCaT cells, which do not show significant transcriptional feedback regulation. The expression levels of the TGFβ receptor, of Smad2 and Smad4, and several kinetic parameters were altered relative to the hepatocyte model to match experimental measurements in HaCaT cells (see panels A–I). More specifically, the reaction parameters for protein synthesis (receptor, Smad2 and Smad4), for endosomal sorting of receptor-ligand complexes, for lysosomal degradation, and for Smad trimer degradation (referred to as parameter set 1 below) were fitted to HaCaT data, as indicated by the black symbols in the model scheme. (A – B) Time courses of Smad2 phosphorylation in whole-cell lysates (A) and nuclear extracts (B). Data points in (A) are taken from [342] and those in (B) from [311,374]. The solid line represents the model trajectory. (C – F) Time courses of cytoplasmic and nuclear Smad2 and Smad4 pools (data points taken from [338]). Note that the same y-axis scaling was used in panels (C) and (D), and for panels (E) and (F). Two model fits are shown in each panel: The grey line represents a model, where only parameter set 1 (see above) was fitted to HaCaT data, while the nuclear import and export reactions of Smad2 and Smad 4 were additionally optimised for the black line (note that all other fits in panels A – B and G – I look very similar for both fits). Numerical simulations of a mutant model devoid of any Smad trimerisation (dotted lines) are also shown in panels (C) and (D). (G – H) Time courses of Smad2 phosphorylation in whole-cell lysates in cells treated with the proteasome inhibitor MG132 (grey line and data points) are compared to control cells (black line and data points). Cells were either continuously incubated
with TGFβ (G), or the ligand was washed out by medium exchange after 30 min (H). Data was taken from [345]. (I) Time courses of Smad2 phosphorylation in whole-cell lysates in cells treated with the TGFβ receptor inhibitor SB431542 (black solid line and circles), treated with the TGFβ receptor inhibitor SB431542 together with a proteasome inhibitor (black dotted line and triangles), or treated with the TGFβ receptor inhibitor SB431542 after a shRNA-mediated knock-down of the nuclear Smad2 phosphatase PPM1A (grey solid line and circles). Data was taken from [342]. Error bars were generally calculated using an error model with typical relative and absolute error values (E_rel = 10% and E_abs = 10%; see Section 6.5.2). Errors bars are not shown in panel (I) for clarity.

Mathematical modelling strongly suggested that transcriptional feedback via SnoN modulates Smad signalling at the level of Smad trimerisation, and thus downstream of Smad phosphorylation. In line with the model prediction, SnoN-depleted cells showed reduced Smad2-Smad4 complex formation when compared to wildtype cells, while Smad phosphorylation was unchanged (Figs. 6.9 and 6.10). Remarkably, the effect of SnoN depletion (Fig. 6.9) on Smad phosphorylation and trimerisation was qualitatively and quantitatively similar to the effect of actinomycin D treatment (Fig. 6.5), thus suggesting that SnoN is the major transcriptional feedback regulator in primary mouse hepatocytes. The functional relevance of SnoN-mediated feedback was further corroborated by the finding that depletion of SnoN enhances Smad-induced gene expression (Fig. 6.13).

A variety of tumors constitutively release large amounts of TGFβ and thereby suppress immune responses or induce angiogenesis, while being themselves insensitive towards the growth-inhibitory effects of autocrine TGFβ stimulation [46]. The SnoN oncogene is commonly overexpressed in tumors, and might contribute to such TGFβ insensitivity by inhibiting the anti-proliferative Smad signalling pathway [366]. Constitutive SnoN overexpression was simulated in the best-fit model by increasing the basal SnoN synthesis rate by 10-fold, while leaving the kinetic parameters of Smad-induced SnoN expression unchanged. The model predicts that constitutive SnoN overexpression has little effect on the amount of SnoN-repressed Smad trimers in response to stimulation (Fig. 6.14B, E and H), and thus on TGFβ-induced gene expression. Previous literature studies indicated that SnoN overexpression can be insufficient to fully suppress TGFβ signalling [343], and thus support the simulation results.

SnoN is subject to incoherent feed-forward regulation in response to TGFβ stimulation, as both the synthesis and the degradation rates of SnoN are strongly enhanced by Smad signalling (each by more than one order of magnitude in the best-fit model). This raises the possibility that the cellular sensitivity towards TGFβ is mainly governed by the Smad-induced synthesis and degradation rates of SnoN, while the basal synthesis and degradation rates play no major role. Therefore, numerical simulations were performed for hypothetical primary mouse hepatocytes devoid of Smad-induced SnoN degradation. More specifically, SnoN in Smad complexes was assumed to be degraded with the same rate constant as free SnoN. The results, shown in Fig. 6.14C, F and I, reveal that Smad-induced SnoN degradation has indeed a large impact on the amount of SnoN-bound repressive Smad trimers (Fig. 6.14I), and thus on Smad-induced gene expression. These simulation results are corroborated by experimental studies showing that overexpression of non-degradable SnoN inhibits TGFβ signalling much more efficiently than overexpression of wildtype SnoN [343].

Taken together, the simulation results in Fig. 6.14 indicate that the Smad-induced (but not the basal) synthesis and degradation rates of SnoN determine the cellular TGFβ sensitivity. Accordingly, it has recently been proposed that a loss of TGFβ-induced SnoN degradation enables various tumours to escape TGFβ-induced growth inhibition [375,376]. The predictions in Fig. 6.14 can be tested for by overexpressing SnoN in primary mouse hepatocytes (panels B, E and H) and by treating cells with proteasome inhibitors (panels C, F and I). It has previously been shown that TGFβ-induced SnoN degradation can be blocked by proteasome inhibitors such as MG132 and Lactacystin [358,359,362]. Given the relatively
slow turnover of Smad proteins and the strong receptor saturation in primary mouse hepatocytes (see above), proteasomal inhibition is expected to mainly affect SnoN degradation, and not other parts of the Smad signalling cycle. Experiments employing a combination of proteasomal inhibitors and actinomycin D (to simultaneously block SnoN induction and degradation) might be used to further test for the model predictions.

It was surprising to see that sub-stoichiometric amounts of SnoN efficiently regulate the much larger pools of Smad2 and Smad4 (Figs. 6.3, 6.9 and 6.10). The mathematical model presented in this chapter indicates that this due to relatively inefficient trimerisation of Smad proteins. The simulations predict that the concentration of nuclear Smad homo- and heterotrimers is approximately 10-fold lower than that of nuclear Smad2 monomers (Fig. 6.12A). As SnoN-mediated feedback specifically targets nuclear Smad trimers (Section 6.4.4), the effective Smad/SnoN ratio shifts in favour of SnoN, so that the interacting populations of nuclear SnoN and nuclear Smad trimers are present at similar levels. The nuclear translocation of Smad4 in response to TGFβ stimulation depends on its ability to heteromerise with Smad2/3 proteins [377] (Fig. 6.4). Given the weak trimer affinity in the model, it was not surprising to see that only low amounts of Smad4 translocate into the nucleus in the best-fit hepatocyte model (not shown). While weak Smad4 translocation in primary mouse hepatocytes remains to be confirmed experimentally, it is well established that significant amounts of Smad4 translocate into the nucleus in other cellular systems such as HaCaT cells [338]. Numerical simulations were therefore performed to investigate whether low Smad trimerisation efficiency is, in principle, consistent with pronounced stimulus-induced Smad4 translocation. More specifically, it was analysed whether a model with weak trimerisation affinity is able to describe quantitative time course measurements in HaCaT cells from the literature (Figs. 6.15A – I).

The HaCaT cell model, schematically depicted in Fig. 6.15, contains most of the reaction steps of the hepatocyte model, and the majority of kinetic parameters were assumed to be equal in both models as well. The differences between the HaCaT and the hepatocyte models are summarised in the following: (i) SnoN-mediated feedback was eliminated from the HaCaT model, as pre-treatment of HaCaT cells with the translation inhibitor cycloheximide does not dramatically affect the dynamics of TGFβ-induced Smad signalling [49,374]. (ii) The expression levels of Smad2 and Smad4 were altered such that they equal values experimentally measured in HaCaT cells [378], and the TGFβ receptor expression level was fitted to match time course data from HaCaT cells (Figs. 6.15A – I). (iii) The cytoplasmic and nuclear volumes were set to values relevant for HaCaT cells [339]. (iv) The constitutive and ligand-induced receptor degradation rates were fitted to match time course data from HaCaT cells (Figs. 6.15A – I). (v) Smad2 degradation was shown to be enhanced upon TGFβ stimulation in HaCaT cells [345], most likely due to recruitment of ubiquitin ligases to Smad trimers. The degradation rates of cytoplasmic and nuclear Smad trimers were therefore assumed to be larger than those of the corresponding monomeric Smad species, and fitted to match time course data from HaCaT cells (Figs. 6.15A – I). Taken together, the kinetics of the core Smad signalling cycle were kept unchanged in the HaCaT cell model, and only some protein synthesis and degradation rates were assumed to be different from the hepatocyte model.
Fig. 6.16 Memory effect in TGFβ signalling.

(A) The time course Smad2 phosphorylation in the cytoplasm in response to TGFβ stimulation for 3 minutes (followed by a ligand washout by medium exchange) was simulated using the best-fit hepatocyte model (solid line; see Tables F.1 – F.3 for kinetic parameters). The model predicts that the cytoplasmic pSmad2 continues to rise even after the ligand has been washed out (see inset), and this prediction could be verified by experimental analyses of primary mouse hepatocytes (data points; measured by Peter Nickel). (B) Same as (A), but a 10 min ligand pulse was simulated and measured, respectively.

The multi-experiment fitting results (solid lines in Fig. 6.15) indicate that the modified model is able to describe the Smad signalling dynamics in HaCaT cells reasonably well. The model quantitatively matches the experimentally measured Smad2 phosphorylation time courses in whole-cell lysates and in nuclear fractions, respectively (Fig. 6.15A). Moreover, measurements of Smad2 and Smad4 in the nucleus and in the cytoplasm are qualitatively reproduced by the model (grey solid lines in Figs. 6.15C – F). In particular, a pronounced translocation of Smad4 into the nucleus is observed, despite weak Smad trimerisation efficiency in the model. However, a quantitative match between model and experiment was not possible, mainly because the nucleo-cytoplasmic distribution of Smad proteins in unstimulated HaCaT cells differs from that in primary mouse hepatocytes [339,364]. This could be corrected by fitting the nuclear import and export rates of Smad2 and Smad4 in addition to the parameter changes (i – v) listed above (black solid lines in Figs. 6.15C – F). Note that all other trajectories in panels A – B and G – I look very similar for both fitting results, with and without import/export optimisation, so that the model with import/export optimisation will be discussed in the following. It has been shown experimentally that a trimerisation-deficient Smad2 mutant does not translocate into the nucleus in TGFβ-treated HaCaT cells [339]. Numerical simulations of a mutant HaCaT model devoid of any Smad trimerisation indeed revealed that Smad2 translocation is abolished (dotted lines in Figs. 6.15C and D), thus further supporting that the shuttling dynamics are correctly described. In conclusion, it seems that inefficient trimerisation is not per se not inconsistent with pronounced Smad4 translocation. HaCaT cells express 5–10 fold higher levels of Smad2 and Smad4 when compared to hepatocytes [378] (Fig. 6.3). This, together with more efficient Smad2 phosphorylation (due to increased TGFβ receptor expression in the HaCaT model), shifts the Smad equilibrium towards trimerisation and thus allows for pronounced Smad4 translocation into the nucleus in the HaCaT cell model. However, measurements of Smad protein complexes under non-denaturising conditions, e.g., fractionation according to the molecular weight by ultracentrifugation or gel filtration, are required to directly prove that Smad trimerisation is relatively inefficient in living cells.
The ligand washout measurements in Figs. 6.6J – L reveal that primary mouse hepatocytes maintain elevated Smad signalling for a couple of hours after TGF\(\beta\) has been removed from the culture medium. Such memory effects might be common to many cellular systems, as slow Smad dephosphorylation was also observed in HaCaT cells subjected to medium exchange (Fig. 6.15 H; black line) [341,345], and in Xenopus oocytes [326]. Experiments in primary mouse hepatocytes and in HaCaT cells reveal that the memory effect arises at the receptor level, and not within the core Smad activation cycle: Smad2 phosphorylation declined rapidly if receptor activation was terminated using a small-molecule inhibitor of the TGF\(\beta\) receptor (Figs. 6.6G – I and Figs. 6.15H and I). Interestingly, the memory effect in HaCaT cells is enhanced in the presence of the proteasome inhibitor MG132 (Fig. 6.15 H; grey line). The dotted line and the triangles in Fig. 6.15I demonstrate that MG132 does not affect the dephosphorylation kinetics of the Smad cycle in response to TGF\(\beta\) receptor inhibition, thus excluding that proteasome inhibitors enhance the memory effect by affecting signalling downstream of the TGF\(\beta\) receptor. This strongly suggests that MG132 prolongs the Smad activation memory by acting at the level of receptors, most likely by blocking proteasome-mediated targeting of TGF\(\beta\) receptors to the lysosomal degradation pathway [326,379]. Taken together, these data support the hepatocyte and the HaCaT models where the memory effect arises from slow degradation of the non-recyclable pool of receptor-ligand receptors in late-endosomes (TR\(_{\text{le}}\)). Accordingly, it has recently been reported for Xenopus oocytes that the basis for memory is the long-lasting residence of active receptor-ligand complexes in the endo-lysosomal pathway [326]. To further investigate the memory effect, pulse stimulation experiments were simulated numerically using the hepatocyte model (Fig. 6.16). For a TGF\(\beta\) washout 3 min after stimulation, the model predicted that cytoplasmic Smad2 phosphorylation continues to increase by approximately 3-fold after ligand removal (solid line in Fig. 6.16A). Preliminary pulse experiments done by Peter Nickel indeed confirmed that such an overshoot in cytoplasmic Smad2 phosphorylation occurs after the ligand had been washed out, thus confirming the model prediction (data points in Fig. 6.16A). By contrast, the model predicted that no pronounced overshoot should be observed for a 10 min pulse, but that the peak activation level observed at 10 min should be roughly maintained for 1 h (solid line in Fig. 6.16B). Again, this prediction was in reasonable agreement with preliminary measurements done by Peter Nickel (data points in Fig. 6.16B). In conclusion, the data presented in Fig. 6.16 strongly suggests that the memory effect arises from a non-recyclable endosomal compartment on the way to lysosomal degradation. Moreover, it seems that a mathematical model using physiologically relevant kinetic parameters for receptor trafficking (Table F.1) is able to quantitatively describe memory formation and receptor-ligand degradation within the endosome. However, more experiments are required to further prove the proposed mechanism: First, the simple medium exchange procedure for ligand washout should be replaced by a more stringent acid wash procedure, which completely removes all ligand possibly remaining bound to cell surfaces or to the culture dish. Second, it should be tested whether the memory can be prolonged by inhibiting receptor trafficking, e.g., by a temperature-shift to 15\(^\circ\)C or by treating cells with inhibitors of lysosomal degradation.

The mathematical model presented in this chapter was useful to investigate whether a large number of quantitative experimental measurements (Figs. 6.5 – 6.8) can be simultaneously described by a single mechanistic model using a set of physiologically relevant kinetic parameters (Tables F.1 – F.3), and thus serves as a consistency check for the present biochemical knowledge. Moreover, the model was used to generate experimentally testable predictions regarding SnoN feedback (Figs. 6.9 – 6.11, 6.14), Smad trimerisation affinity (Fig. 6.12A), Smad-induced gene expression (Fig. 6.12B) and regarding memory effects upon ligand removal (Fig. 6.16). The model also provided mechanistic insights into Smad complex formation within the nucleus that cannot easily be analysed experimentally (Fig. 6.12), and thus might help to minimise the number of expensive and time-consuming measurements.

A future aim is to extend the present model to Hepa-16 liver cancer cells that are closely related to primary mouse hepatocytes. These cells exhibit transient Smad2 phosphorylation
and nuclear translocation (for ~ 180 min), and signalling can be switched from transient to sustained upon incubation with the general transcription inhibitor actinomycin D (Peter Nickel, not shown). This indicates that transcriptional feedback inhibiting signalling at the level of receptor operates in Hepa-16 to terminate Smad2 phosphorylation early after stimulation. Interestingly, it has been reported that the anti-proliferative action of TGFβ requires sustained Smad signalling [49], so that model analyses of Hepa-16 cells might provide insights into how cancer cells escape growth control by TGFβ.

The results presented in Sections 5 and 6 reveal how transcriptional feedbacks shape the dynamics of signal transduction. Intracellular signals typically elicit phenotypic responses by regulating the expression of downstream target genes that are, for example, involved in cell cycle regulation. Sections 7 and 8 are focused on the dynamics of gene expression, and thus might provide insights into how intracellular signals are decoded at the genomic level.

In many cases, specific target gene expression patterns and thus specific phenotypic responses are induced depending on the quantitative aspects of upstream signalling pathway activation [380]. For instance, it is known that the duration of Erk or Smad signalling can determine whether a cell divides or not [49,381]. In Section 7, it is analysed how the dynamics of gene expression are affected by small non-coding RNAs which regulate gene expression at the post-transcriptional level. The results reveal that small RNAs filter out short transcription pulses, and might thus help to explain how cells discriminate transient vs. sustained signalling.

Many of the downstream target genes induced by signalling pathways are transcription factors, thus giving rise to a complex transcriptional regulatory network [67]. In Section 8, a reverse engineering approach is applied to get insights into a transcription factor network that is regulated by oncogenic Ras and involved in the regulation of cell division. The results represent a first step towards an integrated model including both, upstream signal transduction and downstream phenotypic responses.
Part III

Decoding at the level of gene expression
7 Small RNAs Establish Delays and Temporal Thresholds in Gene Expression

This chapter is the result of a collaboration with Dennis Dienst, Annegret Wilde and Ilka Axmann from the Institute for Biochemistry of the Humboldt University Berlin. All experiments were performed by Dennis Dienst and Ilka Axmann in the lab of Annegret Wilde. The Northern Blot quantification was done by Ilka Axmann.

SYNOPSIS

Non-coding RNAs are crucial regulators of gene expression in prokaryotes and eukaryotes, but it remains poorly understood how they affect the dynamics of transcriptional networks. In this Section, the temporal characteristics of the cyanobacterial iron stress response are analysed by mathematical modelling and quantitative experimental analyses. A recently discovered small non-coding RNA, IsrR, is shown to be responsible for a pronounced delay in the accumulation of isiA mRNA encoding the late-phase stress protein, IsiA. Moreover, it is demonstrated that IsrR ensures a rapid decline in isiA levels once external stress triggers are removed. These kinetic properties allow the system to selectively respond to sustained (as opposed to transient) stimuli, and thus establish a temporal threshold, which prevents energetically costly IsiA accumulation under short-term stress conditions. Biological information is frequently encoded in the quantitative aspects of intracellular signals (e.g., amplitude and duration). The simulations presented in this Section reveal that competitive inhibition and regulated degradation allow intracellular regulatory networks to efficiently discriminate transient and sustained inputs.

7.1 Introduction

Non-protein-coding RNA regulators such as microRNAs (miRNA) and short interfering RNAs (siRNA) control diverse processes in metazoans including development, cell differentiation, and cell proliferation [63,64]. Recent research revealed that small non-coding RNAs (sRNAs) also play important roles in bacteria [65,66], where they are mainly involved in the modulation of stress responses. Approximately 80 sRNAs have been identified in E. coli, many of which are evolutionary conserved [382]. sRNAs are typically less than 200 nucleotides in size and are either encoded in cis or in trans. Cis-encoded sRNAs are transcribed from the antisense strand of their target mRNAs, and are thus perfectly complementary, while trans-encoded sRNAs have only limited complementarity [66]. Most sRNAs inhibit gene expression employing a non-catalytic mechanism of action: Base-pairing with target mRNAs either interferes with ribosome binding and thus with target mRNA translation, or even induces degradation of the whole sRNA-target complex [66,383].

Previous mathematical modelling indicated that sRNA-mediated regulation may allow for faster control over gene expression than transcriptional and post-translational regulation [384]. Moreover, it was shown theoretically and experimentally that small non-coding RNAs efficiently suppress steady state target mRNA accumulation if the mRNA transcription rate is low, while they have little impact at higher mRNA transcription rates [385,386]. Levine et al (2007) argue that this phenomenon, termed a “threshold-linear response”, efficiently prevents costly and potentially harmful expression of bacterial stress proteins under normal conditions.
In this Section, the impact of sRNAs on the temporal regulation of gene expression is analysed. Theoretical predictions derived from mathematical modelling are confirmed by quantitative experimental analyses of the iron stress response in a cyanobacterial organism \textit{Synechocystis} sp. PCC 6803). The IsiA (iron stress induced protein A) stress response protein, which is transcriptionally induced upon iron depletion or oxidative stress \cite{387} is controlled by a naturally occurring antisense sRNA, IsrR (iron stress repressed RNA) \cite{388}. By comparing strains expressing different levels of IsrR sRNA, it is shown that IsrR is responsible for a pronounced delay in \textit{isiA} induction. This delay ensures that iron stress proteins are expressed in a temporally ordered manner, with the “emergency” protein IsiA accumulating only if the stress duration exceeds a critical temporal threshold. Moreover, it is shown that IsrR sRNA ensures a rapid decline in \textit{isiA} levels once external stress triggers are removed. IsiA expression must be tightly controlled, as it reduces photosynthesis efficiency in unstressed cells and becomes highly abundant under stress conditions. The results in this chapter reveal how the IsrR sRNA ensures that \textit{isiA} accumulation is restricted to severe, prolonged and ongoing stress conditions.
7.2 Model Implementation

A mathematical model of sRNA-mediated regulation was implemented (schematically depicted in Fig. 7.7.1A). The model comprises synthesis and degradation of target mRNA and sRNA, and also takes reversible association between mRNA and sRNA as well as degradation of the resulting heteroduplex (‘pair’) into account. In contrast to published models [384,385], the heteroduplex concentration was treated as a dynamic variable, mainly because the experimental analyses of the cyanobacterial iron stress response did not allow to distinguish between free and sRNA-bound isiA mRNA. Previous studies indicated that heteroduplex association and dissociation proceed with rapid kinetics when compared to protein synthesis and degradation [389,390]. A rapid-equilibrium assumption was therefore applied, so that the model depicted in Fig. 7.7.1A reduces to a two-variable system (see also Appendix G):

\[
\begin{align*}
\frac{d[T_{tot}]}{dt} &= v_{syn,T} - k_{deg,T} \cdot ([T_{tot}] - [Pair]) - k_{deg,P} \cdot [Pair] \\
\frac{d[S_{tot}]}{dt} &= v_{syn,S} - k_{deg,S} \cdot ([S_{tot}] - [Pair]) - k_{deg,P} \cdot [Pair]
\end{align*}
\]

Here, \([T_{tot}]\) and \([S_{tot}]\) indicate the dynamical variables of the system which represent the total intracellular concentrations of the target mRNA and sRNA, respectively. More specifically, \([T_{tot}]\) equals the sum of the concentrations of free and sRNA-bound target mRNA (i.e., \([T_{tot}] = [Target] + [Pair]\)), and \([S_{tot}] = [sRNA] + [Pair]\) is defined similarly. Owing to the rapid equilibrium assumption, the concentration of the heteroduplex (i.e., \([Pair]\)) is given by (see Appendix G):

\[
[Pair] = \frac{1}{2} \left( \frac{[T_{tot}]+[S_{tot}]+K_{d,P} \cdot \sqrt{([T_{tot}]+[S_{tot}]+K_{d,P})^2 - 4 \cdot [T_{tot}] \cdot [S_{tot}]}}{2} \right) \quad (7.2)
\]

Equations (7.1) and (7.2) constitute a reduced differential equation system that depends on the dynamical variables \([T_{tot}]\) and \([S_{tot}]\) only. From the total RNA concentrations, \([T_{tot}]\) and \([S_{tot}]\), one can calculate back to the individual concentrations (i.e., \([Target]\), \([sRNA]\) and \([Pair]\)) by using Eq. 7.2 and the mass conservation definitions of \([S_{tot}]\) and \([T_{tot}]\). Numerical simulations were generally done using the units nM for intracellular concentrations and h for time. In Figs. 7.1 and 7.2, the concentrations on the y-axes are given in arbitrary units (A.U.), because no absolute experimental quantification of RNA expression was available.

The cyanobacterial stress response protein IsiA is induced by reactive oxygen species (ROS) and by iron depletion, and its expression is further modulated by a small non-coding RNA, IsrR (Fig. 7.1A). The kinetics of isiA regulation was measured to confirm the model predictions (see below). The input functions in the model were therefore chosen such that they reflect those in the cyanobacterial stress response. IsiA expression is known to be transcriptionally regulated by the iron-sensitive Fur (ferric uptake regulator) repressor, while the activity of the promoter controlling the antagonizing sRNA, IsrR, is not affected by cellular stress (Wolfgang Hess, pers. comm.). The synthesis rate of target mRNA (\(v_{syn,T}\)) is therefore modelled to be controlled by external stimuli, as indicated in Fig. 7.1A. For simplicity, it is assumed that the regulation of the isiA promoter activity occurs on a much faster time scale than downstream RNA synthesis and degradation. Thus, extracellular stimulation was simulated by a step-like change of \(v_{syn,T}\) from one value to another. The Michaelis-Menten equation (\(v_{syn,T} = V_{max,synT} \cdot \frac{\text{Stimulus}}{K_{M,synT} + \text{Stimulus}}\)) was employed to simulate dose-response behaviour (Fig. 7.1B).
7.3 Theoretical Analysis of Steady State and Dynamical Behaviour

The steady state and dynamical behaviour of the sRNA circuit was systematically analysed by numerical simulations. The following analysis will mainly focus on the scenario, where the heteroduplex is much less stable than the monomeric sRNA and target species, as such behaviour was recently reported for the cyanobacterial stress response [388]. However, the main conclusions drawn in this chapter remain valid even if the sRNA merely acts as a competitive inhibitor of translation, which does not enhance target degradation (see Section 7.6).

Recent work by Levine et al. [385] revealed that sRNAs establish sharp thresholds in the steady state stimulus-response behaviour of gene regulatory circuits. Similarly, the
simulations in Fig. 7.1B (top, solid line) also yielded an all-or-none expression pattern of the target mRNA if physiologically relevant kinetic parameters were assumed. As discussed previously [385] , sRNA-mediated regulation efficiently suppresses mRNA accumulation as long as the sRNA synthesis rate exceeds that of the target mRNA (i.e., if $v_{syn,T} < v_{syn,S}$). In Fig. 7.1B, the condition $v_{syn,T} = v_{syn,S}$ is fulfilled if the stimulus level equals unity, so that a threshold is observed for stimulus $= 1$. As expected, the threshold disappeared in the absence of sRNA (dashed line) or in case the affinity of the heteroduplex was too low (not shown). Moreover, sharp thresholds were typically accompanied by a mutually exclusive expression pattern of target mRNA and sRNA (solid and dashed-dotted lines in Fig. 7.1B). Previous experimental work revealed that the $isiA$ mRNA and its modulator, the IsrR sRNA, accumulate in a mutually exclusive manner under various stimulation conditions [388] . Further experiments, where amounts of $isiA$ and IsrR were measured for varying iron concentrations in the medium, revealed that IsrR sRNA starts to accumulate only if $isiA$ mRNA falls below a certain level (Appendix G). Finally, a comparison of wild-type cells with IsrR knock-down cells reveals that IsrR completely suppresses residual $isiA$ levels under unstressed conditions (compare circle and square at $t = 0$ in Fig. 7.2A). Taken together, these data strongly suggest that a sharp threshold exists in the cyanobacterial iron stress response, and that the system operates near this threshold under physiological conditions.

The temporal dynamics around the threshold were investigated in order to derive experimentally testable predictions from the model. The analysis was focused on the time required to pass from one steady state to another (‘response time’). Step-like increases in the external stimulus were applied, and the response time $t_{50}$ required to reach 50% of the difference between old and new steady state was calculated for the target mRNA. The dotted line in Fig 7.1B (bottom) reflects the response time of a reference system without sRNA, and is valid for both input scenarios considered: (i) a step-like stimulus increase from no stimulation to the value indicated on the x-axis (‘upregulation’); (ii) a stimulus decrease from the value indicated on the x-axis to no stimulation (‘downregulation’). The simulations of the sRNA circuit revealed that sRNAs speed up the upregulation kinetics for subthreshold stimuli, while they establish a delay upon suprathreshold stimulation (solid line). Moreover, Fig. 7.1B indicates that sRNAs always accelerate target downregulation even for suprathreshold stimuli (dashed line).

In the subthreshold regime the sRNA is present in excess over the target mRNA. Under these conditions, the sRNA can be assumed to be constant, so that mRNA degradation via the Pair intermediate behaves like a first-order decay reaction, which dominates over the much slower direct mRNA degradation step (see Appendix G). The response times of mRNA up- and downregulation are known to be solely determined by the (fastest) first-order decay rate [391] , which explains why sRNA-mediated regulation accelerates target regulation in both directions under subthreshold stimulation conditions. The dynamic phenomena observed upon strong stimulation originate from nonlinear threshold effects in the sRNA circuit. The upregulation time lag arises, because the residual pool of free small RNAs needs to be cleared before the target mRNA can accumulate. Small RNAs accelerate downregulation, since the mRNA concentration quickly falls into a subthreshold range, where sRNA-mediated regulation efficiently degrades residual mRNA. This results in an abrupt termination of the mRNA expression time course, as can be seen from solid line in Fig. 7.2B. Moreover, the initial phase of mRNA down-regulation is accelerated in the presence of small non-coding RNAs (compare solid and dotted lines in Fig. 7.2B), as the pair intermediate is turned over at a higher rate than the free mRNA. Taken together, the simulations suggest that IsrR sRNA delays $isiA$ induction upon stress, but speeds up $isiA$ downregulation when the upstream trigger is removed.
7.4 Experimental Verification of the Simulated Dynamic Behaviour

Quantitative experimental analyses were performed to confirm the simulations in Fig. 7.1B. The impact of sRNA-mediated regulation was investigated by comparing the kinetics of isiA mRNA accumulation in mutant strains expressing different levels of IsrR sRNA. The data points shown in Fig. 7.2A are densitometric quantifications of measurements of isiA induction in response to hydrogen peroxide stress previously published by the Wilde group [388]. These data revealed that the delay in isiA accumulation seen in wild-type cells can be further enhanced in cells overexpressing IsrR, while it is abolished in IsrR-depleted cells. The measured time courses thus agree well with the simulation result that sRNAs decelerate target mRNA induction.

Table 7.1 Kinetic parameters of the best-fit model.

<table>
<thead>
<tr>
<th>Figure</th>
<th>Fig. 7.2A</th>
<th>Fig. 7.2B</th>
</tr>
</thead>
<tbody>
<tr>
<td>vsyn,T [nM h⁻¹]</td>
<td>247.2 (basal level)</td>
<td>996.9 (basal level)</td>
</tr>
<tr>
<td></td>
<td>1136 (after stimulus increase)</td>
<td>41.9 (after stimulus removal)</td>
</tr>
<tr>
<td>kdeg,T [h⁻¹]</td>
<td>0.42</td>
<td>0.42</td>
</tr>
<tr>
<td>vsyn,S [nM h⁻¹]</td>
<td>516.6 (red line)</td>
<td>516.6 (red line)</td>
</tr>
<tr>
<td></td>
<td>706.9 (green line)</td>
<td>145.3 (blue line)</td>
</tr>
<tr>
<td></td>
<td>145.3 (blue line)</td>
<td></td>
</tr>
<tr>
<td>kdeg,S [h⁻¹]</td>
<td>0.35</td>
<td>0.35</td>
</tr>
<tr>
<td>Kd,P [nM]</td>
<td>0.0045</td>
<td>0.0045</td>
</tr>
<tr>
<td>kdeg,P [h⁻¹]</td>
<td>13.75</td>
<td>13.75</td>
</tr>
</tbody>
</table>

The time course of isiA target mRNA downregulation after removal of the stress trigger was also measured (see Appendix G for corresponding Northern Blots). Hydrogen peroxide could not be used as the stimulus in these experiments, since it induces irreversible oxidation of cellular components so that stress might actually persist even after hydrogen peroxide is removed. isiA accumulation was therefore induced by iron depletion (48 h), and isiA expression was subsequently downregulated by iron re-addition (t = 0 in Fig. 7.2B). The time course measurements after relief from iron stress were in accordance with the model predictions, as the decline in isiA levels was faster in wild-type cells when compared to knock-down cells harbouring reduced levels of IsrR sRNA (Fig. 7.2B).

Having established a qualitative agreement between experiments and simulations, it was then analysed whether the model could also quantitatively reproduce the dynamics of isiA target mRNA up- and downregulation. The model parameters were estimated from the time course data shown in Figs. 7.2A and B using PottersWheel, a multi-experiment fitting toolbox for Matlab programmed by Thomas Maiwald (www.potterswheel.de). The trajectories of the best-fit model (solid lines in Figs. 7.2A and B) indicate that all measurements can be accurately described by a single set of kinetic constants. Remarkably, the best-fit parameters, summarised in Table 7.1, suggest that the heteroduplex is rapidly degraded, while the single stranded forms of isiA and IsrR are predicted to be much more stable than typical bacterial RNAs [392]. A long half-life of isiA directly follows from the slow isiA downregulation kinetics in sRNA-depleted cells (Fig. 7.2B; blue line), while the stability of IsrR cannot be as straightforwardly deduced from the time course data. The half-life of IsrR sRNA was therefore directly measured in unstressed cells, which express negligible amounts of isiA target mRNA [388]. Very little degradation of IsrR occurred within a 45 min time interval after incubation of cells with the general transcription inhibitor Rifampicin (see Appendix G for corresponding Northern Blots). These data confirm that IsrR sRNA is stable and thus support the best-fit model.
The above RNA measurements (Figs. 7.2A and B) did not discriminate between IsrR-bound ('pair') and free ('target') isiA RNA. In order to get more direct insights into the regulation of isiA action by IsrR sRNA, time courses of free (i.e., biologically active) isiA target mRNA were simulated using the best-fit parameters. Figure 7.2C shows model responses to a step-like suprathreshold pulse stimulation, and reveals that regulation by small RNAs gives rise to a sharp, spike-like time course (solid line), when compared to a system devoid of sRNA (dashed line). More specifically, target mRNA accumulation is completely suppressed until all sRNA is degraded (indicated by black circle), and the mRNA decline terminates abruptly, once the sRNA (re-)starts to accumulate.

Taken together, the simulations indicate that IsrR-mediated control serves to prevent premature and unnecessarily prolonged isiA synthesis. This is consistent with the hypothesis that IsiA establishes a second line of defence against iron depletion, and with the fact that its expression occurs relatively late during iron stress [393]. The delay established by sRNA-mediated regulation thus enables cells to induce both early- and late-phase stress proteins in response to a single stress trigger in a temporally ordered manner. Two lines of evidence further suggest that spike-like isiA expression (Fig. 7.2C) is beneficial to the cellular energy budget. First, isiA becomes the most abundant transcript in cells subjected to oxidative stress [394]. Second, IsiA expression saves photosynthesis in stressed cells, but decreases photosynthesis efficiency and thus energy production in non-stressed cells [393]. The present analyses indicate that IsrR sRNA-mediated control allows avoiding lavish isiA accumulation unless cells are subjected to severe and prolonged stress.

Further simulations were performed to confirm the hypothesis that sRNA circuits suppress short stimuli, but efficiently transmit prolonged inputs. The time courses of target mRNA expression were simulated for step-like stimulus pulses, and the time course maximum was analysed as a function of pulse duration (Fig. 7.2D). The best-fit model (Table 7.1) indeed responds to the pulse duration in a highly ultrasensitive manner (Hill coefficient $n_H \approx 3.5$), while a much more gradual duration response is seen for a system without sRNA. Here, the Hill coefficient $n_H$ was estimated by using the formula $n_H = \log(81) / \log(D_{90}/D_{10})$, where $D_{90}$ and $D_{10}$ are the stimulus durations required for the target amplitude to reach 90% and 10% of the steady state, respectively [19]. Similar ultrasensitive behaviour was also observed when the integral under the target mRNA time courses was analysed (Fig. 7.2D, inset). This further confirms that sRNAs establish pulse filtering and temporal thresholds in biochemical signalling networks.

**7.6 Conclusions**

Using a combination of mathematical modelling and quantitative experimental analyses, it was shown in this chapter that sRNAs establish delays and (steady state and temporal) thresholds in gene expression. In order to allow for a comparison of the simulations with experimental data from wildtype and knock-down cells, the dynamical behaviour of the wildtype model was mostly compared with that of a model devoid of sRNA. It should be noted that, from a systems theoretical point of view, the dynamical behaviour of both models (+/- sRNA) is not directly comparable for a given stimulus level, as they differ in their steady state dose-response curves. However, it can be seen in Fig. 7.1B that sRNA-mediated regulation still establishes a delay in mRNA upregulation and accelerates mRNA downregulation if both models (+/- sRNA) are compared for a given suprathreshold steady state activation level. In this context, it is important to note that the up- and downregulation response times of the model without sRNA are stimulus-invariant (dashed horizontal line in Fig. 7.1B, bottom).

