

Strategies for High Quality Quantitative Data Generation and Dynamic Modeling of the MAP-Kinase Signaling Cascade

Dissertation

submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

presented by

Diplom-Biologe Marcel Schilling
born in Binningen, Switzerland

Mündliche Prüfung: 24.05.07

Referees:

Prof. Dr. Michael Brunner

PD Dr. Ursula Klingmüller

13:26 *Restate my assumptions.*

*One: **Mathematics** is the language of nature.*

*Two: Everything around us can be represented and understood through **numbers**.*

*Three: If you graph the numbers of any **system**, patterns emerge.*

*Therefore: There are patterns everywhere in **nature**.*

10:18 *Press return.*

π (1998), screenplay by Darren Aronofsky

ACKNOWLEDGEMENTS

Many thanks to all the people who supported me during this work.

First of all, I would like to thank my supervisor PD Dr. Ursula Klingmüller for the opportunity to work in her lab, her continuous support and advice as well as for the interesting collaborations she established.

I thank Prof. Dr. Michael Brunner for acting as second referee for this thesis.

This work would not have been possible without the fantastic help from all the current and former members of the lab. I am grateful for the cheerful and stimulating working environment. I would especially like to thank Verena Becker for being such a harmonious benchmate and for our joint project on endocytosis, Ute Bauman for all the technical help provided, Sebastian Bohl for his work on the quantitative data projects and Julie Bachmann for her contributions to the endocytosis project. Many thanks go to Dr. Andrea C. Pfeifer for counsel as a member of my PhD committee as well as for countless scientific discussions.

I am very grateful to our collaborators from the University of Freiburg, especially Prof. Dr. Jens Timmer for stimulating discussions and continuing support and Thomas Maiwald for close and fruitful collaborations. Furthermore, Stefan Hengl, Dr. Markus Kollmann and Clemens Kreutz provided vital contributions to our projects.

I thank Prof. Dr. Jennifer Reed for critically reading the manuscript on strategies for standardizing quantitative data.

This work was supported by the funding priority “Systems of Life – Systems Biology” (Hepatosys) of the German Federal Ministry of Education and Research (BMBF).

I am grateful to my friends, especially Johanna Engelhard for being the best neighbor in the world, Alexandra Kienast for many late night discussions, Nina Stössinger and Conny Gysin for providing home bases in the east and the south, respectively, and Georg Bandl for everlasting friendship.

Most of all, I would like to thank my sister and my parents to whom I dedicate this work.

INTRODUCTION

1.1 Table of contents

Introduction.....	5
1.1 Table of contents	5
1.2 English summary	6
1.3 Deutsche Zusammenfassung	7
1.4 Summary of the results.....	8
1.4.1 Systems biology	8
1.4.2 Erythropoietin and the hematopoietic lineage	11
1.4.3 The MAP-kinase signaling network	14
1.4.4 New strategies for quantitative immunoblotting.....	19
1.4.5 Modeling of the Epo-induced MAP-kinase network.....	23
1.4.6 Endocytosis of the Epo receptor	28
1.4.7 References.....	33
2. Original Publications	39
2.1 Strategies for standardizing quantitative data.....	39
2.2 Quantitative data generation for systems biology	61
2.3 Cellular decisions predicted by MAPK modeling	70
2.4 Internalization determining Epo receptor activation kinetics	109
3. Appendix.....	141
3.1 Abbreviations.....	141
3.2 Curriculum Vitae	144
3.3 Erklärung	149

1.2 English summary

Systems biology aims at understanding how living organisms function in health and fail in disease by studying how new properties arise from dynamic interactions. Beyond qualitatively analyzing static data of individual components, dynamic quantitative data are combined with mathematical modeling to elucidate systems properties that determine cellular decisions. Such cell fate decisions are taken during erythropoiesis, when erythroid progenitor cells mature to erythrocytes by tightly regulated proliferation and differentiation processes, which are dependent on the cytokine erythropoietin (Epo) and the signaling transduction network activated by its receptor (EpoR).

A major bottleneck in systems biology is the lack of high-quality quantitative data. Therefore, we developed strategies for error reduction and algorithms for automated data processing, establishing the widely used techniques of immunoprecipitation and immunoblotting as highly precise methods for the quantification of protein levels and modifications. By randomized gel-loading we prevented correlated errors and further improved our data using housekeeping proteins or adding purified proteins to immunoprecipitation in combination with criteria-based normalization, enabling the generation of large and accurate sets of quantitative data.

Dysfunctional signaling in erythroid progenitor cells is associated with diseases such as anemia and leukemia, but the effects of interfering with the MAP-kinase signaling network are unknown. To causatively understand cell fate decisions and be able to predictably manipulate growth and maturation of erythroid progenitor cells, we applied a systems biology approach. We monitored components of the Epo-induced MAP-kinase network after stimulation of primary murine erythroid progenitor cells by quantitative immunoblotting. A dynamic mathematical model was compiled and kinetic parameters were estimated by multi-parameter fitting algorithms. We predicted that an increase in expression of a single ERK isoform would lead to feedback-mediated rerouting of signaling, which was confirmed by isoform-specific protein overexpression. The model was extended based on two hypotheses of negative feedback mechanisms. We experimentally confirmed feedback inhibition by phosphorylation as expressing a kinase-defective ERK isoform resulted in similar phenotypes as overexpression of the wild-type isoform. We demonstrated the influence of the integrated response of activated ERK on erythroid proliferation and differentiation, demonstrating that hyperactivation of the MAP-kinase signaling network leads to accelerated erythropoiesis but surprisingly to reduced hemoglobinization.

The input for signaling is critically dependent on the receptor presence on the cell surface. Endocytosis of cell surface receptors was thought to be responsible for long-term adaptation of a cell to a continuous stimulus. However, it remained to be identified what induces the rapid decline in signal transduction after activation of a cell surface receptor. We performed dynamic modeling of EpoR endocytosis, showing that the majority of internalized Epo is recycled to the medium. Sensitivity analysis revealed that the constant turnover of the receptor on the plasma membrane and ligand-induced internalization determine the sharp peak of EpoR activation. Furthermore, we predicted that the binding kinetics, but not the binding affinity determine the strength of EpoR signaling. Surprisingly, receptor internalization is crucial for rapid activation and deactivation of signaling, but irrelevant for long-term desensitization of cells.