The key finding of the present chapter is that sRNA-mediated regulation establishes a sign-sensitive delay for supra-threshold stimulation (Figs. 7.2A and B), and this conclusion does
not depend on the precise dose-response behaviour or on the kinetic parameters chosen. Here, the term ‘sign-sensitive delay’ denotes that a delay is observed exclusively for mRNA upregulation, but not for mRNA downregulation [395]. This sign-sensitive delay is responsible for the pulse-filtering behaviour shown in Figs. 7.2C and D: The delay in mRNA upregulation helps to filter out short supra-threshold transients. However, ultrasensitive pulse-filtering with a Hill coefficient significantly larger than unity additionally requires that no such delay and in particular no memory effects arise if the stimulus is removed. The experimental data presented in the paper rules out that mRNA downregulation occurs with a delay (Fig. 7.2B), and thus strongly supports that ultrasensitive pulse-filtering occurs in the cyanobacterial iron stress response. However, the mRNA downregulation kinetics might depend on the history of the system (e.g., short vs. long mRNA upregulation), so that explicit pulse-stimulation experiments are required to further prove the existence of ultrasensitive pulse-filtering.

The analysis was mainly focused on the scenario, where the heteroduplex is less stable than the single stranded sRNA and target species. However, several bacterial sRNAs do not enhance target degradation, but merely act as competitive inhibitors of translation [383]. Importantly, small RNAs acting in this way still establish sharp thresholds and delays, as translation of subthreshold target mRNA levels is efficiently suppressed by sequestration into the heteroduplex [16] (Appendix G). Eukaryotic miRNA action can also be described by the model scheme depicted in Fig. 7.1A, as miRNAs either competitively inhibit translation or induce mRNA degradation. However, miRNAs frequently remain intact after target degradation, and can guide the recognition and destruction of additional messages [64]. Further numerical analyses revealed that the kinetic properties described above remain valid even if a fraction of the sRNA remains intact during the pair degradation reaction (Appendix G) [385]. Thus, the main results in this chapter apply for eukaryotic systems as well, although this remains to be confirmed in more detailed models of miRNA action [396].

In Figs. 7.1 and 7.2, the dynamic characteristics of mRNA expression were analysed. However, the kinetic properties of the sRNA circuit may be obscured by a slow response at the level of protein expression. It was therefore investigated numerically whether the conclusions regarding pulse-filtering continue to hold at the protein level. The simulation results (shown in Appendix G) demonstrate that for the best-fit parameters pulse-filtering is preserved at the protein level even if IsiA protein is assumed to be relatively stable, with a half-life of 10 h. Up to now, studies focusing on the stability of IsiA protein are missing in the literature. However, half-life measurements of the CP43 photosynthesis protein homologous to IsiA revealed a half-life of about an hour under stress conditions [397], and a similar rapid turnover was also reported for another photosynthesis protein, D1 [398]. These data suggest that IsiA protein is short-lived in the experimental setup chosen here, and that the pulse filtering property discussed here for the RNA level is observed at the level of proteins as well.

Cells face a specificity problem, as broadly overlapping signalling pathways are activated by diverse stimuli. Biological information is therefore encoded in the quantitative aspects of the signal, such as amplitude and duration [380,381]. An important role for the signal duration in the initiation of cell fate decisions was described for various biological networks including MAPK signalling [381], TGFβ signalling [49], cAMP signalling [42], and NF-κB signalling [43]. Previous work indicated that multistep regulation in the form of feed-forward loops [59] and multisite phosphorylation [15,60] allows cells to discriminate transient and sustained stimuli. In this chapter, competitive inhibition and/or regulated degradation were identified as alternative plausible mechanisms for duration decoding (Fig. 7.2D). Further functional analysis of small RNAs might explain why some genes are selectively transcribed upon sufficiently long stimulation [43,58]. The results are likely to be of broader relevance, because regulation by protein-protein interactions frequently involves competitive inhibition and regulated degradation as well [22,151,169,399].
The simulations shown in this chapter demonstrate that mathematical modelling can help to get qualitative and quantitative insights into the functioning of small transcriptional regulatory units. However, these simple transcriptional motifs typically interact in complex transcriptional regulatory networks, many of which remain poorly understood. In the following chapter (Section 8), it is described how systems biological approaches allow deriving the topology of such transcriptional regulatory networks from high-throughput and knock-down data.
8 Modular Response Analysis of a Ras-regulated Transcription Factor Network

This chapter is the result of a collaboration with Iwona Stelniec, Oliver Raudies, Oleg Tschernitsa and Reinhold Schäfer from the Institute for Pathology at the Charité Berlin. The knock-down experiments were performed by Iwona Stelniec in the lab of Reinhold Schäfer, and the array measurements were done by Iwona Stelniec and Oleg Tschernitsa. Pål Westermark helped with the implementation of the modular response analysis code.

SYNOPSIS

Gene expression profiling studies revealed that extracellular stimuli or oncogenic activation events typically induce hundreds of genes. Many of these induced genes are themselves transcription factors, giving rise to a complex network of transcriptional regulation that cannot be understood intuitively. In this chapter, modular response analysis (MRA), a recently proposed reverse engineering method, is employed to derive the topology of a transcription factor network that mediates proliferation downstream of oncogenic Ras. Statistical analyses of the MRA results strongly suggest the existence of a linear HMGA2 → Fra-1 → Copb → JunB transcription factor cascade. Experiments that allow verifying the predicted regulatory interactions are proposed, and the possibility to extend the MRA calculations to downstream phenotypic responses (i.e., cell growth) is discussed.

8.1 Introduction

Biochemical and cell biological analyses are typically focussed on specific parts or even on single proteins of the intracellular regulatory network. In many cases, systems biological approaches help to integrate such detailed biological knowledge, and allow gathering quantitative insights into the functioning at the network level. A less well developed field of systems biology is reverse engineering, where regulatory interactions are deduced from high-throughput data using a top-down approach.

Modular response analysis (MRA), a reverse engineering approach recently proposed in [68], allows to reconstruct the network structure from an experimentally measured global response matrix by assuming a linear model. As schematically depicted in Fig. 8.1A, the global response specifies the experimentally measurable response of a species (B, C, or D in Fig 8.1A) towards a perturbation in another species (A in Fig 8.1A). The global response considers both direct and indirect regulatory interactions, while the so-called local response (Fig 8.1B) considers only direct interactions, and thus specifies the network structure. MRA allows to compute the local response from experimental data provided that the global responses have been measured for all possible perturbations (i.e., if the full matrix of global responses is available). In Fig. 8.1A, this means that each of the species A – D needs to be perturbed, e.g., by RNAi-mediated knock-down, and that the responses of A – D need to be measured for each of these perturbations.

An inherent weakness of the top-down strategy underlying MRA is that the predicted regulatory interactions cannot straightforwardly be tested using conventional biochemical assays. For example, it cannot be deduced from MRA calculations whether the inhibition of B by A in Fig. 8.1 occurs by direct binding or through a multi-step chain of regulatory events. In other words, direct interactions in the sense of MRA are not necessarily direct in a biochemical sense. Moreover, MRA gives no clue about the biochemical mechanism of regulation (e.g., protein-protein interactions, promoter regulation, covalent modification), and this further complicates experimental verification. Accordingly, a positive feedback loop
predicted by MRA in a recent study focussing on MAPK signalling in PC12 neuronal precursors could only be confirmed by indirect experimental evidence [24].

Figure 8.1: Deriving experimentally testable predictions from modular response analysis (MRA).

(A) The global response specifies the response of a species (e.g., species C) towards a perturbation in another species (e.g., species A), arising from both direct and indirect regulation (the latter occurring via species B). (B) The local response considers only direct interactions, and thus specifies the network structure. MRA allows to calculate the local response from experimental data provided that the global responses have been measured for all possible perturbations (i.e., if the matrix of global responses is available). This means that each of the species A – D needs to be perturbed, e.g., by RNAi-mediated knock-down, and that the responses of A – D need to be measured for each of these perturbations. (C) A comparison of the local and global response coefficients provides further insights into the network structure. Similar local and global responses indicate that weak indirect regulation relative to direct regulation (although more complicated explanations are conceivable), while significantly different local and global responses reveal an important role for indirect regulation. For example, the A globally inhibits C (grey bar) despite local activation (black bar), as the indirect inhibitory pathway via species B dominates over direct activation. (D) Thus, the global inhibition of C by A should be convertible into a global activation if indirect regulation via species B is blocked, and this prediction can be confirmed in B knock-out cells. The regulatory interaction A → D seems to be less well suited for experimental verification of MRA, as multiple indirect pathways may require a simultaneous knock-out of species B and C.

In this chapter, MRA is applied to a Ras-regulated transcription factor network. Constitutive Ras activation is observed in ~30% of all tumours, and allows cells to proliferate in the absence of external growth stimuli (see Section 2). The signal transduction networks activated by oncogenic Ras have been extensively studied, but the transcription factor networks ultimately leading to Ras-induced changes in gene expression remain poorly understood. Using customised arrays, Oliver Raudies found that ~50 transcription factors are transcriptionally upregulated after transfection of normal ROSE 199 rat epithelial cells with constitutively active Ras (these Ras overexpressing cells will be termed ROSE A5/2 cells hereafter). Seven of these transcription factors (i.e., Gfi-1, Otx-1, RelA, Fra1, Copeb, JunB and HMGA2) were chosen for further analyses, and were individually knocked down by Iwona Stelniec in ROSE A5/2 cells using an RNAi approach. Array measurements in each of
the knock-down cell lines yielded a matrix of global responses, and thus allowed to calculate the network structure using MRA.

Table 8.1: Transcription Factor Expression in Response to siRNA-mediated knock-down.

<table>
<thead>
<tr>
<th>Transcription Factor</th>
<th>Gfi-1</th>
<th>Otx-1</th>
<th>RelA</th>
<th>Fra1</th>
<th>Copb</th>
<th>JunB</th>
<th>HMGA2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(overexp)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KRas</td>
<td>1,3836</td>
<td>2,2965</td>
<td>0,395</td>
<td>1,963</td>
<td>1,176</td>
<td>0,893</td>
<td>2,6798</td>
</tr>
<tr>
<td></td>
<td>6%</td>
<td>12%</td>
<td>18%</td>
<td>3%</td>
<td>2%</td>
<td>5%</td>
<td>3%</td>
</tr>
<tr>
<td></td>
<td>(siRNA 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gfi-1</td>
<td>-1,3405</td>
<td>-0,414</td>
<td>-0,3716</td>
<td>-0,2992</td>
<td>-0,9826</td>
<td>-0,047</td>
<td>1,5741</td>
</tr>
<tr>
<td></td>
<td>7%</td>
<td>20%</td>
<td>20%</td>
<td>20%</td>
<td>3%</td>
<td>20%</td>
<td>3%</td>
</tr>
<tr>
<td></td>
<td>(siRNA 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gfi-1</td>
<td>-1,0455</td>
<td>0,0973</td>
<td>0,4137</td>
<td>-0,4758</td>
<td>-0,494</td>
<td>0,3014</td>
<td>1,0287</td>
</tr>
<tr>
<td></td>
<td>9%</td>
<td>20%</td>
<td>20%</td>
<td>20%</td>
<td>5%</td>
<td>20%</td>
<td>6%</td>
</tr>
<tr>
<td></td>
<td>(siRNA 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Otx-1</td>
<td>0</td>
<td>-2,0879</td>
<td>-0,3022</td>
<td>0,7863</td>
<td>-0,7082</td>
<td>-0,2263</td>
<td>1,049</td>
</tr>
<tr>
<td></td>
<td>0%</td>
<td>4%</td>
<td>20%</td>
<td>3%</td>
<td>5%</td>
<td>20%</td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td>(siRNA 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Otx-1</td>
<td>0</td>
<td>-1,4339</td>
<td>-0,2746</td>
<td>0,707</td>
<td>-1,346</td>
<td>-0,2031</td>
<td>0,771</td>
</tr>
<tr>
<td></td>
<td>20%</td>
<td>10%</td>
<td>20%</td>
<td>3%</td>
<td>8%</td>
<td>20%</td>
<td>7%</td>
</tr>
<tr>
<td></td>
<td>(siRNA 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RelA</td>
<td>0</td>
<td>-0,28</td>
<td>-0,804</td>
<td>0,9055</td>
<td>1,01</td>
<td>0</td>
<td>0,4404</td>
</tr>
<tr>
<td></td>
<td>20%</td>
<td>20%</td>
<td>6%</td>
<td>6%</td>
<td>5%</td>
<td>20%</td>
<td>20%</td>
</tr>
<tr>
<td></td>
<td>(siRNA 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RelA</td>
<td>0</td>
<td>-0,905</td>
<td>-1,3767</td>
<td>1,23</td>
<td>0</td>
<td>0</td>
<td>0,4137</td>
</tr>
<tr>
<td></td>
<td>20%</td>
<td>20%</td>
<td>10%</td>
<td>14%</td>
<td>9%</td>
<td>20%</td>
<td>9%</td>
</tr>
<tr>
<td></td>
<td>(siRNA 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fra1</td>
<td>1,089</td>
<td>0,9655</td>
<td>-0,865</td>
<td>-1,947</td>
<td>-0,974</td>
<td>-1,346</td>
<td>0,836</td>
</tr>
<tr>
<td></td>
<td>10%</td>
<td>12%</td>
<td>13%</td>
<td>5%</td>
<td>6%</td>
<td>11%</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td>(siRNA 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fra1</td>
<td>0,8265</td>
<td>0,5017</td>
<td>-0,48</td>
<td>-1,368</td>
<td>-0,687</td>
<td>-1,068</td>
<td>0,538</td>
</tr>
<tr>
<td></td>
<td>11%</td>
<td>15%</td>
<td>14%</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td>(siRNA 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Copb</td>
<td>0</td>
<td>0,215</td>
<td>1,01</td>
<td>-0,94</td>
<td>-0,772</td>
<td>0,783</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20%</td>
<td>20%</td>
<td>20%</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td>(siRNA 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Copb</td>
<td>0</td>
<td>0</td>
<td>-0,1237</td>
<td>0,8647</td>
<td>-0,8541</td>
<td>-0,805</td>
<td>0,763</td>
</tr>
<tr>
<td></td>
<td>20%</td>
<td>20%</td>
<td>20%</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
<td>7%</td>
</tr>
<tr>
<td></td>
<td>(siRNA 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JunB</td>
<td>-0,731</td>
<td>0</td>
<td>0,023</td>
<td>0,2359</td>
<td>-0,293</td>
<td>-0,985</td>
<td>0,54</td>
</tr>
<tr>
<td></td>
<td>11%</td>
<td>0</td>
<td>20%</td>
<td>20%</td>
<td>20%</td>
<td>9%</td>
<td>12%</td>
</tr>
<tr>
<td></td>
<td>(siRNA 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMGA2</td>
<td>0</td>
<td>-0,0014</td>
<td>0,0657</td>
<td>-1,205</td>
<td>-0,868</td>
<td>-0,907</td>
<td>-2,342</td>
</tr>
<tr>
<td></td>
<td>20%</td>
<td>20%</td>
<td>20%</td>
<td>5%</td>
<td>9%</td>
<td>9%</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>(siRNA 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMGA2</td>
<td>0</td>
<td>0,069</td>
<td>-0,481</td>
<td>-1,107</td>
<td>-0,702</td>
<td>-1,105</td>
<td>-1,748</td>
</tr>
<tr>
<td></td>
<td>20%</td>
<td>20%</td>
<td>20%</td>
<td>5%</td>
<td>5%</td>
<td>10%</td>
<td>7%</td>
</tr>
</tbody>
</table>

The first row specifies the alterations in transcription factor expression in ROSE 199 cells induced by an overexpression of mutationally activated K-Ras. The resulting Ras-transformed cells (ROSE A5/2) were then perturbed by siRNA-mediated transcription factor knock-down against individual transcription factors as indicated in the following rows (transfections with different siRNA constructs against the same target are indicated by 1 and 2). Each column contains the average perturbation-induced changes in mRNA expression measured by microarray 48 h after transfection on a log2-scale (upper number), and lower number the corresponding standard deviation ($\sigma$ of mean out of 6 replicates) in per cent (lower number). Bold font refers to significant expression changes, where the log2-fold change is larger than the cut-off of 0.7 or smaller than -0.7 (lower cut-offs were used for the K-Ras overexpression experiment due to a larger number of replicates). Thin black and thin grey font indicates that the expression changes are not significant, either because they are less than 0.7 (thin black) or because the dye swap did not work (thin grey). Based on experimental data, the non-significant responses are generally assumed to have a standard deviation of 20%. All array measurements were performed six times and the results were independently confirmed by RT-PCR.

The following strategy was used to derive experimentally testable predictions from the MRA calculations. The regulatory interactions were sub-classified into two groups depending on whether their local and global responses are significantly different or not. Significantly different local and global responses imply that indirect regulation is important in addition to direct regulation. For example, Fig. 8.1C indicates that an experimental perturbation in A results in an inhibition of C (global response), while MRA predicts that A directly activates C (local response). It can be seen in Fig. 8.1A that this discrepancy in the local and global responses emerges from a dominance of indirect inhibition (via B) over direct activation, resulting in an inhibitory global response. Thus, the global inhibition of C by A should be convertible into a global activation if indirect regulation via B is blocked (Fig. 8.1D), and this prediction can be confirmed experimentally in B knock-out cells. It was the aim of this chapter to identify relatively simple network alterations such as a knock-out of a single species (B in Fig. 8.1D) that allow testing for the results of MRA. The analysis revealed the existence of transcription factor cascade that is likely to play a key role for Ras-mediated transformation, and that can be tested for by further experimental analyses.
Figure 8.2: Modular Response Analysis of a Ras-regulated Transcription Factor Network.

Each bar graph indicates how a single transcription factor (indicated on the top) is controlled by the other species of the system (as indicated on the x-axis). The y-axis corresponds to the strength of regulatory control (i.e., to the response coefficient). The grey bars indicate the median local response coefficients obtained by MRA (Eq. 8.2), which quantify direct regulatory effects only, and thus specify the network structure. The black bars refer to the normalised global response coefficients ($R_{ij} = \frac{\text{dln}x_i}{\text{dln}x_j}$), which quantify the net effect from both direct and indirect regulatory interactions. Normalisation of global response coefficients was done dividing each column of the global response matrix $R_{ij}$ (Eq. 8.1) by its diagonal element $R_{jj}$. The error bars indicate the 68% confidence intervals (see Section 8.2).

8.2 Methodology

Table 8.1 shows how siRNA-mediated knock-down of each transcription factor affects the mRNA expression levels of all other transcription factors as measured by customised arrays in ROSE A5/2 cells. These experimentally measured mRNA expression changes arise from both direct and indirect regulation and are thus global responses (Fig. 8.1A), so that Table 8.1 could be used as the input for MRA.

The change in the steady state concentration ($x_i$) of each component in the network in response to a perturbation ($p_j$) is given by the global response matrix $R_{ij} = \frac{\text{dln}x_i}{\text{dln}p_j}$, whose elements can be estimated by the formula [24,68,400]
\[
R_{ij} \approx 2 \cdot \frac{x_i^{(1)}/x_i^{(0)} - 1}{x_i^{(1)}/x_i^{(0)} + 1}.
\]  

(8.1)

where \(x_i^{(0)}\) and \(x_i^{(1)}\) are the steady-state concentrations of component \(x_i\) before and after perturbation \(p_j\), respectively. Thus, the matrix \(R_{ij}\) is only a reformulation of Table 8.1.

The network structure is determined by the local response matrix \(r_{ij}\), whose elements are given by:

\[
\begin{align*}
  r_{ij} &= \left(\frac{\text{d} \ln x_i}{\text{d} \ln x_j}\right)_{x_k = \text{const}} \\
  &= \frac{\text{const}}{x_k = \text{const}} \\
  &= (8.2)
\end{align*}
\]

Thus, each element of \(r_{ij}\) quantifies how a change in the steady state concentration of \(x_j\) affects the steady state concentration of \(x_i\). Importantly, \(r_{ij}\) refers to the hypothetical case where all other intermediates in the network (\(x_k\) with \(k \neq i, j\)) are kept constant, so that the elements \(r_{ij}\) quantify only direct regulatory interactions. This is in contrast to the matrix of global responses \(R_{ij}\) which additionally takes indirect regulatory effects (via \(x_k\) with \(k \neq i, j\)) into account. In other words, the local response matrix gives a hint about the network structure [68]. In this work, \(r_{ij}\) was calculated from the matrix of global responses \(R_{ij}\) by a total least squares (TLS) approach [400]. TLS incorporates biological replicas (i.e., different siRNA-transfections) and optimizes the solution \(r_{ij}\) such that it is in maximal agreement with all available biological replicas.

Technical errors arising during the microarray measurements were taken into account as described previously [24]: Random realisations of the fold expression changes were drawn from a normal distribution with mean and standard deviation equal to those measured experimentally (given in Table 8.1). Local response matrices \(r_{ij}\) were calculated for \(10^6\) random realisations and the histograms of the resulting local response coefficients \(r_{ij}\) were analysed. The median of these histograms was taken as the average response coefficient. The 68% confidence interval (which corresponds to one standard deviation of a normal distribution) was also calculated from the histograms in order to determine the error of response coefficient estimation. The lower bound of the confidence interval was estimated by the value below which 16% of the realisations were found, and the upper bound corresponds to the value below which 84% of the realisations occur.

In addition to TLS, the local response matrix \(r_{ij}\) and the confidence intervals were also calculated by the simpler, but (in case of noisy data) less accurate matrix inversion method [68,400]. This control implementation gave qualitatively similar results as TLS (not shown).

Figure 8.3: Interactions with high average local response coefficients.

The local responses given in Fig. 8.2 (grey bars) were classified based on their median response coefficients. The solid lines indicate strong interactions with median response coefficients over 1 or below -1. Dashed lines indicate medium interactions with median response coefficients over 0.5 or below -0.5. Grey blunted arrows indicate inhibition, while black arrows denote...
activation. The black dots on the arrows mark interactions with high average response coefficients, which are characterised by high confidence response coefficients (termed HA-HC interactions; see main text and compare Fig. 8.4).

8.3 Calculating the network structure using modular response analysis (MRA)

Identification of key regulatory interactions: MRA was performed for a transcription factor network comprising Copb, RelA, Fra1, Gfi-1, HMGA2, JunB, and Otx-1. The knock-down matrix given in Table 8.1 allowed calculating the global response matrix (using Eq. 8.1) which was used as the input for MRA (see Section 8.2). Figure 8.2 summarises the results of MRA (i.e., the elements of the local response matrix \( r_{ij} \)). Each Subplot shows how a single transcription factor is controlled by the other players of the systems (see x-axis). The response coefficient indicated on the y-axis corresponds to the strength of regulatory control. A value of zero means no control, while positive and negative values correspond to activation and inhibition, respectively. A value of one refers to linear control, a value of two means quadratic control, and so on.

![Figure 8.4: Interactions with high confidence local response coefficients.](image)

The local responses given in Fig. 8.2 (grey bars) were classified based on the confidence intervals of their response coefficients. The solid lines indicate high confidence interactions whose complete error bar is above 0.5 or below -0.5. The dashed lines indicate medium confidence interactions with complete error bars over 0.2 or below -0.2. Grey blunted arrows indicate inhibition, while black arrows denote activation. The black dots on the arrows and the numbers mark high confidence interactions which are additionally characterised by a high average response coefficients (termed HA-HC interactions; see main text and compare Fig. 8.4).

The grey bars in Fig. 8.2 are the elements of the local response matrix \( r_{ij} \), and thus quantify the network structure, i.e., if and to what extent a transcription factor is directly controlled by the species of the system (Fig. 8.1A). For example, Fig. 8.2A implies that Gfi-1 has essentially no direct control over Copb, while Copb is strongly induced by Otx-1. In order to identify the most relevant regulatory interactions in the network, the local responses were classified based on their median response coefficients using cut-offs. Interactions with median response coefficients below 0.5 were neglected, while those with median response coefficients above 0.5 and 1 are schematically depicted in the network diagram in Fig. 8.3 as medium and strong interactions, respectively.

Noisy experimental data might give rise to uncertainties in the network structure predicted by MRA. The local response coefficients were therefore classified according to their error bar ranges in order to identify high confidence interactions in the network (Fig. 8.4). A regulatory interaction was defined to be reliable if its complete response coefficient error bar is above 0.2 (or below -0.2). This implies that the corresponding local response coefficients have a 84% probability to be above 0.2. These reliable interactions were further sub-classified into medium and high confidence interactions, and are accordingly drawn into the network diagram in Fig. 8.4.

A comparison of Figs. 8.3 and 8.4 reveals that regulatory interactions with high average response coefficients often also exhibit high reliability, as may be expected. These HA-HC (high average - high confidence) interactions are expected to be particularly relevant from a physiological point of view, and the corresponding arrows are thus marked by a dot on the
The HA-HC interactions were chosen for further robustness analyses and were used to predict further experiments as outlined in the following.

Table 8.2: Structural robustness of the network structure predicted by MRA.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>c</td>
<td>m</td>
<td>c</td>
<td>m</td>
<td>c</td>
<td>m</td>
<td>c</td>
<td>m</td>
<td>c</td>
<td>m</td>
</tr>
<tr>
<td>Cbp siRNA</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Fra1 siRNA</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>*</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>*</td>
<td>-</td>
<td>*</td>
</tr>
<tr>
<td>Gfi-1 siRNA</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>*</td>
</tr>
<tr>
<td>HMGA2 siRNA</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>JunB siRNA</td>
<td>+</td>
<td>+</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Otx-1 siRNA</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>RelA siRNA</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

The table summarises whether the HA-HC interactions derived from Fig. 8.3 and 8.4 retain high mean and/or high confidence even if less knock-down data is included in the calculations. MRA was repeated for 7 different reduced global response matrices, each of which lacks one of the siRNA perturbation experiments (i.e., all of its replicas). Each row corresponds to such a calculation, and the siRNA experiment not considered is indicated on the left. The column numbering on the top corresponds to the reaction numbers given in Fig. 8.4. The abbreviations m and c refer to mean response coefficient and to the confidence level, respectively. The symbols in the table indicate if the numbered reactions in Fig. 8.4 are still highly significant if less knock-down data is used for calculation. A (+)-symbol indicates that a reaction is still characterised by a large mean local response coefficient (absolute value > 0.5) or that the whole error bar is still above 0.2 or below -0.2. The notation (-) indicates that the mean local response coefficient has the same sign as in Fig. 8.3 or that the whole error bar range has still the same sign as in Fig. 8.4. The (~)-symbol indicates that even these conditions are no longer fulfilled, and thus indicate that the corresponding regulatory interaction might not be structurally robust. Finally, the notation (*) indicates that no results could be obtained for the corresponding reaction, because one of the participating molecules (i.e., the responses to its knock-down) was not included in the calculations.

Structural Robustness Analysis: In the previous paragraphs, the robustness of the MRA results with respect to experimental error and to biological variability was analysed by calculating error bars. It is also possible that the interactions predicted by MRA are highly sensitive to the number of input experiments, that is, to the number of siRNA experiments made. To empirically test whether the method is sensitive towards the structure of the input matrix, MRA was repeated for seven reduced versions of Table 8.1, each of which lacks one of the siRNA knock-down experiments. As a measure for robustness, it was analysed whether the HA-HC interactions retain high average and high confidence response coefficients even if reduced matrices are used for calculation. More specifically, it was determined whether the HA-HC interactions still fulfil the cut-off criteria defined in the caption of Figs. 8.3 and 8.4.

The results, summarised in Table 8.2, indicate that the MRA indeed yields quite robust predictions for the regulatory interactions within the network. From seven reduced matrices, a total of 50 calculation results were obtained for the HA-HC interactions. 38 of the results retained both high average and high confidence response coefficients as indicated by the (+)-notation in Table 8.2, and thus exhibit strong robustness. In 4 of the remaining 12 calculations both the high average and the high confidence of the response coefficient were lost simultaneously, thus indicating structural sensitivity. Nevertheless, Table 8.2 reveals that the average response coefficients are in general a structurally robust property of the HA-HC interactions, because in all 50 cases they retain at least the same sign when compared to Fig. 8.2 (~) notations in Table 8.2). In contrast, the error bars show more prominent changes in the reduced calculations. In 6 out of 50 cases the error bar extends into the range of response coefficients with opposite sign (~ notation in Table 8.2), so that a high confidence distinction between activation and inhibition is in principle no longer possible. However, more detailed analyses of the reduced calculations indicated that these error bars only slightly cross the zero point (not shown), and are thus qualitatively similar to those shown in Fig. 8.2.
Table 8.3: Regulatory Interactions with significantly different local and global response coefficients.

<table>
<thead>
<tr>
<th>Regulatory Interaction</th>
<th>Global (Median)</th>
<th>Global (Range)</th>
<th>Local (Median)</th>
<th>Local (Range)</th>
<th>z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fra-1 → JunB</td>
<td>0.74</td>
<td>0.65 to 0.88</td>
<td>-0.29</td>
<td>-0.5 to -0.13</td>
<td>4.12</td>
</tr>
<tr>
<td>HMGA2 → Copb</td>
<td>0.44</td>
<td>0.38 to 0.5</td>
<td>-0.21</td>
<td>-0.35 to -0.05</td>
<td>2.95</td>
</tr>
<tr>
<td>Copb → Fra-1</td>
<td>-1.01</td>
<td>-1.22 to -0.84</td>
<td>0.45</td>
<td>-0.11 to 1.08</td>
<td>2</td>
</tr>
<tr>
<td>Otx-1 → JunB</td>
<td>0.13</td>
<td>-0.05 to 0.35</td>
<td>-0.47</td>
<td>-0.75 to -0.24</td>
<td>1.46</td>
</tr>
<tr>
<td>RelA → JunB</td>
<td>0.00</td>
<td>-0.38 to 0.38</td>
<td>0.81</td>
<td>0.56 to 1.12</td>
<td>1.29</td>
</tr>
<tr>
<td>Gfi-1 → Copb</td>
<td>0.49</td>
<td>0.38 to 0.83</td>
<td>0.04</td>
<td>-0.25 to 0.3</td>
<td>1.22</td>
</tr>
<tr>
<td>Fra-1 → HMGA2</td>
<td>-0.42</td>
<td>-0.52 to -0.34</td>
<td>0.13</td>
<td>-0.25 to 0.6</td>
<td>1.2</td>
</tr>
<tr>
<td>Otx-1 → HMGA2</td>
<td>-0.56</td>
<td>-0.68 to -0.44</td>
<td>0.15</td>
<td>-0.42 to 0.89</td>
<td>1.03</td>
</tr>
</tbody>
</table>

The local and the global response coefficients shown in Fig. 8.2 (grey and black bars, respectively) were compared for each regulatory interaction. Listed are those interactions where the whole error bar of the local response coefficient is completely non-overlapping with the error bar of the normalised global response coefficient. The list is sorted according to the z-score, which is given by $z = \frac{|\text{median}_{\text{global}} – \text{median}_{\text{local}}|}{\left(\sigma_{\text{global}} + \sigma_{\text{local}}\right)}$, and thus expresses the difference between the local and the global responses as multiples of the standard deviation. The local standard deviation was estimated using the formula $\sigma_{\text{local}} = \left|\text{median}_{\text{local}} – \max(\text{range}_{\text{local}})\right|$ if $\text{median}_{\text{local}} < \text{median}_{\text{global}}$ or by $\sigma_{\text{local}} = \left|\text{median}_{\text{local}} – \min(\text{range}_{\text{local}})\right|$ if $\text{median}_{\text{local}} > \text{median}_{\text{global}}$. The standard deviation $\sigma_{\text{global}}$ was calculated similarly.

Taken together, the results summarised in Table 8.2 indicate that the HA-HC interaction are fairly robust towards the number of siRNA input experiments used for MRA. Especially, reactions 1, 2, 3, 4 and 9 in Fig. 8.4 retain a high average response coefficients and high confidence for all calculations (Table 8.2), and are thus expected to be particularly reliable.

### 8.4 Deriving experimentally testable predictions from modular response analysis (MRA)

**Comparison of local and global responses:** In many cases, the computed direct regulatory interactions (grey bars in Fig. 8.2) will trivially reflect the experimentally observed global response (given in Table 8.1). However, both quantities can differ if indirect regulation plays a significant role in addition to direct regulation (Fig. 8.1C). As discussed in the context of Fig. 8.1D, these differences between local and global responses could be used to derive non-trivial, experimentally testable predictions from MRA. Therefore, the local and the global responses within the transcription factor network will be systematically compared in the following.

The experimental results in Table 8.1 were processed such that they are directly comparable to the local response coefficients obtained by MRA. This was done by normalizing each column of the global response matrix $R_{ij}$ (Eq. 8.1) by its diagonal element $R_{jj}$. The resulting normalised global response matrix has the same form as the local response matrix (Eq. 8.2), but the condition that all other network species ($x_k$ with $k \neq i, j$) are kept constant does no longer hold. The normalised global response coefficients are plotted in Fig. 8.2 (black bars), and the error bars were again estimated from the standard deviations given in Table 8.1.
Table 8.4: Indirect pathways explain differences in local and global responses.

<table>
<thead>
<tr>
<th>#</th>
<th>Regulatory Interaction</th>
<th>Local Effect</th>
<th>Global Effect</th>
<th>Indirect Pathways</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Fra1 → JunB</td>
<td>Weak Inhibition (-0.29)</td>
<td>Strong Activation (0.74)</td>
<td>Fra1 → Copb → JunB</td>
</tr>
<tr>
<td>2</td>
<td>HMGA2 → Copb</td>
<td>Weak Inhibition (-0.21)</td>
<td>Activation (0.44)</td>
<td>HMGA2 → Fra1 → Copb</td>
</tr>
<tr>
<td>3</td>
<td>Cobp → Fra1</td>
<td>Activation (0.45)</td>
<td>Strong Inhibition (-1.01)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Otx-1 → JunB</td>
<td>Inhibition (-0.47)</td>
<td>Weak Activation (0.13)</td>
<td>Otx1 → Copb → JunB</td>
</tr>
<tr>
<td>5</td>
<td>RelA → JunB</td>
<td>Strong Activation (0.81)</td>
<td>No effect (0.00)</td>
<td></td>
</tr>
</tbody>
</table>

The five regulatory interactions with the most significant differences between local and (normalised) global response coefficients according to Table 8.3 are listed. The local and global effects of the corresponding transcription factor knock-downs on target expression are summarised in the 2nd and 3rd columns, and the response coefficients are given in parentheses. The right column contains possible indirect pathways which according to Figs. 8.3 and 8.4 might be responsible for the observed differences in the local and global responses. Pathways in parentheses do not completely consist of highly reliable HA-HC interactions.

It should be noted that the black bars in Fig. 8.2 also allow for direct comparison among siRNA replicate transfections (‘biological replicas’). The black bars in Fig. 8.2 for a given regulatory interaction are usually very similar which indicates good reproducibility of the perturbation experiments. In some cases, the degree of target mRNA downregulation differs significantly from one siRNA replicate to another (see, e.g., siRNA for Otx-1 in Table 8.1). Even in these cases the normalized global response coefficients (black bars in Fig. 8.2) are typically very similar among replicas. This suggests that the key assumption of MRA (i.e., a linearisation in logarithmic space) is indeed fulfilled within the typical range of siRNA-mediated downregulation.

Table 8.3 shows a compilation of those regulatory interactions whose local response coefficients significantly differ from the normalised global response coefficients. Listed are those interactions where the whole error bar of the local response coefficient (grey bars in Fig. 8.2) is completely non-overlapping with the error bar of the normalised global response coefficient (black bars in Fig. 8.2). The list is sorted according to a z-score, which expresses the difference between the local and the global responses as multiples of the standard deviation. Thus, the regulatory interactions on the top of Table 8.3 exhibit particularly significant differences between the local and the normalised global response coefficients.

As discussed in the context of Fig. 8.1, indirect regulatory pathways are likely to be responsible for the differences between the local and global response coefficients given in Table 8.3. Such indirect regulatory routes within the transcription factor network that might compensate for direct local effects are summarised in Table 8.4. For example, the three-step pathway Fra1 → Copb → JunB consisting of HA-HC interactions only could explain why Fra1 does not affect JunB locally, while it does so globally.
**Figure 8.5: Experimental Verification by Elimination of Alternative Pathways.**

(A) Shown are the response coefficients for the two top regulatory interactions in Table 8.4 (i.e., Fra1 → JunB and HMGA2 → Copb). The black bars are the normalised global response coefficients derived from the experimental data given in Table 8.1 (already shown in Fig. 8.2). The dark grey bars are the local response coefficients calculated for wildtype cells (already shown in Fig. 8.2), and the light grey bars are the normalised global response coefficients for hypothetical Copb and Fra1 knock-out cell lines, respectively (calculated by using Eq. 8.3). The light grey and black bars are significantly different from each other, so that the knock-out cell line can be used to verify the predicted network structure. (B) The pathway HMGA2 → JunB can also be used to verify the network structure. The predicted (normalised) global response coefficients of Fra-1 and Copb knock-out cells (light grey bars) are significantly smaller than the (normalised) global response coefficients of wildtype cells.

**Experimental verification by elimination of alternative pathways:** Table 8.4 allows proposing new experiments which might verify the transcription factor network depicted in Figs. 8.3 and 8.4. The indirect pathways listed in Table 8.4 are likely to be responsible for the differences in local and normalised global response coefficients. As discussed in the context of Fig. 8.1, it should be possible to convert the experimentally observed global response coefficients (black bars in Fig. 8.2) into the calculated local response coefficients (grey bars in Fig. 8.2) by knocking out intermediates in the indirect pathways in Table 8.4. For example, a knock-out of Copb is expected to abolish the experimentally observed global activation of JunB by Fra-1, and to convert it into a weak inhibition (see regulatory interaction #1 in Table 8.4).

This strategy was further confirmed by calculating the global response matrix $R_{ij}$ (Eq. 8.1) for hypothetical knock-out cell lines (e.g., for Copb knock-out cells for the example above). Then, a comparison $R_{ij}$ calculated for knock-out cells with the experimentally retrieved wildtype $R_{ij}$ (black bars in Fig. 8.1) reveals whether the indirect pathways listed in Table 8.4 are indeed responsible for the differences of local and normalised global response coefficients. The global response matrix $R_{ij}$ with the elements given in Eq. 8.1 can be calculated back from the local response matrix $r_{ij}$ by using the following formula [68]

$$R = r_{ij}^{-1} \cdot (dgr_p).$$  \hspace{1cm} (8.3)

Here, $dgr_p$ is a matrix with all off-diagonal elements equal to zero. The diagonal elements of $dgr_p$ equal the fractional change in transcription factor expression induced by siRNA knock-down against the very same transcription factor, and thereby quantify the efficiency of siRNA-mediated knock-down. Importantly, $dgr_p$ quantifies only direct knock-down effects, and neglects any network feedback effects (i.e., the module is considered to be in isolation). Due to the linearity assumption underlying MRA, the diagonal elements of $dgr_p$ can be chosen arbitrarily, because the analysis was anyway restricted to *normalised* global response coefficients (see Figs. 8.2 and 8.5). The previously calculated local response matrix $r_{ij}$ (grey bars in Fig. 8.2) was used as an input for Eq. 8.3. Hypothetical knock-out cell lines were simulated by eliminating the depleted species from $r_{ij}$. In order to estimate confidence
intervals for the global response coefficients, the calculations incorporated the $10^6$ realisations for $r_{ij}$ which were obtained during the calculation of Fig. 8.2 (see Section 8.2). Unfortunately, most of the calculations yielded large error bars for the global response coefficients in hypothetical knock-out cell lines, so that the proposed experimental verification method is not suitable for these cases.

Two exceptions, shown in Fig. 8.5A, are characterised by significantly different global response coefficients between wildtype and hypothetical knock-out cell lines. As expected from Table 8.4, the global effect of a Fra-1 knock-down on JunB can indeed be weakened by a Copb knock-out. Also, the global response coefficient in Copb knock-out cells is predicted to be much more similar to the local response coefficient, thus suggesting that Copb is the major mediator of Fra1-mediated JunB activation. Likewise, Fra-1 seems to play a key role for HMGA2-mediated Copb activation, and this prediction can be tested in Fra-1 knock-out cells. The predictions in Fig. 8.5A are based on the simplest assumption regarding differences in local vs. global responses, but more complicated relationships between local and global responses are conceivable. It was therefore systematically analysed (using Eq. 8.3) whether any hypothetical single knock-out cell gives rise to a significant shift in any of the global response relative to wildtype cells. It turned out that global response of the pathway HMGA2 $\rightarrow$ JunB is significantly lowered in Fra1 and Copb knock-out cells (Fig. 8.5B). This seems to be due to the fact that strong indirect activation of JunB by HMGA2 via the pathways already identified in Fig. 8.5A (i.e., HMGA2 $\rightarrow$ Copb and Fra1 $\rightarrow$ JunB) is lost in Fra1 or Copb knock-out cells. Moreover, indirect inhibition of JunB by HMGA2 (via RelA, Gfi-1 and/or Otx-1) appears to outweigh direct activation of JunB by HMGA2 (dark grey bar in Fig. 8.5B), thus giving rise to a significant shift in the global response.

In conclusion, the analysis presented in this section suggest the existence of a multistep regulatory pathway (HMGA2 $\rightarrow$ Fra-1 $\rightarrow$ Copb $\rightarrow$ JunB) which appears to play a key functional role for the induction of JunB by HMGA2.

8.5 Conclusions and outlook

By applying MRA, the network interaction map was calculated from the knock-down data set shown in Table 8.1. The strongest and most significant interactions are shown in Figs. 8.3 and 8.4, and their relevance was further confirmed using the structural robustness analysis in Table 8.2. However, the regulatory interactions predicted by MRA cannot straightforwardly be tested using conventional biochemical assays, because it cannot be distinguished whether transcriptional regulation occurs by proximate promoter binding or through a multi-step chain of regulatory events. Using the strategy schematically depicted in Fig. 8.1C and D, experimentally testable predictions were derived from MRA. Regulatory interactions with significantly different local and global responses imply that indirect regulation via other species within the network is important in addition to direct regulation. Inhibition of the corresponding indirect regulatory pathways should therefore shift the experimentally measurable global response towards the local response predicted by MRA. Figure 8.5 shows that experimental blockade of indirect regulatory pathways via Copb and Fra1 can indeed be used to test for network structure. This network verification strategy, however, requires that knock-out cell lines for Copb and/or Fra1 can be established. Here, the term knock-out does not necessarily mean that Copb/Fra1 expression needs be totally abolished. More importantly, Copb/Fra1 must lose their ability to transmit relative changes in transcription factor expression, e.g., by being insensitive to inputs. Then, the system leaves the linear regulatory range, and thus corresponds to a knock-out cell line in the context of MRA. There are several ways to experimentally establish such MRA knock-out cell lines as summarised in the following. A homozygous knock-out of the gene locus completely abolishes expression, and thus completely removes the protein from the network. In some cases, essentially the same effect can be achieved by overexpression of a dominant-negative construct. Alternatively, a strong siRNA-mediated knock-down might be sufficient to drop
transcription factor protein expression below a certain threshold, thus kinetically removing the protein from the network. Also, exogenous overexpression will render the transcription factor pool insensitive towards endogenous induction/repression mechanisms, and thereby prevents signal transmission. Finally, the transcription factor can be removed from the network by incubating cells using a small-molecule inhibitor, which blocks transcription factor action at the protein level.

The analysis presented in this chapter strongly supports the existence of a linear transcription factor cascade consisting of HMGA2 → Fra-1 → Copb → JunB. According to Figs. 8.3 and 8.4, all interactions within the cascade are characterised by high average and high confidence local response coefficients, and are shown to be structurally robust (Table 8.2). Interestingly, all feed-forward regulatory interactions between the cascade species (i.e., HMGA2 → Copb, HMGA2 → JunB, Fra1 → JunB) are small relative to those of the interactions within the cascade (i.e., HMGA2 → Fra1, Fra1 → Copb, and Copb → JunB). Moreover, the upstream cascade species (HMGA2, Fra1, and Copb) do not strongly regulate the external network species Otx-1, RelA and Gfi-1 (Figs. 8.3 and 8.4). This suggest a linear multi-step cascade topology devoid of direct or indirect feed-forward regulation (although feedback from Copb and JunB might play a role). The published literature supports the existence of a HMGA2 → Fra-1 → Copb → JunB transcription factor cascade: HMGA2 expression is known to enhance Fra1 and JunB transcription [401,402], and Fra1 expression was shown to induce JunB transcription [403]. Accordingly, Fra1 and JunB have been reported to be co-expressed at the mRNA level in various tumour cells [402], and HMGA2 and Fra1 protein levels correlate in multiple cell lines under different stimulation conditions (Christine Sers, unpublished reverse-phase protein array data). Copb (also known as KLF6) is a tumour suppressor involved in growth inhibition [404] that is known to be induced upon extracellular stimulation [405,406,407], much like HMGA2 [408], Fra-1 [409], and JunB [410]. Recent analyses, suggest that Ras activation promotes alternative splicing of Copb [411], and that the alternatively spliced Copb transcript does not exert tumour suppressor functions, but instead acts as a positive regulator of cell division [412]. Likewise, it is known that HMGA2, JunB, and Fra1 positively regulate cell cycle progression, at least in part by activation of the cyclin A promoter [403,413,414]. Taken together, these data might explain why siRNA-mediated knock-down of HMGA2, Fra1, Copb or JunB abolishes Ras induced cell growth (Iwona Stelniec, data not shown).

Interestingly, siRNA-mediated knock-down of HMGA2, Fra1, Copb or JunB inhibits anchorage-independent cell growth in soft agar ('3D growth'), but does not affect growth on a cell culture dish ('2D growth'). This is in contrast to knock-down of Otx-1, RelA or Gfi-1, all of which inhibit both 2D and 3D growth (Iwona Stelniec, data not shown). The mechanisms underlying such perturbation-specific regulation of downstream 2D vs. 3D growth behaviour remain to be identified. It has been reported that Cyclin A induction is dependent on cell adhesion, and that ectopic Cyclin A expression confers anchorage-independent growth to fibroblasts [415]. Given that HMGA2, Fra1 and JunB are known to induce cyclin A (see above), it might be worthwhile to complement the theoretical MRA analyses on cell growth by experimental measurements of cyclin A expression. The key transcription factors regulating growth can, in principle, be calculated using MRA, but a simplifying assumption is required: It is technically impossible to perturb the 2D or the 3D growth rate of cells, and to measure the resulting response of the transcription factor network. Thus, the full matrix of global responses required for MRA is not available. For simplicity, one might assume that cellular growth does not affect the transcription factor network under consideration, and therefore set the corresponding global responses to zero. Moreover, the subsequent MRA calculations need to be constrained in a bottom-up sense such that the elements of the resulting local response matrix corresponding to transcription factor regulation by cell growth are always zero (i.e., the corresponding arrows in the network structure need to be excluded).

Constrained MRA calculations similar to those mentioned in the last paragraph might also be used to include the transcription factor responses towards upstream Ras overexpression (1st
row in Table 8.1). It seems reasonable to assume that the downstream transcription factor network does not affect the amount of ectopically expressed, constitutively active Ras protein, so that the corresponding local and global response coefficients can be set zero as well. A further aim is to complement the mRNA data given in Table 8.1 by protein measurements of the transcription factor network, which are currently in preparation by Iwona Stelniec in the lab of Reinhold Schäfer. A reasonable constraint for MRA calculations including mRNA and protein data might be to assume that regulation of protein levels occurs by transcriptional regulation (i.e., via the mRNA species) only. Once these network extensions included in the calculations, the MRA results will be compared with those of other reverse engineering methods [416,417,418] to further confirm the model predictions. Finally, the MRA calculations will be constrained such that the number regulatory interactions becomes minimal (by setting multiple local response coefficients to zero in all possible combinations). This seems reasonable, as most biochemical networks are not fully connected, and will enable us to derive a minimal network topology describing the experimentally measured global responses.
Part IV

Conclusions
9 General conclusions and outlook

9.1 Experimental data for systems biology

This thesis provides insights into the dynamics of signal transduction by combining theoretical modelling with various experimental approaches including high throughput data (Sections 5, 6 and 8), small-scale time-resolved measurements (Section 6 and 7), perturbation by siRNA-mediated knock-down (Sections 6, 7 and 8), and of half-life data (Sections 5, 6 and 7). However, further experimental measurement techniques would be required to further verify the conclusions drawn and are generally required for comprehensive systems biological analyses of living cells.

All measurements mentioned above analyse cellular behaviour on the population level, that is, cellular heterogeneity is neglected and population data is assumed to reflect the behaviour of a hypothetical average cell. In many cases, however, single cells differ significantly from each other. For example, it is known that single-cell caspase activation is complete within minutes after initiation of the apoptotic process, and that cellular heterogeneity gives rise to a much more gradual time course in population measurements [30,419]. Likewise, all-or-none MAPK activation at the single-cell level during Xenopus oocyte maturation was shown to be obscured at the population level [136]. Thus, single cell measurements are required to prove all-or-none behaviour with respect to time and/or input concentration. Ideally, the single-cell techniques are non-invasive, that is, cellular behaviour is essentially unaffected by the measurement process, so that an individual cell can be tracked over multiple time points. Antibody-based measurements typically do not fulfil this requirement, as the cells often need be fixed prior to detection. Therefore, it seems that single-cell measurements based on fluorescent reporter substrates [28,30] or on FRET probes [149] are advantageous for time-resolved systems biological analyses. However, some nodes of the network are typically not measurable at the single cell level or only by invasive single-cell methods. In such cases, an integrated modelling approach incorporating population date and invasive as well as non-invasive single-cell measurements has proven to be useful to build a predictive model of apoptosis signalling (Peter Sorger, talk at SBMC 2006).

Another limitation of the experimental data included in this thesis is the lack of quantitative knowledge on sub-cellular compartmentalisation. More specifically, the cytoplasmic, nuclear and whole-cell measurements of the Smad proteins in Section 6 are not directly comparable, because the relative efficiency of the fractionation protocols is unknown. Moreover, SnoN protein could be detected in both, the nucleus and the cytoplasm (Peter Nickel, data not shown), so that the model assumption regarding its predominant localisation to the nucleus remains speculative. Immunofluorescence measurements of have proven to be useful to detect the sub-cellular distribution of SnoN and Smad proteins [338,349]. In particular, the Smad shuttling dynamics can be analysed in real-time by microscopic analysis of cells expressing low levels of GFP-tagged Smad proteins [339]. Unfortunately, these techniques are not applicable in primary mouse hepatocytes, as these cells show strong background fluorescence signals, most likely due to high expression of fluorophores. Additionally, the behaviour of primary mouse hepatocytes differs from preparation to preparation (i.e., from one donor animal to another), and they cannot easily be manipulated by transfection, which complicates systems biological analyses. Taken together, it seems that other cellular systems are superior over primary mouse hepatocytes when it comes to systems biology. In particular, the main argument for analyses of primary mouse hepatocytes, their close relationship to the in vivo situation, seems questionable: Ho et al. [420] reported that GRK2 is an important transcriptional feedback regulator of TGFβ signalling in primary mouse hepatocytes, while we did not find any evidence for such GRK2-mediated feedback (Section 6). Thus, it seems that the behaviour of primary mouse hepatocytes depends on the isolation

111
protocol, implying that most experiments in hepatocytes are unlikely to fully reflect the in vivo situation as discussed previously [421].

For the analysis presented in Section 5, protein half-life measurements had to be collected manually from the literature (Appendix E). Given that the protein half-life is expected to vary with the cell type, the culture conditions and with the experimental protocol, the analysis presented in Section 5 can only be considered to be a first step towards an integrated understanding of gene expression and turnover data. Global measurements of all protein half-lives under standard experimental conditions are required to further confirm the results obtained in Section 5. Moreover, such data would generally simplify kinetic modelling of transcriptional networks, as the time course of protein expression in response to transcriptional regulation is known to be determined by the mRNA and protein half-lives [391]. In fact, large-scale protein half-life data is already available for yeast [287], and turnover measurements should also be possible in mammalian, e.g., by using mass spectrometry [422]. Additional large-scale data resources are required to boost systems biological research on protein complex formation as summarised in the following. The relative abundances of different proteins in a given cell are central to the control of many phenomena including protein complex formation [423], enzyme catalysis [15], transcriptional control [424], and feedback regulation (Sections 3 and 4). Thus, measurements of absolute or relative protein concentrations for at least one standard cell line are expected to greatly improve the predictive power of mathematical models. Such data is already available for yeast [425], and a mass spectrometry-based quantification of several thousand proteins has also been reported for mammalian neurons [108]. Ideally, these protein quantifications should be complemented by measurements on competitive binding and on relative protein-protein affinities, at least for central parts of the regulatory network. The results presented in Sections 3, 4 and 7, and those obtained by others [399,426] suggest that competitive vs. simultaneous binding of proteins to common binding partners strongly affects the qualitative behaviour of biochemical networks. A central problem underlying mathematical models of receptor tyrosine kinase (RTK) signalling via multi-adaptors such as insulin receptor substrates (IRS) is the lack of knowledge regarding competitive vs. simultaneous binding, thus making any RTK model uncertain. Experimentally, competitive binding should be measurable on a high-throughput scale by protein microarrays or by comparing yeast two-hybrid interactions with and without co-expression of a hypothetical competitor. Competitive binding can only be significant if the competitor-target interaction affinity is sufficiently high, so that quantitative interaction data is required to investigate competition effects. Moreover, physiologically relevant parameter ranges generally need to be known to implement and to fit realistic mathematical models (discussed in Section 6.7). Protein-protein affinities and enzyme-substrate interactions are typically characterised in vitro using truncated protein constructs, so that measured kinetic parameter values will, in many cases, not reflect the in vivo situation, and are not comparable among each other. More specifically, it is known that the affinity of protein-protein interactions is highly dependent on the buffer conditions [427], and that protein truncation can profoundly alter the characteristics of protein-protein interactions, e.g., due to loss of multivalent binding [428] or due to loss of intramolecular autoinhibition [429]. One solution to this dilemma might be to measure the binding of multiple proteins to a common binding partner in parallel as recently reported for EGF receptor signalling [430] and for cyclin-dependent kinases [431], thus standardizing the buffer conditions and the protein constructs used. More in general, it seems more practical to define physiologically relevant ranges for each type of biochemical reaction (e.g., for kinase-mediated catalysis) based on measurements using full-length proteins in vitro (e.g., [432]) and/or obtained in vivo (e.g., [149,339]), as described in Appendix F for the TGFβ model. A public consensus database containing such parameter ranges would greatly simplify the implementation of physiologically relevant mathematical models, especially for newcomers to the field. Taken together, the data sets discussed in this section would help to reduce kinetic models of cellular processes, as specific reactions could be neglected a priori based on reasoning on time scale (protein half-life data; database containing typical parameter values) or based on competition arguments (lack of competition, low abundance or low relative
affinity). In other words, these high throughput data would allow for a more mechanistic simplification of mathematical models, thus decreasing the need for ad hoc assumptions (see also Section 9.3).

9.2 Testing the quantitative model of signal specificity

It is often claimed that systems biology will simplify drug discovery by enabling for in silico identification of potential drug targets (e.g., [433]). However, such therapeutical applications require that the present systems biological models can be extended such that they describe long-term phenotypic responses downstream of cellular signalling. So far, successful predictions of stimulus-induced cell fate decisions were only obtained using linear approaches that ‘simply’ correlate the activation profiles of intracellular signalling molecules with phenotypic responses [434]. However, these approaches, though valuable, neglect feedback loops from the downstream gene expression machinery involved in cell fate decisions to the upstream signalling network. Moreover, they do not provide mechanistic insights into how the multiple signalling pathways actually interact by multiple crosstalk mechanisms to bring about a stimulus-specific response. The long-term goal is therefore to construct mechanistic mathematical models based on high throughput and time-resolved protein array data covering many nodes within the network. As phenotypic commitment of cells typically requires ongoing signalling for several hours (Appendix A), protein array measurements should include both the modification state of signalling proteins and their absolute abundance. Based on time-resolved protein array data, we are currently trying to understand differential response of a cancer cell line towards different cancer drugs (in collaboration with Christine Sers, Charite Berlin). It turned out that parallel microarray-based expression profiling at the mRNA level is necessary to mechanistically understand the connections between the network species, i.e., to distinguish alternative paths of transcriptional regulation and to separate transcriptional regulation from regulation by proteasomal degradation.

A first step towards an integrated model incorporating upstream signal transduction and downstream phenotypic responses might be to analyse signal decoding at the level of gene expression for a specific signalling pathway. As summarised in Section 1.3, the quantitative aspects of intracellular signals such as amplitude and duration are thought to determine downstream gene expression and phenotypic responses. However, a rigorous experimental proof for this quantitative model signalling specificity is lacking for most signalling pathways and cell types. In most cases, conclusions regarding quantitative encoding were postulated if the time courses for a given signalling molecule (e.g., phospho-Erk) differed between two ligands, each of which induces a specific phenotypic response. Importantly, possible contributions of other signalling pathways were typically not investigated. Thus, a more systematic approach is required, where the amplitude and the duration of a specific intracellular signalling molecule can be varied selectively. Such an analysis should, for example, be possible in cells expressing a Raf construct (ΔRaf-ER) that can rapidly be activated by treating cells with estradiol [435]. Titration with different doses of estradiol should elicit a gradual increase in Erk activation in fibroblasts [141], and the corresponding dose-response can be measured using the monoclonal antibody against double phosphorylated, active Erk. Using the quantitatively measured Erk signal as the input, one can analyse gene expression or phenotypic responses as a readout, and thus relate the quantitative aspects of the signal to downstream responses. Importantly, the duration of the input Erk signal can easily be controlled by addition of the Mek inhibitor U0126 which terminates Erk signalling within minutes. Thus, a three-dimensional map can be generated showing phenotypic or gene expression responses as a function of Erk amplitude and duration. For instance, it can be investigated in detail whether c-fos indeed acts as an ultrasensitive decoder of Erk dynamics [436]. Often two intracellular signals interact in a nonlinear fashion to bring about synergistic [437,438], mutually exclusive [140,439] or even more complex [440] downstream responses. For example, it is known that Raf and Akt
cooperate to bring about much stronger cell cycle progression than either signal alone at least in some cells [437] . The interactions between these pathways can be investigated in cells expressing estrogen-sensitive Akt and androgen-sensitive Erk constructs [437] , and the signals can be terminated using small molecule inhibitors against Akt and Mek, thus allowing to generate a four-dimensional input map (Akt amplitude, Erk amplitude, Akt duration, Erk duration).

9.3 Modelling approaches in systems biology

Systems biology comes in various approaches, ranging from integration of large-scale data sets to a mechanistic description of small subsystems. Three classes of mathematical models can be distinguished: First, small mechanistic models are used to understand well-characterised sub-networks ('motifs'). Second, complex reactions networks are simulated using small fuzzy models if experimental data is limiting ('top-down' or 'mesoscale' models). Third, large mechanistic models are used to understand the effects of molecular perturbations on complex reaction networks (e.g., to simulate the effects of drugs acting at specific sites of the network). All models aim to conceptualise experimental results, and often focus on the explanation of non-trivial observations. Further aims might include the cost-effective design of new experiments and/or the discrimination of alternative reaction mechanisms [441] .

Unfortunately, it is hard to objectively compare all these approaches based on the published literature. This is due to the fact that, in many cases, some of the predicted data sets were most likely known prior to model finalisation, thus giving rise to a bias in model development. Moreover, there is often no clear distinction between predicted measurements, and data sets used to adjust the model. Especially for large mechanistic models all available data sets are typically used for model training, and all simulations are discussed as model ‘predictions’. This indicates that large mechanistic models typically do not allow to reliably predicting new experiments unless enormous amounts of experimental data are available. For instance, the relative contributions of redundant pathways cannot be identified unless experimental measurements are performed for each of the parallel branches. Although such comprehensive experimental coverage of signalling pathways might be available in the near future if protein array measurements become more reliable, the value of large mechanistic models remains unclear. More specifically, they often do not provide a qualitative understanding of the system behaviour, but merely serve as ‘dumb’ in silico replicates of the intracellular regulatory network that can be used to plan new experiments. Taken together, large mechanistic models that are generated based on sufficient amounts of experimental data might contribute to the design of drugs or experiments targeting specific interactions within the network. However, they will probably contribute little to our qualitative understanding of the intracellular regulatory network unless new model analysis tools are developed.

We recently described a mesoscale model describing the initiation of DNA synthesis during liver regeneration [442] . Due to limited experimental data, a simplistic approach was chosen: Post-translational regulation via signalling cascades was assumed to be fast relative to transcriptional regulation. Thus, only species induced at the transcriptional level were implemented as dynamical variables, while nonlinearities arising in fast signal transduction were included implicitly in the rate laws. The model was trained based on knock-out data, and the kinetic parameters and rate laws were chosen such that they match these data. Moreover, many of the nodes considered in the model were chosen based on the availability of corresponding knock-out data. In the retro-perspective, the model of liver regeneration reveals several weak points of mesoscale modelling: The simplifications made are often completely arbitrarily due to the lack of sufficient experimental data, and the same will be true for other mesoscale models (see also Section 9.1). In particular, the rate laws used to describe whole signal transduction pathways may or (more likely) may not reflect biological
reality. Moreover, it is impossible to choose a physiologically relevant range for the lumped kinetic parameters, because they have no direct physical meaning, thus giving rise to an unreasonably large parameter space. Owing to these pronounced topological and mechanistic uncertainties, one does not expect that mesoscale models will be predictive at present. Recently, two groups applied engineering control theory to get insights into the input-output processing by whole signal transduction pathways [443,444]. Such black-box approaches may provide a more solid basis for lumping together many molecular steps in future mesoscale models.

Based on the above arguments, I believe that mechanistic modelling of small modules or motifs will be the most successful application of mathematical modelling in biology (e.g., [445]). Biological sub-networks are often characterised in much detail, thus allowing for mechanistic modelling with parameters that have direct physical meaning (Sections 2, 3, 4, 6 and 8). The resulting models and their possible behaviours can, in many cases, be understood in qualitative terms, and the models are expected to have predictive power due to their mechanistic basis. However, modelling on a small scale requires that motifs can actually be identified within the complex intracellular regulatory network, and that they can be analysed experimentally without significant contribution from external interactions. Motif identification and separation from the network is particularly relevant for applications such as model discrimination and experimental design which are expected to fail for more complex reaction networks. Motif identification as the first step of modelling is not straightforward, as, for example, the MAPK cascade, which is often considered to be a module, is far from being isolated in living cells. More specifically, it is known that upstream stimuli not only increase Raf activity, but that they also directly regulate Mek [446] and that they control the phosphatase activity within the pathway [53]. High throughput analyses suggest that a large number of additional MAPK regulators exist in living cells [447], thus casting further doubt on the modularity assumption underlying many MAPK models. The following section briefly summarises experimental and theoretical analyses of small regulatory motifs described in the literature, and delineates a bottom-up in vitro approach that may be suited to identify and to analyse motifs in biochemical signalling networks.

9.4 Motifs in signal transduction - a bottom-up approach

Extensive analyses of the E. coli transcriptional regulatory network revealed the existence of recurrent regulation patterns, termed network motifs [59,62]. The properties of common transcription motifs including autoregulation and feed-forward loops were then investigated numerically, and the predicted behaviour was subsequently confirmed experimentally [62,395,448,449]. Likewise, another recurrent motif of gene expression networks, sRNA-mediated regulation, was analysed in Section 7, and the model predictions regarding sign-sensitive delay behaviour were subsequently confirmed experimentally.

Unfortunately, the motifs found in protein-protein interaction networks are of little value for mathematical modelling, as the data sets do not distinguish between different types of interactions (e.g., inhibition vs. activation and catalysis vs. simple binding). Published kinetic modelling studies were therefore mainly focussed on basic motifs that have long been known to be relevant for signal transduction such as the (de)phosphorylation cycle [14], multisite phosphorylation [167] and protein kinase cascades [280]. In Section 5, an additional recurrent motif of mammalian signal transduction cascades was identified: transcriptional autoregulation of signalling cascades generally results in negative feedback, and occurs by inhibitor (RFI) upregulation, while transducer downregulation plays no role. The simulations in Section 6 reveal another twist in TGFβ-mediated regulation of SnoN: Smad signalling enhances SnoN degradation in addition to its transcriptional upregulation, thus giving rise to a combination of negative feedback and incoherent feed-forward regulation. Intriguingly, such dual induction and degradation has been reported for other RFIs including MKP-3 [450], Sprouty [451], Smad7 [452] and SOCS-3 [453]. The simulations in Section 6.7 suggest
that strong Smad-induced induction and degradation of SnoN ensures that SnoN selectively acts as a feedback inhibitor (‘homologous desensitisation’), while its induction by other stimuli (‘heterologous desensitisation’) is predicted to be without significant effect on TGFβ signalling. However, more detailed theoretical analyses are required to get further insights into the kinetic implications of RFI action.

A recent study in Science revealed that three cyanobacterial circadian proteins, termed KaiA, KaiB and KaiC, oscillate with a 24h period when the respective recombinant proteins were mixed in vitro [454]. Subsequent mathematical modelling studies provided detailed mechanistic insights into the functioning of the cyanobacterial clock (e.g., [168]). The theoretical analyses were greatly simplified and corroborated by the fact that potential additional regulatory factors present in living cells could be excluded. Moreover, in vitro mixtures of recombinant proteins are advantageous, because they behave like single cells, so that heterogeneity effects, typically arising in population measurements, can be neglected. In conclusion, it seems that in vitro reconstitution systems provide valuable tools for the quantitative analysis of regulatory network motifs. In collaboration with Lee Bardwell (University of California, Irvine), we are currently investigating the behaviour of the Erk (de)phosphorylation cycle by mixing the kinase (Mek), the substrate (Erk) and the phosphatase (MKP-3). Our preliminary results indicate that the Erk phosphorylation level responds to changes in the Mek and MKP-3 concentration in a highly ultrasensitive fashion. Another well-established in vitro system is the caspase cascade comprising Apaf-1, caspase-9, caspase-3 and XIAP that was analysed by mathematical modelling in Section 4 [201]. The relevance of experimental and theoretical results obtained in vitro needs to be confirmed in living cells, where multiple additional regulatory factors might play a role. This can be done in a step-wise manner: Experiments in cytosolic extracts can easily be manipulated by addition of recombinant proteins or by antibody-based protein depletion, and heterogeneity effects can again be neglected. Thus, extract experiments are directly comparable to the results obtained in vitro, and such a comparison might allow verifying the central assumption that a simple regulatory motif under consideration is separable from the rest of the network. Such experiments in cytosolic extracts already proved to be useful for analyses of the MAPK cascade in Xenopus oocytes [12] and of the apoptotic caspase cascade in mammalian cells [175]. As a final step, the results obtained in vitro and in extracts might be compared to those in living cells which are harder to manipulate (e.g., by overexpression or knock-down). The proposed approach might fail for regulatory motifs that require transcription and translation, even though protein biosynthesis can, in principle, be maintained in vitro and in extracts. In such cases, synthetic biology approaches (e.g., [455]) might help to analyse motif behaviour in isolation by expression of motif constituents in heterologous organisms. In conclusion, it seems that one of the major benefits of synthetic biology could be the separation of functional modules from the rest of the intracellular regulatory network, and thus to understand them in more detail.
Part V

Appendix
## A Additional material on time scales of signal transduction

### Table A.1 Duration of intracellular signals.

<table>
<thead>
<tr>
<th>Signal Event</th>
<th>Stimulus</th>
<th>Condition</th>
<th>Signal Duration</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erk phosphorylation</td>
<td>Growth Factors</td>
<td>Cell culture</td>
<td>10’ – 24h</td>
<td>[456], [36]</td>
</tr>
<tr>
<td>Erk phosphorylation</td>
<td>Antigenic peptides</td>
<td>Thymocyte selection (in vivo)</td>
<td>&gt; 70 h</td>
<td>[123]</td>
</tr>
<tr>
<td>Erk phosphorylation</td>
<td>Developmental Signals</td>
<td>Mouse development (in vivo)</td>
<td>up to 72 h</td>
<td>[457]</td>
</tr>
<tr>
<td>PI3K activity</td>
<td>insulin</td>
<td>Cell culture</td>
<td>10’ - 8h</td>
<td>[458,459]</td>
</tr>
<tr>
<td>Akt phosphorylation</td>
<td>NGF</td>
<td>Cell culture</td>
<td>10’ - 15h</td>
<td>[459,460]</td>
</tr>
<tr>
<td>cAMP increase or PKA activation</td>
<td>Isoproterenol</td>
<td>Cell culture</td>
<td>5’ - 2 h</td>
<td>[37,461]</td>
</tr>
<tr>
<td>Smad2 phosphorylation</td>
<td>TGFβ</td>
<td>Cell culture</td>
<td>2 h – 24 h</td>
<td>[49,350]</td>
</tr>
<tr>
<td>Nuclear Smad2-Smad4 complex formation</td>
<td>TGFβ</td>
<td>Cell culture</td>
<td>2 h – 6 h</td>
<td>[49]</td>
</tr>
<tr>
<td>Translocation of Smad2</td>
<td>TGFβ</td>
<td>Cell culture</td>
<td>2 h – 6 h</td>
<td>[49]</td>
</tr>
<tr>
<td>Smad2 DNA binding</td>
<td>TGFβ</td>
<td>Liver regeneration (in vivo)</td>
<td>&gt; 120h</td>
<td>[300]</td>
</tr>
<tr>
<td>STAT activation</td>
<td>Cytokines</td>
<td>Cell culture</td>
<td>Minutes to several hours</td>
<td>[462,463]</td>
</tr>
<tr>
<td>STAT3 DNA binding</td>
<td>Cytokines</td>
<td>Liver regeneration (in vivo)</td>
<td>8 h</td>
<td>[464]</td>
</tr>
<tr>
<td>NF-kB DNA binding</td>
<td>Cytokines</td>
<td>Cell culture</td>
<td>1 h – 6 h</td>
<td>[40,43]</td>
</tr>
<tr>
<td>NF-kB DNA binding</td>
<td>Cytokines</td>
<td>Liver regeneration (in vivo)</td>
<td>~ 10 h</td>
<td>[465]</td>
</tr>
</tbody>
</table>

The time a given signal remains significantly elevated over the basal level was estimated from literature studies. The data reveal that signalling pathways frequently remain active over several hours under physiological conditions.
<table>
<thead>
<tr>
<th>Signalling event</th>
<th>Stimulus</th>
<th>Cell Type</th>
<th>Inhibitor</th>
<th>Duration (no inhibitor)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>NF-κB binding to DNA</td>
<td>IL1beta</td>
<td>epithelial</td>
<td>ActD</td>
<td>~320'</td>
<td>[466]</td>
</tr>
<tr>
<td>NF-κB binding to DNA</td>
<td>IL1beta</td>
<td>epithelial</td>
<td>CHX</td>
<td>~320'</td>
<td>[466]</td>
</tr>
<tr>
<td>Smad2 phosphorylation</td>
<td>TGFbeta</td>
<td>Hepa1-6</td>
<td>ActD</td>
<td>~300'</td>
<td>Peter Nickel, pers. comm.</td>
</tr>
<tr>
<td>Smad1/5 phosphorylation</td>
<td>TGFbeta</td>
<td>BAECs</td>
<td>CHX</td>
<td>~120'</td>
<td>[370]</td>
</tr>
<tr>
<td>STAT3 binding to DNA</td>
<td>IL-10</td>
<td>Macrophages</td>
<td>ActD</td>
<td>~60'</td>
<td>[467]</td>
</tr>
<tr>
<td>STAT5 phosphorylation</td>
<td>G-CSF</td>
<td>32D</td>
<td>ActD</td>
<td>~20'</td>
<td>[468]</td>
</tr>
<tr>
<td>STAT5 phosphorylation</td>
<td>G-CSF</td>
<td>32D</td>
<td>CHX</td>
<td>~20'</td>
<td>[468]</td>
</tr>
<tr>
<td>STAT5 binding to DNA</td>
<td>G-CSF</td>
<td>32D</td>
<td>ActD</td>
<td>~20'</td>
<td>[468]</td>
</tr>
<tr>
<td>STAT5 binding to DNA</td>
<td>G-CSF</td>
<td>32D</td>
<td>CHX</td>
<td>~20'</td>
<td>[468]</td>
</tr>
<tr>
<td>Erk phosphorylation</td>
<td>EGF</td>
<td>Swiss 3T3</td>
<td>CHX</td>
<td>~45'</td>
<td>[469]</td>
</tr>
<tr>
<td>Erk phosphorylation</td>
<td>serum</td>
<td>fibroblasts</td>
<td>CHX</td>
<td>~120'</td>
<td>[470]</td>
</tr>
<tr>
<td>Erk phosphorylation</td>
<td>EGF</td>
<td>HeLa</td>
<td>CHX</td>
<td>~30'</td>
<td>[67]</td>
</tr>
<tr>
<td>Erk phosphorylation</td>
<td>HRG</td>
<td>MCF7</td>
<td>CHX</td>
<td>~120'</td>
<td>[471]</td>
</tr>
<tr>
<td>Nuclear Erk phosphorylation</td>
<td>serum</td>
<td>CCL39</td>
<td>ActD</td>
<td>?</td>
<td>[472]</td>
</tr>
<tr>
<td>Nuclear Erk phosphorylation</td>
<td>serum</td>
<td>CCL39</td>
<td>CHX</td>
<td>?</td>
<td>[472]</td>
</tr>
<tr>
<td>Erk kinase activity</td>
<td>Angiotensin II</td>
<td>VSMC</td>
<td>ActD</td>
<td>~60'</td>
<td>[473]</td>
</tr>
<tr>
<td>JNK phosphorylation</td>
<td>EGF</td>
<td>HeLa</td>
<td>CHX</td>
<td>~30'</td>
<td>[67]</td>
</tr>
<tr>
<td>JNK phosphorylation</td>
<td>H2O2</td>
<td>293T</td>
<td>ActD</td>
<td>~240'</td>
<td>[39]</td>
</tr>
<tr>
<td>JNK phosphorylation</td>
<td>H2O2</td>
<td>293T</td>
<td>CHX</td>
<td>~240'</td>
<td>[39]</td>
</tr>
<tr>
<td>JNK kinase activity</td>
<td>TNFα</td>
<td>Mesangial</td>
<td>ActD</td>
<td>~20'</td>
<td>[474]</td>
</tr>
<tr>
<td>JNK kinase activity</td>
<td>TNFα</td>
<td>Mesangial</td>
<td>CHX</td>
<td>~20'</td>
<td>[474]</td>
</tr>
<tr>
<td>p38 phosphorylation</td>
<td>EGF</td>
<td>HeLa</td>
<td>CHX</td>
<td>~30'</td>
<td>[67]</td>
</tr>
</tbody>
</table>