In conclusion we employed mathematical modeling based on high-quality quantitative data, providing computational models of Epo-induced receptor endocytosis and MAP-kinase activation. Our systems biology approaches provided counterintuitive results that could not be obtained by conventional methods. For example, overexpression of an ERK isoform leads to rerouting of signaling and endocytosis of the EpoR is not required for long-term termination, but for rapid activation and deactivation of signaling. Furthermore, the mathematical models enable the identification of general systems properties and the sensitivity analyses predict targets for efficient interventions. In the future, this information can be used for drug design, opening new possibilities for treatments of anemia and leukemia.

1.3 Deutsche Zusammenfassung

Im Rahmen der Systembiologie werden neue Eigenschaften untersucht, die durch dynamische Wechselwirkungen von Einzelkomponenten entstehen. Dazu werden nicht nur qualitative statische Daten von Einzelmolekülen eingesetzt, sondern dynamische quantitative Daten in Kombination mit mathematischen Modellen untersucht. Auf diese Weise können entscheidungsbestimmende Systemeigenschaften identifiziert werden und Einblick in die Funktionsweisen von Lebewesen und in die Entstehung von Krankheiten erhalten werden. Präzise regulierte Proliferations- und Differenzierungsprozesse ermöglichen während der Erythropoese eine Expansion und Reifung der erythroiden Vorläuferzellen zu Erythrozyten. Diese Prozesse werden entscheidend von dem Zytokin Erythropoietin (Epo) und dem Signaltransduktionsnetzwerk des dazugehörigen Rezeptors (EpoR) bestimmt.

Gegenwärtig mangelt es in der Systembiologie an qualitativ hochwertigen quantitativen Daten. Wir haben Strategien zur Fehlerreduktion und Algorithmen zur automatischen Datenverarbeitung entwickelt und Immunpräzipitation und Immunoblot als hochpräzise Quantifizierungsmethoden für Proteinmengen und Proteinmodifikationen etabliert. Das randomisierte Laden der Proben auf Proteingele verhindert korrelierte Fehler. Die Datenqualität konnte durch kriteriengestützte Normierung basierend auf konstant exprimierten Proteine oder hinzugefügten Proteinstandards deutlich verbessert werden. Auf diese Weise können umfangreiche und genaue quantitative Datensätze erzeugt werden.

Eine defekte Signalleitung in erythroiden Vorläuferzellen kann Krankheiten wie Anämien und Leukämien auslösen. Ein Eingreifen in das MAP-Kinase-Signalleitungsnetzwerk könnte unvorhersehbare Effekte hervorrufen. Um zelluläre Entscheidungsfindungen ursächlich zu verstehen und um gezielt Wachstum und Reifung von erythroiden Vorläuferzellen beeinflussen zu können, haben wir einen systembiologischen Ansatz zur Analyse des von Epo aktivierten MAP-Kinase-Signalleitungsnetzwerk eingesetzt. Nach Stimulation von primären murinen erythroiden Vorläuferzellen wurden Bestandteile des MAP-Kinase-Signalweges mittels quantitativem Immunoblot untersucht. Ein dynamisches mathematisches Modell wurde etabliert und die kinetischen Parameter durch Mehrparameter-Anpassungsalgorithmen geschätzt. Modellverhersagen ergaben, dass eine Erhöhung des Expressionslevels einer einzelnen Isoform von ERK zu einer rückkoppelungsvermittelten Umleitung des Signalweges führen würde. Dies konnte durch isoform-spezifische Proteinüberexpression experimentell bestätigt werden. Das datenbasierte Modell wurde um zwei Hypothesen über den negativen Rückkoppelungsmechanismus erweitert. Die Rückkoppelungsinhibierung durch Phosphorylierung konnte experimentell bestätigt werden, da die Überexpression von ERK-Isoformen sowohl ohne als auch mit Kinaseaktivität zu gleichartigen Phänotypen bezüglich erythroider Proliferation und Differenzierung führte. Dabei konnten wir nachweisen, dass eine Überaktivierung des MAP-Kinase-Signalleitungsnetzwerkes zu schnellerer Differenzierung, überraschenderweise aber auch zu einer reduzierten Hämoglobinisierung führt.

Die Signalumsetzung hängt entscheidend von der Rezeptoranzahl an der Zelloberfläche ab. Es wird angenommen, dass die Endozytose von Zelloberflächenrezeptoren verantwortlich für die langfristige Anpassung einer Zelle an einen anhaltenden Stimulus ist. Durch dynamische Modellierung der Endozytose des EpoR konnte nachgewiesen werden, dass der Grossteil des internalisierten Epo in das Medium zurückgeschleust wird. Sensitivitätsanalysen zeigten, dass der konstante Umsatz des Rezeptors an der Plasmamembran und die ligandeninduzierte Internalisierung den steilen Anstieg und die schnelle Abnahme der Rezeptoraktivität in der Anfangsphase bestimmen. Außerdem wurde vorausgesagt, dass die Bindungskinetik und nicht die Bindungsaffinität die Signalstärke des EpoR ausmachen. Da der internalisierte Rezeptor nach anfänglicher Abnahme in erheblichem Ausmaß wieder an die Oberfläche zurückkehrt, ist die Rezeptorinternalisierung überraschenderweise ausschlaggebend für die rasche Signalaktivierung und -deaktivierung, aber unerheblich für die langfristige Desensibilisierung der Zelle.

Zusammenfassend wurden qualitativ hochwertige quantitative Daten mit mathematischen Modellen kombiniert und dabei Computermodelle der Epo-induzierten Endozytose und MAP-Kinase Signalaktivierung erstellt. Die systembiologischen Ansätze lieferten überraschende Resultate, die nicht mit herkömmlichen Methoden erzielt werden konnten: Die Überexpression einer ERK-Isoform führt zur Umleitung des Signalweges und die Endozytose des EpoR beeinflusst die rasche Aktivierung und Deaktivierung der Signalleitung. Die datenbasierten Modelle ermöglichen außerdem die Identifizierung von allgemeingültigen Systemeigenschaften und durch Sensitivitätsanalysen die Vorhersage von Angriffspunkte für effiziente Interventionen. In Zukunft können solche datenbasierte Modelle für die gezielte Entwicklung von Medikamenten eingesetzt werden und eröffnen auf diese Weise neue Möglichkeiten zur Therapie von Anämien und Leukämien.