Listed are literature studies where time courses of signalling protein activity was measured in the presence and in the absence of a transcription inhibitor (actinomycin D = ActD) and/or a translation inhibitor (cycloheximide = CHX). In all studies actinomycin and cycloheximide treatment prolonged signalling which supports the relevance of transcriptional negative feedback regulation in the mammalian signalling network. The column ‘signal duration (no inhibitor)’ indicates the time of signal termination in cells not treated with inhibitor and thus gives a hint about the time scale of transcriptional feedback regulation.
<table>
<thead>
<tr>
<th>Signal Event</th>
<th>Cellular Response</th>
<th>Stimulus (cell type)</th>
<th>Method</th>
<th>Threshold time</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mek activity</td>
<td>S-Phase Entry</td>
<td>FGF (NIH3T3)</td>
<td>Inhibitor addition</td>
<td>14 h&lt;sup&gt;a&lt;/sup&gt;</td>
<td>[456]</td>
</tr>
<tr>
<td>Mek activity</td>
<td>S-Phase Entry</td>
<td>PDGF (NIH3T3)</td>
<td>Inhibitor addition</td>
<td>&gt; 8 h&lt;sup&gt;d&lt;/sup&gt;</td>
<td>[475]</td>
</tr>
<tr>
<td>Mek activity</td>
<td>S-Phase Entry</td>
<td>PDGF (IIC9)</td>
<td>Inhibitor addition</td>
<td>&gt; 4 h&lt;sup&gt;a&lt;/sup&gt;</td>
<td>[122]</td>
</tr>
<tr>
<td>PI3K activity</td>
<td>S-Phase Entry</td>
<td>insulin (Rat1)</td>
<td>Inhibitor addition or antibody injection</td>
<td>13 h&lt;sup&gt;a&lt;/sup&gt;</td>
<td>[458]</td>
</tr>
<tr>
<td>Mek activity</td>
<td>Cell Motility</td>
<td>HGF (MDCK)</td>
<td>Inhibitor addition</td>
<td>&gt; 30&lt;sup&gt;b&lt;/sup&gt;</td>
<td>[476]</td>
</tr>
<tr>
<td>Mek activity</td>
<td>Cell Motility</td>
<td>EGF (SCC-11F)</td>
<td>Inhibitor addition</td>
<td>4 h&lt;sup&gt;b&lt;/sup&gt;</td>
<td>[477]</td>
</tr>
<tr>
<td>Mek activity</td>
<td>Differentiation</td>
<td>TPA (K562)</td>
<td>Inhibitor addition</td>
<td>18 h&lt;sup&gt;b&lt;/sup&gt;</td>
<td>[478]</td>
</tr>
<tr>
<td>Mek activity</td>
<td>Differentiation</td>
<td>M-CSF (myeloid)</td>
<td>Inhibitor addition</td>
<td>&gt; 24 h&lt;sup&gt;b&lt;/sup&gt;</td>
<td>[479]</td>
</tr>
<tr>
<td>Mek activity</td>
<td>Thymocyte positive selection</td>
<td>Antigenic peptides (in vivo)</td>
<td>Inhibitor addition</td>
<td>&gt; 24 h&lt;sup&gt;b&lt;/sup&gt;</td>
<td>[123]</td>
</tr>
<tr>
<td>TGFbeta receptor activity</td>
<td>Growth arrest</td>
<td>TGFbeta (HaCaT)</td>
<td>Inhibitor addition</td>
<td>12-14 h&lt;sup&gt;a&lt;/sup&gt;</td>
<td>[49]</td>
</tr>
<tr>
<td>Calcineurin activity</td>
<td>Thymocyte lineage commitment</td>
<td>PMA + Ionomycin (thymocytes)</td>
<td>Inhibitor addition</td>
<td>~8 h&lt;sup&gt;a&lt;/sup&gt;</td>
<td>[480]</td>
</tr>
<tr>
<td>PKA activity</td>
<td>Astroglial Differentiation</td>
<td>Isoprotenerol (astrocytes)</td>
<td>Inhibitor addition</td>
<td>&lt; 2 h&lt;sup&gt;e&lt;/sup&gt;</td>
<td>[461]</td>
</tr>
<tr>
<td>Mek activity</td>
<td>Late-phase Gene Expression</td>
<td>HGF (MDCK)</td>
<td>Inhibitor addition</td>
<td>&gt; 30&lt;sup&gt;b&lt;/sup&gt;</td>
<td>[476]</td>
</tr>
<tr>
<td>Mek activity</td>
<td>MMP-9 expression</td>
<td>EGF (SCC-11F)</td>
<td>Inhibitor addition</td>
<td>4 h&lt;sup&gt;b&lt;/sup&gt;</td>
<td>[477]</td>
</tr>
<tr>
<td>Erk phosphorylation</td>
<td>Late-phase Gene Expression</td>
<td>LPA (Rat-1)</td>
<td>Stimulus-strength specific signal duration</td>
<td>~1 h&lt;sup&gt;b,c&lt;/sup&gt;</td>
<td>[35]</td>
</tr>
</tbody>
</table>

Listed are literature studies where activity of signalling intermediates was blocked at different stimulation times (stimulus addition at t = 0 h) by incubating cells with rapidly-acting small-molecule inhibitors. The threshold times indicate the signal duration required to irreversibly commit the cell population to the given phenotypic response (stimulation occurred at t = 0 h). Importantly, the signalling events are primary in the sense that they activated within minutes after stimulation, that is, the threshold time reflects the true signal duration. The data suggests that decision making in mammalian cells will be profoundly affected by slowly acting transcriptional feedback mechanisms. Superscript legend: (a) half of the cells perform cellular response if signal is terminated at the threshold time; (b) no response observed if signal is terminated at the threshold time; (c) Erk dependency of late-phase gene expression verified by Erk inhibitor; (d) half-maximal thymidine incorporation is observed if signal is terminated at the threshold time; (e) complete commitment occurred before the indicated threshold time. Note: One study [35] employed a different experimental approach: low and high doses of LPA induce transient and sustained Erk phosphorylation, respectively, and some Erk-dependent downstream genes (e.g., Fra1) were shown to be selectively expressed upon sustained Erk activation.
Table A.4: Commitment Times for Extracellular Stimulation.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Cellular Response</th>
<th>Cell type</th>
<th>Method</th>
<th>Threshold time</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGF</td>
<td>S-Phase Entry</td>
<td>MDCK</td>
<td>Medium exchange</td>
<td>~6 h&lt;sup&gt;a&lt;/sup&gt;</td>
<td>[481]</td>
</tr>
<tr>
<td>EGF</td>
<td>S-Phase Entry</td>
<td>BT20</td>
<td>Medium exchange</td>
<td>~6 h&lt;sup&gt;a&lt;/sup&gt;</td>
<td>[481]</td>
</tr>
<tr>
<td>Serum</td>
<td>S-Phase Entry</td>
<td>MDCK</td>
<td>Medium exchange</td>
<td>~6 h&lt;sup&gt;a&lt;/sup&gt;</td>
<td>[481]</td>
</tr>
<tr>
<td>Serum</td>
<td>S-Phase Entry</td>
<td>BT20</td>
<td>Medium exchange</td>
<td>~6 h&lt;sup&gt;a&lt;/sup&gt;</td>
<td>[481]</td>
</tr>
<tr>
<td>Thrombin</td>
<td>S-Phase Entry</td>
<td>CCL34</td>
<td>Rapid Ligand Removal by Competitor</td>
<td>8 h&lt;sup&gt;b&lt;/sup&gt;</td>
<td>[482]</td>
</tr>
<tr>
<td>Forskolin</td>
<td>S-Phase Entry</td>
<td>Thyrocytes</td>
<td>Medium exchange</td>
<td>20 h&lt;sup&gt;b&lt;/sup&gt;</td>
<td>[483]</td>
</tr>
<tr>
<td>PDGF</td>
<td>S-Phase Entry</td>
<td>NIH3T3</td>
<td>Medium exchange (acid wash)</td>
<td>9 h&lt;sup&gt;c&lt;/sup&gt;</td>
<td>[475]</td>
</tr>
<tr>
<td>Ionomycin + PMA</td>
<td>Thymocyte lineage commitment</td>
<td>Thymocytes</td>
<td>Medium exchange</td>
<td>~8 h&lt;sup&gt;a&lt;/sup&gt;</td>
<td>[480]</td>
</tr>
<tr>
<td>FGF deprivation</td>
<td>Differentiation</td>
<td>Myoblasts</td>
<td>Medium exchange</td>
<td>~2.5 h&lt;sup&gt;a&lt;/sup&gt;</td>
<td>[484]</td>
</tr>
<tr>
<td>NGF deprivation</td>
<td>Apoptosis</td>
<td>Sympathetic Neurons</td>
<td>Medium exchange</td>
<td>~24 h&lt;sup&gt;a&lt;/sup&gt;</td>
<td>[31]</td>
</tr>
<tr>
<td>IL-3 deprivation</td>
<td>Apoptosis</td>
<td>myeloid</td>
<td>Medium exchange</td>
<td>~24 h&lt;sup&gt;a&lt;/sup&gt;</td>
<td>[485]</td>
</tr>
<tr>
<td>High Glucose</td>
<td>Late-phase Gene Expression</td>
<td>Min6</td>
<td>Medium exchange</td>
<td>2-3 h&lt;sup&gt;b&lt;/sup&gt;</td>
<td>[58]</td>
</tr>
<tr>
<td>PDGF</td>
<td>c-myc expression</td>
<td>NIH3T3</td>
<td>Medium exchange (acid wash)</td>
<td>30'&lt;sup&gt;b&lt;/sup&gt;</td>
<td>[475]</td>
</tr>
<tr>
<td>TNF-alpha</td>
<td>RANTES expression</td>
<td>Fibroblasts</td>
<td>Medium exchange</td>
<td>~2 h&lt;sup&gt;b&lt;/sup&gt;</td>
<td>[43]</td>
</tr>
</tbody>
</table>

Listed are literature studies where extracellular stimuli were removed (or re-added) at different stimulation times, e.g., by medium exchange. The threshold times indicate the signal duration required to irreversibly commit the cell population to the given phenotypic response (stimulation occurred at t = 0 h). Importantly, the signalling events are primary in the sense that they activated within minutes after stimulation, that is, the threshold time reflects the true signal duration. The data suggests that decision making in mammalian cells will be profoundly affected by slowly acting transcriptional feedback mechanisms. Superscript legend: (a) half of the cells perform cellular response if signal is terminated at the threshold time; (b) no response observed if signal is terminated at the threshold time; (c) half-maximal thymidine incorporation is observed if signal is terminated at the threshold time.
B Additional material on kinetic mechanisms for tumour suppression and oncogene cooperation

B.1 Default model without Ras sequestration

The mathematical model used for the numerical simulations in Figs. 2.1B and 2.2A was previously described by Stites et al. [10]. The model is schematically depicted in Fig. 2.1A and the differential equations are given in Equations B.1.

Table B.1 Kinetic parameters.

<table>
<thead>
<tr>
<th>Parameter [unit]</th>
<th>Description</th>
<th>Intrinsic Model</th>
<th>GEF-GAP model</th>
</tr>
</thead>
<tbody>
<tr>
<td>GTP [mol/l]</td>
<td>Cellular GTP conc.</td>
<td>1.8 · 10^-4</td>
<td>1.8 · 10^-4</td>
</tr>
<tr>
<td>GTP [mol/l]</td>
<td>Cellular GDP conc.</td>
<td>1.8 · 10^-4</td>
<td>1.8 · 10^-4</td>
</tr>
<tr>
<td>Ras [mol/l]</td>
<td>Cellular Ras conc.</td>
<td>4 · 10^-7</td>
<td>4 · 10^-7</td>
</tr>
<tr>
<td>GEF [mol/l]</td>
<td>Basally active GEF conc.</td>
<td>0</td>
<td>10^-10</td>
</tr>
<tr>
<td>GAP [mol/l]</td>
<td>Basally active GAP conc.</td>
<td>0</td>
<td>5 · 10^-9</td>
</tr>
</tbody>
</table>

d(RasGDP)/dt = k_hyd · RasGTP - k_d,GDP · RasGDP + ka,GDP · Ras · GDP - (GEF · (k_cat,GDP / (KM,GDP / D) · RasGDP - k_cat,GTP / (KM,GTP / D) · RasGTP) / (1 + RasGDP / (KM,GDP / D) + RasGTP / (KM,GTP / D))) + (GAP · k_cat,GAP · RasGTP / (KM,GAP / D) + RasGTP)

Listed are the kinetic parameters used for the simulations depicted in Figs. 2.1B and 2.2A (see Equations B.1 for differential equations). The column ‘Intrinsic model’ indicates the parameters used to calculate the grey line in Fig. 2.1B, while the column ‘GEF-GAP model’ lists the parameter set underlying the black lines in Fig. 2.1B and Fig. 2.2A (some parameters were varied as indicated in the figure captions).

In contrast to Stites et al., competitive binding of an effector to RasGTP was neglected, and the concentration of RasGTP was used as the readout. The GTPase reaction intrinsic to Ras was modelled with first-order kinetics. GTP and GDP were assumed to be present in excess, so that the spontaneous nucleotide exchange reactions follow linear kinetics as well. The GEF- and GAP-catalysed reactions were modelled by using the reversible and irreversible Michaelis-Menten equations, respectively.
\[
dd{\text{Ras}}{t} = \text{k}_{\text{RasGDP}} \cdot \text{Ras} \cdot \text{GDP} - \text{k}_{\text{RasGTP}} \cdot \text{Ras} \cdot \text{GTP}
\]
\[
dd{\text{RasGTP}}{t} = - \text{k}_{\text{RasGTP}} \cdot \text{RasGTP} - \text{ka}_{\text{GTP}} \cdot \text{RasGTP} + \text{ka}_{\text{GDP}} \cdot \text{RasGDP} + \text{k}_{\text{GAP}} \cdot \text{RasGTP} - \text{k}_{\text{GTP}} \cdot \text{RasGTP}
\]

Equations B.1 Differential equations of the Stites model.

The grey terms indicate intrinsic Ras GTPase activity and spontaneous nucleotide exchange, while the black terms refer to GEF- and GAP-catalysed reactions. The membrane correction factor, D, takes into account that protein-protein interactions are strongly enhanced if both interaction partners are localised to the membrane, and was set to 250 [10]. The kinetic parameters used for numerical simulations are given in Table B.1.

B.2 Model including Ras sequestration

In Section B.1, the GEF- and GAP-catalysed reactions were modelled using the Michaelis-Menten equation. This implies that the concentrations of the GEF-Ras and GAP-Ras complex are assumed to be much smaller than the concentrations of RasGTP and RasGDP, so that sequestration of Ras in enzyme-substrate complexes is insignificant. In contrast, the numerical simulations in Fig. 2.3 were obtained using an elementary step description of the GEF- and GAP-catalysed reactions. In other words, the enzyme-substrate complexes are included as dynamic variables (Equations B.2), so that Ras sequestration becomes significant provided that the GEF and/or GAP concentrations are sufficiently large. The sequestration model neglects intrinsic GTPase and nucleotide exchange reactions, and also assumes that the GEF irreversibly catalyses exchange of GDP for GTP. The kinetic parameters (given in Table B.2) were chosen similarly to those in Table B.1.

Table B.2 Kinetic parameters.

<table>
<thead>
<tr>
<th>Parameter [unit]</th>
<th>Description</th>
<th>Parameter value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ras [mol/l]</td>
<td>Cellular Ras conc.</td>
<td>4 \cdot 10^{-7}</td>
</tr>
<tr>
<td>GEF [mol/l]</td>
<td>Basally active GEF conc.</td>
<td>3 \cdot 10^{-8} (varied)</td>
</tr>
<tr>
<td>GAP [mol/l]</td>
<td>Basally active GAP conc.</td>
<td>7 \cdot 10^{-7}</td>
</tr>
<tr>
<td>\text{k}_{\text{cat,GEF}} [s^{-1}]</td>
<td>\text{k}_{\text{cat}} for exchange (RasGDP \rightarrow RasGTP)</td>
<td>3.9</td>
</tr>
<tr>
<td>\text{K}_{\text{M,GEF}} [mol/l]</td>
<td>\text{K}_{\text{M}} for exchange (RasGDP \rightarrow RasGTP)</td>
<td>3.86 \cdot 10^{-6}</td>
</tr>
<tr>
<td>\text{k}_{\text{off,GEF}} [s^{-1}]</td>
<td>\text{k}_{\text{off}} for exchange (RasGDP \rightarrow RasGTP)</td>
<td>10</td>
</tr>
<tr>
<td>\text{k}_{\text{cat,GAP}} [s^{-1}]</td>
<td>\text{k}_{\text{cat}} for GAP activity on Ras</td>
<td>5.4</td>
</tr>
<tr>
<td>\text{K}_{\text{M,GAP}} [mol/l]</td>
<td>\text{K}_{\text{M}} for GAP activity on Ras</td>
<td>2.3 \cdot 10^{-7}</td>
</tr>
<tr>
<td>\text{k}_{\text{off,GAP}} [s^{-1}]</td>
<td>\text{k}_{\text{off}} for GAP activity on Ras</td>
<td>10</td>
</tr>
</tbody>
</table>

Listed are the kinetic parameters used for the simulations depicted in Figs. 2.3A (see Equations B.2 for differential equations). Most parameters were taken from Table B.1 and the off-rates of enzyme-substrate complexes were estimated to be 10 s^{-1}. 
\[
\begin{align*}
\frac{d(RasGDP)}{dt} &= v_{cat,GAP} - v_{on,GEF} + v_{off,GEF} \\
\frac{d(GEF)}{dt} &= -v_{on,GEF} + v_{off,GEF} + v_{cat,GEF} \\
\frac{d(RasGDP\cdot GEF)}{dt} &= v_{on,GEF} - v_{off,GEF} - v_{cat,GEF} \\
\frac{d(RasGTP)}{dt} &= v_{cat,GEF} - v_{on,GAP} + v_{off,GAP} \\
\frac{d(GAP)}{dt} &= -v_{on,GAP} + v_{off,GAP} + v_{cat,GAP} \\
\frac{d(RasGTP\cdot GAP)}{dt} &= v_{on,GAP} - v_{off,GAP} - v_{cat,GAP} \\
\end{align*}
\]

\[
\begin{align*}
v_{on,GEF} &= \frac{(k_{off,GEF} + k_{cat,GEF})}{(KM,GEF / D)} \cdot RasGDP \cdot GEF \\
v_{off,GEF} &= k_{off,GEF} \cdot RasGDP\cdot GEF \\
v_{cat,GEF} &= k_{cat,GEF} \cdot RasGDP\cdot GEF \\
v_{on,GAP} &= \frac{(k_{off,GAP} + k_{cat,GAP})}{(KM,GAP / D)} \cdot RasGTP \cdot GAP \\
v_{off,GAP} &= k_{off,GAP} \cdot RasGTP\cdot GAP \\
v_{cat,GAP} &= k_{cat,GAP} \cdot RasGTP\cdot GAP \\
\end{align*}
\]

Equations B.2 Differential equations of the sequestration model.

The enzyme-substrate complexes (RasGDP→GEF and RasGTP→GAP) are included as dynamic variables, so that Ras sequestration becomes significant provided that the GEF and/or GAP concentrations are sufficiently large. The intrinsic GTPase and nucleotide exchange reactions are neglected, and the GEF is assumed to irreversibly catalyse the exchange of GDP for GTP. The membrane correction factor D was again chosen to be 250 (see Section B.1).

B.3 Michaelis-Menten constants for mammalian Ras-GEFs

The tumour suppression effect described in the main text requires that the Michaelis-Menten constant of the GEF-catalysed reaction (\(K_{M,GEF}\)) is close to or even below the intracellular Ras concentration. The \(K_{M,GEF}\) in the default model was measured for Cdc25Mm on unprocessed Ras [98]. However, it is well established that post-translational Ras farnesylation drastically enhances the Ras-GEF interaction [99,100,101]. Moreover, it is known that Michaelis-Menten constants of other GEFs than Cdc25Mm are lower, especially if Ras is farnesylated [99,101,102,103]. Kinetic measurements taken from the literature (Table B.3) indicate that the \(K_{M,GEF}\) will often be two or more orders of magnitude lower than that used in the default Stites model, so that the \(K_{M,GEF}\) will be close to or even below the intracellular Ras concentration.

Table B.3 Experimentally measured \(K_{M}\)-values of GEF-catalysed nucleotide exchange.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Substrate</th>
<th>(K_{M})</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>mouse CDC25Mm</td>
<td>H-Ras (unprocessed)</td>
<td>&gt; 11 µM</td>
<td>[486]</td>
</tr>
<tr>
<td>mouse CDC25Mm</td>
<td>H-Ras (unprocessed)</td>
<td>386 µM</td>
<td>[98]</td>
</tr>
<tr>
<td>mouse CDC25Mm</td>
<td>Ras2p-[GDP] (yeast, unprocessed)</td>
<td>70 nM</td>
<td>[102]</td>
</tr>
<tr>
<td>hSos1</td>
<td>K-Ras-[GDP] (prenylated)</td>
<td>225 nM</td>
<td>[101]</td>
</tr>
<tr>
<td>hSos1</td>
<td>H-Ras-[GTP] (unprocessed)</td>
<td>2.2 µM</td>
<td>[103]</td>
</tr>
<tr>
<td>Ras GRF</td>
<td>K-Ras-[GDP] (processed)</td>
<td>680 nM</td>
<td>[99]</td>
</tr>
<tr>
<td>smg GDS</td>
<td>K-Ras-[GDP] (processed)</td>
<td>220 nM</td>
<td>[99]</td>
</tr>
</tbody>
</table>
C Additional material on competing docking interactions can bring about bistability in the MAPK cascade

C.1 Sequestration Model Used for Calculation of Figs. 3.2 and 3.3

The sequestration model which is depicted in Fig. 3.1D includes the previously described standard MAPK model ('basic model') [150] , and additionally takes Mek sequestration by un-/monophosphorylated Erk into account. The impact of such Mek sequestration was investigated by altering the sequestration factor S, which enters the association rates of the Mek sequestration complexes (Mek–Erk, pMek–pErk, pMek–pErk and pMek–ppErk). The sequestration factor was set to S = 0 in Fig. 3.2A (black line) in order to simulate the behaviour of the basic model devoid of Mek sequestration. S was set to unity in Figs. 3.2A (grey line), 3.2B and 3.3 in order to Mek sequestration into account.

i) Differential Equations


\[ pMek' = kcat_{Raf\_Mek} \cdot pRaf\_Mek - kon_{Raf\_Mek} \cdot pMek \cdot (pRaf_{tot}-pRaf\_Mek-pRaf\_pMek) + koff_{Raf\_Mek} \cdot pRaf\_pMek + (Vmax_{PPase\_Mek} \cdot ppMek/(KM_{PPase\_Mek}+ppMek)) - (Vmax_{PPase\_Mek} \cdot ppMek/(KM_{PPase\_Mek}+ppMek)) \cdot kon_{Mek\_Erk} \cdot pMek' \cdot Erk < pMek\_Erk-ppMek\_Erk-ppMek\_pErk-ppMek\_pErk-ppMek\_pErk) + koff_{Mek\_Erk} \cdot pMek\_pErk-ppMek\_pErk > pMek\_pErk \]

\[ pRaf\_pMek' = kon_{Raf\_Mek} \cdot pMek \cdot (pRaf_{tot}-pRaf\_Mek-pRaf\_pMek) - koff_{Raf\_Mek} \cdot pRaf\_pMek - kcat_{Raf\_Mek} \cdot pRaf\_pMek \]

\[ ppMek' = kcat_{Raf\_Mek} \cdot pRaf\_pMek - (Vmax_{PPase\_Mek} \cdot ppMek/(KM_{PPase\_Mek}+ppMek)) - kon_{Mek\_Erk} \cdot ppMek \cdot (pRaf_{tot}-pRaf\_Mek-pRaf\_pMek) + koff_{Mek\_Erk} \cdot ppMek\_pErk + kcat_{Mek\_Erk} \cdot ppMek\_pErk \]

\[ ppMek\_pErk' = kon_{Mek\_Erk} \cdot ppMek\_pErk - koff_{Mek\_Erk} \cdot ppMek\_pErk - kcat_{Mek\_Erk} \cdot ppMek\_pErk \]

\[ ppErk' = kcat_{Mek\_Erk} \cdot ppMek\_pErk - (Vmax_{PPase\_Erk} \cdot ppErk/(KM_{PPase\_Erk}+ppErk)) - kon_{Mek\_Erk} \cdot ppMek\_pErk \cdot (pRaf_{tot}-pRaf\_Mek-pRaf\_pMek) + koff_{Mek\_Erk} \cdot ppMek\_pErk \]

\[ ppMek\_pErk' = kon_{Mek\_Erk} \cdot ppMek\_pErk - koff_{Mek\_Erk} \cdot ppMek\_pErk - kcat_{Mek\_Erk} \cdot ppMek\_pErk \]

\[ ppErk' = kcat_{Mek\_Erk} \cdot ppMek\_pErk - (Vmax_{PPase\_Erk} \cdot ppErk/(KM_{PPase\_Erk}+ppErk)) \]

\[ Mek\_Erk' = kon_{Mek\_Erk} \cdot S \cdot (Mektot-pRaf\_Mek-pRaf\_pMek-ppMek\_Mek\_Erk-ppMek\_Erk-ppMek\_pErk-ppMek\_pErk) \cdot (pRaf_{tot}-pRaf\_Mek-pRaf\_pMek) \cdot \text{Mektot} \]

\[ Mek\_pErk' = kon_{Mek\_Erk} \cdot S \cdot (Mektot-pRaf\_Mek-pRaf\_pMek-ppMek\_Mek\_Erk-ppMek\_Erk-ppMek\_pErk-ppMek\_pErk) \cdot (pRaf_{tot}-pRaf\_Mek-pRaf\_pMek) \cdot \text{Mektot} \]

\[ pMek\_pErk' = kon_{Mek\_Erk} \cdot S \cdot (Mektot-pRaf\_Mek-pRaf\_pMek-ppMek\_Mek\_Erk-ppMek\_Erk-ppMek\_pErk-ppMek\_pErk) \cdot (pRaf_{tot}-pRaf\_Mek-pRaf\_pMek) \cdot \text{Mektot} \]

\[ pMek\_pErk' = kon_{Mek\_Erk} \cdot S \cdot (Mektot-pRaf\_Mek-pRaf\_pMek-ppMek\_Mek\_Erk-ppMek\_Erk-ppMek\_pErk-ppMek\_pErk) \cdot (pRaf_{tot}-pRaf\_Mek-pRaf\_pMek) \cdot \text{Mektot} \]

ii) Mass Conservation Relationships

\[ pRaf = (pRaf_{tot}-pRaf\_Mek-pRaf\_pMek) \]


iii) Kinetic Parameters / Initial Concentrations
pRaftot: varied for calculation of stimulus-response
Mektot = 1 (in Fig. 3.2 and 3.3B) or Mektot varied in Fig. 3.3A
Erktot = 1 (in Fig. 3.2 and 3.3B) or Erktot varied in Fig. 3.3A
Vmax_PPase_Mek = 0.001 (in Fig. 3.2 and 3.3A) or Vmax_PPase_Mek varied in Fig. 3.3B
Vmax_PPase_Erk = 0.04 (in Fig. 3.2 and 3.3A) or Vmax_PPase_Erk varied in Fig. 3.3B
S = 0 (Fig. 3.2A, black line) or S = 1 (Figs. 3.2A (grey line), 3.2B and 3.3)

\[
\begin{align*}
\text{kon}_\text{Raf}_\text{Mek} &= 0.65; \\
\text{koff}_\text{Raf}_\text{Mek} &= 0.065; \\
\text{kcat}_\text{Raf}_\text{Mek} &= 0.18; \\
\text{KM}_\text{PPase}_\text{Mek} &= 0.1; \\
\text{kon}_\text{Mek}_\text{Erk} &= 0.88; \\
\text{koff}_\text{Mek}_\text{Erk} &= 0.088; \\
\text{kcat}_\text{Mek}_\text{Erk} &= 0.22; \\
\text{KM}_\text{PPase}_\text{Erk} &= 0.5; \\
\end{align*}
\]

C.2 Extended Sequestration Model Used for Calculation of Fig. 3.4

Figure 3.4 shows that significant competition between Raf and Erk for binding to Mek is required for bistability to arise from the proposed mechanism. Additionally, bistability requires that bisphosphorylated Erk hardly binds to Erk, i.e., that product inhibition does not occur in Mek-mediated Erk phosphorylation.

Non-competitive binding of Raf and Erk to Mek was simulated by considering ternary complexes between Raf, Mek and Erk, and Raf-mediated catalysis on these ternary complexes (see Fig. C.1). More specifically, pRaf–Mek complexes with Erk (Fig. C.1A), with pErk (Fig. C.1B) and with ppErk (Fig. C.1C) were taken into account. The association rate constants of these ternary complexes equal those of the binary Raf–Mek complexes (kon,Raf–Mek) of the Mek–Erk complexes (kon,Mek–Erk) multiplied by the competition-factor c. The degree of competition was varied by changing c. A value of c = 0 means pure competitive binding, while a value of c = 1 implies pure non-competitive binding. The dissociation rate constants and the catalytic rates of the ternary complexes equal those of the binary complexes (koff,Raf–Mek, koff,Mek–Erk and kcat,Raf–Mek).

![Figure C.1](image.png)

Product inhibition in Mek-mediated Erk phosphorylation was simulated by considering complexes of ppErk with Mek, pMek, and ppMek (see Fig. C.1C). The association rates of these additional complexes equal those of the other Mek–Erk complexes (kon,Mek–Erk) multiplied by the product-inhibition-factor p. The degree of product inhibition was varied by changing p. A value of p = 0 means no product inhibition, while a value of p = 1 implies strong product inhibition. The dissociation rate constants of the product inhibition complexes equal those of the other Mek–Erk complexes (koff,Mek–Erk).
i) Differential Equations

\[
pMek' = \text{kon}_\text{Mek} \cdot \text{ppMek} \cdot \text{pErk} - \text{koff}_\text{Mek} \cdot \text{ppMek} \cdot \text{pErk} - \text{koff}_\text{Mek} \cdot \text{ppMek} \cdot \text{pErk} + \text{ppMek} \cdot \text{pErk} \cdot \text{ppMek} \cdot \text{pErk} - \text{koff}_\text{Mek} \cdot \text{ppMek} \cdot \text{pErk} + \text{koff}_\text{Mek} \cdot \text{ppMek} \cdot \text{pErk} - \text{koff}_\text{Mek} \cdot \text{ppMek} \cdot \text{pErk} + \text{koff}_\text{Mek} \cdot \text{ppMek} \cdot \text{pErk} - \text{koff}_\text{Mek} \cdot \text{ppMek} \cdot \text{pErk}
\]

\[
pRaf \cdot \text{ppMek} \cdot \text{pErk} = \text{kon}_\text{Raf} \cdot \text{ppMek} \cdot \text{pErk} - \text{koff}_\text{Raf} \cdot \text{ppMek} \cdot \text{pErk} + \text{ppMek} \cdot \text{pErk} \cdot \text{ppMek} \cdot \text{pErk} - \text{ppMek} \cdot \text{pErk} \cdot \text{ppMek} \cdot \text{pErk} - \text{ppMek} \cdot \text{pErk} \cdot \text{ppMek} \cdot \text{pErk} - \text{ppMek} \cdot \text{pErk} \cdot \text{ppMek} \cdot \text{pErk} - \text{ppMek} \cdot \text{pErk} \cdot \text{ppMek} \cdot \text{pErk}
\]

\[
pRaf \cdot \text{ppMek} \cdot \text{pErk} = \text{kon}_\text{Raf} \cdot \text{ppMek} \cdot \text{pErk} - \text{koff}_\text{Raf} \cdot \text{ppMek} \cdot \text{pErk} - \text{koff}_\text{Raf} \cdot \text{ppMek} \cdot \text{pErk} - \text{ppMek} \cdot \text{pErk} \cdot \text{ppMek} \cdot \text{pErk} - \text{ppMek} \cdot \text{pErk} \cdot \text{ppMek} \cdot \text{pErk} - \text{ppMek} \cdot \text{pErk} \cdot \text{ppMek} \cdot \text{pErk} - \text{ppMek} \cdot \text{pErk} \cdot \text{ppMek} \cdot \text{pErk}
\]

\[
pMek' = \text{kon}_\text{Mek} \cdot \text{ppMek} \cdot \text{pErk} - \text{koff}_\text{Mek} \cdot \text{ppMek} \cdot \text{pErk} + \text{koff}_\text{Mek} \cdot \text{ppMek} \cdot \text{pErk} - \text{koff}_\text{Mek} \cdot \text{ppMek} \cdot \text{pErk} + \text{koff}_\text{Mek} \cdot \text{ppMek} \cdot \text{pErk} - \text{koff}_\text{Mek} \cdot \text{ppMek} \cdot \text{pErk} - \text{koff}_\text{Mek} \cdot \text{ppMek} \cdot \text{pErk} - \text{koff}_\text{Mek} \cdot \text{ppMek} \cdot \text{pErk}
\]

\[
pRaf \cdot \text{ppMek} \cdot \text{pErk} = \text{kon}_\text{Raf} \cdot \text{ppMek} \cdot \text{pErk} - \text{koff}_\text{Raf} \cdot \text{ppMek} \cdot \text{pErk} - \text{ppMek} \cdot \text{pErk} \cdot \text{ppMek} \cdot \text{pErk} - \text{ppMek} \cdot \text{pErk} \cdot \text{ppMek} \cdot \text{pErk} - \text{ppMek} \cdot \text{pErk} \cdot \text{ppMek} \cdot \text{pErk} - \text{ppMek} \cdot \text{pErk} \cdot \text{ppMek} \cdot \text{pErk}
\]

\[
pMek' = \text{kon}_\text{Mek} \cdot \text{ppMek} \cdot \text{pErk} - \text{koff}_\text{Mek} \cdot \text{ppMek} \cdot \text{pErk} + \text{koff}_\text{Mek} \cdot \text{ppMek} \cdot \text{pErk}
\]
\[ p_{\text{Mek}_p\text{Erk}} = \frac{p_{\text{Mek}} p_{\text{Erk}} k_{\text{off}}}{p_{\text{Mek}} k_{\text{cat}} p_{\text{Raf}} M_{\text{Mek}_p\text{Erk}}} \]
\[ p_{\text{Raf}_{\text{Mek}_p\text{Erk}}} = \frac{k_{\text{off}}}{k_{\text{cat}} M_{\text{Raf}_p\text{Mek}_p\text{Erk}}} \]
\[ p_{\text{Mek}_p\text{Erk}} = \frac{p_{\text{Mek}} p_{\text{Erk}} k_{\text{off}}}{p_{\text{Mek}} k_{\text{cat}} p_{\text{Raf}} M_{\text{Mek}_p\text{Erk}}} \]
\[ p_{\text{Raf}_{\text{Mek}_p\text{Erk}}} = \frac{k_{\text{off}}}{k_{\text{cat}} M_{\text{Raf}_p\text{Mek}_p\text{Erk}}} \]
\[ p_{\text{Mek}_p\text{Erk}} = \frac{p_{\text{Mek}} p_{\text{Erk}} k_{\text{off}}}{p_{\text{Mek}} k_{\text{cat}} p_{\text{Raf}} M_{\text{Mek}_p\text{Erk}}} \]
\[ p_{\text{Raf}_{\text{Mek}_p\text{Erk}}} = \frac{k_{\text{off}}}{k_{\text{cat}} M_{\text{Raf}_p\text{Mek}_p\text{Erk}}} \]
\[ p_{\text{Mek}_p\text{Erk}} = \frac{p_{\text{Mek}} p_{\text{Erk}} k_{\text{off}}}{p_{\text{Mek}} k_{\text{cat}} p_{\text{Raf}} M_{\text{Mek}_p\text{Erk}}} \]
\[ p_{\text{Raf}_{\text{Mek}_p\text{Erk}}} = \frac{k_{\text{off}}}{k_{\text{cat}} M_{\text{Raf}_p\text{Mek}_p\text{Erk}}} \]
\[ p_{\text{Mek}_p\text{Erk}} = \frac{p_{\text{Mek}} p_{\text{Erk}} k_{\text{off}}}{p_{\text{Mek}} k_{\text{cat}} p_{\text{Raf}} M_{\text{Mek}_p\text{Erk}}} \]
\[ p_{\text{Raf}_{\text{Mek}_p\text{Erk}}} = \frac{k_{\text{off}}}{k_{\text{cat}} M_{\text{Raf}_p\text{Mek}_p\text{Erk}}} \]

**ii) Mass Conservation Relationships**

\[ p_{\text{Raf}} = \frac{p_{\text{Raf}} p_{\text{Mek}} p_{\text{Erk}}}{p_{\text{Raf}} p_{\text{Mek}} p_{\text{Erk}} p_{\text{Mek}_p\text{Erk}}} \]
\[ p_{\text{Mek}} p_{\text{Erk}} = \frac{p_{\text{Mek}} p_{\text{Erk}} k_{\text{off}}}{p_{\text{Mek}} k_{\text{cat}} p_{\text{Raf}} M_{\text{Mek}_p\text{Erk}}} \]
\[ p_{\text{Raf}_{\text{Mek}_p\text{Erk}}} = \frac{k_{\text{off}}}{k_{\text{cat}} M_{\text{Raf}_p\text{Mek}_p\text{Erk}}} \]

**iii) Kinetic Parameters / Initial Concentrations**

C = (Competition factor) varied
P = (Product inhibition factor) varied
S = 1 = (Sequestration factor)
kon_Raf_Mek = 0.65
koff_Raf_Mek = 0.065
kcat_Raf_Mek = 0.18
V_{max,PPase_{Mek}} = 0.001
K_{M,PPase_{Mek}} = 0.1
kon_Mek_Erk = 0.66
koff_Mek_Erk = 0.066
kcat_Mek_Erk = 0.22
V_{max,PPase_{Erk}} = 0.04
K_{M,PPase_{Erk}} = 0.5
C.3 Model Used for Calculation of Fig. 3.5

Figure 3.5 shows that the bistability mechanism proposed in Section 3.3 synergises with that proposed by Markevich et al. [23], which arises from enzyme depletion effects in Erk double phosphorylation. The models used for calculation of Fig. 3.5 exhibit the topology depicted in Fig. 3.1D.