1.4 Summary of the results

1.4.1 Systems biology

Systems biology aims at understanding how living organisms function in health and fail in disease. The new insight is that the important properties of life occur as a result of the connections between individuals, from molecules to cells. Thus, systems biology studies how new properties that are functionally important for life, arise in interactions (Alberghina and Westerhoff 2005).

The term 'systems' is derived from the general systems theory (von Bertalanffy 1968) or more specifically from systems dynamics. Attempts for system-level understanding have a long tradition, dating back to Norbert Wiener (Wiener 1948). This level of understanding has gained new interest due to the explosive progress of genome sequencing projects and the massive amounts of data generated by high-throughput experiments in genomics, proteomics, and metabolomics. It is becoming increasingly evident that certain aspects of biology can only be understood at the system-level (Kitano 2001). While an understanding of genes and proteins is the basis for systems biology, the focus is on understanding the network's structure and dynamics. Because a system is not just an assembly of genes and proteins, its properties cannot be fully understood merely by drawing diagrams of their interconnections (Kitano 2002). Or, as Olaf Wolkenhauer put it in a modified quote by Henri Poincaré: "A cell is built up of molecules, as a house is with stones. But a soup of molecules is no more a cell than a heap of stones is a house" (Wolkenhauer et al. 2005). To understand how a particular system works, knowing the components and interactions is merely the first step. More importantly, the strength of the interactions, the mode of signal encoding and the robustness of the signal against noise and perturbations have to be determined. And, the way the system reacts if a malfunction occurs as in cancer or other diseases has to be identified. The techniques employed include quantitative measurements, modeling, reconstruction and theory (Kirschner 2005). If we know the design principles and circuit patterns, we can learn how to modify the network to improve system performance and interfere against malfunctions.

To understand a biological network at the systems-level, four key properties can be analyzed: (I) System structure. The analysis of gene interactions networks, the protein interactosome and biochemical pathway interactions and of the mechanisms these interactions regulate. (II) System dynamics. The analysis of a network behavior over time. Methods include metabolic analysis, sensitivity analysis, phase plane and bifurcation analysis. (III) Control method. By identifying mechanisms that control the state of a cell,

potential drug targets can be identified. Also known as metabolic control analysis (MCA). (IV) Design method. The modification and construction of biological systems having desired properties based on design principles and simulations. Also known as synthetic biology, an area of research that combines science and engineering in order to design and build novel biological functions and systems.

These properties can be analyzed by hypothesis-driven research (Fig. 1). Through close collaborations between theoreticians and experimentalists, an iterative research process is established. This iterative cycle begins with a review of biological knowledge and the selection of contradictory issues. Then, new experiments are devised by hypothesis-driven experimental design generating quantitative data of pathway components. Mathematical models are built that can describe the data. Using this model, *in silico* predictions are obtained that are validated or falsified experimentally. If the predictions are not fulfilled, the model has to be changed accordingly. Based on the mathematical model, new experiments are designed, creating novel biological knowledge.

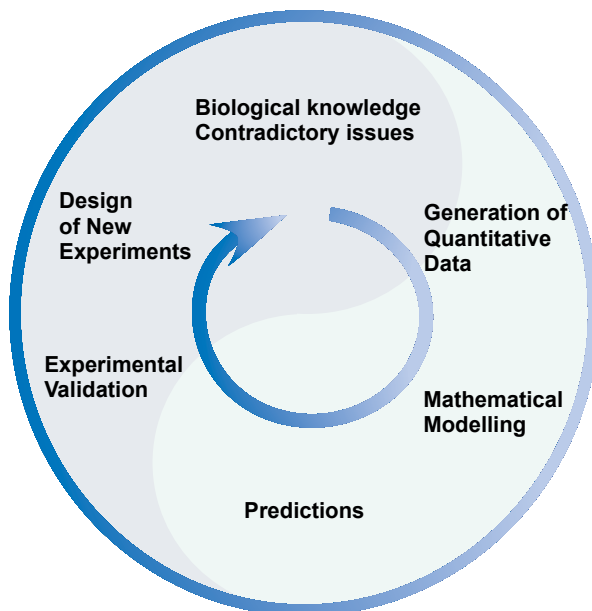


Figure 1: Systems biology approaches consist of an iterative process between experiments and modeling. In a close collaboration between experimentalists and theoreticians, contradictory issues in biology are answered in an iterative cycle of quantitative data generation, mathematical modeling, *in silico* predictions, experimental validation and design of new experiments. The advancement of research in computational science, analytical methods, technologies for measurements and improved high-throughput methodologies will gradually transform biological research to become a more systematic and hypothesis-driven science, as represented by this iterative cycle process.

A property of biological system that has gained new interest is robustness. Robustness is an essential property of biological systems (Csete and Doyle 2002). Three phenomena can be classified that are a result of robustness: (I) Adaptation, the ability to cope with environmental changes. This allows the cell to respond to relative changes in stimuli from the medium. (II) Parameter insensitivity, the ability to cope with changes in internal kinetic parameters. This also comprises robustness against intracellular noise, i.e. variance in the expression level of system components. Furthermore, threshold behavior against low stimulus concentrations prevent signaling in the absence of ligand. (III) Fault-tolerance, also known as graceful degradation. This property allows the system to continue operating properly in the event of the failure of some of its components. If the system's performance

drops, the decrease is proportional to the severity of the failure, as opposed to minor fluctuations causing complete breakdown.

In engineering science, robustness is critical for operation of the system and obtained using the following strategies (I) Negative feedback and feed-forward control stabilizing the system. (II) Redundancy, several components sharing the same function act as backup. (III) Structural stability, intrinsic mechanisms promote robustness. (IV) Modularity, subsystems are functionally or physically separated, in order to prevent failure in one subsystem from spreading to other parts of the system.

These tactics employed in engineering are discovered in biological systems as well. Bacterial chemotaxis is a good example, featuring several of the phenomena and strategies conferring robustness (Kollmann et al. 2005). It is a perfect adaptation system, intrinsically insensitive to intracellular noise and features several feedback loops stabilizing the system.

Systems biology promises new biological insights by mathematical modeling of complex cellular networks based on experimental data. Currently, this branch of science still suffers from a lack of quantitative data. Technical innovations in experimental devices and data processing are critical aspects of systems biology research. Furthermore, standardized computational tools for data acquisition, storage and analysis are needed.

Mathematical models integrating various signal transduction cascades and molecules will provide pharmaceutical industries with mechanism-based drug discovery strategies (Noble 2002). Using this models, the effects of drugs as well as side effects can be predicted. This may lead to unforeseen results. For example, pharmaceutical industries have focused their research on oncogenes. However, if an oncogene is mutated, its activity or expression level are elevated to an extent that inhibiting this protein has little effect on the system. On the other hand, some of the nonmutated genes may become more important compared to the healthy situation. This implies that signaling proteins encoded by nonmutated genes should represent better drug targets against cancer (Hornberg and Westerhoff 2006). Thus, although systems biology is still in its infancy, its potential benefits are considerable in both basic and translational research. By combining quantitative data with mathematical modeling, systems biology can provide new insights for biology and medicine.

1.4.2 Erythropoietin and the hematopoietic lineage

The process of hematopoiesis is highly regulated by both humoral growth factors and cytokines as well as intracellular contacts in the stem cell niche in bone marrow. Besides erythrocytes, the hematopoietic stem cell gives rise to the other blood cells of the lymphoid and myeloid lineage (Weissman et al. 2001). Several highly adaptive developmental processes are regulated by cytokines and their cognate receptors, the cytokine receptor superfamily. A well-studied example of cytokine signaling is erythropoietin (Epo) signaling through the erythropoietin receptor (EpoR) involved in erythropoiesis, the continuous renewal of red blood cells from hematopoietic stem cells.

Erythrocytes provide the important function of oxygen transport from the lung into tissue. Multi-level control of erythropoiesis ensures both constant replenishment of erythrocytes and fast adaptation in case of blood loss. Failure of these control processes leads to anemia or leukemia. Red blood cell production in the mouse is classified into distinct stages in development. Primitive erythropoiesis occurs in the yolk sac 7.5 days post conception, producing large nucleated erythrocytes. At day 12, production of red blood cells shifts to the fetal liver and enucleated definitive erythrocytes are generated. Around day 15 to 16 post conception, definitive erythropoiesis is relocated to the bone marrow and spleen where it occurs for the remaining life of the mouse (Hoffman et al. 1995). Mice are born at 19 to 21 days post conception. In human embryos, the yolk sac serves as the initial site of primitive erythropoiesis from week 3 to 6 post conception. The fetal liver functions as the site of definitive erythropoiesis from week 6 to 22. After week 22 of gestation, the bone marrow becomes the predominant and lifelong site of blood cell production (Palis and Yoder 2001). Humans are usually born around 40 weeks post conception.

The definitive erythropoietic lineage in mice is depicted in Fig. 2. Hematopoietic stem cells in the bone marrow give rise to colony forming units granulocytes, erythrocytes, monocytes, macrophages (CFU-GEMM) that commit to the erythroid lineage by differentiating into burst forming units erythroid (BFU-E). These cells develop into colony forming units erythroid (CFU-E), which express erythropoietin receptors (EpoR) on their surface and require erythropoietin (Epo) for survival and differentiation. This has been verified by EpoR knockout mice that show severe anemia due to a block in maturation of CFU-E cells and thus die at midgestation (Wu et al. 1995). Triggered by Epo, CFU-E mature into proerythroblasts and erythroblasts, gradually downregulating EpoR expression, increasing hemoglobin content and the expression of surface markers such as Ter119 and reducing cell size. After ejecting the nucleus, reticulocytes enter the blood stream and terminally differentiate into the typical oxygen-transporting doughnut-shaped erythrocytes.

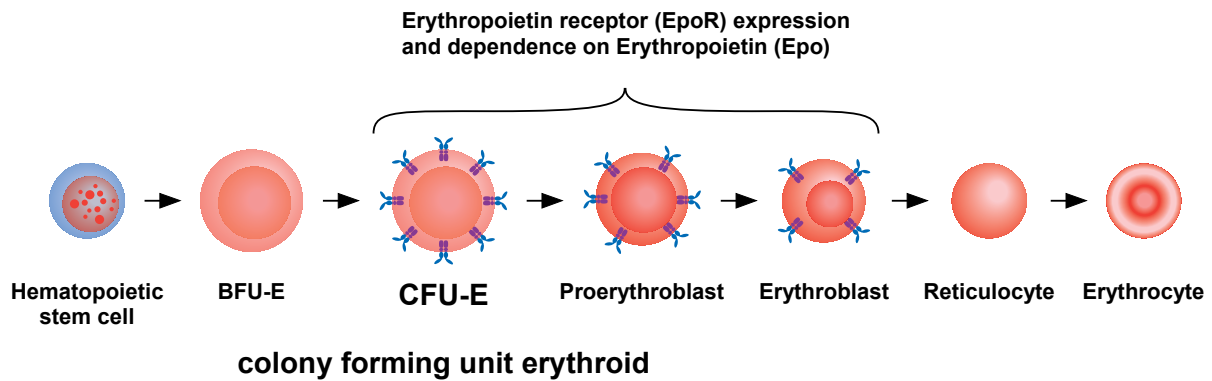


Figure 2: Erythropoietic lineage in mice. Hematopoietic stem cells give rise to erythrocytes undergoing several developmental stages. Erythropoietin (Epo) is the key regulator of red blood cell production, being necessary for survival and differentiation of erythroid progenitor cells from the colony forming unit erythroid to the erythroblast stage. At the late erythroblast stage, cells eject their nucleus and successively mature into erythrocytes with their characteristic doughnut-shaped appearance.

The murine EpoR, which was isolated from an expression library of murine erythroleukemia cells (D'Andrea et al. 1989), is a member of the type I cytokine receptor superfamily with a single transmembrane domain. Crystallographic studies demonstrated that the EpoR exists as a preformed dimer at the cell surface (Livnah et al. 1999). Binding of an Epo molecule to the receptor induces a conformational switch (Remy et al. 1999) that activates the associated kinase Janus kinase (JAK) 2. JAK2 is prebound to the receptor dimer by the proline-rich EpoR motifs Box 1 and Box 2. Upon Epo-binding, JAK2 undergoes transphosphorylation and thereby gets activated. Activated JAK2 phosphorylates eight tyrosine residues within the EpoR cytoplasmic domain that serve as docking sites for Src homology 2 (SH2)-containing signaling proteins. JAK2 is essential for definitive erythropoiesis, as revealed by the analysis of JAK2 knock-out mice having similar phenotypes as EpoR knock-out mice (Neubauer et al. 1998).