The models used for the calculation of curves 1 and 2 in Fig. 3.5 exhibit positive cooperativity in Mek-mediated Erk phosphorylation, so that the (isolated) Erk cycle exhibits weak bistability [153]. More specifically, it was assumed that the first and the second phosphorylation steps of ppMek-mediated Erk phosphorylation proceed with different kinetic constants. Relatively high affinity (i.e., a low $K_{M,1}$), but slow catalysis (i.e., a low $k_{cat,1}$) for the first phosphorylation step. By contrast, low affinity (i.e., a high $K_{M,2}$), but very fast catalysis (i.e., a high $k_{cat,2}$) was assumed for the second phosphorylation step. No phosphorylation-site-specific kinetic differences in phosphatase-mediated Erk dephosphorylation were assumed (see maximal velocities, $V_{m,pp1} = V_{m,pp2}$, and Michaelis-Menten constants, $K_{M,pp1} = K_{M,pp2}$ below), but dephosphorylation was modelled by a more complex Michaelis-Menten Mechanism, which takes enzyme competition effects into account [23].

Curve 1 in Fig. 3.5 shows that relatively narrow range of bistability arises from the Markevich mechanism alone. In these simulations, the feedback mechanism proposed in Section 3.3 (i.e., Mek sequestration in Raf-inaccessible complexes) was switched off by abolishing the formation unproductive sequestration complexes (Mek~Erk, pMek~Erk, Mek~pErk and pMek~pErk). This was accomplished by setting $k_{on,Mek-Erk} = 0$.

Curve 2 in Fig. 3.5 demonstrates that the feedback mechanism proposed in Section 3.3 brings about a broad range of bistability when combined with that proposed by Markevich et al. [23]. These simulations correspond to a system, which includes Mek sequestration into Raf-inaccessible complexes (black and grey arrows in Fig. 3.1D). The association rate constant of these complexes was set to the measured value $k_{on,Mek-Erk} = 0.88$ (Table 3.1). This value implies that the affinity of the ppMek~Erk-complex, which is the enzymatic intermediate on the first Erk phosphorylation site, equals that of the unproductive sequestration complexes (Mek~Erk, pMek~Erk, Mek~pErk and pMek~pErk).

Curve 3 in Fig. 3.5 shows that a relatively narrow range of bistability arises from the feedback mechanism proposed in Section 3. These simulations correspond to a sequestration model (black and grey arrows in Fig. 3.1D), where positive cooperativity and enzyme depletion effects were eliminated from the Erk cycle. This was done by assuming the same catalytic rate constant for the first and the second steps of Mek-mediated Erk phosphorylation ($k_{cat,1} = k_{cat,2} = 1$).

\[ \text{i) Differential Equations} \]

\[
p_{Raf,Mek}' = k_{on,Raf,Mek}*(Mektot-p_{Raf,Mek}-p_{Mek}-p_{Praf,Mek}-p_{Mek-Mek,Erk}-p_{Mek,Erk}-pp_{Mek,Erk}-pp_{Mek,pErk}-Mek_pErk-p_{Praf,Mek})-k_{off,Raf,Mek} * p_{Raf,Mek} - k_{cat,Raf,Mek} * p_{Raf,Mek}
\]

\[
p'_{Mek} = k_{cat,Raf,Mek} * p_{Raf,Mek} - k_{on,Mek} * p_{Mek} - k_{off,Mek} * p_{Mek} + (V_{max,PPase} * pp_{Mek}/(K_{M,PPase,ppMek} + pp_{Mek})) - (V_{max,PPase} * pp_{Mek}/(K_{M,PPase,ppMek} + pp_{Mek}))*kon_{Mek} * p_{Mek}*(Erk_{tot} - p_{Mek,Erk} - pp_{Mek,Erk} - pp_{Mek,pErk} - Mek_pErk - p_{Mek,pErk})
\]

\[
+k_{off,Mek} * p_{Mek,Erk} + k_{on,Mek} * p_{Mek,Erk} * p_{Erk} + k_{off,Mek} * p_{Mek,pErk}
\]

\[
p_{Mek,pErk}' = k_{on,Mek} * p_{Mek} * (p_{Raf,Mek} - p_{Mek} - p_{Praf,Mek} - p_{Mek-Mek,Erk}) - k_{off,Mek} * p_{Mek,pErk} - k_{cat,Mek} * p_{Mek,pErk}
\]

\[
p_{Erk}' = k_{cat,Mek} * p_{Mek,pErk} - (V_{max,PP1} * pp_{Erk}/(1 + pp_{Erk}/K_{M,PP1})) - (V_{max,PP2} * pp_{Erk}/(1 + pp_{Erk}/K_{M,PP2})) - (k_{off,Mek} * p_{Mek} - k_{on,Mek} * p_{Mek} - k_{off,Mek} * p_{Mek,pErk} - k_{on,Mek} * p_{Mek,pErk}) + k_{cat,Mek} * p_{Mek,pErk}
\]

\[
Mek_Erk' = k_{on,Mek} * p_{Mek}*(Mektot-p_{Raf,Mek}-p_{Mek}-p_{Praf,Mek}-p_{Mek-Mek,Erk}-p_{Mek,Erk}-pp_{Mek,Erk}-pp_{Mek,pErk}-Mek_pErk-p_{Praf,Mek})-k_{off,Mek} * p_{Mek,Erk} - k_{cat,Mek} * p_{Mek,Erk}
\]

\[
p_{Erk}' = k_{cat,Mek} * p_{Mek,pErk} - (V_{max,PP1} * pp_{Erk}/(1 + pp_{Erk}/K_{M,PP1})) - (V_{max,PP2} * pp_{Erk}/(1 + pp_{Erk}/K_{M,PP2})) - (k_{off,Mek} * p_{Mek} - k_{on,Mek} * p_{Mek} - k_{off,Mek} * p_{Mek,pErk} - k_{on,Mek} * p_{Mek,pErk}) + k_{cat,Mek} * p_{Mek,pErk}
\]

\[
p_{Mek,pErk}' = k_{on,Mek} * p_{Mek} * (p_{Raf,Mek} - p_{Mek} - p_{Praf,Mek} - p_{Mek-Mek,Erk}) - k_{off,Mek} * p_{Mek,pErk} - k_{cat,Mek} * p_{Mek,pErk}
\]

\[
p_{Erk}' = k_{cat,Mek} * p_{Mek,pErk} - (V_{max,PP1} * pp_{Erk}/(1 + pp_{Erk}/K_{M,PP1})) - (V_{max,PP2} * pp_{Erk}/(1 + pp_{Erk}/K_{M,PP2})) - (k_{off,Mek} * p_{Mek} - k_{on,Mek} * p_{Mek} - k_{off,Mek} * p_{Mek,pErk} - k_{on,Mek} * p_{Mek,pErk}) + k_{cat,Mek} * p_{Mek,pErk}
\]

\[
p_{Mek,pErk}' = k_{on,Mek} * p_{Mek} * (p_{Raf,Mek} - p_{Mek} - p_{Praf,Mek} - p_{Mek-Mek,Erk}) - k_{off,Mek} * p_{Mek,pErk} - k_{cat,Mek} * p_{Mek,pErk}
\]

\[
p_{Erk}' = k_{cat,Mek} * p_{Mek,pErk} - (V_{max,PP1} * pp_{Erk}/(1 + pp_{Erk}/K_{M,PP1})) - (V_{max,PP2} * pp_{Erk}/(1 + pp_{Erk}/K_{M,PP2})) - (k_{off,Mek} * p_{Mek} - k_{on,Mek} * p_{Mek} - k_{off,Mek} * p_{Mek,pErk} - k_{on,Mek} * p_{Mek,pErk}) + k_{cat,Mek} * p_{Mek,pErk}
\]

\[
p_{Mek,pErk}' = k_{on,Mek} * p_{Mek} * (p_{Raf,Mek} - p_{Mek} - p_{Praf,Mek} - p_{Mek-Mek,Erk}) - k_{off,Mek} * p_{Mek,pErk} - k_{cat,Mek} * p_{Mek,pErk}
\]
\[ p_{\text{Mek,Erk}} = k_{\text{on, Mek,Erk}} \cdot p_{\text{Mek}} \cdot (Erk_{\text{tot}} - p_{\text{Mek,Erk}} - p_{\text{Mek,Erk}} - pp_{\text{Mek,Erk}} - k_{\text{off, Mek,Erk}} \cdot p_{\text{Mek,Erk}}) \]

\[ Mek_{\text{p,Erk}} = k_{\text{on, Mek,Erk}} \cdot Mek_{\text{tot}} - p_{\text{Raf, Mek}} - p_{\text{Mek}} - p_{\text{Raf, pMek}} - p_{\text{Mek,Erk}} - pp_{\text{Mek}} - k_{\text{off, Mek,Erk}} \cdot Mek_{\text{p,Erk}} \]

\[ p_{\text{Mek, pErk}} = k_{\text{on, Mek,Erk}} \cdot p_{\text{Mek}} \cdot p_{\text{Erk}} - k_{\text{off, Mek,Erk}} \cdot p_{\text{Mek, pErk}} \]

**ii) Mass Conservation Relationships**

\[ p_{\text{Raf}} = (p_{\text{Raf, tot}} - p_{\text{Raf, Mek}} - p_{\text{Raf, pMek}}) \]

\[ Mek = (Mek_{\text{tot}} - p_{\text{Raf, Mek}} - p_{\text{Mek}} - p_{\text{Raf, pMek}} - pp_{\text{Mek}} - Mek_{\text{Erk}} - pp_{\text{Mek,Erk}} - pp_{\text{Mek, pErk}} - Mek_{\text{p,Erk}} - p_{\text{Mek, pErk}}) \]

\[ Erk = (Erk_{\text{tot}} - Mek_{\text{Erk}} - pp_{\text{Mek,Erk}} - pp_{\text{Mek, pErk}} - pp_{\text{Erk}} - Mek_{\text{p,Erk}} - p_{\text{Mek, pErk}}) \]

**iii) Kinetic Parameters / Initial Concentrations**

\[ p_{\text{Raf, tot}}: \text{varied for calculation of stimulus-response} \]

\[ Mek_{\text{tot}} = 1 \]

\[ Erk_{\text{tot}} = 10 \]

\[ k_{\text{on, Raf, Mek}} = 0.65 \]

\[ k_{\text{off, Raf, Mek}} = 0.065 \]

\[ k_{\text{cat, Raf, Mek}} = 0.18 \]

\[ V_{\text{max, PPase, Mek}} = 0.001 \]

\[ K_{\text{M, PPase, Mek}} = 0.01 \]

\[ k_{\text{on, Mek, Erk}} = 0.88 \text{ (curves 2 and 3) or } k_{\text{on, Mek, Erk}} = 0 \text{ (curve 1)} \]

\[ k_{\text{off, Mek, Erk}} = 0.1 \]

\[ k_{\text{cat, M1}} = 0.2 \text{ (curves 1 and 2) or } k_{\text{cat, M1}} = 1 \text{ (curve 3)} \]

\[ K_{\text{MM1}} = 0.35 \]

\[ k_{\text{cat, M2}} = 10 \text{ (curves 1 and 2) or } k_{\text{cat, M2}} = 1 \text{ (curve 3)} \]

\[ K_{\text{MM2}} = 1 \]

\[ V_{\text{mPP1}} = 0.4 \]

\[ V_{\text{mPP2}} = 0.4 \]

\[ K_{mPP1} = 0.5 \]

\[ K_{mPP2} = 0.5 \]
D Additional material on simultaneous inhibition of initiator and effector caspases by XIAP establishes implicit bistability in caspase activation

D.1 Differential equations for the wild-type and non-competitive models

In the following, the differential equations of the wildtype model are given, and it is explained how the wildtype model was extended in Fig. 4.4H and Fig. 4.5 in order to take non-competitive binding of Casp3 and Casp9 to XIAP into account ("Non-competitive model").

Wildtype Model: The velocities numbered according to Fig. 4.1B and the differential equations of the model are:

\[
\begin{align*}
  v_1 & = k_1 \cdot A^* \cdot C9 - k_{-1} \cdot A^* C9 \\
  v_2 & = k_2 \cdot C3 \cdot C9 \\
  v_3 & = k_3 \cdot C3 \cdot A^* C9 \\
  v_4 & = k_4 \cdot C9 \cdot C3^* \\
  v_5 & = k_5 \cdot A^* C9 \cdot C3^* \\
  v_6 & = k_6 \cdot C3 \cdot C9^* \\
  v_7 & = k_7 \cdot C3 \cdot A^* C9^* \\
  v_8 & = k_8 \cdot C9^* \cdot A^* - k_{-8} \cdot A^* C9^* \\
  v_9 & = k_9 \cdot C9 \cdot X - k_{-9} \cdot C9 X \\
  v_{10} & = k_{10} \cdot A^* C9 \cdot X - k_{-10} \cdot A^* C9 X \\
  v_{11} & = k_{11} \cdot C9^* \cdot X - k_{-11} \cdot C9^* X \\
  v_{12} & = k_{12} \cdot A^* C9^* \cdot X - k_{-12} \cdot A^* C9^* X \\
  v_{13} & = k_{13} \cdot C9 X \cdot A^* - k_{-13} \cdot A^* C9 X \\
  v_{14} & = k_{14} \cdot C9 X \cdot A^* - k_{-14} \cdot A^* C9^* X \\
  v_{15} & = k_{15} \cdot C3^* \cdot X - k_{-15} \cdot C3^* X \\
  v_{16} & = k_{-13} - k_{16} \cdot A^* \\
  v_{17} & = k_{-17} - k_{17} \cdot C9 \\
  v_{18} & = k_{-18} - k_{18} \cdot X \\
  v_{19} & = k_{19} \cdot C9 X \\
  v_{20} & = k_{20} \cdot A^* C9 X \\
  v_{21} & = k_{21} \cdot A^* C9 \\
  v_{22} & = k_{22} - k_{22} \cdot C3 \\
  v_{23} & = k_{23} \cdot C3^* \\
  v_{24} & = k_{24} \cdot C3^* X \\
  v_{25} & = k_{25} \cdot C9^* X \\
  v_{26} & = k_{26} \cdot C9^* \\
  v_{27} & = k_{27} \cdot A^* C9^* \\
  v_{28} & = k_{28} \cdot A^* C9^* X
\end{align*}
\]
Non-competitive model: XIAP-mediated feedback requires that XIAP binds caspases competitively at least to some extent, since otherwise Casp3 cannot sequester XIAP away from Casp9 (see main text). In order to analyse the role of competitive binding in more detail, the wildtype model (see Fig. 4.1B) was extended according to Fig. D.1. Casp3, Casp9 and XIAP form a ternary complex in this ‘non-competitive model’ either by binding of Casp3 to Casp9-associated XIAP (reactions 33 – 36) or by binding of Casp9 to Casp3-associated XIAP (reactions 37-40). While some differential equations of the wildtype model (see above) were retained in the non-competitive model, those describing the species marked in red in Fig. D.1 were altered.

Non-competitive model: XIAP-mediated feedback requires that XIAP binds caspases competitively at least to some extent, since otherwise Casp3 cannot sequester XIAP away from Casp9 (see main text). In order to analyse the role of competitive binding in more detail, the wildtype model (see Fig. 4.1B) was extended according to Fig. D.1. Casp3, Casp9 and XIAP form a ternary complex in this 'non-competitive model' either by binding of Casp3 to Casp9-associated XIAP (reactions 33 – 36) or by binding of Casp9 to Casp3-associated XIAP (reactions 37-40). While some differential equations of the wildtype model (see above) were retained in the non-competitive model, those describing the species marked in red in Fig. D.1 were altered.

The wildtype model (Fig. 4.1B) was extended by the black reactions (29 – 42). Accordingly, the differential equations of the red intermediates were altered in comparison to the wildtype model (see text), and those of the black intermediates were newly added in the non-competitive model.
The full model reads:

\[
\begin{align*}
\frac{dA^*}{dt} &= -v_4 - v_{15} - v_{13} - v_{14} + v_{16} - v_{17} - v_{18} \\
\frac{dC0}{dt} &= -v_9 - v_{13} - v_{19} - v_{20} \\
\frac{dC0X}{dt} &= v_9 - v_{13} - v_{19} - v_{20} \\
\frac{dX}{dt} &= -v_9 - v_{13} - v_{15} - v_{11} - v_{12} + v_{18} \\
\frac{dA^*C0X}{dt} &= v_{19} + v_{13} - v_{26} - v_{34} \\
\frac{dA^*C0}{dt} &= v_{14} - v_{10} - v_{5} - v_{21} - v_{34} \\
\frac{dC3}{dt} &= -v_2 - v_3 - v_6 - v_7 + v_{22} \\
\frac{dC3^*}{dt} &= v_9 + v_3 - v_{16} + v_6 + v_7 - v_6 - v_3 - v_{31} - v_{34} - v_{26} - v_{26} \\
\frac{dc3X}{dt} &= v_{13} + v_{24} - v_{37} - v_{34} - v_{19} - v_{40} \\
\frac{dc3^*X}{dt} &= v_{11} - v_{14} - v_{26} - v_{35} \\
\frac{dc3^*}{dt} &= v_4 - v_9 - v_{11} - v_{20} - v_{39} \\
\frac{dA^*c3^*}{dt} &= v_5 + v_3 - v_{12} - v_{27} - v_{40} \\
\frac{dA^*c3^*X}{dt} &= v_{12} + v_{14} - v_{28} - v_{50}
\end{align*}
\]

Additionally, four complexes between Casp3, Casp9 and XIAP were considered in the non-competitive model (black intermediates in Fig. D.1). The corresponding differential equations read:

\[
\begin{align*}
\frac{dc0XC3^*}{dt} &= -v_{20} + v_{25} + v_{37} - v_{31} \\
\frac{dA^*c0XC3^*}{dt} &= -v_{26} + v_{34} + v_{38} + v_{41} \\
\frac{dc9XC3^*}{dt} &= -v_{31} + v_{35} + v_{38} - v_{42} \\
\frac{dA^*c9XC3^*}{dt} &= -v_{39} + v_{36} + v_{40} + v_{42}
\end{align*}
\]

The new velocities (reactions 29 – 42) in the non-competitive model are given by:

\[
\begin{align*}
\upsilon_{29} &= k_{29} \cdot C0XC3^* \\
\upsilon_{30} &= k_{30} \cdot A^*C0XC3^* \\
\upsilon_{31} &= k_{31} \cdot C0XC3^* \\
\upsilon_{32} &= k_{32} \cdot A^*C0XC3^* \\
\upsilon_{33} &= k_{33} \cdot C3^* \cdot C0X - k_{33} \cdot C0XC3^* \\
\upsilon_{34} &= k_{34} \cdot C3^* \cdot A^*C0X - k_{34} \cdot A^*C0XC3^* \\
\upsilon_{35} &= k_{35} \cdot C3^* \cdot C0X - k_{35} \cdot C0XC3^* \\
\upsilon_{36} &= k_{36} \cdot C3^* \cdot A^*C0X - k_{36} \cdot C0XC3^* \\
\upsilon_{37} &= k_{37} \cdot C0X \cdot A^*C0X - k_{37} \cdot C0XC3^* \\
\upsilon_{38} &= k_{38} \cdot A^*C0X \cdot C3^* - k_{38} \cdot A^*C0XC3^* \\
\upsilon_{39} &= k_{39} \cdot C0X \cdot C3^* - k_{39} \cdot C0XC3^* \\
\upsilon_{40} &= k_{40} \cdot A^*C0X \cdot C3^* - k_{40} \cdot A^*C0XC3^* \\
\upsilon_{41} &= k_{41} \cdot C0XC3^* \cdot A^* - k_{41} \cdot A^*C0XC3^* \\
\upsilon_{42} &= k_{42} \cdot C0XC3^* \cdot A^* - k_{42} \cdot A^*C0XC3^*
\end{align*}
\]
The values chosen for the additional kinetic parameters in the non-competitive model are summarised in Table D.1. The association and dissociation rate constants of the ternary Casp9-XIAP-Casp3 complexes (reactions 33 – 40 in Fig. D.1) were assumed to equal those of the simple caspase-XIAP complexes, Casp3-XIAP (k_{15}, k_{-15}) and Casp9-XIAP (k_9, k_{-9}), and multiplied by the factors a and d (see Table D.1). Thus, a and d quantify to what extent Casp3 and Casp9 bind competitively to XIAP. In Fig. 4.4H, both a and d were assumed to equal unity, so that competition does not occur. The competition ratio varied in Fig. 4.5 is given by \( \alpha = a / d \), and increasing caspase competition for XIAP (i.e., decreasing \( \alpha \)) was modelled by simultaneously decreasing a and increasing d to the same extent (fold-changes).

### Table D.1: Additional kinetic parameters used in the non-competitive model.

<table>
<thead>
<tr>
<th>#</th>
<th>Reaction</th>
<th>( k_+ )</th>
<th>k.</th>
<th>( K_D = k_- / k_+ )</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>C9XC3* ( \rightarrow )</td>
<td>( 10^2 ) s(^{-1} )</td>
<td>-</td>
<td>-</td>
<td>k, as #16</td>
</tr>
<tr>
<td>30</td>
<td>A<em>C9XC3</em> ( \rightarrow )</td>
<td>( 10^2 ) s(^{-1} )</td>
<td>-</td>
<td>-</td>
<td>k, as #16</td>
</tr>
<tr>
<td>31</td>
<td>C9<em>XC3</em> ( \rightarrow )</td>
<td>( 10^2 ) s(^{-1} )</td>
<td>-</td>
<td>-</td>
<td>k, as #16</td>
</tr>
<tr>
<td>32</td>
<td>A<em>C9</em>XC3* ( \rightarrow )</td>
<td>( 10^2 ) s(^{-1} )</td>
<td>-</td>
<td>-</td>
<td>k, as #16</td>
</tr>
<tr>
<td>33</td>
<td>C9X + C3* ( \leftrightarrow ) C9XC3*</td>
<td>( a \cdot k_{15} )</td>
<td>( d \cdot k_{15} )</td>
<td>( d / a \cdot k_{15} / k_{15} = 1 / \alpha \cdot k_{15} / k_{15} )</td>
<td>-</td>
</tr>
<tr>
<td>34</td>
<td>A<em>C9X + C3</em> ( \leftrightarrow ) A<em>C9XC3</em></td>
<td>( a \cdot k_{15} )</td>
<td>( d \cdot k_{15} )</td>
<td>( d / a \cdot k_{15} / k_{15} = 1 / \alpha \cdot k_{15} / k_{15} )</td>
<td>-</td>
</tr>
<tr>
<td>35</td>
<td>C9X + C3* ( \leftrightarrow ) C9XC3*</td>
<td>( a \cdot k_{15} )</td>
<td>( d \cdot k_{15} )</td>
<td>( d / a \cdot k_{15} / k_{15} = 1 / \alpha \cdot k_{15} / k_{15} )</td>
<td>-</td>
</tr>
<tr>
<td>36</td>
<td>A<em>C9X + C3</em> ( \leftrightarrow ) A<em>C9XC3</em></td>
<td>( a \cdot k_{15} )</td>
<td>( d \cdot k_{15} )</td>
<td>( d / a \cdot k_{15} / k_{15} = 1 / \alpha \cdot k_{15} / k_{15} )</td>
<td>-</td>
</tr>
<tr>
<td>37</td>
<td>C9 + C3X ( \leftrightarrow ) C9XC3*</td>
<td>( a \cdot k_{9} )</td>
<td>( d \cdot k_{9} )</td>
<td>( d / a \cdot k_{9} / k_{9} = 1 / \alpha \cdot k_{9} / k_{9} )</td>
<td>-</td>
</tr>
<tr>
<td>38</td>
<td>A<em>C9 + C3X ( \leftrightarrow ) A</em>C9XC3*</td>
<td>( a \cdot k_{9} )</td>
<td>( d \cdot k_{9} )</td>
<td>( d / a \cdot k_{9} / k_{9} = 1 / \alpha \cdot k_{9} / k_{9} )</td>
<td>-</td>
</tr>
<tr>
<td>39</td>
<td>C9* + C3X ( \leftrightarrow ) C9XC3*</td>
<td>( a \cdot k_{9} )</td>
<td>( d \cdot k_{9} )</td>
<td>( d / a \cdot k_{9} / k_{9} = 1 / \alpha \cdot k_{9} / k_{9} )</td>
<td>-</td>
</tr>
<tr>
<td>40</td>
<td>A<em>C9</em> + C3X ( \leftrightarrow ) A<em>C9XC3</em></td>
<td>( a \cdot k_{9} )</td>
<td>( d \cdot k_{9} )</td>
<td>( d / a \cdot k_{9} / k_{9} = 1 / \alpha \cdot k_{9} / k_{9} )</td>
<td>-</td>
</tr>
<tr>
<td>41</td>
<td>C9XC3* + A* ( \leftrightarrow ) A<em>C9XC3</em></td>
<td>( 2 \times 10^8 ) M(^{-1}) s(^{-1} )</td>
<td>0.1 s(^{-1} )</td>
<td>( 5 \times 10^{-8} ) M</td>
<td>as #1</td>
</tr>
<tr>
<td>42</td>
<td>C9XC3* + A* ( \leftrightarrow ) A<em>C9XC3</em></td>
<td>( 2 \times 10^8 ) M(^{-1}) s(^{-1} )</td>
<td>0.1 s(^{-1} )</td>
<td>( 5 \times 10^{-8} ) M</td>
<td>as #1</td>
</tr>
</tbody>
</table>

The reactions numbered according to Fig. D.1 (Column "#") are listed and the corresponding reactants and products are indicated (Column "Reaction"). The rate constants "k.\(^+-\)" describe the reactions from left to right, while "k.\(^-\)" are the rate constants for the opposite direction (for reversible reactions). Additionally, the dissociation constants \( K_D \) are indicated for reversible bimolecular reactions. The degradation rates (29 – 31) of the ternary complexes were set equal to those of all other molecular species in the model, and association between Casp9 and active Apaf-1 (41 – 42) was assumed to be unaffected by ternary complex formation (see "Notes").
D.2 Bifurcation analysis of XIAP-mediated feedback

In this Section, detailed simulations are provided, which show that XIAP-mediated feedback contributes to bistability and irreversibility in the wildtype model as schematically depicted in Fig. 4.3.

**Fig. D.2: Bifurcation analysis of XIAP-mediated feedback.**

The steady state concentrations of free Casp3 (C3*; black line in A), of XIAP-associated Casp3 (C3*X; grey line in A), of free Casp9 (C9 + A*C9 + C9* + A*C9*; black line in B) and of XIAP-associated Casp9 (C9X + A*C9X + C9*X + A*C9*X; grey line in B) are plotted as a function of the stimulus, A*_tot. The wildtype model (depicted in Fig. 4.1B) with the parameters given in Table 4.1 was used for simulations. The solid and dashed lines indicate stable and unstable steady states, respectively.

Upon weak stimulation the vast majority of Casp9 molecules is inhibited by excess XIAP (black and grey lines in B), so that Casp3 activation is negligible (black line in A). As the stimulus strength is increased above the threshold (~ 3 nM), XIAP is suddenly redistributed from Casp9 to Casp3 (grey lines in A and B), and Casp3 and Casp9 activities are switched on (black lines in A and B). Caspase activity is maintained even if the stimulus is removed, as Casp3, once activated, retains XIAP (grey line in A), and thereby prevents Casp9 inhibition (black line in B).

D.3 An *in vitro* Test for XIAP-mediated Feedback

In Section 4, it was predicted that XIAP mediates positive feedback and bistability in the intrinsic pathway. This Section contains a description of an *in vitro* experiment designed to confirm that sequestration of XIAP by Casp3 indeed results in feedback amplification.

XIAP is known to suppress the activity of apoptosome-activated Casp9 *in vitro* [201], but according to the results given in the main text excess pro-Casp3 should reverse this inhibition by sequestering XIAP away from Casp9. The proposed experiments (‘Experiment 1’ and ‘Experiment 2’) are summarised in Table D.2: The pro-Casp9 mutant D330A, which is refractory towards feedback cleavage by Casp3, should be activated *in vitro* by adding cyto c, dATP, and Apaf-1 [201]. Full-length XIAP is added to both ‘Experiment’ reaction mixtures, while pro-Casp3 is present in Experiment 2 (feedback on), but absent in Experiment 1 (feedback on). Then, the Casp9 activities (E₁ and E₂) of both reaction mixtures should be measured as a readout.
Table D.2: An *in vitro* test experiment for XIAP-mediated feedback

<table>
<thead>
<tr>
<th>Substance</th>
<th>Experiment 1</th>
<th>Control 1</th>
<th>Control 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>cyto c</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>dATP</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Apaf-1</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>pro-Casp9 (D330A)</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>full-length XIAP</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BIR1-BIR2 (XIAP fragment)</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>BIR3-RING (XIAP fragment)</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>pro-Casp3</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Readout (Casp9 activity)</td>
<td>E₁</td>
<td>E₂</td>
<td>C₁</td>
</tr>
</tbody>
</table>

The predicted results are shown in Fig. D.3. Here, the ratio of Casp9 activities with and without Casp3 (E₂ / E₁) as a measure of (XIAP-mediated) feedback strength is plotted against the XIAP concentration. As expected, XIAP-mediated feedback is especially pronounced for intermediate XIAP concentrations, where C₉tot < XIAPₜot < C₃tot (C₉tot = 20 nM in these simulations). Additionally, an increase in the Casp3 concentration improves feedback strength, and also widens the range of XIAP concentrations, where feedback is observed. Therefore, the *in vitro* experiments should be done using low Casp9 concentrations, high Casp3 concentrations and intermediate XIAP concentrations (XIAPₜot ≈ ¾ C₃tot; see Fig. D.3). Additional simulations revealed that the results shown in Fig. D.3 are independent of the concentrations of active Apaf-1, but sufficiently high Apaf1-levels should be chosen in order to minimise errors.

Figure D.3: Predicted results of an *in vitro* test experiment for XIAP-mediated feedback.

*In vitro* behaviour of the caspase cascade with mutant Casp9 (D330A) was modelled by eliminating Casp3-mediated feedback cleavage (v₄ and v₅) and protein synthesis/degradation reactions (v₁₆ - v₂₈) from the differential equations given in Section D.1 (see text for details).
As a control experiment, the whole procedure should be repeated with the XIAP fragments, BIR1-BIR2 (specific for Casp3) and BIR3-RING (specific for Casp9), instead of full-length XIAP (see Table D.3). These controls mimic non-competitive caspase inhibition, so that the feedback strength \( \frac{C_2}{C_1} \) is predicted to equal unity regardless of the protein concentrations chosen.

### D.4 Casp3-induced Degradation of XIAP Does not Result in Bistability

Experimental evidence suggests that Casp3 activation may result in XIAP cleavage and/or degradation, although this seems to be a cell-type-specific phenomenon (see Section 4.6). In the following, it is demonstrated that Casp3-induced XIAP degradation (‘inhibition of inhibition’ = positive circuit) does not result in physiologically relevant bistability in the absence of other feedback amplification loops.

![Figure D.4: Model of Casp3-induced XIAP Degradation](image)

In order to obtain analytical results, a simplified model of Casp3-induced XIAP degradation comprising only the black reactions in Fig. D.4 is analysed first. The corresponding differential equations read:

\[
\frac{dC_3}{dt} = k_1 - (k_2 + k_3) \cdot C_3 \\
\frac{dC_3^*}{dt} = k_3 \cdot C_3 - (k_4 + k_5 \cdot X) \cdot C_3^* + k_{-5} \cdot C_3^* X \\
\frac{dC_3^* X}{dt} = k_5 \cdot X \cdot C_3^* - (k_{-5} + k_6) \cdot C_3^* X \\
\frac{dX}{dt} = k_7 + k_{-5} \cdot C_3^* X - (k_5 \cdot C_3^* + k_8 + k_9 \cdot C_3^*) \cdot X
\]

The steady state condition of active Casp3 \( \frac{dC_3^*}{dt} = 0 \) can be written as a quadratic equation in \( C_3^* \), which implies that the steady state of \( C_3^* \) cannot be bistable.

A more realistic model (black and grey arrows in Fig. D.4) considering that XIAP inhibits both Casp3 and Casp9 was then analysed numerically. Here, the feedback loops discussed in Section 4 (i.e., Casp3-mediated feedback cleavage of Casp9 and XIAP-mediated feedback) were assumed to be inactive in order to focus on the role of Casp3-mediated XIAP...
degradation. Hence, XIAP was assumed to bind to Casp3 and Casp9 in a non-competitive manner.

All kinetic parameters were chosen according to Tables 4.1 and D.1 (with $\alpha = 1$). Additionally, Casp3-mediated XIAP degradation (reaction 9 in Fig. D.4) was modelled as an irreversible second-order process with the rate constant $k_9 = 3 \times 10^6 \text{M}^{-1}\text{s}^{-1}$, which is the value measured for high-affinity substrates of Casp3 [487,488].

These studies revealed that Casp3-induced XIAP degradation does not result in bistability for experimentally measured protein concentrations (Casp3 = 200 nM; Casp9 = 20 nM; XIAP = 40 nM). Although some bistability could be observed for significantly different expression levels (e.g., for XIAP > 100 nM), hysteresis was restricted to a very small stimulus range, so that the stimulus-response was virtually indistinguishable from that of a monostable system.

Importantly, the physiological feedback strength (i.e., $k_9$ in Fig. D.4) is likely to be lower than that assumed here, as most Casp3 substrates do not match the optimal Casp3 target sequence, DEVD [487]. For example, Casp3 cleaves XIAP at a suboptimal site [259]. Likewise, Akt, a protein kinase that mediates XIAP stabilisation unless it is processed and thereby inactivated by Casp3 [264], also does not contain the optimal DEVD-target sequence [489]. As high feedback strength is required for bistability [197], it can be concluded that Casp3-induced XIAP degradation is unlikely to result in physiologically relevant bistability as long as other feedback amplification loops are absent.

### D.5 The Shared Inhibitor Motif

In Section 4, it was shown that competitive inhibition of Casp3 and Casp9 by XIAP can bring about positive feedback and bistability in the intrinsic apoptosis pathway (see Fig. 4.2B, grey line; Fig. 4.4F). Similar conclusions regarding positive feedback and bistability also hold in general if an inhibitory protein competitively inhibits two consecutive intermediates in signal transduction cascades. This ‘shared inhibitor motif’ is schematically depicted in Fig. D.5. A stimulus, $S$, activates the upstream intermediate, $U$, which then in turn catalyses the activation of the downstream intermediate, $D$. Both active intermediates, $U^*$ and $D^*$, are subject to negative regulation by the shared inhibitor, $I$. As indicated in Fig. D.5, the shared inhibitor can either be a stoichiometric inhibitor of the intermediates (black arrows in Fig. D.5) or alternatively catalyses their deactivation, e.g., dephosphorylation (black and grey arrows in Fig. D.5).

![Figure D.5: The Shared Inhibitor Motif](image)
In general, bistability can arise if the shared inhibitor binds the intermediates competitively at least to some extent. Furthermore, bistability requires that only the active downstream intermediate, D*, but not its precursor, D, binds to the inhibitor, I. In addition to these ‘structural’ requirements, the downstream intermediate, D, needs to be significantly more abundant than the inhibitor, I, which in turn must exceed the upstream intermediate, U (see Section 4). Finally, the inhibitor, I, mediates particularly strong positive feedback if the downstream intermediate exceeds the dissociation constant (or the Michaelis-Menten constant) of the D*I-complex.