The most direct pathway to induce gene expression is represented by signal transducer and activator of transcription (STAT) 5. STAT5 binds to phosphorylated tyrosines Y343 and Y401, and to a minor extent Y429 and Y431 (Klingmüller et al. 1996). STAT5 is phosphorylated by JAK2, dimerizes, translocates into the nucleus and activates target gene expression, including the anti-apoptotic protein Bcl-X_L (Socolovsky et al. 1999) and the negative regulator cytokine-inducible SH2 domain-containing (CIS) protein (Matsumoto et al. 1997). STAT5 cycles between nucleus and cytoplasm, thereby acting as a remote sensor constantly monitoring receptor activation (Swameye et al. 2003).

Another protein network activated after Epo stimulation is the phosphoinositide 3-kinase (PI3K) pathway. The regulatory subunit p85 of PI3K can either bind to phosphorylated Y429 of the EpoR (Klingmüller et al. 1997) or to GRB2 associated binding protein (Gab) 1 and 2, which are associated with the phosphorylated EpoR via the scaffold proteins GRB2 and SHC (Ravichandran et al. 1995). PI3K signaling leads to phosphorylation of protein

kinase B (PKB/Akt) and protein kinase C (PKC) ϵ . PI3K activation was proposed to trigger proliferation of primary erythroid cells (Zhang and Lodish 2004).

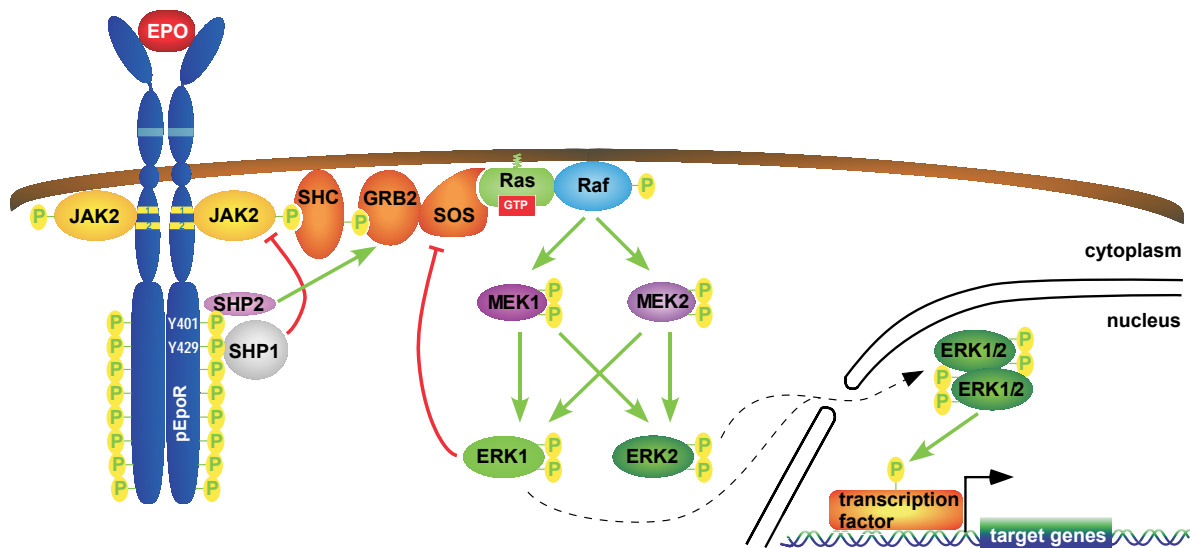


Figure 3: The erythropoietin-induced MAP-kinase signaling network. Phosphorylated EpoR leads to the recruitment of the GRB2-SOS complex to the plasma membrane. Activation of Ras triggers the Raf/MEK/ERK phosphorylation cascade. Activated ERK can dimerize, enter the nucleus and induce gene transcription. SHP1 and ERK act as negative feedback loops, inhibiting JAK2 and GRB2-SOS, respectively.

The recruitment of the complex GRB2 - son of sevenless (SOS) to the receptor triggers activation of the mitogen-activated protein (MAP)-kinase pathway (Fig. 3). GRB2 directly interacts with phosphorylated Y464 of the EpoR (Barber et al. 1997) or indirectly by binding to either SH2 inositol 5-phosphatase (SHIP)1 (Mason et al. 2000) or SH2 containing phosphatase (SHP)-2 (Tauchi et al. 1996). Furthermore, SHC can bind to phosphorylated JAK2 (He et al. 1995) and serve as adaptor protein for GRB2. In every case, EpoR activation leads to recruitment of the guanine-nucleotide exchange factor SOS to the membrane. Ras is loaded with GTP, which leads to activation of Raf, the first protein of the MAP-kinase cascade (see next chapter). An alternative route for MAP-kinase activation has been proposed through activation of PKC via the PI3K pathway (Karnitz et al. 1995).

Signal termination is mediated by negative feedback loops. Activated SHP-1 inhibits JAK2 activity when bound to phosphorylated Y429 of the EpoR. ERK1/2 activation leads to phosphorylation of SOS, triggering dissociation of the Grb2-SOS complex from the receptor. Furthermore, endocytosis of activated receptor complexes was proposed to terminate signaling by depleting the cell surface of receptors.

The MAP-kinase cascade is critical for the differentiation of erythroid progenitor cells as K-ras^{-/-} mice die between 12 and 14 days of gestation with fetal liver defects and evidence of anemia (Johnson et al. 1997). Furthermore, c-Raf knock-out studies suggested that Raf delays terminal differentiation (Kolbus et al. 2002) and overexpression of constitutively active

Ras blocks terminal erythroid differentiation (Zhang and Lodish 2004) in cells stimulated with both Epo and tyrosine receptor kinase ligands.

1.4.3 The MAP-kinase signaling network

The mitogen activated protein (MAP)-kinase pathway is one of the best studied signal transduction pathways. It is conserved from yeast to humans. The MAP-kinase cascade consists of three modules. Upon a certain stimulus, a MAP-kinase kinase kinase (MAPKKK) is activated that phosphorylates a MAP-kinase kinase (MAPKK) on two serine residues. MAPKK is a dual specific kinase, phosphorylating MAP-kinase (MAPK) on a threonine and a tyrosine residue. Now, MAPK can serine/threonine phosphorylate several target proteins, eliciting a specific biological response (Fig. 4 left panel).

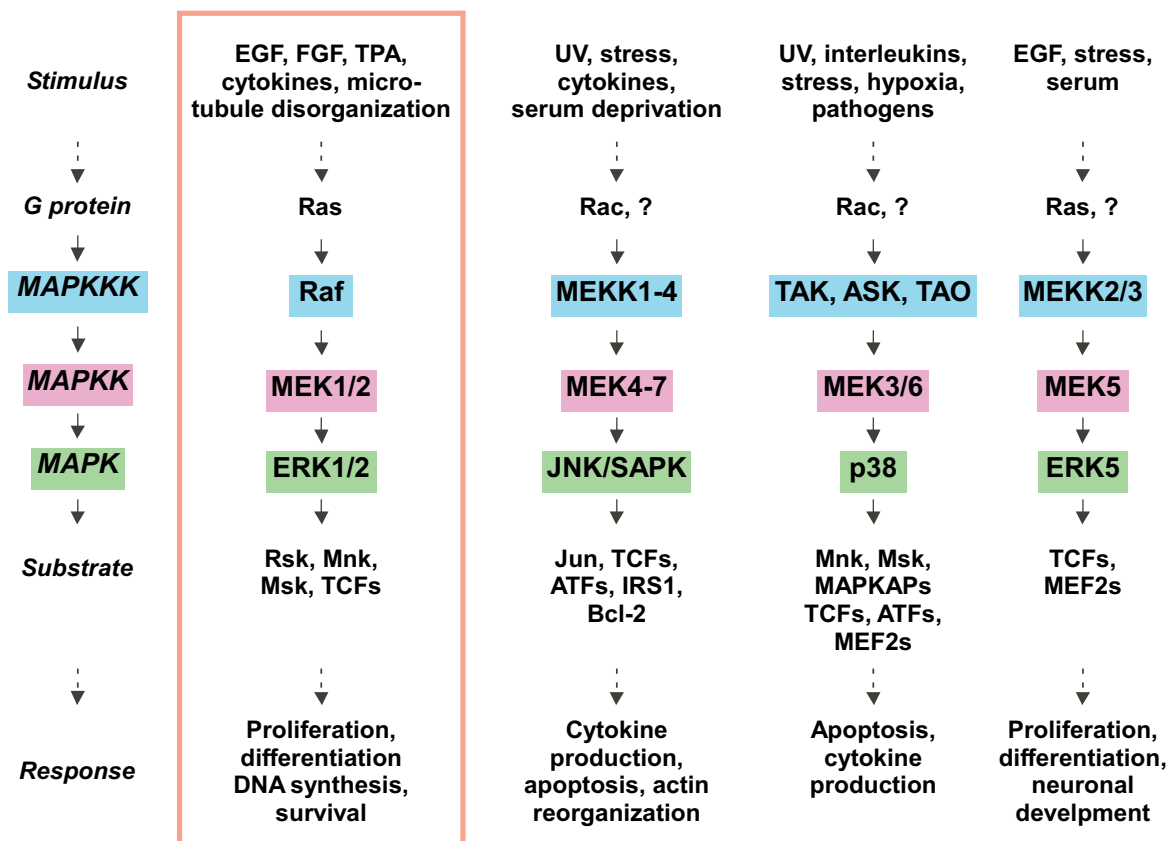


Figure 4: MAP-kinase modules in mammals. A stimulus leads to activation of a small G protein, which activates a MAPKKK. Subsequently, MAPKK and MAPK are phosphorylated. Phosphorylation of several substrate proteins elicits various biological responses. In mammals, five MAP-kinase modules are known, of which the Raf/MEK/ERK cascade is best studied. See appendix for abbreviations.

In budding yeast, five MAP-kinase modules are known, controlling various biological responses including mating, filamentation, osmolyte synthesis, cell wall remodeling and sporulation (Pouyssegur and Lenormand 2003). In mammalian cells (Fig. 4), besides the

classical Raf/MEK/ERK module, JNK/SAPK, p38 and the MEK5/ERK5 cascade have been described (Chen et al. 2001).

Considering the multitude of MAP-kinase modules, the question arises how specificity can be assured and cross talk between these conserved proteins can be prevented. This issue is particularly important in yeast, where e.g. the MAPKKK STE11 is involved in the mating, osmolarity as well as filamentation pathways. One explanation lies in scaffolding proteins. The scaffold STE5 binds the MAP-kinases responsible for mating, while the MAPKK PBS2 binds STE11 and the MAPK HOG1, triggering osmolyte synthesis (Elion 1998). In mammals, the role of scaffolding proteins is less obvious. Several scaffolding proteins have recently been described, including kinase suppressor of Ras (KSR), MEK-partner 1 (MP1) and β -arrestins for the Ras/MEK/ERK pathway, β -arrestin-2 and JNK interacting proteins (JIP1/2/3) for the JNK cascade, and JIP2, JIP4 and osmosensing scaffold for MEKK3 (OSM) for the p38 MAPK module (Dard and Peter 2006). Besides the functions described for yeast, these scaffolds are likely to increase signaling efficiency by orienting the kinases and restrict signaling to specific subcellular locations.

MAP-kinase activation has a prominent role in cancer as it frequently promotes cellular proliferation. Indeed, the MAP-kinase pathway is hyperactivated in 30% of human cancers. Ras and Raf are both well-known proto-oncogenes, Ras mutations were found in 90% of pancreas adenocarcinoma, 50% of thyroid tumors and 30% of myeloid leukemia (Bos 1989). Constitutive activation of ERK1/2 has been found in 20 – 45% of human tumor cell lines and solid tumors (Hoshino et al. 1999). Many attempts have been made to target this pathway for cancer therapy. Promising results were obtained *in vitro* with the second-generation MEK1/2 inhibitor PD184352 (CI-1040) (English and Cobb 2002) and its derivative PD0325901. Since the results in the clinic were disappointing, combinations with other therapeutics or targeting genetically defined tumor subtypes that are MEK-dependent (Solit et al. 2006) might be a solution.

In the rat pheochromocytoma cell line PC12, the MAP-kinase pathway is activated by both EGF and NGF or FGF. However, upon EGF stimulation, these cells proliferate, while upon NGF or FGF, the cells differentiate into neuronal-like cells with neurites (Tsuji et al. 2001). It was demonstrated that the same pathway can elicit two fundamentally different responses by using different activation kinetics (Marshall 1995). EGF leads to a transient MAP-kinase activation, whereas NGF/FGF triggers sustained MAPK activation (Yamada et al. 2004). Only sustained MAP-kinase activation can induced nuclear translocation of ERK (Pouyssegur and Lenormand 2003) and induce gene transcription. A set of immediate early gene products contain docking site for ERK, FXFP (DEF) domains. These proteins are post-translationally stabilized by activated ERK, thus sensing sustained MAP-kinase activation (Murphy et al. 2004).