Table D.3: The ‘Shared Inhibitor motif’, a recurrent motif in cellular signal transduction

<table>
<thead>
<tr>
<th>Upstream (U)</th>
<th>Downstream (D)</th>
<th>Shared Inhibitor(s) (I)</th>
<th>Type</th>
<th>Refs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caspase-9</td>
<td>Caspase-3</td>
<td>X-IAP, c-IAP1/2, Survivin</td>
<td>Inhib.</td>
<td>[490]</td>
</tr>
<tr>
<td>Caspase-9</td>
<td>Caspase-7</td>
<td>X-IAP, c-IAP1/2, Survivin</td>
<td>Inhib.</td>
<td>[490]</td>
</tr>
<tr>
<td>FGFR</td>
<td>Mek</td>
<td>Sef</td>
<td>Inhib.</td>
<td>[157]</td>
</tr>
<tr>
<td>Grb-2</td>
<td>Raf</td>
<td>Sprouty</td>
<td>Inhib.</td>
<td>[157]</td>
</tr>
<tr>
<td>CyclinD / CDK4/6</td>
<td>Cyclin E / CDK2</td>
<td>p21-CIP1, p27-KIP1, p57-KIP2</td>
<td>Inhib.</td>
<td>[491]</td>
</tr>
<tr>
<td>Daxx (-&gt; Ask1)</td>
<td>Cytochrome C</td>
<td>Hsp27</td>
<td>Inhib.</td>
<td>[492,493]</td>
</tr>
<tr>
<td>Cdc42</td>
<td>Rac</td>
<td>10 different GAPs reported</td>
<td>GAP</td>
<td>[494,495]</td>
</tr>
<tr>
<td>Rac</td>
<td>Rho</td>
<td>4 different GAPs reported</td>
<td>GAP</td>
<td>[494,495]</td>
</tr>
<tr>
<td>Cdc42</td>
<td>Rho</td>
<td>5 different GAPs reported</td>
<td>GAP</td>
<td>[494,495]</td>
</tr>
<tr>
<td>insulin receptor</td>
<td>IRS-1</td>
<td>LAR</td>
<td>PP</td>
<td>[496]</td>
</tr>
<tr>
<td>EGF receptor</td>
<td>HGF receptor</td>
<td>LAR</td>
<td>PP</td>
<td>[497,498]</td>
</tr>
<tr>
<td>CaMK II</td>
<td>AMPAR</td>
<td>PP1</td>
<td>PP</td>
<td>[198]</td>
</tr>
<tr>
<td>Aurora-B kinase</td>
<td>Histone 3</td>
<td>PP1</td>
<td>PP</td>
<td>[499]</td>
</tr>
<tr>
<td>Aurora-B kinase</td>
<td>Ndc10</td>
<td>PP1</td>
<td>PP</td>
<td>[499]</td>
</tr>
<tr>
<td>NMDA receptor</td>
<td>CaMK II</td>
<td>PP1</td>
<td>PP</td>
<td>[499]</td>
</tr>
<tr>
<td>p38</td>
<td>Caspase-3</td>
<td>PP2A</td>
<td>PP</td>
<td>[500]</td>
</tr>
<tr>
<td>Mek</td>
<td>Erk</td>
<td>PP2A</td>
<td>PP</td>
<td>[501]</td>
</tr>
<tr>
<td>PAK</td>
<td>CREB</td>
<td>PP2A</td>
<td>PP</td>
<td>[148,502,503]</td>
</tr>
<tr>
<td>Akt</td>
<td>CREB</td>
<td>PP2A</td>
<td>PP</td>
<td>[148,502,503]</td>
</tr>
<tr>
<td>CaMK IV</td>
<td>CREB</td>
<td>PP2A</td>
<td>PP</td>
<td>[148,502,503]</td>
</tr>
<tr>
<td>SEK1</td>
<td>JNK</td>
<td>PP2A</td>
<td>PP</td>
<td>[148,502,503]</td>
</tr>
<tr>
<td>Akt</td>
<td>p70S6K</td>
<td>PP2A</td>
<td>PP</td>
<td>[148,503,504]</td>
</tr>
<tr>
<td>PAK</td>
<td>Mek/Erk</td>
<td>PP2A</td>
<td>PP</td>
<td>[148,503,505]</td>
</tr>
<tr>
<td>PKCα</td>
<td>Mek/Erk</td>
<td>PP2A</td>
<td>PP</td>
<td>[21,503,506]</td>
</tr>
<tr>
<td>MKK6</td>
<td>p38</td>
<td>PP2C-α</td>
<td>PP</td>
<td>[507]</td>
</tr>
<tr>
<td>insulin receptor</td>
<td>IRS-1</td>
<td>PTP-1B</td>
<td>PP</td>
<td>[496]</td>
</tr>
<tr>
<td>Epo receptor</td>
<td>Jak2</td>
<td>PTP-1B</td>
<td>PP</td>
<td>[506,508]</td>
</tr>
<tr>
<td>Epo receptor</td>
<td>STAT5a/b</td>
<td>PTP-1B</td>
<td>PP</td>
<td>[506,509]</td>
</tr>
<tr>
<td>EGFR</td>
<td>IGF-1 receptor</td>
<td>PTP-1B</td>
<td>PP</td>
<td>[497,510]</td>
</tr>
<tr>
<td>IGF-1 receptor</td>
<td>EGFR</td>
<td>PTP-1B</td>
<td>PP</td>
<td>[497,511]</td>
</tr>
<tr>
<td>Jak2</td>
<td>EGFR</td>
<td>PTP-1B</td>
<td>PP</td>
<td>[497,508,512]</td>
</tr>
<tr>
<td>insulin receptor</td>
<td>Jak2</td>
<td>PTP-1B</td>
<td>PP</td>
<td>[497,508,513,514]</td>
</tr>
<tr>
<td>insulin receptor</td>
<td>STAT5a/b</td>
<td>PTP-1B</td>
<td>PP</td>
<td>[497,509,513,514]</td>
</tr>
<tr>
<td>PDGF receptor</td>
<td>Jak2</td>
<td>PTP-1B</td>
<td>PP</td>
<td>[497,508,515]</td>
</tr>
<tr>
<td>PDGF receptor</td>
<td>STAT5a/b</td>
<td>PTP-1B</td>
<td>PP</td>
<td>[509,512,515]</td>
</tr>
<tr>
<td>CAKβ</td>
<td>p130(CAS)</td>
<td>PTP-PEST</td>
<td>PP</td>
<td>[516]</td>
</tr>
<tr>
<td>CAKβ</td>
<td>Paxillin</td>
<td>PTP-PEST</td>
<td>PP</td>
<td>[516]</td>
</tr>
<tr>
<td>EGFR receptor</td>
<td>p52Shc</td>
<td>TCPTP</td>
<td>PP</td>
<td>[517]</td>
</tr>
</tbody>
</table>

In general, bistability can arise if the shared inhibitor binds the intermediates competitively at least to some extent. Furthermore, bistability requires that only the active downstream intermediate, D*, but not its precursor, D, binds to the inhibitor, I. In addition to these ‘structural’ requirements, the downstream intermediate, D, needs to be significantly more abundant than the inhibitor, I, which in turn must exceed the upstream intermediate, U (see Section 4). Finally, the inhibitor, I, mediates particularly strong positive feedback if the downstream intermediate exceeds the dissociation constant (or the Michaelis-Menten constant) of the D*I-complex.
It should be noted that shared inhibitors, which function enzymatically (black and grey arrows in Fig. D.5), can be efficiently sequestered by the downstream intermediate, D*, and thereby mediate positive feedback and bistability even if the D*I is only transiently formed and then broken down by catalysis (as long as the Michaelis-Menten constant is low enough).

Table D.3 gives an overview on signal transduction pathways, where the ‘shared inhibitor motif’ has been reported to occur. Inhibitory proteins were sub-classified into three groups, according to their biochemical mechanism of action: (i) stoichiometric inhibitors (Inh.), which reversibly sequester proteins away from their cellular targets; (ii) GTPase-activating proteins (GAP), that stimulate the intrinsic GTPase activity of small G proteins, and thereby catalyse their deactivation; (iii) Phosphatases (PP), which antagonise protein-phosphorylation cascades, are the most prominent group in Table D.3.

Available experimental data suggest that bistability due to sequestration of a shared phosphatase can occur in vivo. Most phosphatases exhibit a single active site, i.e., they bind their substrates in a competitive manner. Additionally, they usually recognise only phosphorylated, but not non-phosphorylated, substrates, so that the structural requirements mentioned above are fulfilled. Quantitative measurements of protein abundance in the MAPK cascade revealed that the downstream intermediates in this system are (much) more abundant when compared to their upstream activators [16,158]. Finally, many phosphatases exhibit Michaelis-Menten constants in the sub-micromolar range [154,432], which suggests that strong feedback can be established (see above).

The feedback mechanism proposed in this paper may, for example, contribute to bistability in the JNK cascade [135], since PP2A was shown to dephosphorylate both JNK and its upstream activator SEK1 (see Table D.3).
E Additional material on recurrent design patterns in transcriptional feedback regulation of signalling cascades

E.1 Protein Decay Rates

This Section contains the protein half-live data collected from the published literature which was used to generate Figs. 5.1 and 5.2. Two types of half-lives were distinguished: those measured under unstimulated conditions ($t_{1/2}$ (basal)), and those measured under stimulation conditions ($t_{1/2}$ (stim)). Some protein half-lives had been measured both under stimulated and unstimulated conditions with different results. In this case, the half-life under unstimulated conditions was used for non-induced proteins, while the half-life upon stimulation was taken for transcriptionally induced proteins. The underlying rationale follows: The unstimulated state was assumed to be the default situation for the cell, and thus it is the basal protein turnover rate of non-induced (i.e., permanently expressed) proteins that should be relevant for the cellular energy budget. On the other hand, rapid induction of feedback mediators requires short mRNA/protein half-lives, so that the protein half-life measured in stimulated cells was taken for induced proteins. In the light of the energy hypothesis, these rapidly induced factors can still contribute to free-energy dissipation minimisation if they are much more stable in the basal state when compared to stimulation conditions. Qualitatively similar results were obtained if the half-lives for stimulated cells were taken for transcriptionally non-induced proteins.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Alternative Name</th>
<th>Function</th>
<th>$t_{1/2}$ (basal)</th>
<th>$t_{1/2}$ (stim)</th>
<th>$t_{1/2}$ (used)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGFBR1</td>
<td>TGFRI receptor</td>
<td></td>
<td>4-6h</td>
<td>~12h</td>
<td>~12h</td>
<td>8.5h</td>
</tr>
<tr>
<td>TGFBR2</td>
<td>TGFRII receptor</td>
<td></td>
<td>~1h</td>
<td>~0.75h</td>
<td>1h</td>
<td>[518] [316]</td>
</tr>
<tr>
<td>SMAD2</td>
<td>Smad2 TF</td>
<td></td>
<td>&gt;&gt;4h</td>
<td>~10h</td>
<td>~3h</td>
<td>9h</td>
</tr>
<tr>
<td>SMAD3</td>
<td>Smad3 TF</td>
<td></td>
<td>&gt;&gt;4.5h</td>
<td>~6h</td>
<td>~4.5h</td>
<td>7.5h</td>
</tr>
<tr>
<td>SMAD4</td>
<td>Smad4 TF</td>
<td></td>
<td>&gt;12h</td>
<td>&gt;&gt;9h</td>
<td>&gt;12h</td>
<td>12h</td>
</tr>
<tr>
<td>TGIF1</td>
<td>TGIF Inhibitor</td>
<td></td>
<td>~1h</td>
<td>~0.5h</td>
<td>0.75h</td>
<td></td>
</tr>
<tr>
<td>SMAD7</td>
<td>Smad7 Inhibitor</td>
<td></td>
<td>~3h</td>
<td>~1h</td>
<td>1h</td>
<td></td>
</tr>
<tr>
<td>SKIL</td>
<td>SnoN Inhibitor</td>
<td></td>
<td>~1h</td>
<td>~0.2h</td>
<td>0.5h</td>
<td></td>
</tr>
<tr>
<td>SKI</td>
<td>Ski Inhibitor</td>
<td></td>
<td>~1.6h</td>
<td>~0.5h</td>
<td>1.6h</td>
<td></td>
</tr>
</tbody>
</table>

TGFβ/Activin pathway (Table E.1)
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Alternative Names</th>
<th>Function</th>
<th>$t_{1/2}$(basal)</th>
<th>$t_{1/2}$(stim)</th>
<th>$t_{1/2}$(used)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>INSR</td>
<td>Insulin receptor</td>
<td>Receptor</td>
<td>20-40h</td>
<td>~7h</td>
<td>7.5h</td>
<td>[532]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8h</td>
<td>~12h</td>
<td></td>
<td>[532]</td>
</tr>
<tr>
<td>EGFR</td>
<td>EGFR Receptor 1</td>
<td>Receptor</td>
<td>20h</td>
<td></td>
<td>4h</td>
<td>12h</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8h</td>
<td></td>
<td></td>
<td>[535]</td>
</tr>
<tr>
<td>MET</td>
<td>HGF receptor</td>
<td>Receptor</td>
<td>~0.7h</td>
<td>~6h</td>
<td>5h</td>
<td>0.5-2h</td>
</tr>
<tr>
<td>PDGFRB</td>
<td>PDGF receptor</td>
<td>Receptor</td>
<td>&gt;&gt;4h</td>
<td></td>
<td>0.5-2h</td>
<td>3.5h</td>
</tr>
<tr>
<td>IGF1R</td>
<td>IGF-1 receptor</td>
<td>Receptor</td>
<td>~8h</td>
<td></td>
<td>8h</td>
<td>0.7-8h</td>
</tr>
<tr>
<td>PTPN2</td>
<td>TC-PTP Phosphatase</td>
<td>~12h</td>
<td>12h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTPN11</td>
<td>SHP-2 Ras activator</td>
<td>~19h</td>
<td>19h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PKRCA</td>
<td>PKCα Ras activator</td>
<td>~24h</td>
<td>26h</td>
<td></td>
<td>0.8h</td>
<td></td>
</tr>
<tr>
<td>RAPGEF2</td>
<td>cnRasGEF Ras activator</td>
<td>~10h</td>
<td>10h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOS1</td>
<td>mSos1 RasGEF</td>
<td>&gt;18h</td>
<td>20h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOS2</td>
<td>mSos2 RasGEF</td>
<td>&gt;18h</td>
<td>3h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RASGRF2</td>
<td>Ras-GRF1 RasGEF</td>
<td>&gt;18h</td>
<td>8h</td>
<td></td>
<td></td>
<td>[552]</td>
</tr>
<tr>
<td>NRAS</td>
<td>N-Ras MAP4K</td>
<td>~24h</td>
<td>24h</td>
<td></td>
<td></td>
<td>[553]</td>
</tr>
<tr>
<td>HRAS</td>
<td>H-Ras MAP4K</td>
<td>~19h</td>
<td>19.5h</td>
<td></td>
<td></td>
<td>[554]</td>
</tr>
<tr>
<td>NF1</td>
<td>NF1 RasGAP</td>
<td>~24h</td>
<td>60-80h</td>
<td>24h</td>
<td></td>
<td>[556]</td>
</tr>
<tr>
<td>RAF1</td>
<td>Raf-1 MAP3K</td>
<td>17.5h</td>
<td>14.25h</td>
<td></td>
<td></td>
<td>[557]</td>
</tr>
<tr>
<td>MAP3K2</td>
<td>Mekk-2 MAP3K</td>
<td>~1h</td>
<td>1h</td>
<td></td>
<td></td>
<td>[558]</td>
</tr>
<tr>
<td>MAP2K6</td>
<td>MKK6 MAP2K</td>
<td>&gt;24h</td>
<td>24h</td>
<td></td>
<td></td>
<td>[559]</td>
</tr>
<tr>
<td>MAP2K1/2</td>
<td>Mek1/2 MAP2K</td>
<td>&gt;6h</td>
<td>10h</td>
<td></td>
<td></td>
<td>[560]</td>
</tr>
<tr>
<td>MAPK3</td>
<td>Erk1 MAPK</td>
<td>&gt;60h</td>
<td>22h</td>
<td></td>
<td></td>
<td>[561]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;20h</td>
<td></td>
<td></td>
<td></td>
<td>[564]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;8</td>
<td></td>
<td></td>
<td></td>
<td>[566]</td>
</tr>
<tr>
<td>MAPK1</td>
<td>Erk2 MAPK</td>
<td>&gt;60h</td>
<td>10h</td>
<td></td>
<td></td>
<td>[561]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;8</td>
<td></td>
<td></td>
<td></td>
<td>[564]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;8</td>
<td></td>
<td></td>
<td></td>
<td>[566]</td>
</tr>
<tr>
<td>MAPK15</td>
<td>Erk7 MAPK</td>
<td>~1.7h</td>
<td>1.7h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAPK14</td>
<td>p38α MAPK</td>
<td>&gt;24h</td>
<td>24h</td>
<td></td>
<td></td>
<td>[559]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;20h</td>
<td></td>
<td></td>
<td></td>
<td>[561]</td>
</tr>
<tr>
<td>MAPK8</td>
<td>JNK1 MAPK</td>
<td>&gt;7h</td>
<td>9h</td>
<td></td>
<td></td>
<td>[566]</td>
</tr>
<tr>
<td>DUSP1</td>
<td>MKP-1 MAPK PPhase</td>
<td>2h</td>
<td>0.75h</td>
<td>0.75h</td>
<td>0.75h</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>~0.75h</td>
<td>~0.75h</td>
<td>~0.75h</td>
<td>~0.75h</td>
<td>[567]</td>
</tr>
<tr>
<td>DUSP4</td>
<td>MKP-2 MAPK PPhase</td>
<td>1.2h</td>
<td>0.75h</td>
<td>1h</td>
<td></td>
<td>[572]</td>
</tr>
<tr>
<td>DUSP6</td>
<td>MKP-3 MAPK PPhase</td>
<td>0.25h</td>
<td>0.75h</td>
<td>0.9h</td>
<td></td>
<td>[573]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2h</td>
<td>1h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUSP10</td>
<td>MKP-5 MAPK PPhase</td>
<td>4h</td>
<td>4h</td>
<td></td>
<td></td>
<td>[572]</td>
</tr>
<tr>
<td>DUSP16</td>
<td>MKP-7 MAPK PPhase</td>
<td>1.5h</td>
<td>1.5h</td>
<td></td>
<td></td>
<td>[572]</td>
</tr>
<tr>
<td>DUSP8</td>
<td>M3/6 MAPK PPhase</td>
<td>2h</td>
<td>1.9h</td>
<td></td>
<td></td>
<td>[574]</td>
</tr>
<tr>
<td>DUSP3</td>
<td>VHR MAPK PPhase</td>
<td>0.5h</td>
<td>0.5h</td>
<td></td>
<td></td>
<td>[576]</td>
</tr>
<tr>
<td>SPRY1</td>
<td>Sprouty Inhibitor</td>
<td>&gt;1h</td>
<td>~0.5h</td>
<td>0.5h</td>
<td></td>
<td>[451]</td>
</tr>
</tbody>
</table>
### STAT pathway (Table E.3)

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Alternative Name</th>
<th>Function</th>
<th>$t_{1/2}$ (basal)</th>
<th>$t_{1/2}$ (stim)</th>
<th>$t_{1/2}$ (used)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>IFNGR1</td>
<td>IFN-γ receptor</td>
<td>receptor</td>
<td>-2.5h</td>
<td>-2.5h</td>
<td>5h</td>
<td>[577]</td>
</tr>
<tr>
<td>IL6ST</td>
<td>gp130</td>
<td>receptor</td>
<td>-2.5h</td>
<td>-2.5h</td>
<td>2.5h</td>
<td>[546]</td>
</tr>
<tr>
<td>IL6R</td>
<td>Gp80</td>
<td>receptor</td>
<td>-2.5h</td>
<td>-2.5h</td>
<td>2.5h</td>
<td>[579]</td>
</tr>
<tr>
<td>EPOR</td>
<td>EpoR</td>
<td>receptor</td>
<td>3h</td>
<td>3h</td>
<td>3h</td>
<td>[580]</td>
</tr>
<tr>
<td>JAK1</td>
<td>JAK1 Kinase</td>
<td></td>
<td>-3.2h</td>
<td>-3.2h</td>
<td>3.2h</td>
<td>[546]</td>
</tr>
<tr>
<td>JAK2</td>
<td>JAK2 Kinase</td>
<td></td>
<td>-2h,-3h,-6h,-10h, -1h, &gt;5h</td>
<td>4.5h</td>
<td>[546]</td>
<td>[581]</td>
</tr>
<tr>
<td>STAT1</td>
<td>STAT1 TF</td>
<td></td>
<td>-16h, &gt;24h, &gt;12h, -4h, -1.6h, &gt;24h</td>
<td>15h</td>
<td>[546]</td>
<td></td>
</tr>
<tr>
<td>STAT2</td>
<td>STAT2 TF</td>
<td></td>
<td>&gt;24h</td>
<td></td>
<td>24h</td>
<td>[588]</td>
</tr>
<tr>
<td>STAT3</td>
<td>STAT3α, STAT3β, STAT3</td>
<td>TF</td>
<td>-8.5h, -4.5h, &gt;&gt;4h, -1.6h, -1.6h</td>
<td>4.8h</td>
<td>[546]</td>
<td></td>
</tr>
<tr>
<td>STAT5</td>
<td>STAT5 TF</td>
<td></td>
<td>&gt;&gt;12h</td>
<td></td>
<td>16h</td>
<td>[584]</td>
</tr>
<tr>
<td>CISH</td>
<td>CIS Inhibitor</td>
<td></td>
<td>-1h</td>
<td></td>
<td>1h</td>
<td>[590]</td>
</tr>
<tr>
<td>SOCS1</td>
<td>SOCS-1 Inhibitor</td>
<td></td>
<td>-1.5h, -1.5h, 2-3h, -5h</td>
<td>2h</td>
<td>[546]</td>
<td></td>
</tr>
<tr>
<td>SOCS2</td>
<td>SOCS-2 Inhibitor</td>
<td></td>
<td>-1h</td>
<td></td>
<td>1h</td>
<td>[593]</td>
</tr>
<tr>
<td>SOCS3</td>
<td>SOCS-3 Inhibitor</td>
<td></td>
<td>-1.6h, -8h, -4h, -0.5h, -1h</td>
<td>1h</td>
<td>[546]</td>
<td></td>
</tr>
<tr>
<td>SOCS5</td>
<td>SOCS-5 Inhibitor</td>
<td></td>
<td>-1h</td>
<td></td>
<td>1h</td>
<td>[595]</td>
</tr>
</tbody>
</table>

### PI3K pathway (Table E.4)

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Alternative Name</th>
<th>Function</th>
<th>$t_{1/2}$ (basal)</th>
<th>$t_{1/2}$ (stim)</th>
<th>$t_{1/2}$ (used)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRS1</td>
<td>IRS-1</td>
<td>Adaptor</td>
<td>~8h, &gt;24h, &gt;15h, &gt;8h, -20h</td>
<td>~2h, ~9h</td>
<td>15h</td>
<td>[596]</td>
</tr>
<tr>
<td>IRS2</td>
<td>IRS-2</td>
<td>Adaptor</td>
<td>~2h, &gt;24h</td>
<td></td>
<td>13h</td>
<td>[563]</td>
</tr>
<tr>
<td>PIK3R1</td>
<td>P85α</td>
<td>PI-3'-Kinase</td>
<td>&gt;8h, -9h</td>
<td></td>
<td>9.5h</td>
<td>[565]</td>
</tr>
<tr>
<td>PIK3CB</td>
<td>p110κ</td>
<td>PI-3'-Kinase</td>
<td>~7.3h</td>
<td></td>
<td>7.3h</td>
<td>[599]</td>
</tr>
<tr>
<td>PTEN</td>
<td>PTEN</td>
<td>PI-3'-PPase</td>
<td>~30h, &gt;&gt;4h, ~10h, -2h</td>
<td>~3h</td>
<td>9h</td>
<td>[600]</td>
</tr>
<tr>
<td>AKT1</td>
<td>Akt/PKB</td>
<td>Kinase</td>
<td>&gt;&gt;6h, ~30h, -2h, &gt;8h</td>
<td>~3h</td>
<td>10h</td>
<td>[604]</td>
</tr>
<tr>
<td>TSC2</td>
<td>Tuberin</td>
<td>GAP</td>
<td>6h</td>
<td></td>
<td>6h</td>
<td>[607]</td>
</tr>
<tr>
<td>FRAP1</td>
<td>mTOR</td>
<td>Kinase</td>
<td>&gt;&gt;4h</td>
<td></td>
<td>8h</td>
<td>[608]</td>
</tr>
<tr>
<td>GSK3A</td>
<td>GSK3α</td>
<td>Kinase</td>
<td>&gt;&gt;12h</td>
<td></td>
<td>16h</td>
<td>[609]</td>
</tr>
<tr>
<td>GSK3B</td>
<td>GSK3βi</td>
<td>Kinase</td>
<td>&gt;&gt;12h</td>
<td></td>
<td>16h</td>
<td>[609]</td>
</tr>
</tbody>
</table>

---

144
E.2 Estimate of total protein concentrations

The analyses in Section 5 suggested that a considerable evolutionary pressure exists in signal transduction networks, which keeps the turnover of non-induced signalling molecules low. One such evolutionary pressure might be energy cost, as the energy expenditure of mammalian cells is dominated by protein synthesis and degradation.

In order to show that this energy minimisation is indeed physiologically relevant, quantitative measurements of intracellular signalling protein concentrations were collected (see Table E.6), where available from non-transformed cell lines. The cellular protein concentrations of membrane tyrosine kinase receptors and their downstream signalling molecules are typically in the range of 10,000 – 1,000,000 molecules per cell [426,622]. The collection of protein abundances listed below is in accordance with this estimate. However, depending on the cell type, receptor levels are even lower than 10,000 per cell, which is consistent with a more general expression range of 1,000 – 1,000,000 molecules per cell for signalling proteins [622]. Additionally, the list reveals that several phosphatases (e.g., PP1, PP2C, SHP-1 and CD45), several nucleo-cytoplasmically shuttling transcription factors (e.g., STAT2, RelA and the Smads) and several caspases (e.g., Caspases 3, 8 and 9) are also expressed at high levels (>10,000 per cell). Finally, Table E.6 demonstrates that there are several outliers which can be even (much) more abundant than 1,000,000 per cell (e.g., PP2A, PP2B, Ras and Caspase-8).

Taken together, these data suggest that a range of 10,000 – 1,000,000 molecules per cell seems to be reasonable for most signalling proteins. Given the large number of kinases, phosphatases and receptors in the human genome [3,4,623], these data suggest that signalling proteins make up a significant portion of the total cellular protein, so that their production constitutes a central energy sink in the cell.
A more quantitative estimate can be derived if one sums up all measured protein concentrations listed in the Table below (using the maximal values if a range of protein expression is given). This yields ~1.2 10^8 molecules per cell. In this context, it has been measured that there are ~300 pg total protein per cell [624] *. Assuming an average protein molecular weight of 60 kDa, one arrives at ~2 10^8 total protein molecules per cell *. Thus, even the few quantitatively measured signalling proteins listed below can (theoretically) make up as much as ~6% of total cellular protein. Similarly, the weights of the measured molecules add up to 14 pg, corresponding to 5% of the cellular protein mass.

To get an independent estimate of the amount of signalling proteins in cells, high-throughput proteomics data were also analysed. A meta-analysis of the peptide atlas [625] revealed that about 5% of the small sequenced peptides in the atlas stem from annotated signalling proteins (according the gene ontology annotation of the corresponding Ensembl peptide). If one assumes that the probability to find and sequence a peptide in such high-throughput proteomics measurements are proportional to the abundance of a protein one can conclude that signalling proteins constitute about 5% of the cellular protein mass.

Table E.6 Expression of intracellular signalling proteins.

<table>
<thead>
<tr>
<th>Name</th>
<th>Pathway</th>
<th># per cell</th>
<th>Cells</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDGFR</td>
<td>RTK</td>
<td>90.000</td>
<td>NIH3T3 fibroblasts</td>
<td>[626]</td>
</tr>
<tr>
<td>EGFRI</td>
<td>RTK</td>
<td>20.000-200.000</td>
<td>Normal body tissue</td>
<td>[627]</td>
</tr>
<tr>
<td>EGFRII</td>
<td>RTK</td>
<td>5.000-60.000</td>
<td>Human mammary epithelial cells</td>
<td>[628]</td>
</tr>
<tr>
<td>InsulinR</td>
<td>RTK</td>
<td>7.000-250.000</td>
<td>3T3-L1 Adipocytes **</td>
<td>[629]</td>
</tr>
<tr>
<td>IGFRI</td>
<td>RTK</td>
<td>2.800-1.200.000</td>
<td>Primary fibroblasts **</td>
<td>[630]</td>
</tr>
<tr>
<td>Grb2</td>
<td>RTK</td>
<td>23.000</td>
<td>HeLa cells ** ***</td>
<td>[16]</td>
</tr>
<tr>
<td>Ras</td>
<td>MAPK</td>
<td>20.000-11.000.000 [400.000]</td>
<td>[3T3 fibroblasts]</td>
<td>[150]</td>
</tr>
<tr>
<td>Raf</td>
<td>MAPK</td>
<td>10.000</td>
<td>Cos cells</td>
<td>[16]</td>
</tr>
<tr>
<td>Mek</td>
<td>MAPK</td>
<td>360.000-20.000.000 [800.000] [1.800.000]</td>
<td>HeLa cells ** *** [CHO cells] [Rat 1 cells]</td>
<td>[16] [150] [22]</td>
</tr>
<tr>
<td>Erk</td>
<td>MAPK</td>
<td>750.000-20.000.000 [1.800.000] [2.700.000]</td>
<td>HeLa cells ** *** [CHO cells] [Rat 1 cells]</td>
<td>[16] [150] [22]</td>
</tr>
<tr>
<td>PI3K</td>
<td>PI3K</td>
<td>10.000</td>
<td>NIH3T3 fibroblasts</td>
<td>[634]</td>
</tr>
<tr>
<td>Protein Kinase B</td>
<td>600.000</td>
<td>PC12 cells</td>
<td>[635]</td>
<td></td>
</tr>
<tr>
<td>PP1</td>
<td>misc</td>
<td>500.000</td>
<td>Skeletal muscle</td>
<td>[154]</td>
</tr>
<tr>
<td>PP2A</td>
<td>misc</td>
<td>5.000.000 -20.000.000</td>
<td>Various primary tissues</td>
<td>[636]</td>
</tr>
<tr>
<td>PP2B</td>
<td>misc</td>
<td>1.000.000-20.000.000</td>
<td>Brain **</td>
<td>[154]</td>
</tr>
<tr>
<td>PP2C</td>
<td>misc</td>
<td>20.000</td>
<td>Muscle</td>
<td>[154]</td>
</tr>
<tr>
<td>PTP1-B</td>
<td>misc</td>
<td>10.000</td>
<td>Placenta</td>
<td>[154]</td>
</tr>
<tr>
<td>SHP-1</td>
<td>misc</td>
<td>60.000-800.000</td>
<td>Thymocytes **</td>
<td>[637]</td>
</tr>
<tr>
<td>CD45</td>
<td>TCR</td>
<td>&gt;100.000</td>
<td>Jurkat T cells</td>
<td>[638]</td>
</tr>
<tr>
<td>TGFR</td>
<td>TGFβ</td>
<td>1.500-80.000</td>
<td>Swiss 3T3 fibroblasts **</td>
<td>[305]</td>
</tr>
<tr>
<td>Smad2</td>
<td>TGFβ</td>
<td>100.000-450.000</td>
<td>HaCaT keratinocytes **</td>
<td>[639] [378]</td>
</tr>
<tr>
<td>Smad3</td>
<td>TGFβ</td>
<td>15.000-450.000</td>
<td>HaCaT keratinocytes **</td>
<td>[639] [378] [338]</td>
</tr>
<tr>
<td>Smad4</td>
<td>TGFβ</td>
<td>100.000-900.000</td>
<td>HaCaT keratinocytes **</td>
<td>[639] [378]</td>
</tr>
<tr>
<td>Calmodulin</td>
<td>Ca2+</td>
<td>5.000.000-10.000.000</td>
<td>PC12 cells **</td>
<td>[640]</td>
</tr>
<tr>
<td>Bcl-2</td>
<td>Apoptosis</td>
<td>50.000</td>
<td>Peripheral Blood</td>
<td>[641]</td>
</tr>
</tbody>
</table>
About 4000 genes are annotated with the term signal transduction in Ensembl. Assuming a typical signalling protein expression of 20,000 molecules per cell (see Table and [426,622]), this would sum up to ~4% of total protein. Taken together, these data strongly suggest that the energy optimisation principle discussed in Section 5 is physiologically relevant.

* See also:
Alvis Brazma, Helen Parkinson, Thomas Schiltt and Mohammadreza Shojatalab: Basic Biology (http://www.ebi.ac.uk/2can/biology/molecules_proteins5.html)


### E.3 RFIs act at multiple levels of the network

The analysis in Section 5 suggested that energy minimisation might have contributed to the dominance of negative feedback regulation via RFI induction over negative feedback via signal transducer downregulation. More specifically, it was reasoned that rapid transcriptional negative feedback regulation requires unstable proteins. Therefore, signal transducer downregulation is an energetically unfavourable mode of negative regulation, as constitutive expression of unstable signal transducers consumes large amounts of energy in the basal state. In contrast, rapidly turning over an inhibitor that is selectively induced stimulation will only require energy over short time scales.

Of course, this energetic advantage of RFI induction over transducer downregulation will be more pronounced the more nodes in the signalling network are affected by transcriptional...
feedback. The results in Section 5 already indicated that RFI-mediated feedback operates at multiple levels (Fig. 5.2). However, the analysis was limited to published microarray results (and thus to particular cell types), and only proteins, where both half-life and gene expression data was available, were considered. Thus, cell-type specific and less well studied feedback regulators might have been missed in the analysis.

To further substantiate the energy argument, known transcriptional negative feedback regulators of growth-factor-induced MAPK signalling and of TGFβ-induced Smad signalling were collected from the literature. The results, summarised in Tables E.7 and E.8, reveal that transcriptional negative feedback operates at almost all levels of the considered cascade, and even single layers are regulated in different ways (e.g., feedback modulates both receptor activity and degradation). Moreover, it seems that transcriptional negative feedback loops often operate in a cell-type-specific manner.

Taken together, Tables E.7 and E.8 reveal that multiple nodes are subject to transcriptional feedback regulation for each signalling pathway. Flexible, cell-type-specific negative feedback regulation by signal transducer downregulation would thus require that most of the transducers are unstable, and therefore would represent a strong energy burden. This suggests that an energy minimisation principle might have contributed to an evolutionary selection pressure favouring transcriptional feedback regulation via RFI induction.

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Mechanism of action</th>
<th>Cell line</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lefty-2</td>
<td>Ligand inhibition</td>
<td>Development</td>
<td>[651]</td>
</tr>
<tr>
<td>Follistatin</td>
<td>Ligand inhibition</td>
<td>HepG2</td>
<td>[652]</td>
</tr>
<tr>
<td>BAMBI</td>
<td>Receptor inhibition</td>
<td>HepG2</td>
<td>[653,654]</td>
</tr>
<tr>
<td>Smad7</td>
<td>Receptor inhibition</td>
<td>various</td>
<td>[655]</td>
</tr>
<tr>
<td>TMEPAI</td>
<td>Receptor inhibition</td>
<td>AML, NMuMG</td>
<td>*</td>
</tr>
<tr>
<td>Smad7</td>
<td>Receptor degradation</td>
<td>various</td>
<td>[316]</td>
</tr>
<tr>
<td>Smurf2</td>
<td>Receptor degradation</td>
<td>HepG2</td>
<td>[656]</td>
</tr>
<tr>
<td>GRK2</td>
<td>Regulation of Smad2/3/4 accessibility</td>
<td>Primary hepatocytes</td>
<td>[420]</td>
</tr>
<tr>
<td>CM184</td>
<td>Regulation of Smad2/3/4 shuttling</td>
<td>Mv1Lu cells</td>
<td>[657]</td>
</tr>
<tr>
<td>SnoN</td>
<td>Regulation of Smad2/3/4 shuttling</td>
<td>Various primary cells</td>
<td>[343,349]</td>
</tr>
<tr>
<td>Ski</td>
<td>Co-Repressor for Smad2/3/4</td>
<td>Fibroblasts</td>
<td>[658]</td>
</tr>
<tr>
<td>TGIF</td>
<td>Co-Repressor for Smad2/3/4</td>
<td>HuT78</td>
<td>[659]</td>
</tr>
<tr>
<td>SnoN</td>
<td>Co-Repressor for Smad2/3/4</td>
<td>various</td>
<td>[343,349]</td>
</tr>
<tr>
<td>Smad7</td>
<td>Competition with Smad2/3/4 for DNA</td>
<td>Hep3B</td>
<td>[660]</td>
</tr>
<tr>
<td>Smurf2</td>
<td>Smad2/3/4 degradation</td>
<td>HepG2</td>
<td>[656]</td>
</tr>
</tbody>
</table>

Listed are literature studies which reported that negative regulators of TGFβ signalling are transcriptionally induced by TGFβ. The mechanism of negative feedback regulation is indicated as well. The reference marked by a star (*) is available online: Nature Precedings <http://hdl.handle.net/10101/npre.2007.1403.1> (2007).
Table E.8: Transcriptional negative feedback loops in growth factor signalling via the Erk-MAPK cascade.