Activation of the classical Raf/MEK/ERK MAP-kinase pathway involves the following steps. Upon binding of a ligand to a cytokine receptor or a receptor tyrosine kinase, the receptor is phosphorylated on tyrosine residues. The SH2-domain containing adaptor molecules SHC and GRB2 can bind, recruiting the guanine exchange factor SOS to the plasma membrane. SOS catalyzes the exchange of Ras-bound GDP for GTP, leading to Ras activation. Ras activates Raf, which in turn phosphorylates MEK. MEK phosphorylates ERK. ERK can dimerize and activate transcription factors (Fig.5).

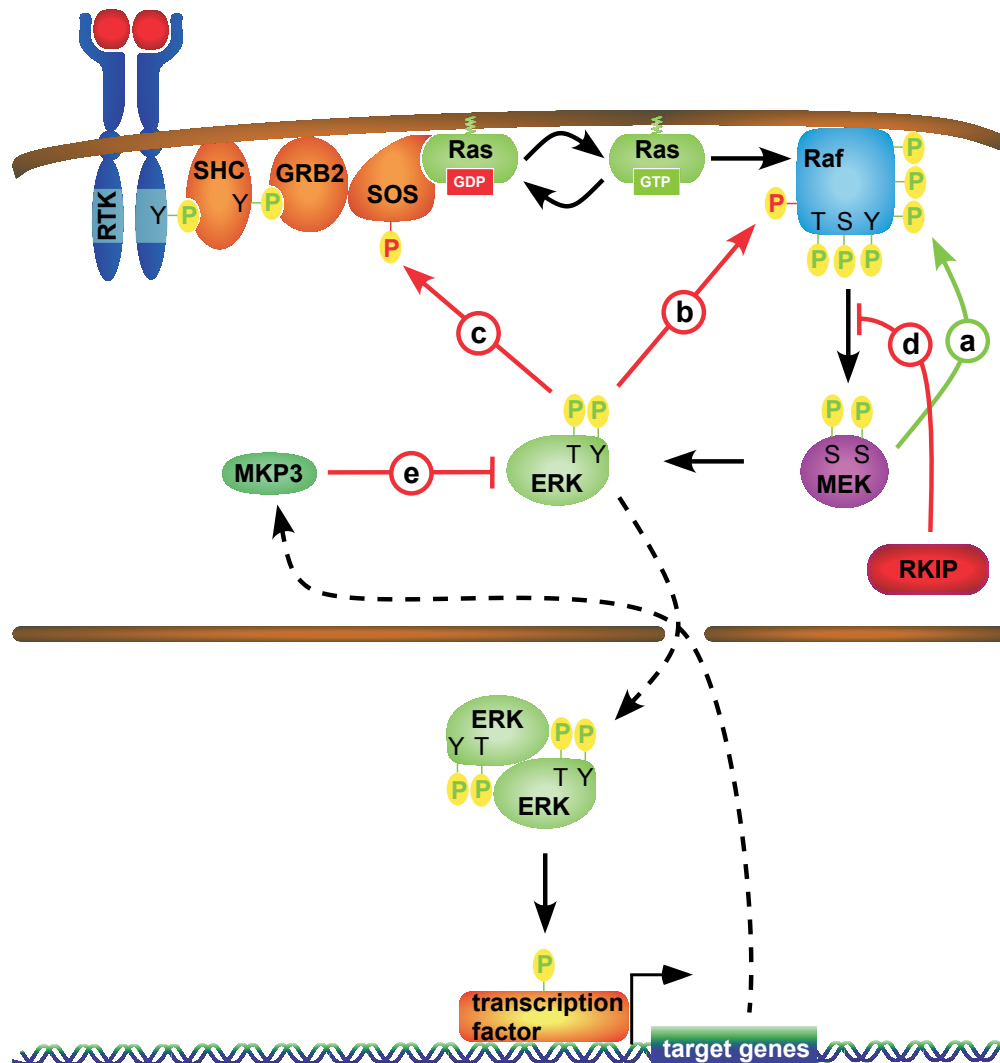


Figure 5: Feedback loops in MAP-kinase signaling. Positive and negative feedback loops include a) activation by phosphorylation, b) inhibition by phosphorylation, c) disruption of protein interaction by phosphorylation, d) competitive disruption of protein interaction and e) inactivation by dephosphorylation.

Several positive and negative feedback loops fine-tune this system. Activated MEK1 mediates hyperphosphorylation of c-Raf, which increases c-Raf kinase activity (Fig. 5a) (Zimmermann et al. 1997). ERK2 phosphorylates serine/threonine residues at the C-terminus of B-Raf, thereby reducing its biological activity (Fig. 5b) (Brummer et al. 2003). Activation of MAP kinases through Ras results in phosphorylation of SOS, uncoupling the

SOS/GRB2 complex from tyrosine kinase substrates (Fig. 5c) (Buday et al. 1995). RKIP competitively disrupts the interaction between c-Raf and MEK (Fig 3d) (Yeung et al. 1999) and NGF-stimulation of PC12 cells induces expression of MKP3, a dual specificity phosphatase that inactivates ERK1/2 (Fig. 5) (Camps et al. 1998). MKP3/DUSP6 is among the genes showing the highest transcriptional upregulation after prolonged ERK activation by constitutive active Ras (Nils Blüthgen, personal communication).

There is a wide variety of mathematical models of the MAP-kinase pathway which have led to novel insights and predictions as to how this system functions (Orton et al. 2005). The first task for modeling is to determine the stoichiometry and concentration of signaling molecules, an important information that has often been neglected. Fig. 6 summarizes the data compiled by Ferrell (1996) for frog oocytes and isoform-specific data of CFU-E and BaF3 cells that were obtained by us using the methods described in the next chapter. It became evident from this data that the cytoplasmic proteins MEK and ERK are much more abundant in the cell than the membrane associated proteins Ras and Raf.