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Mechanism of action</th>
<th>Cell line</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argos</td>
<td>Ligand inhibition</td>
<td>Drosophila Development</td>
<td>[661]</td>
</tr>
<tr>
<td>RALT</td>
<td>Receptor inhibition</td>
<td>32D</td>
<td>[662]</td>
</tr>
<tr>
<td>Sef</td>
<td>Receptor inhibition</td>
<td>various</td>
<td>[663,664]</td>
</tr>
<tr>
<td>LRIG1</td>
<td>Receptor degradation</td>
<td>HeLa</td>
<td>[535]</td>
</tr>
<tr>
<td>Sprouty1 &amp; 2</td>
<td>Adaptor (Grb2) inhibition</td>
<td>various</td>
<td>[665,666]</td>
</tr>
<tr>
<td>Sprouty2 &amp; 4</td>
<td>Raf inhibition</td>
<td>various</td>
<td>[665,667,668]</td>
</tr>
<tr>
<td>Sef</td>
<td>Inhibition of Mek-mediated Erk phosphorylation</td>
<td>various</td>
<td>[669]</td>
</tr>
<tr>
<td>MKP-1</td>
<td>Erk dephosphorylation (nucleus)</td>
<td>various</td>
<td>[670]</td>
</tr>
<tr>
<td>MKP-3</td>
<td>Erk dephosphorylation (cytosol)</td>
<td>various</td>
<td>[670]</td>
</tr>
<tr>
<td>Sef</td>
<td>Erk sequestration in the cytosol</td>
<td>various</td>
<td>[671]</td>
</tr>
<tr>
<td>MKP-1/2 (?)</td>
<td>Sequestration of inactive Erk in the nucleus</td>
<td>various</td>
<td>[672]</td>
</tr>
</tbody>
</table>

Listed are literature studies which reported that negative regulators of RTK-Ras-MAPK signalling are transcriptionally induced by growth factor stimulation. The mechanism of negative feedback regulation is indicated as well.
### E.4 Graphical summary of the data used

#### Figure E.1 Gene expression profiles and half-lives of growth factor signalling via the Erk-MAPK cascade.

Gene expression data from [67,471,673] was expressed as log₂-fold change and the time courses were plotted as heatmaps. mRNA half-lives were taken from genome-wide studies [278,279] and protein half-lives were taken from Tables E.1 – E.5. The star (*) indicates that Sprouty1 was shown to be rapidly induced upon Erk activation [674]. Sprouty 1 was thus marked as induced in the pathway map (Fig. 5.2), but was considered unchanged in the statistical analysis and in Fig. 5.1.
Figure E.2 Gene expression profiles and half-lives of TGFβ signalling via Smad transcription factors.

Gene expression data from [675] and unpublished gene expression data (from GEO repository) was expressed as log₂-fold change and the time courses were plotted as heatmaps. mRNA half-lives were taken from genome-wide studies [278,279] and protein half-lives were taken from Tables E.1 – E.5. The star (*) indicates that TGIF was shown to be rapidly induced upon TGFβ stimulation [659], and (**) indicates that BAMBI was shown to be rapidly induced upon TGFβ stimulation [654]. TGIF and BAMBI were thus marked as induced in the pathway map (Fig. 5.2), but were considered unchanged in the statistical analysis and in Fig. 5.1.
**Figure E.3** Gene expression profiles and half-lives of growth factor signalling via the PI3K-Akt cascade.

Gene expression data from [67, 471, 676] was expressed as log₂-fold change and the time courses were plotted as heatmaps. mRNA half-lives were taken from genome-wide studies [278, 279] and protein half-lives were taken from Tables E.1 – E.5.
Figure E.4 Gene expression profiles and half-lives of JAK/STAT signalling.

Gene expression data from [677, 678, 679, 680] was expressed as log₂-fold change and the time courses were plotted as heatmaps. mRNA half-lives were taken from genome-wide studies [278, 279] and protein half-lives were taken from Tables E.1 – E.5. The star (*) indicates that SOCS-5 was shown to be rapidly induced upon stimulation of JAK/STAT signalling [595, 681]. SOCS-5 was thus marked as induced in the pathway map (Fig. 5.2), but was considered unchanged in the statistical analysis and in Fig. 5.1.
Figure E.5 Gene expression profiles and half-lives of cAMP signalling.

Gene expression data from [682] was expressed as log2-fold change and the time courses were plotted as heatmaps. mRNA half-lives were taken from genome-wide studies [278,279] and protein half-lives were taken from Tables E.1 – E.5.
Figure E.6 Gene expression profiles of cAMP signalling in yeast.

Log fold-change after oxidative stress induced by cumene hydroperoxide. Data from Sha, Martins, Laubenbacher, Mendes, Shulaev (submitted), available as GSE7645. The inhibitor RGS2 is upregulated on a time scale below cell-cycle time. S.I. is short for signal inhibitor.
F Additional material on transcriptional feedback regulation of TGFβ signalling

F.1 Differential equations of the model

\[
\begin{align*}
S2 \text{ mRNA}^*: & \quad \dot{x}_1 = -v_1 \\
S4 \text{ mRNA}^*: & \quad \dot{x}_2 = -v_3 \\
R \text{ mRNA}^*: & \quad \dot{x}_3 = -v_5 \\
\text{SnoN mRNA}^*: & \quad \dot{x}_4 = -v_{43} \\
S2: & \quad \dot{x}_5 = v_2 - v_{16} - v_{21} + v_{31} - v_{33} \\
S4: & \quad \dot{x}_6 = v_4 - v_{17} + v_{19} - v_{25} + v_{22} - v_{41} \\
R_4: & \quad \dot{x}_7 = v_6 + v_{11} - v_{12} - v_{15} \\
T: & \quad \dot{x}_8 = -v_7 + v_8 \\
R: & \quad \dot{x}_9 = -v_7 + v_8 + v_{10} - v_{11} + v_{12} \\
TR: & \quad \dot{x}_{10} = v_7 - v_8 - v_9 \\
TR_4: & \quad \dot{x}_{11} = v_9 - v_{10} - v_{13} \\
TR_6: & \quad \dot{x}_{12} = v_{13} - v_{14} \\
pS2: & \quad \dot{x}_{13} = v_{16} - v_{17} - v_{18} - v_{18} - v_{17} + v_{19} + v_{20} + v_{20} - v_{22} - v_{35} \\
S4\sim(pS2)_2: & \quad \dot{x}_{14} = v_{17} - v_{19} - v_{23} - v_{30} \\
(pS2)_3: & \quad \dot{x}_{15} = v_{18} - v_{20} - v_{24} - v_{37} \\
nS2: & \quad \dot{x}_{16} = v_{21} + v_{30} - v_{31} - v_{34} \\
npS2: & \quad \dot{x}_{17} = v_{22} + v_{26} + v_{26} + v_{27} + v_{27} + v_{28} - v_{28} - v_{29} - v_{29} - v_{30} - v_{36} \\
nS4\sim(npS2)_2: & \quad \dot{x}_{18} = v_{24} - v_{26} + v_{28} - v_{40} + v_{22} + v_{33} + v_{54} \\
(npS2)_3: & \quad \dot{x}_{19} = v_{23} - v_{25} + v_{29} - v_{30} - v_{55} + v_{56} \\
nS4: & \quad \dot{x}_{20} = v_{25} + v_{36} - v_{38} - v_{42} \\
\text{SnoN:} & \quad \dot{x}_{21} = v_{44} + v_{50} - v_{51} - v_{52} + v_{53} - v_{55} + v_{56} \\
\text{SnoN Delay1}: & \quad \dot{x}_{22} = v_{45} - v_{46} \\
\text{SnoN~nS4\sim(npS2)_2}: & \quad \dot{x}_{23} = v_{52} - v_{53} - v_{54} \\
\text{SnoN\sim(npS2)_3}: & \quad \dot{x}_{24} = v_{55} - v_{56} - v_{57} \\
\text{SnoN Delay2}: & \quad \dot{x}_{25} = v_{46} - v_{47} \\
\text{SnoN Delay3}: & \quad \dot{x}_{26} = v_{47} - v_{48} \\
\text{SnoN Delay4}: & \quad \dot{x}_{27} = v_{48} - v_{49} \\
\text{SnoN Delay5}: & \quad \dot{x}_{28} = v_{49} - v_{50}
\end{align*}
\]


The molecular species depicted in Figure 6.4 and nine additional auxiliary species (asterisks) are listed as variables \(x_1 - x_{28}\), and the corresponding ODEs are indicated. For definition of reaction velocities \(v\), see Equations F.2. The auxiliary variables are necessary to model inhibition of transcription by actinomycin D and to account for a delay in SnoN protein biosynthesis. In the model, all species were formally rescaled to cytoplasmic volume, which explains why the nuclear, cytoplasmic and medium volumes enter some of the velocities in Eq. F.2 [883]. The initial concentrations of the auxiliary mRNA species were set to unity (i.e., \(x_9 = x_9 = x_4 = 1\)), the initial amount of TGFβ is set by the input function \(u_1\) (\(x_8 = u_1\); see Section 6.5.4), and the initial concentrations of \(x_6, x_8, x_9, x_{16}, x_{20}\) and \(x_{21}\) are determined by the kinetic parameters (Section F.3).
Equations F.2. Reaction velocities.

The rates used for the formulation of the differential equations (see Equations F.1) are listed as $v_1 - v_{57}$. The parameters $k_6 - k_{41}$ are fitted and defined in Tables F.1 – F.3. The remaining parameters ($k_1 - k_5$) are not subject to fitting and are defined as:

- ($k_1$) degree of SnoN depletion (see Section 6.5.4),
- ($k_2$) cytoplasmic volume (set to 10 pl),
- ($k_3$) nuclear volume (set to 0.5 pl),
- ($k_4$) volume of cell culture medium per cell (set to 1000 pl),
- ($k_5$) degradation of cell-surface receptors (set to 0).
### F.2 Kinetic parameters and ranges used for fitting

Table F.1: Parameters for the receptor trafficking module.

<table>
<thead>
<tr>
<th>#</th>
<th>Reaction(s)</th>
<th>Parameter description</th>
<th>Experimentally measured value</th>
<th>References</th>
<th>Range used for fitting</th>
<th>Value in best-fit model</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>$T + R \leftrightarrow TR$</td>
<td>Dissociation constant ($k_{off} / k_{on}$)</td>
<td>0.001 - 0.22 nM</td>
<td>[314] [313] [305]</td>
<td>0.001 - 0.22 nM (in the model rescaled to cytoplasmic volume)</td>
<td>0.0098 nM</td>
</tr>
<tr>
<td>22</td>
<td>$T + R \leftrightarrow TR$</td>
<td>Off-rate ($k_{off}$) (typical range for receptor-ligand off-rates)</td>
<td>0.01 - 1 min$^{-1}$</td>
<td>[684]</td>
<td>0.01 - 1 min$^{-1}$</td>
<td>0.1125 min$^{-1}$</td>
</tr>
<tr>
<td>30</td>
<td>$TR \rightarrow TR_e$ $- R \rightarrow R_e$</td>
<td>Rate constant for endocytosis ($k_w$)</td>
<td>$\sim$ 0.2 min$^{-1}$</td>
<td>[312] [310]</td>
<td>0.1 - 0.5 min$^{-1}$</td>
<td>0.4040 min$^{-1}$</td>
</tr>
<tr>
<td>30</td>
<td>$TR_e \rightarrow R$ $- R_e \rightarrow R$</td>
<td>Rate constant for recycling ($k_r$)</td>
<td>$\sim$ 0.1 min$^{-1}$ $\sim$ 50 - 90% of the total cellular TGFβ receptor pool resides in the endosome</td>
<td>[312] [309] [310]</td>
<td>Recycling rate constant is 1-10x smaller than the endocytosis rate constant ($1 &lt; k_21 &lt; 10$)</td>
<td>0.4040 min$^{-1}$</td>
</tr>
<tr>
<td>19</td>
<td>$R_e \rightarrow$</td>
<td>Constitutive receptor degradation</td>
<td>0.0083 min$^{-1}$ 0.0067 min$^{-1}$</td>
<td>[685] [686]</td>
<td>0.003 - 0.03 min$^{-1}$</td>
<td>0.0124 min$^{-1}$</td>
</tr>
<tr>
<td>24</td>
<td>$TR_e \rightarrow TR_w$</td>
<td>1st order rate constant for progression from early to late endosomes</td>
<td>$\sim$ 0.2 min$^{-1}$ (value measured for LDL and for α2M)</td>
<td>[687]</td>
<td>0.03 - 1 min$^{-1}$</td>
<td>0.6667 min$^{-1}$</td>
</tr>
<tr>
<td>26</td>
<td>$TR_e \rightarrow$</td>
<td>1st order rate constant for targeting to lysosome and thus for degradation</td>
<td>$\sim$ 0.007 min$^{-1}$ (apparent rate constant for lysosomal TGFβ degradation in HepG2 and Hep3B cells)</td>
<td>[314]</td>
<td>0.001 - 0.05 min$^{-1}$</td>
<td>0.0051 min$^{-1}$</td>
</tr>
<tr>
<td>18</td>
<td>$R_e \rightarrow R_e$</td>
<td>Total number of TGFβ receptors per cell (determines receptor synthesis rate)</td>
<td>Cell-surface receptor number: 2,000 per rat hepatocyte 4,500 per HepG2 cell 1,500 per Hep3B cell $\sim$ 50 - 90% of the total cellular TGFβ receptor pool resides in the endosome</td>
<td>[290] [314] [310]</td>
<td>3,000 - 30,000 per cell</td>
<td>3104 per cell</td>
</tr>
<tr>
<td>17</td>
<td>$R_e \rightarrow R_e$</td>
<td>Half-life of TGFβ receptor mRNA (determines actinomycin effect on receptor synthesis)</td>
<td>0.0024 min$^{-1}$</td>
<td>[685]</td>
<td>0.0005 - 0.0015 min$^{-1}$</td>
<td>0.0005 min$^{-1}$</td>
</tr>
</tbody>
</table>

The kinetic parameters are numbered according to Equations F.1 and F.2 (column #). Based on experimentally measured values, a parameter range allowed during the fitting procedure was defined. The parameter value in the best-fit model is given on the right.
Table F.2: Parameters for the Smad activation module.

<table>
<thead>
<tr>
<th>#</th>
<th>Reaction(s)</th>
<th>Parameter description</th>
<th>Experimentally measured value</th>
<th>References</th>
<th>Range used for fitting</th>
<th>Value in best-fit model</th>
</tr>
</thead>
<tbody>
<tr>
<td>k26</td>
<td>S2 → pS2</td>
<td>Catalytic rate constant ((k_{cat})) (typical (k_{cat}) of kinases for full-length substrates)</td>
<td>1 - 600 min(^{-1})</td>
<td>([159,688,689])</td>
<td>1 - 600 min(^{-1})</td>
<td>1.1391 min(^{-1})</td>
</tr>
</tbody>
</table>
| k27 | S2 → pS2    | Michaelis-Menten constant \((K_M)\) (typical \(K_M\) of kinases for full-length substrates) | 0.2 - 50 µM | SARA recruits Smad2 to the membrane, thereby increasing Smad2 affinity for the transmembrane receptor by \(
\sim 1,000\)-fold | \([159,688,689]\) | 0.2 - 50,000 nM \((\mu M)\) | 0.2006 nM |
| k28 | 2 pS2 + S4 ↔ \((pS2)_2S4\) 2 npS2 + nS4 ↔ \((npS2)_2nS4\) | Monomer-heterotrimer dissociation constant \((k_{off} / k_{on})\) | \(~ 1,500 \) nM\(^2\) | \([365]\) \([334]\) | 100 - 10,000 nM\(^2\) | 9983.91 nM\(^2\) |
| k29 | 2 pS2 + S4 ↔ \((pS2)_2S4\) 2 npS2 + nS4 ↔ \((npS2)_2nS4\) | Monomer-heterotrimer dissociation rate constant \((k_{off})\) | 0.1 - 600 min\(^{-1}\) (typical range for off-rates of intracellular protein-protein interactions) | \([254,690,691]\) | 0.1 - 600 min\(^{-1}\) | 0.1029 min\(^{-1}\) |
| k30 | 3 pS2 ↔ \((pS2)_3\) 3 npS2 ↔ \((npS2)_3\) | Monomer-homotrimer dissociation constant \((k_{off} / k_{on})\) | Heterotrimer formation is slightly more efficient than homotrimer formation in vitro Heterotrimer formation can be significantly enhanced by Smad4 ubiquitination in vivo | \([334]\) \([336]\) \([335]\) \([337]\) | Affinity of homotrimer is 1-10x weaker than that of heterotrimerisation \((1 < k_{30} < 10)\) | 99830.9 nM\(^2\) |
| k31 | 3 pS2 ↔ \((pS2)_3\) 3 npS2 ↔ \((npS2)_3\) | Monomer-homotrimer dissociation rate constant \((k_{off})\) | Homotrimer dissociation is assumed to proceed with the same rate constant as heterotrimer dissociation | - | - | 0.1029 min\(^{-1}\) |
| k32 | S2 → nS2 - pS2 → npS2 \((pS2)_2S4 → (npS2)_2nS4\) \((pS2)_2 \rightarrow (npS2)_2\) | 1st-order import rate constant | \(0.16 \) min\(^{-1}\) (based on FRAP measurements in HaCaT cells) | \([339]\) \([336]\) \([340]\) | 0.07 - 0.7 min\(^{-1}\) | 0.6973 min\(^{-1}\) |
| k33 | nS2 → S2 | 1st-order export rate constant | \(1 \) min\(^{-1}\) (based on FRAP measurements in unstimulated HaCaT cells, where 85% of Smad2 is cytosolic) Smad2 is roughly equally distributed between cytoplasm and nucleus in primary mouse hepatocytes | \([311,339,364]\) | 1 - 10x import rate \((1 < k_9 < 10)\) | 1.1993 min\(^{-1}\) |
| $k_{16}$ | $\text{S4} \rightarrow \text{nS4}$ | 1st-order import rate constant | $0.08 \text{ min}^{-1}$ (based on FRAP measurements in unstimulated HaCaT cells, where 90% of Smad4 is cytosolic) | [311,339] | 0.03 - 0.3 min$^{-1}$ | 0.0301 min$^{-1}$ |
| $k_{15}$, $k_{18}$ | $\text{nS4} \rightarrow \text{S4}$ | 1st-order export rate constant | $0.5 \text{ min}^{-1}$ (based on FRAP measurements in HaCaT cells) | [311,339] | 1 - 10x import rate ($1 < k_{15} < 10$) | 0.3009 min$^{-1}$ |
| $k_{31}$ | $\text{npS2} \rightarrow \text{S2}$ | $V_{\text{max}} / K_{\text{M}}$ of nuclear Smad2 dephosphorylation | Smad dephosphorylation after administration of a TGFβ receptor inhibitor occurs with a half-life of 16 - 20 min - 0.5 - 600 min$^{-1}$ (typical $V_{\text{max}}/K_{\text{M}}$ reported for other phosphatases) | [342] [374] [154] [149] | 0.05 - 500 min$^{-1}$ | 0.3601 min$^{-1}$ |
| $k_{7}$ | $\rightarrow \text{S2}$ | Total number of Smad2 molecules per cell (determines Smad2 synthesis rate) | 80,000 per primary mouse hepatocyte | This study | 50,000 - 120,000 per cell | 51269 per cell |
| $k_{8}$ | $\text{S2} \rightarrow$ | Half-life of Smad2 mRNA (determines actinomycin effect on Smad2 synthesis) | 0.0049 min$^{-1}$ 0.0023 min$^{-1}$ 0.0079 min$^{-1}$ 0.0068 min$^{-1}$ 0.0185 min$^{-1}$ | [278] [279] | 0.003 - 0.03 min$^{-1}$ | 0.0273 min$^{-1}$ |
| $k_{9}$ | $\text{S2} \rightarrow$ | Half-life of cytosolic Smad2-containing species | 0.0017 - 0.0056 min$^{-1}$ $< 0.0019$ min$^{-1}$ | [345] [338] | 0.0005 - 0.005 min$^{-1}$ | 0.0024 min$^{-1}$ |
| $k_{10}$ | $\text{nS2} \rightarrow$ | Half-life of nuclear Smad2-containing species | 0.0017 - 0.0056 min$^{-1}$ $< 0.0019$ min$^{-1}$ | [345] [338] | 0.0005 - 0.005 min$^{-1}$ | 0.0005 min$^{-1}$ |
| $k_{13}$ | $\rightarrow \text{S4}$ | Total number of Smad4 molecules per cell (determines Smad4 synthesis rate) | 80,000 per primary mouse hepatocyte | This study | 50,000 - 120,000 per cell | 51478.7 per cell |
| $k_{12}$ | $\rightarrow \text{S4}$ | Half-life of Smad4 mRNA (determines actinomycin effect on Smad4 synthesis) | 0.0245 min$^{-1}$ 0.0216 min$^{-1}$ 0.0256 min$^{-1}$ 0.0245 min$^{-1}$ | [278] [279] | 0.005 - 0.05 min$^{-1}$ | 0.0049 min$^{-1}$ |
| $k_{14}$ | $\text{S4} \rightarrow$ | Half-life of Smad4 protein | $< 0.0019$ min$^{-1}$ 0.0026 min$^{-1}$ $< 0.0007$ min$^{-1}$ | [338] [521] [337] | 0.0005 - 0.005 min$^{-1}$ | 0.0005 min$^{-1}$ |

The kinetic parameters are numbered according to Equations F.1 and F.2 (column #). Based on experimentally measured values, a parameter range allowed during the fitting procedure was defined. The parameter value in the best-fit model is given on the right. The superscript (a) indicates that the anchor protein SARA is assumed to recruit Smad2 to the membrane [327] which increases the affinity of Smad2 for the TGFβ receptor by approximately a factor of 1000 [329].
Table F.3: Parameters for the SnoN feedback module.

<table>
<thead>
<tr>
<th>#</th>
<th>Reaction(s)</th>
<th>Parameter description</th>
<th>Experimentally measured value</th>
<th>References</th>
<th>Range used for fitting</th>
<th>Value in best-fit model</th>
</tr>
</thead>
<tbody>
<tr>
<td>k33</td>
<td>→ SnoN</td>
<td>Total number of SnoN molecules per cell (determines SnoN synthesis rate)</td>
<td>600 per (unstimulated) primary mouse hepatocyte</td>
<td>This study</td>
<td>300 - 1,500 per cell</td>
<td>1489 per cell</td>
</tr>
<tr>
<td>k36</td>
<td>→ SnoNprecursor</td>
<td>$K_m$ of SnoN Induction by Smad Trimmers</td>
<td>Affinity of Smad trimers for target DNA sequences is likely to be in the nM range or even in the subnanomolar range</td>
<td>[353,354,355, 356]</td>
<td>0.1 - 100 nM (in the model rescaled to cytoplasmic volume)</td>
<td>1.0365 nM</td>
</tr>
<tr>
<td>k38</td>
<td>→ SnoNprecursor</td>
<td>$V_{max}$ of SnoN Induction by Smad trimmers</td>
<td>Maximal protein biosynthesis rate is 150,000 / cell / min for Actin (= 30 nM min$^{-1}$)</td>
<td>[692]</td>
<td>0.003 - 300 nM min$^{-1}$</td>
<td>0.0298 nM min$^{-1}$</td>
</tr>
<tr>
<td>k35</td>
<td>→ SnoNprecursor</td>
<td>Half-life of SnoN mRNA (determines actinomycin effect on SnoN synthesis)</td>
<td>$-0.01$ min$^{-1}$</td>
<td>This study (not shown)</td>
<td>0.005 - 0.02 min$^{-1}$</td>
<td>0.005 min$^{-1}$</td>
</tr>
<tr>
<td>k37</td>
<td>SnoNprecursor → SnoN</td>
<td>Delay in SnoN production due to protein biosynthesis</td>
<td>Protein biosynthesis takes 1 - 25 min</td>
<td>Estimation</td>
<td>Modelled as a multistep process with the delay $\tau = 5k - k = 0.2 - 5$ min$^{-1}$</td>
<td>0.2000 min$^{-1}$</td>
</tr>
<tr>
<td>k34</td>
<td>SnoN → SnoN</td>
<td>SnoN Degradation rate constant in unstimulated cells</td>
<td>0.017 min$^{-1}$, 0.0033 min$^{-1}$</td>
<td>[358], [343]</td>
<td>0.003 - 0.02 min$^{-1}$</td>
<td>0.0039 min$^{-1}$</td>
</tr>
<tr>
<td>k38</td>
<td>SnoN + (npS2)$_2$nS4 ↔ (npS2)$_2$nS4~SnoN</td>
<td>Dissociation constant of the heterotrimer-SnoN complex ($k_{off}$ / $k_{on}$)</td>
<td>0.3 – 1,000 nM (typical dissociation constants of intracellular high-affinity protein-protein interactions)</td>
<td>[254,690,691]</td>
<td>0.3 - 1,000 nM (in the model rescaled to cytoplasmic volume)</td>
<td>5.3413 nM</td>
</tr>
<tr>
<td>k39</td>
<td>SnoN + (npS2)$_2$nS4 ↔ (npS2)$_2$nS4~SnoN</td>
<td>Dissociation rate constant of the heterotrimer-SnoN complex ($k_{offs}$)</td>
<td>0.1 - 600 min$^{-1}$ (typical range for off-rates of intracellular protein-protein interactions)</td>
<td>[254,690,691]</td>
<td>0.1 - 600 min$^{-1}$</td>
<td>1.0466 min$^{-1}$</td>
</tr>
<tr>
<td>k41</td>
<td>SnoN + (npS2)$_3$</td>
<td>Dissociation constant of the homotrimer-SnoN complex ($k_{offs}$)</td>
<td>Affinity of SnoN–homotrimer 0.3 - 3x different from the affinity of SnoN–heterotrimer (0.3 &lt; $k_{offs}$ &lt; 30)</td>
<td>[360]</td>
<td>15.1479 nM</td>
<td>15.1479 nM</td>
</tr>
<tr>
<td>k40</td>
<td>(npS2)$_3$~SnoN → SnoN + (npS2)$_2$nS4 ↔ (npS2)$_2$nS4</td>
<td>SnoN dissociation from homotrimers is assumed to proceed with the same rate constant as dissociation from heterotrimers</td>
<td>SnoN-related Ski protein binds Smad homo- and heterotrimers with similar affinity</td>
<td>[360]</td>
<td>Affinity of SnoN–homotrimer 0.3 - 3x different from the affinity of SnoN–heterotrimer (0.3 &lt; $k_{offs}$ &lt; 30)</td>
<td>15.1479 nM</td>
</tr>
<tr>
<td>k32</td>
<td>SnoN + (npS2)$_3$ ↔ (npS2)$_2$nS4~SnoN</td>
<td>Dissociation rate constant of the homotrimer-SnoN complex ($k_{offs}$)</td>
<td>SnoN dissociation from homotrimers is assumed to proceed with the same rate constant as dissociation from heterotrimers</td>
<td>-</td>
<td>-</td>
<td>1.0466 min$^{-1}$</td>
</tr>
<tr>
<td>k34</td>
<td>SnoN + (npS2)$_3$ ↔ (npS2)$_2$nS4~SnoN</td>
<td>SnoN degradation rate constant in stimulated cells (reflects degradation rate constant of Smad-SnoN complexes)</td>
<td>0.083 min$^{-1}$, 0.0333 min$^{-1}$</td>
<td>[358], [343]</td>
<td>0.083 - 300 nM min$^{-1}$</td>
<td>0.1406 min$^{-1}$</td>
</tr>
</tbody>
</table>
The kinetic parameters are numbered according to Equations F.1 and F.2 (column #). Based on experimentally measured values, a parameter range allowed during the fitting procedure was defined. The parameter value in the best-fit model is given on the right.

F.3 Modelling Basal Protein Expression

Several protein species in the model (S2, nS2, S4, nS4, SnoN, R and Re) are expressed at non-zero levels even in the absence of stimulation. The basal protein concentrations of these proteins are determined by the kinetic parameters. For instance, the expression levels of nuclear and cytosolic Smad2 depend on the nuclear import and export rates, and on the nuclear and the cytosolic degradation rate, respectively.

**Figure F.1. Schematic representation of protein expression in the basal state.**

The species C1 is subject to synthesis and degradation, and is additionally interconverted into the species C2, which is then eventually degraded. The TGFβ model species corresponding to C1 and C2 in the are given in Table F.4.

**Table F.4: Basally expressed proteins in the TGFβ model.**

<table>
<thead>
<tr>
<th>Protein</th>
<th>C1</th>
<th>C2</th>
<th>$k_{\text{syn}}$</th>
<th>$k_{\text{deg,C1}}$</th>
<th>$k_{C1\rightarrow C2}$</th>
<th>$k_{C2\rightarrow C1}$</th>
<th>$k_{\text{deg,C2}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smad2</td>
<td>S2</td>
<td>nS2</td>
<td>Synthesis Rate</td>
<td>Cytosolic Degradation</td>
<td>Nuclear Import</td>
<td>Nuclear Export</td>
<td>Nuclear Degradation</td>
</tr>
<tr>
<td>Smad4</td>
<td>S4</td>
<td>nS4</td>
<td>Synthesis Rate</td>
<td>Cytosolic Degradation</td>
<td>Nuclear Import</td>
<td>Nuclear Export</td>
<td>Nuclear Degradation</td>
</tr>
<tr>
<td>Receptor</td>
<td>Re</td>
<td>R</td>
<td>Synthesis Rate</td>
<td>Endosomal Degradation</td>
<td>Recycling</td>
<td>Endocytosis</td>
<td>- (a)</td>
</tr>
<tr>
<td>SnoN</td>
<td>SnoN</td>
<td>-</td>
<td>Synthesis Rate</td>
<td>Nuclear Degradation</td>
<td>- (a)</td>
<td>- (a)</td>
<td>- (a)</td>
</tr>
</tbody>
</table>

Listed are the species in the TGFβ model corresponding to C1 and C2 in Fig. F.1, and the respective kinetic parameters are also given. The superscript (a) indicates that no cell-surface receptor degradation occurs in the model, and superscript (b) indicates that the cytoplasmic SnoN pool was neglected in the model.

**Figure F.1** is a general representation of protein synthesis and degradation, which applies to all basally expressed proteins in the model. Table F.4 summarises the relationship of the parameters given in Fig. F.1 to those in the TGFβ model for each protein. The steady state solution of the model depicted in Fig. F.1 is given by:

$$C_1 = C_\text{tot} \cdot \frac{(k_{\text{deg,C2}} + k_{C2\rightarrow C1})}{k_{\text{deg,C2}} + k_{C2\rightarrow C1} + k_{C1\rightarrow C2}} \quad \text{and} \quad C_2 = C_\text{tot} \cdot \frac{k_{C1\rightarrow C2}}{k_{\text{deg,C2}} + k_{C2\rightarrow C1} + k_{C1\rightarrow C2}} \quad (F.3)$$

In these equations, the synthesis rate $k_{\text{syn}}$ (Fig. F.1) was replaced by the total protein expression level ($C_\text{tot} = C_1 + C_2$), which was observed experimentally for most of the proteins (Fig. 6.3). Equation F.3 and Table F.4 were used to calculate the initial conditions (i.e., the basal protein expression) for the TGFβ model. The protein synthesis rate in the differential equations of the model is given by the following relation:

162
\[ k_{\text{eq}} = \frac{C_{\text{int}} \cdot (k_{\text{arg,C1}} \cdot k_{\text{deg,C2}} + k_{\text{arg,C1}} \cdot k_{\text{C1,C2}} - k_{\text{deg,C1}} \cdot k_{\text{C2,C1}})}{k_{\text{arg,C2}} \cdot k_{\text{C2,C1}} + k_{\text{deg,C2}}} \]  

(F.4)

### F.4 Modelling TGFβ signalling in HaCaT cells

The HaCaT cell model, schematically depicted in Fig. 6.15, contains most of the reaction steps of the hepatocyte model, and the majority of kinetic parameters was assumed to be equal in both models as well. SnoN-mediated feedback described by the velocities \( v_{44} \rightarrow v_{57} \) in Eq. F.1 was completely eliminated from the HaCaT model, as HaCaT cells seem to be devoid of transcriptional feedback (see main text). Some protein synthesis and degradation rates in the HaCaT model were assumed to be different from the hepatocyte model, as summarised in Table F.5. Washout and TGFβ receptor inhibitor experiments in Fig. 6.15 were modelled as described for the hepatocyte model (Section 6.5.4). MG132 treatment was modelled by setting the degradation rates for all Smad2, Smad4 and receptor species to zero.

### Table F.5: Fitted parameters in the HaCaT cell model.

<table>
<thead>
<tr>
<th>#</th>
<th>Reaction(s)</th>
<th>Parameter description</th>
<th>Range used for fitting</th>
<th>Value in HaCaT model 1</th>
<th>Value in HaCaT model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>k19</td>
<td>( R_e \rightarrow )</td>
<td>Constitutive receptor degradation</td>
<td>0.003 - 0.035 min(^{-1})</td>
<td>0.035 min(^{-1})</td>
<td>0.024 min(^{-1})</td>
</tr>
<tr>
<td>k25</td>
<td>( TR_e \rightarrow )</td>
<td>1st order rate constant for targeting to lysosome and thus for degradation</td>
<td>0.001 – 0.05 min(^{-1})</td>
<td>0.002 min(^{-1})</td>
<td>0.0029 min(^{-1})</td>
</tr>
<tr>
<td>k18</td>
<td>( \rightarrow R_e )</td>
<td>Total number of TGFβ receptors per cell (determines receptor synthesis rate)</td>
<td>3,000 - 300,000 per cell</td>
<td>266,564 per cell</td>
<td>297,994 per cell</td>
</tr>
<tr>
<td>k7</td>
<td>( \rightarrow S2 )</td>
<td>Total number of Smad2 molecules per cell (determines Smad2 synthesis rate)</td>
<td>300,000 - 700,000 per cell</td>
<td>582,328 per cell</td>
<td>485,508 per cell</td>
</tr>
<tr>
<td>k13</td>
<td>( \rightarrow S4 )</td>
<td>Total number of Smad4 molecules per cell (determines Smad4 synthesis rate)</td>
<td>300,000 - 700,000 per cell</td>
<td>408,061 per cell</td>
<td>301,353 per cell</td>
</tr>
<tr>
<td>k9</td>
<td>( S2 \rightarrow pS2 \rightarrow )</td>
<td>Half-life of cytosolic Smad2 monomers</td>
<td>0.0005 - 0.005 min(^{-1})</td>
<td>same as hepatocyte model</td>
<td>same as hepatocyte model</td>
</tr>
<tr>
<td>k9’</td>
<td>( (pS2)_2 S4 \rightarrow (pS2)_3 \rightarrow )</td>
<td>Half-life of cytosolic Smad2 trimers</td>
<td>0.0005 - 0.005 min(^{-1})</td>
<td>0.0135 min(^{-1})</td>
<td>0.001 min(^{-1})</td>
</tr>
<tr>
<td>k10</td>
<td>( nS2 \rightarrow npS2 \rightarrow )</td>
<td>Half-life of nuclear Smad2 monomers</td>
<td>0.0005 - 0.005 min(^{-1})</td>
<td>same as hepatocyte model</td>
<td>same as hepatocyte model</td>
</tr>
<tr>
<td>Parameter</td>
<td>Reaction</td>
<td>Rate Constant</td>
<td>Value in HaCaT cell model 1</td>
<td>Value in HaCaT cell model 2</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>---------------</td>
<td>-----------------------------</td>
<td>-----------------------------</td>
<td></td>
</tr>
<tr>
<td>$k_{10}^*$</td>
<td>$(npS2)\cdot nS4 \rightarrow (npS2)\cdot nS4$</td>
<td>Half-life of nuclear Smad2 trimers</td>
<td>0.0005 - 0.005 min$^{-1}$</td>
<td>0.0049 min$^{-1}$</td>
<td>0.0059 min$^{-1}$</td>
</tr>
<tr>
<td>$k_{11}$</td>
<td>S2 $\rightarrow$ nS2 - pS2 $\rightarrow$ npS2 - $(pS2)S4 \rightarrow (npS2)nS4$ - $(pS2)_3 \rightarrow (npS2)_3$</td>
<td>1st-order import rate constant</td>
<td>0.07 - 0.7 min$^{-1}$</td>
<td>same as hepatocyte model</td>
<td>0.464 min$^{-1}$</td>
</tr>
<tr>
<td>$k_{9} - k_{11}$</td>
<td>nS2 $\rightarrow$ S2</td>
<td>1st-order export rate constant</td>
<td>1 - 10$\times$ import rate (1 $&lt; k_9 &lt; 10$)</td>
<td>same as hepatocyte model</td>
<td>2.412 min$^{-1}$</td>
</tr>
<tr>
<td>$k_{16}$</td>
<td>S4 $\rightarrow$ nS4</td>
<td>1st-order import rate constant</td>
<td>0.03 - 0.3 min$^{-1}$</td>
<td>same as hepatocyte model</td>
<td>0.124 min$^{-1}$</td>
</tr>
<tr>
<td>$k_{15} - k_{16}$</td>
<td>nS4 $\rightarrow$ S4</td>
<td>1st-order export rate constant</td>
<td>1 - 10$\times$ import rate (1 $&lt; k_{15} &lt; 10$)</td>
<td>same as hepatocyte model</td>
<td>0.329 min$^{-1}$</td>
</tr>
</tbody>
</table>

Listed are those parameter that were allowed to differ between the HaCaT and hepatocyte models, and were thus fitted to the HaCaT time course data in Fig. 6.15. The column “value in HaCaT cell model 1” indicates the parameter values obtained without fitting the Smad im-/export rates to the HaCaT time course data in Fig. 6.15 (the optimised cytoplasmic and nuclear volumes were 1.77 pl and 0.23 pl, respectively). The column “value in HaCaT cell model 2” indicates the parameter values obtained with fitting the Smad im-/export rates to the HaCaT time course data in Fig. 6.15 (the optimised cytoplasmic and nuclear volumes were 3.1 pl and 0.38 pl, respectively). See caption of Fig. 6.15 for details.
G Additional material on small RNAs Establish Delays and Temporal Thresholds in Gene Expression

G.1 Model derivation

Based on the scheme depicted in Fig. 7.1A, a mathematical model for sRNA-mediated regulation was implemented. Experimental studies indicated that heteroduplex association and dissociation proceed with rapid kinetics when compared to RNA synthesis and degradation [389,390]. The model was therefore simplified by applying a rapid-equilibrium assumption as described in the following.

The differential equations of the model depicted in Fig. 7.1A read:

\[
\begin{align*}
\frac{d\text{[Target]}}{dt} &= v_{\text{syn,T}} - k_{\text{deg,T}} \cdot \text{[Target]} - k_{\text{on}} \cdot \text{[Target]} \cdot \text{[sRNA]} + k_{\text{off}} \cdot \text{[Pair]} \\
\frac{d\text{[sRNA]}}{dt} &= v_{\text{syn,S}} - k_{\text{deg,S}} \cdot \text{[sRNA]} - k_{\text{on}} \cdot \text{[Target]} \cdot \text{[sRNA]} + k_{\text{off}} \cdot \text{[Pair]} \\
\frac{d\text{[Pair]}}{dt} &= k_{\text{on}} \cdot \text{[Target]} \cdot \text{[sRNA]} - k_{\text{off}} \cdot \text{[Pair]} - k_{\text{deg,P}} \cdot \text{[Pair]}
\end{align*}
\]  

(G.1)

Summing up these differential equations, and using the relationships \[\text{[T tot]} = \text{[Target]} + \text{[Pair]}\] and \[\text{[S tot]} = \text{[sRNA]} + \text{[Pair]}\] for the total amounts of target RNA and small RNA yields:

\[
\begin{align*}
\frac{d\text{[T tot]}}{dt} &= v_{\text{syn,T}} - k_{\text{deg,T}} \cdot ([\text{T tot}] \cdot \text{[Pair]}) - k_{\text{deg,P}} \cdot \text{[Pair]} \\
\frac{d\text{[S tot]}}{dt} &= v_{\text{syn,S}} - k_{\text{deg,S}} \cdot ([\text{S tot}] \cdot \text{[Pair]}) - k_{\text{deg,P}} \cdot \text{[Pair]}
\end{align*}
\]  

(G.2)

The association/dissociation reactions of the pair are assumed to proceed much faster than all other steps in the model. Thus, the model species are related by the following equilibrium:

\[
K_{d,P} = \frac{k_{\text{off,P}}}{k_{\text{on,P}}} = \frac{([\text{T tot}] \cdot \text{[Pair]}) \cdot ([\text{S tot}] \cdot \text{[Pair]})}{[\text{Pair}]}
\]  

(G.3)

Solving for the pair concentration yields:

\[
[\text{Pair}] = \frac{1}{2} \left( \frac{[\text{T tot}] + [\text{S tot}] + K_{d,P} \cdot \sqrt{([\text{T tot}] + [\text{S tot}] + K_{d,P})^2 - 4 \cdot [\text{T tot}] \cdot [\text{S tot}]} \right)
\]  

(G.4)

Equations (G.2) and (G.4) constitute a reduced form of the differential equation system (G.1) for the case that association/dissociation reactions of the pair proceed much faster than all other steps (‘rapid equilibrium approximation’). This reduced system was used for all numerical simulations shown in Section 7 and in Appendix G.

**Analytical approximation for the subthreshold regime:** The reduced differential equation system comprising Equations (G.2) and (G.4) was further simplified to derive an analytical expression for the response time in the subthreshold regime \([v_{\text{syn,T}} < v_{\text{syn,S}}]\). Under subthreshold conditions, one can assume that the sRNA is present in vast excess over the target mRNA, so that \([\text{S tot}]\) is approximately constant. Then, Eq. G.4 for the Pair intermediate simplifies to

\[
[\text{Pair}] = \frac{[\text{T tot}] \cdot [\text{S tot}]}{K_{d,P} + [\text{S tot}]},
\]  

(G.5)

and Eq. G.2 reduces to
\[ \frac{d[T_{tot}]}{dt} = v_{syn,T} - \left( k_{deg,T} \cdot \left( \frac{[S_{tot}]}{k_{d,p} + [S_{tot}]} \right) + k_{deg,P} \cdot \frac{[S_{tot}]}{k_{d,p} + [S_{tot}]} \right) \cdot [T_{tot}] \]  

\[ \text{(G.6)} \]

Thus, direct mRNA degradation and degradation via the Pair intermediate behave as two competing first-order decay terms in the subthreshold regime. For sufficiently large sRNA expression (i.e., \([S_{tot}] > K_{d,p}\)) and/or if the degradation of the pair is faster than that of the monomeric mRNA \((k_{deg,P} > k_{deg,T})\) the system further simplifies to

\[ \frac{d[T_{tot}]}{dt} = v_{syn,T} - k_{deg,P} \cdot [T_{tot}] \]  

\[ \text{(G.7)} \]

Thus, the response time \(t_{50}\) for mRNA up- and downregulation is proportional to the inverse of the degradation rate of the pair [391].

G.2 Half-life measurements

![Figure G.1: Degradation of transcripts from the isiAB operon.](image)

(A) Wildtype (Wt) and mutant (Mu) cells lacking the antisense RNA, IsrR, were cultivated under iron-deplete conditions. At time zero (0 hours) an iron pulse was given (red arrow), and cells were grown further for 0–24 hours. The isiAB dicistronic precursor transcript and the isiA mRNA are degraded directly after iron induction (0–4 hours), whereby the degradation rate of isiA mRNA is higher for Wt cells expressing the antisense RNA in comparison to the Mu cells lacking IsrR. (B) Cells were grown under iron-replete conditions. Although at time zero (0 min) rifampicin was added to non-specifically arrest transcription, accumulation of IsrR transcripts is not affected compared to an internal standard (5S rRNA) for at least 45 min, thus indicating a half-life > 45 min for IsrR. Note that longer measurements are not possible because sustained incubation with rifampicin induces cell death. The RNA transcripts were separated on 1.3% formaldehyde-agarose gels and hybridised with an RNA-specific DNA fragment.

G.3 Dose-Response Measurements

Figure 7.1B and additional numerical simulations indicated that sRNAs might establish sharp thresholds in gene expression (compare [385]). These thresholds are typically accompanied by a mutually exclusive expression pattern of sRNA and target mRNA at least if the pair intermediate is significantly less stable than the monomeric RNAs. Previous experimental work revealed that the mRNA of the cyanobacterial iron stress protein, isiA, and its modulator, the IsrR sRNA, are expressed in mutually exclusive manner under various stimulation conditions [388]. Moreover, a comparison of wild-type cells with IsrR knock-down cells revealed that IsrR completely suppresses residual isiA expression under unstressed conditions (compare circle and square at t = 0 in Fig. 7.2A). These data suggest that a sharp threshold exists in the cyanobacterial iron stress response. Dose-response experiments were performed to further confirm the existence of such thresholds. Cells were cultivated under iron-starving conditions for 48 h to induce isiA expression, and then treated with different doses of iron to induce downregulation of isiA transcription. The corresponding measurements of isiA target RNA and IsrR sRNA (shown in
Fig. G.2 reveal that IsrR does not accumulate unless the isiA concentration falls a critical level (∼0.05 in A and B). This conclusion holds in the time domain and in the concentration domain. First, the isiA level 2 h after iron re-addition (Fig. G.2A) is higher when compared to 4 h after iron re-addition (Fig. G.2B), and accordingly IsrR expression is selectively induced at 4 h (Fig. G.2C and D). Second, IsrR expression at 4 h requires that the iron level exceeds a certain threshold (of approximately 0.25 in Fig. G.2B and D). Taken together, these data indicate that a sharp threshold exists in the cyanobacterial iron stress response, and that the system operates near this threshold under physiological conditions.

![Figure G.2: Dose-Response behaviour of the cyanobacterial iron stress response.](image)

Cells were cultivated under iron-starving conditions for 48 h to induce isiA expression, and then treated with different doses of iron to induce downregulation of isiA transcription. The expression levels of the isiA target RNA (A and B) or of the IsrR sRNA (C and D) were measured after 2 h (A and C) and after 4 h (B and D).

G.4 Kinetic parameters used for modelling

This section summarises the kinetic parameters used for the numerical simulations in Figs. 7.1, 7.2, G.3 and G.4. The synthesis of target RNA (vsyn,T) is modelled to be controlled by external stress stimuli (Fig. 7.1A). A simple change of the synthesis rate from one value to another was used in most simulations (Figs. 7.2A – D, G.3B and G.4B), while the Michaelis-Menten equation (vsyn,T = Vmax,synT ⋅ Stimulus / (KM,synT + Stimulus)) was employed to simulate dose-response behaviour (Figs. 7.1B, G.3A, and G.4A).


### Table G.1a

<table>
<thead>
<tr>
<th>Figure</th>
<th>Fig. 7.1B</th>
<th>Fig. 7.2A</th>
<th>Fig. 7.2B</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v_{\text{syn},T}$ [nM h$^{-1}$]</td>
<td>Modelled by Michaelis-Menten equation</td>
<td>247.2 (basal level)</td>
<td>996.9 (basal level)</td>
</tr>
<tr>
<td>$V_{\text{max},\text{syn},T}$ [nM h$^{-1}$]</td>
<td>1.25</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$K_{\text{M},\text{syn},T}$ [nM]</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$k_{\text{deg},T}$ [h$^{-1}$]</td>
<td>1</td>
<td>0.42</td>
<td>0.42</td>
</tr>
<tr>
<td>$v_{\text{syn},S}$ [nM h$^{-1}$]</td>
<td>0.625</td>
<td>516.6 (solid line)</td>
<td>516.6 (solid line)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>706.9 (grey line)</td>
<td>145.3 (dashed line)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>145.3 (dashed line)</td>
<td>-</td>
</tr>
<tr>
<td>$k_{\text{deg},S}$ [h$^{-1}$]</td>
<td>1</td>
<td>0.35</td>
<td>0.35</td>
</tr>
<tr>
<td>$K_{D,P}$ [nM]</td>
<td>0.001</td>
<td>0.0045</td>
<td>0.0045</td>
</tr>
<tr>
<td>$k_{\text{deg},P}$ [h$^{-1}$]</td>
<td>10</td>
<td>13.75</td>
<td>13.75</td>
</tr>
</tbody>
</table>

### Table G.1b

<table>
<thead>
<tr>
<th>Figure</th>
<th>Fig. 7.2C</th>
<th>Fig. 7.2D</th>
<th>Fig. G.1A</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v_{\text{syn},T}$ [nM h$^{-1}$]</td>
<td>0 (basal level)</td>
<td>0 (basal level)</td>
<td>Modelled by Michaelis-Menten equation</td>
</tr>
<tr>
<td></td>
<td>1136 (after stimulus increase)</td>
<td>1136 (after stimulus increase)</td>
<td></td>
</tr>
<tr>
<td>$V_{\text{max},\text{syn},T}$ [nM h$^{-1}$]</td>
<td>-</td>
<td>-</td>
<td>1.25</td>
</tr>
<tr>
<td>$K_{\text{M},\text{syn},T}$ [nM]</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>$k_{\text{deg},T}$ [h$^{-1}$]</td>
<td>0.42</td>
<td>0.42</td>
<td>1</td>
</tr>
<tr>
<td>$v_{\text{syn},S}$ [nM h$^{-1}$]</td>
<td>516.6 (solid line)</td>
<td>516.6 (solid line)</td>
<td>0.625</td>
</tr>
<tr>
<td></td>
<td>0 (dashed line)</td>
<td>0 (dashed line)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{\text{deg},S}$ [h$^{-1}$]</td>
<td>0.35</td>
<td>0.35</td>
<td>1</td>
</tr>
<tr>
<td>$K_{D,P}$ [nM]</td>
<td>0.0045</td>
<td>0.0045</td>
<td>0.001</td>
</tr>
<tr>
<td>$k_{\text{deg},P}$ [h$^{-1}$]</td>
<td>13.75</td>
<td>13.75</td>
<td>1</td>
</tr>
</tbody>
</table>

### Table G.1c

<table>
<thead>
<tr>
<th>Figure</th>
<th>Fig. G.1B</th>
<th>Fig. G.2A</th>
<th>Fig. G.2B</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v_{\text{syn},T}$ [nM h$^{-1}$]</td>
<td>0 (basal level)</td>
<td>Modelled by Michaelis-Menten equation</td>
<td>0 (basal level)</td>
</tr>
<tr>
<td></td>
<td>1.25 (after stimulus increase)</td>
<td></td>
<td>5 (after stimulus increase)</td>
</tr>
<tr>
<td>$V_{\text{max},\text{syn},T}$ [nM h$^{-1}$]</td>
<td>-</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td>$K_{\text{M},\text{syn},T}$ [nM]</td>
<td>-</td>
<td>7</td>
<td>-</td>
</tr>
<tr>
<td>$k_{\text{deg},T}$ [h$^{-1}$]</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$v_{\text{syn},S}$ [nM h$^{-1}$]</td>
<td>0.625</td>
<td>0.625</td>
<td>0.625</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.625</td>
<td></td>
</tr>
<tr>
<td>$k_{\text{deg},S}$ [h$^{-1}$]</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$K_{D,P}$ [nM]</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>$k_{\text{deg},P}$ [h$^{-1}$]</td>
<td>10</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

### G.5 Supplemental Numerical Simulations

**Competition-only Model:** Section 7 was mainly focused on the scenario, where the heteroduplex is much less stable than the target RNA. In this case, the sRNA plays a dual role in target RNA regulation, as it competitively inhibits translation and additionally induces RNA degradation. However, several prokaryotic and eukaryotic non-coding RNAs merely act as competitive inhibitors of translation, but do not affect target RNA degradation [64,383,693]. In the “competition-only model”, which will be analysed in the following, the target RNA and the pair heteroduplex are degraded with the same rate constant (see Table G.1 for kinetic parameters).
Figure G.3: System behaviour in case that the sRNA competitively inhibits translation, but does not enhance target degradation (“competition-only model”).

(A) Steady state (top) and dynamical (bottom) response to varying stimulus strength similar to Fig. 7.1B. The steady state dose-response and the response time here were calculated for the free sRNA, while the response times in Fig. 7.1B were calculated for the total RNAs (i.e., the sum of free RNA and pair). See Table G.1 for kinetic parameters. (B) Time course of free target RNA in response to step-like pulse stimulation similar to Fig. 7.2C. See Table G.1 for kinetic parameters.

Figure G.3 shows numerical simulations similar to those in Figs. 7.1B and 7.2C. It turned out that the competition-only model also shows the kinetic features discussed in the main text when analysed at the level of free sRNA and free target RNA (Fig. G.3). However, these features were no longer observed at the level of total sRNA and total target RNA (i.e., the sum of free and pair species), as pair formation does not induce RNA degradation, but only sequestration. For example, the free RNAs are expressed in a mutually exclusive manner in the competition-only model, while the total levels of target RNA and sRNA show overlapping expression patterns (not shown).

**Partial-Degradation Model:** In the default model, it was assumed that both, the sRNA and target RNA, are degraded during the decay reaction of the pair heteroduplex. Eukaryotic miRNAs frequently remain intact after target degradation, and can guide the recognition and destruction of additional messages [64]. The default model was therefore extended such that it takes such a more catalytic mode of action into account (‘partial-degradation model’). The differential equations of the partial-degradation model are given by:

\[
\begin{align*}
\frac{dT_{\text{tot}}}{dt} &= v_{\text{syn,T}} - k_{\text{deg,T}} \cdot (T_{\text{tot}} \cdot [\text{Pair}]) - k_{\text{deg,P}} \cdot [\text{Pair}] \\
\frac{dS_{\text{tot}}}{dt} &= v_{\text{syn,S}} - k_{\text{deg,S}} \cdot (S_{\text{tot}} \cdot [\text{Pair}]) - (1-f) \cdot k_{\text{deg,P}} \cdot [\text{Pair}]
\end{align*}
\]

(G.8)

where the concentration of the pair is still given by Eq. (G.4). The new parameter \( f \) specifies the fraction of the sRNA that remains intact during the decay of the pair heteroduplex. For \( f = 0 \), the system reduces to the default model (Eq. (G.2)), while the sRNA is not degraded via the pair intermediate in case that \( f = 1 \).
Figure G.4: System Behaviour in case that the sRNA remains intact during pair degradation ("partial-degradation model"). The simulations were done by integrating the extended differential equation system (G.8) and f was set to 0.8.

(A) Steady State (top) and Dynamical (bottom) Response to Varying Stimulus Strength similar to Fig. 7.1B. See Table G.1 for kinetic parameters. (B) Time course of free target RNA in response to step-like pulse stimulation similar to Fig. 7.2C. See Table G.1 for kinetic parameters.

Figure G.4 shows simulation results for f = 0.8 (i.e., 80% of the sRNA remains intact during pair degradation). The kinetic characteristics reported in the main text are still observed, and thus apply for eukaryotic miRNAs as well. Note that the threshold in Fig. G.4A is shifted to higher values when compared to Fig. 7.1B, as expected ($v_{syn,T}$ has the same value for a stimulus of 1 in both simulation; compare Table G.1).

G.6 Simulation of pulse-filtering at the protein level

The analysis presented in Section 7 suggests that ultrasensitive pulse-filtering occurs at the mRNA level. However, it remained to be determined whether pulse-filtering is preserved at the level of protein expression. The differential equation system given in Eq. (G.2) was therefore modified, so that it takes by protein synthesis and degradation into account:

\[
\begin{align*}
\frac{d[T_{tot}]}{dt} &= v_{syn,T} - k_{deg,T} \cdot ([T_{tot}] \cdot [Pair]) - k_{deg,P} \cdot [Pair] \\
\frac{d[S_{tot}]}{dt} &= v_{syn,S} - k_{deg,S} \cdot ([S_{tot}] \cdot [Pair]) - k_{deg,P} \cdot [Pair] \\
\frac{d[Protein]}{dt} &= k_{syn,Protein} \cdot ([T_{tot}] \cdot [Pair]) - k_{deg,Protein} \cdot [Protein]
\end{align*}
\] (G.9)

Equation G.9, together with Eq. G.4, constitutes a model for sRNA-mediated regulation of protein expression, which was analysed by numerical simulations. Duration-response curves were calculated and the maximum of the protein time course was taken as the response. Figure G.5 shows the protein time course amplitude as a function of the stimulus pulse duration (normalised by the half-maximal pulse duration). The duration-response curve was calculated for different protein half-lives as indicated in the legend of Fig. G.5, and the protein synthesis rate was taken to be $k_{syn,Protein} = k_{deg,Protein}$ to ensure the same steady state protein level is achieved regardless of the degradation rate chosen.
Figure G.5: Pulse filtering is preserved at the protein level.

The time course maximum of protein accumulation in response to step-like stimuli of different duration (but of constant strength) was simulated using the differential equation system (G.9). The duration-response curve was calculated for different protein half-lives as indicated in the legend, and the protein synthesis rate was taken to be \( k_{\text{syn,Protein}} = k_{\text{deg,Protein}} \) to ensure the same steady state protein level is achieved regardless of the degradation rate chosen. All other kinetic parameters were chosen as in Fig. 7.2D. The x-axis was normalised by the pulse duration required to reach the half-maximal protein amplitude to allow for comparison among duration-response curves. The Hill coefficient of each curve was calculated as described in the Materials and Methods section, and is indicated in the legend.

The simulations in Fig. G.5 demonstrate that pulse-filtering at the mRNA level (solid line in Fig. 7.2D) is obscured at the protein level if the protein half-life is 100 or 1000 hours (solid and dashed black lines). However, pulse-filtering is preserved at the protein level even if the protein has a half-life of 10 hours (Fig. G.5, dashed grey line). The corresponding Hill coefficient of 2.37 implies that a ~6-fold change in the pulse duration is sufficient to switch the time course from 10% to 90% of the steady state activation level, and thus indicates ultrasensitive pulse filtering. More importantly, the simulations demonstrate that a system with 10 hour protein half-life completely filters out short pulse durations (Fig. G.5, dashed grey line). As may be expected, pulse filtering is more pronounced for even shorter protein half-lives (grey solid and black dashed lines in Fig. G.5), and ultrasensitivity at the protein and mRNA levels is similar (compare black solid line in Fig. 7.2D).

Degradation studies for the IsiA protein are missing in the literature so far. However, half-life measurements of the CP43 photosynthesis protein homologous to IsiA revealed a half-life of about an hour under stress conditions [397], and a similar rapid turnover was also reported for another photosynthesis protein, D1 [398]. These data suggest that IsiA protein is short-lived in the experimental setup chosen in the present analyses, and that the pulse filtering property discussed here for the RNA level is observed at the level of proteins as well.
References


Deshaies, R. J. und Ferrell, J. E., Jr. (2001): Multisite phosphorylation and the countdown to S phase, Cell (Band 107), Nr. 7, Seite 819-22

Millat, T.; Sreenath, S. N.; Soebiyanto, R. P.; Avva, J.; Cho, K. H. und Wolkenhauer, O. (2008): The role of dynamic simulation pattern in the analysis of bistable intracellular networks, Biosystems (Band 92), Nr. 3, Seite 270-81


Hanahan, D. und Weinberg, R. A. (2000): The hallmarks of cancer, Cell (Band 100), Nr. 1, Seite 57-70


Buday, L. und Downward, J. (1993): Epidermal growth factor regulates the exchange rate of guanine nucleotides on p21ras in fibroblasts, Mol Cell Biol (Band 13), Nr. 3, Seite 1903-10


Chevallier-Multon, M. C.; Schweighoffer, F.; Barlat, I.; Baudouy, N.; Fath, I.; Duchesne, M. und Tocque, B. (1993): Saccharomyces cerevisiae CDC25 (1028-1589) is a guanine nucleotide releasing factor for mammalian ras proteins and is oncogenic in NIH3T3 cells, J Biol Chem (Band 268), Nr. 15, Seite 11113-8


Ricketts, M. H. und Levinson, A. D. (1988): High-level expression of c-H-ras1 fails to fully transform rat-1 cells, Mol Cell Biol (Band 8), Nr. 4, Seite 1460-8


Roose, J. P.; Mollenauer, M.; Ho, M.; Kuroski, T. und Weiss, A. (2007): Unusual interplay of two types of Ras activators, RasGRP and SOS, establishes sensitive and robust Ras activation in lymphocytes, Mol Cell Biol (Band 27), Nr. 7, Seite 2732-45

Coughlin, J. J.; Stang, S. L.; Dower, N. A. und Stone, J. C. (2005): RasGRP1 and RasGRP3 regulate B cell proliferation by facilitating B cell receptor-Ras signaling, J Immunol (Band 175), Nr. 11, Seite 7179-84


Lenzen, C.; Cool, R. H.; Prinz, H.; Kuhlmann, J. und Wittinghofer, A. (1998): Kinetic analysis by fluorescence of the interaction between Ras and the catalytic domain of the guanine nucleotide exchange factor Cdc25Mm, Biochemistry (Band 37), Nr. 20, Seite 7420-30


Takekawa, M.; Tatebayashi, K. und Saito, H. (2005): Conserved docking site is essential for activation of mammalian MAP kinase kinases by specific MAP kinase kinase kinases, Mol Cell (Band 18), Nr. 3, Seite 295-306

Bardwell, A. J.; Flateau, L. J.; Matsuukuma, K.; Thorner, J. und Bardwell, L. (2001): A conserved docking site in MEKs mediates high-affinity binding to MAP kinases and cooperates with a scaffold protein to enhance signal transmission, J Biol Chem (Band 276), Nr. 13, Seite 10374-86


Xia, Y.; Wu, Z.; Su, B.; Murray, B. und Karin, M. (1998): JNKK1 organizes a MAP kinase module through specific and sequential interactions with upstream and downstream components mediated by its amino-terminal extension, Genes Dev (Band 12), Nr. 21, Seite 3369-81


Bardwell, A. J.; Abdollahi, M. und Bardwell, L. (2003): Docking sites on mitogen-activated protein kinase (MAPK) kinases, MAPK phosphatases and the Elk-1 transcription factor compete for MAPK binding and are crucial for enzymatic activity, Biochem J (Band 370), Nr. Pt 3, Seite 1077-85


Bardwell, A. J.; Abdollahi, M. und Bardwell, L. (2003): Docking sites on mitogen-activated protein kinase (MAPK) kinases, MAPK phosphatases and the Elk-1 transcription factor compete for MAPK binding and are crucial for enzymatic activity, Biochem J (Band 370), Nr. Pt 3, Seite 1077-85


Datta, R.; Oki, E.; Endo, K.; Ren, J. und Kufe, D. (2000): XIAP regulates DNA damage-inducible cytochrome c release from mitochon-
dria, Proc. Natl. Acad. Sci. USA (Band 97), Nr. 20, Seite 10940-5

Clayton, R.; Clark, J. B. und Sharpe, M. (2005): Cytochrome c release from rat brain mitochondria is proportional to the mitochondrial functional deficit: implications for apoptosis and neurodegenerative disease, J Neurochem (Band 92), Nr. 4, Seite 840-9


Binds and internalization of transforming growth factor-beta 1 by human hepatoma cells: evidence for receptor recycling, Hepatology (Band 14), Nr. 2, Seite 287-95


Kavask, P.; Rasmussen, R. K.; Causing, C. G.; Bonni, S.; Zhu, H.; Thomsen, G. H. und Wrana, J. L. (2000): Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGF beta receptor for degradation, Mol Cell (Band 6), Nr. 6, Seite 1365-75


Runyan, C. E.; Schnaper, H. W. und Poncetale, A. C. (2005): The role of internalization in transforming growth factor beta1-induced Smad2 association with Smad anchor for receptor activation (SARA) and Smad2-dependent signaling in human mesangial cells, J Biol Chem (Band 280), Nr. 9, Seite 8300-8

Wilkes, M. C. und Leof, E. B. (2006): Transforming growth factor beta activation of c-Abl is independent of receptor internalization and regulated by phosphatidylinositol 3-kinase and PAK2 in mesenchymal cultures, J Biol Chem (Band 281), Nr. 38, Seite 27846-54


Panopoulou, E.; Gillooly, D. J.; Wrana, J. L.; Zerial, M.; Stenmark, H.; Murphy, C. und Fotsis, T. (2002): Early endosomal regulation of Smad-dependent signaling in endothelial cells, J Biol Chem (Band 277), Nr. 20, Seite 18046-52


Tsukazaki, T.; Chiang, T. A.; Davison, A. F.; Attisano, L. und Wrana, J. L. (1998): SARA, a FYVE domain protein that recruits Smad2 to the TGFbeta receptor, Cell (Band 95), Nr. 6, Seite 779-91


Kholodenko, B. N.; Hoek, J. und Westerhoff, H. V. (2000): Why cytoplasmic signalling proteins should be recruited to cell membranes, Trends Cell Biol (Band 10), Nr. 5, Seite 173-8

Schmierer, B. und Hill, C. S. (2005): Kinetic analysis of Smad nucleocytoplasmic shuttling reveals a
Xu, L.; Kang, Y.; Col, S. und Massague, J. (2002): Smad2 nucleocytoplasmic shuttling by nucleoporins
Chen, H. B.; Rud, J. G.; Lin, K. und Xu, L. (2005): Nuclear targeting of transforming growth factor-beta-
Mavrakis, K. J.; Andrew, R. L.; Lee, K. L.; Petropoulos, C.; Dixon, J. E.; Navaratnam, N.; Norris, D. P.
Lin, X.; Liang, M. und Feng, X. H. (2000): Smurf2 is a ubiquitin E3 ligase mediating proteasome-
Correia, J. J.; Chacko, B. M.; Lam, S. S. und Lin, K. (2001): Sedimentation studies reveal a direct role of
Brunicardi, F. C.; Shi, Y.; Chen, Y. G.; Meng, A. und Feng, X. H. (2006): PPM1A functions as a Smad
Seite 1318-32
Seite 271-82
Seite 1847-54
Seite 611-7
Seite 20709-16
Seite 9431-41


[470] Sun, H.; Charles, C. H.; Lau, L. F. und Tonks, N. K. (1993): MKP-1 (3CH134), an immediate early gene product, is a dual specificity phosphatase that dephosphorylates MAP kinase in vivo, Cell (Band 75), Nr. 3, Seite 487-93


191


[519] Bourillot, P. Y.; Garrett, N. und Gurdon, J. B. (2002): A changing morphogen gradient is interpreted by continuous transduction flow, Development (Band 129), Nr. 9, Seite 2167--2180


[523] Xin, Hong; Xu, Xiailian; Li, Linyu; Ning, Hongxiu; Rong, Yu; Wang, Yinyin; Fu, Xin-Yuan und Chang, Zhijie (2005): CHIP controls the sensitivity of transforming growth factor-beta signaling by modulating the basal level of Smad3 through ubiquitin-mediated degradation, J Biol Chem (Band 280), Nr. 21, Seite 20842--20850

[524] Dupont, Sirio; Zacchigna, Luca; Cordenonsi, Michelangelo; Soligo, Sandra; Adorno, Maddalena; Rugge, Massimo und Piccolo, Stefano (2005): Germ-layer specification and control of cell growth by Ectodermin, a Smad4 ubiquitin ligase, Cell (Band 121), Nr. 1, Seite 87--99


[526] Wang, Mei; Tang, Yi; Tytler, Ewan M.; Lu, Chongyuan; Jin, Bingwen; Vickers, Selwyn M.; Yang, Lei; Shi, Xingming und Cao, Xu (2004): Smad4 protein stability is regulated by ubiquitin ligase SCF-Trip1, J Biol Chem (Band 279), Nr. 15, Seite 14484--14487

[527] Liang, Min; Liang, Yao-Yun; Wrighton, Katharine; Ungermannova, Dana; Wang, Xiao-Ping; Brunicardi, F. Charles; Liu, Xuedong; Feng, Xin-Hua und Lin, Xia (2004): Ubiquitination and proteolysis of cancer-derived Smad4 mutants by SCFSkp2, Mol Cell Biol (Band 24), Nr. 17, Seite 7524--7537


[529] Wang, Mei; Cao, Xuesong; Wu, Yalei; Bai, Shuting; Wu, Liyu; Shi, Xingming; Wang, Ning und Cao, Xu (2002): Jab1 antagonizes TGF-beta signaling by inducing Smad4 degradation, EMBO Rep (Band 3), Nr. 2, Seite 171--176

[530] Seo, Su Ryeon; Lallemant, Francois; Ferrand, Nathalie; Pessah, Marcia; L'Hoste, Sebastien; Camonis, Jacques und Atti, Azeddine (2004): The novel E3 ubiquitin ligase Tiu11 associates with TGIF to target Smad2 for degradation, EMBO J (Band 23), Nr. 19, Seite 3780--3792

[531] Lo, R. S.; Wotton, D. und Massague, J. (2001): Epidermal growth factor signaling via Ras controls the Smad transcriptional co-repressor TGIF, EMBO J (Band 20), Nr. 1-2, Seite 128--136

[532] Reed, B. C. und Lane, M. D. (1980): Insulin receptor synthesis and turnover in differentiating 3T3-L1 preadipocytes, Proc Natl Acad Sci U S A (Band 77), Nr. 1, Seite 285--289

[533] Reed, B. C.; Ronnett, G. V. und Lane, M. D. (1981): Role of glycosylation and protein synthesis in insulin receptor metabolism by 3T3-L1 mouse adipocytes, Proc Natl Acad Sci U S A (Band 78), Nr. 5, Seite 2908--2912


[535] Gur, Gal; Rubin, Chanan; Katz, Menachem; Amit, Ido; Citri, Ami; Nilsson, Jonas; Amariglio, Ninette; Henriksson, Roger; Rechavi, Gideon; Hedman, Hakan; Wides, Ron und Yarden, Yosef (2004): LRG1 restricts growth factor signaling by enhancing receptor ubiquitylation and degradation, EMBO J (Band 23), Nr. 16, Seite 3270--3281
[560] Abella, Jasmine V.; Peschard, Pascal; Naujokas, Monica A.; Lin, Tong; Saucier, Caroline; Urbe, Sylvie und Park, Morag (2005): Met/Hepatocyte growth factor receptor ubiquitination suppresses transformation and is required for Hrs phosphorylation, Mol Cell Biol (Band 25), Nr. 21, Seite 9632–9645


[562] Coulombe, Philippe; Rodier, Genevieve; Pelletier, Stephane; Pellerin, Johanne und Meloche, Sylvain (2003): Rapid turnover of extracellular signal-regulated kinase 3 by the ubiquitin-proteasome pathway defines a novel paradigm of mitogen-activated protein kinase regulation during cellular differentiation, Mol Cell Biol (Band 23), Nr. 13, Seite 4542–4558

[563] Lingohr, Melissa K.; Briaud, Isabelle; Dickson, Lorna M.; McCuaig, Jill F.; Alarcón, Cristina; Wicksted, Barton L. und Rhodes, Christopher J. (2006): Specific regulation of IRS-2 expression by glucose in rat primary pancreatic islet beta-cells, J Biol Chem (Band 281), Nr. 23, Seite 15884–15892

[564] Buschbeck, Marcus; Hofbauer, Sebastian; Croce, Luciano Di; Keri, Gyorgy und Ullrich, Axel (2005): Abl-kinase-sensitive levels of ERK5 and its intrinsic basal activity contribute to leukaemia cell survival, EMBO Rep (Band 6), Nr. 1, Seite 63–69


[567] Lin, Yun-Wei und Yang, Jia-Ling (2006): Cooperation of ERK and SCF53k2 for MKP-1 destruction provides a positive feedback regulation of proliferating signaling, J Biol Chem (Band 281), Nr. 2, Seite 915–926


[576] Rahmouni, Souad; Cerignoli, Fabio; Alonso, Andres; Tsutji, Yoshiya; Henkens, Rachel; Zhu, Changjun; Sully, Christine Louis-dit; Moutschen, Michel; Jiang, Wei und Mustelin, Tomas (2008): Loss of the VHR dual-specific phosphatase causes cell-cycle arrest and senescence, Nat Cell Biol (Band 8), Nr. 5, Seite 524–531


[580] Walrafen, Pierre; Verdier, Frederique; Kadri, Zahra; Chretien, Stany; Lacombe, Catherine und Mayeux, Patrick (2005): Both proteasomes and lysosomes degrade the activated erythropoietin receptor, Blood (Band 105), Nr. 2, Seite 600–608


[582] Ungureanu, Daniela; Saharinen, Pipsa; Junttila, Ilkka; Hilton, Douglas J. und Silvennoinen, Olli (2002): Regulation of Jak2 through the ubiquitin-proteasome pathway involves phosphorylation of Jak2 on Y1007 and interaction with SOCS-1, Mol Cell Biol (Band 22), Nr. 10, Seite 3316–3326


Shi, W.; Sun, C.; He, B.; Xiong, W.; Shi, X.; Yao, D. and Cao, X. (2004): GADD34-PP1c recruited by Smad7 dephosphorylates TGFbeta type I receptor, J Cell Biol (Band 164), Nr. 2, Seite 291-300


Minowada, G.; Jarvis, L. A.; Chi, C. L.; Neubuser, A.; Sun, X.; Hachen, N.; Krasnow, M. A. and Martin, G. R. (1999): Vertebrate Sprouty genes are induced by FGF signaling and can cause chondrodysplasia when overexpressed, Development (Band 126), Nr. 20, Seite 4465-75

Hanafusa, H.; Torii, S.; Yasunaga, T. und Nishida, E. (2002): Sprouty1 and Sprouty2 provide a control mechanism for the Ras/MAPK signalling pathway, Nat Cell Biol (Band 4), Nr. 11, Seite 850-8


Pouyssegur, J.; Volmat, V. und Lenormand, P. (2002): Fidelity and spatio-temporal control in MAP kinase (ERKs) signalling, Biochem Pharmacol (Band 64), Nr. 5-6, Seite 755-63

Tullai, John W.; Schaffer, Michael E.; Mullenbrock, Steven; Kasif, Simon and Cooper, Geoffrey M. (2004): Identification of Transcription Factor Binding Sites Upstream of Human Genes Regulated by the Phosphatidylinositol 3-Kinase and MEK/ERK Signaling Pathways 10.1074/jbc.M309260200, J. Biol. Chem. (Band 279), Nr. 19, Seite 20167-20177
Acknowledgements

First of all, I would like to thank Hanspeter Herzel for letting me work in his group, and for supervising this thesis as well as giving personal support. Thanks to Jens Timmer and Olaf Wolkenhauer for acting as reviewers for this thesis.

I had lots of fun and exciting discussions with my room-mates, especially with Nils, Pål, Adrian, Kasia, and James. In particular, Nils contributed a large part to this thesis.

Thanks to all other people in Hanspeters group, especially Ilka, Christian, Branka, Patrick, Anuradha, Bertram, and Maciej. It was fun to work with you. Thanks also to members of the Kollmann group for discussions on signal transduction (Ralf, Kajetan and Markus).


Thanks to Andreas Hantschmann for help with computers, and to Karin Winklhöfer as well as Elvira Lauterbach for helping to organise things.

Finally, thanks to my parents and to my sister for support.