Yeast (molecules/cell)¹

MAPKKK	STE11	unknown
MAPKK	STE7	< 200
MAPK	FUS3	500

Xenopus oocytes (molecules/cell)¹

MAPKKK	Mos	1×10 ⁹
MAPKK	Mek1	350×10 ⁹
MAPK	Erk2	100×10 ⁹

Mammalian cells (molecules/cell)¹

	NIH 3T3	Cos
Ras	2000	
MAPKKK	c-Raf	1000
MAPKK	MEK1/2	360 000

Mammalian cells (molecules/cell)²

		CFU-E	BaF3
MAPKK	MEK1	110 000	55 000
	MEK2	237 000	250 000
MAPK	ERK1	73 500	82 000
	ERK2	209 000	250 000

Figure 6: Stoichiometry of the MAP-kinase cascade. Several hundred molecules of MAP-kinase signaling components are present per yeast cell. In the large frog oocytes, billions of molecules of each proteins have been determined. In mammalian cells, signaling components range from thousand to more than a million molecules per cell. ¹ Ferrell (1996) ² Schilling et al, manuscript in preparation

By a combination of quantitative experimental data and mathematical modeling, several systems properties of the MAP-kinase pathway were demonstrated. Ultrasensitivity and positive feedback loops can trigger switch-like responses in *Xenopus* oocytes (Ferrell and Machleder 1998). In some mammalian cells, the MAP-kinase signaling network can switch from a bistable system with a sharp threshold to a monostable proportional response system (Bhalla et al. 2002). As switch-like response systems are intrinsically sensitive to noise, the robustness of the MAP-kinase was analyzed, revealing that the building cascades composed of molecules activated by multisite phosphorylation provides a robust method to obtain switch-like behavior and that this kind of cascade is robust to changes of most parameters (Blüthgen and Herzog 2003). The combination of negative feedbacks and

ultrasensitivity could induce oscillations in the MAP-kinase pathway, a theory that was demonstrated theoretically (Kholodenko 2000) but awaits experimental validation.

Computational MAP-kinase models including receptor activation have revealed novel insights into cellular signaling. For PC12 cells, an ODE-based mathematical model of the EGF signal-transduction pathway was developed to investigate the factors influencing the kinetics of ERK cascade activation. The model consisted of 30 reactions involving 29 species (Brightman and Fell 2000). The kinetic constants of reactions and initial concentrations of molecules were largely based on a range of measured or estimated values published in the existing scientific literature. The analysis of this model indicated that negative feedback inhibition of the ERK cascade by phosphorylation of SOS was the most important factor in determining whether the cascade activation was transient or sustained. Sustained activation of ERK was achieved by increasing SOS dephosphorylation. These differences in feedback regulation were likely to explain the characteristic patterns of EGF- and NGF-induced ERK activation in PC12 cells.

An ODE-based mathematical model describing the dynamics of the EGF-induced signal-transduction pathway consisting of 125 reactions involving 94 species was proposed by Schoeberl et al. (2002). They included receptor internalization via two distinct routes in this model. Kinetic parameters were again based on values published in the scientific literature, and initial concentrations were either compiled from the literature or based on laboratory experiments. A major conclusion of the modeling process predicted that the velocity of receptor activation, which depends on ligand binding kinetics, determines the cellular response to EGF. Therefore, a ligand with higher affinity would result in faster ERK activation than a ligand with lower affinity.

An extensive ODE-based model of EGF- and NGF-induced MAP-kinase activation in PC12 cells was published by Sasagawa et al. (2005). The model consisted of 22 molecules and 106 rate constants and is among the most comprehensive to date, as it includes both the Ras and Rap1 pathways to ERK activation and combines theoretical predictions with experimental validation. The model was used to investigate how EGF encodes transient ERK activation, while NGF encodes sustained dynamics of ERK activation. The analysis indicated that transient activation of ERK was dependent on the change of EGF levels rather than the absolute concentration of EGF. In contrast, NGF-induced sustained ERK activation was dependent on the final concentration of NGF, but not on relative changes in NGF levels. Thus, the signaling networks can distinguish between absolute ligand abundance or changes in ligand concentrations. EGF leads to the activation of Ras, while NGF primarily activates Rap1. Whereas Ras is inactivated by Ras-GAP, which is recruited to the membrane upon receptor activation, Rap1 is inactivated by constitutively active Rap1-GAP. Therefore, EGF leads to transient signaling, whereas NGF triggers sustained ERK activation. Thus, systems

biology could provide an answer to the long-lasting questions regarding MAP-kinases induced proliferation versus differentiation in PC12 cells.

1.4.4 New strategies for quantitative immunoblotting

Accurate quantitative data are the basis for mathematical models. A major bottleneck in current systems biology is generation of high quality quantitative data. While new high-throughput technologies have produced large amounts of data, the quality of these data sets are often insufficient for mathematical modeling because many high-throughput methods sacrifice specificity for scale (Troyanskaya et al. 2003). Therefore, new strategies to improve data quality are needed.

Cellular signals are transduced by changes in amount and modifications of proteins. The most common used method for protein detection is immunoblotting, also known as Western blotting. For immunoblotting, the protein mixture is separated in an SDS polyacrylamide gel by electrophoresis (SDS-PAGE). The proteins are subsequently transferred to a membrane and the protein of interest is revealed with specific antibodies. The protein amount can be determined if the antibody is detectable by fluorescence or chemiluminescence. However, conventional chemiluminescence detection by x-ray films offers only limited linearity. Therefore, CCD-camera based detection with a high linear dynamic range is preferable. The relative protein amount can be calculated from the light intensity, being directly proportional to the amount of the epitope recognized by the antibody.

The most common strategy for normalization of blotting data involves the determination of the ratio of the protein of interest relative to a constant protein (Albeck et al. 2006). However, this often increases the error, since both proteins are not exactly quantifiable and are present in different concentrations and molecular weights.

We have shown that the errors of immunoblotting data are not randomly distributed, but, due to the structure of the acrylamide polymer, appear as correlated errors on the membrane. This becomes manifest in a wavelike concentration gradient for intrinsically constant proteins. The reason for this are irregularities in the polyacrylamide gel and inhomogeneities in transfer to the membrane. Furthermore, we could demonstrate that the data quality improves if the samples of a time series experiment are not loaded chronologically onto the electrophoresis gel, but in a randomized fashion (Schilling et al. 2005b). This does not prevent the errors, but eliminates their autocorrelation. The blotting error is therefore transformed into Gaussian noise, which can be corrected using mathematical methods (Fig. 7).

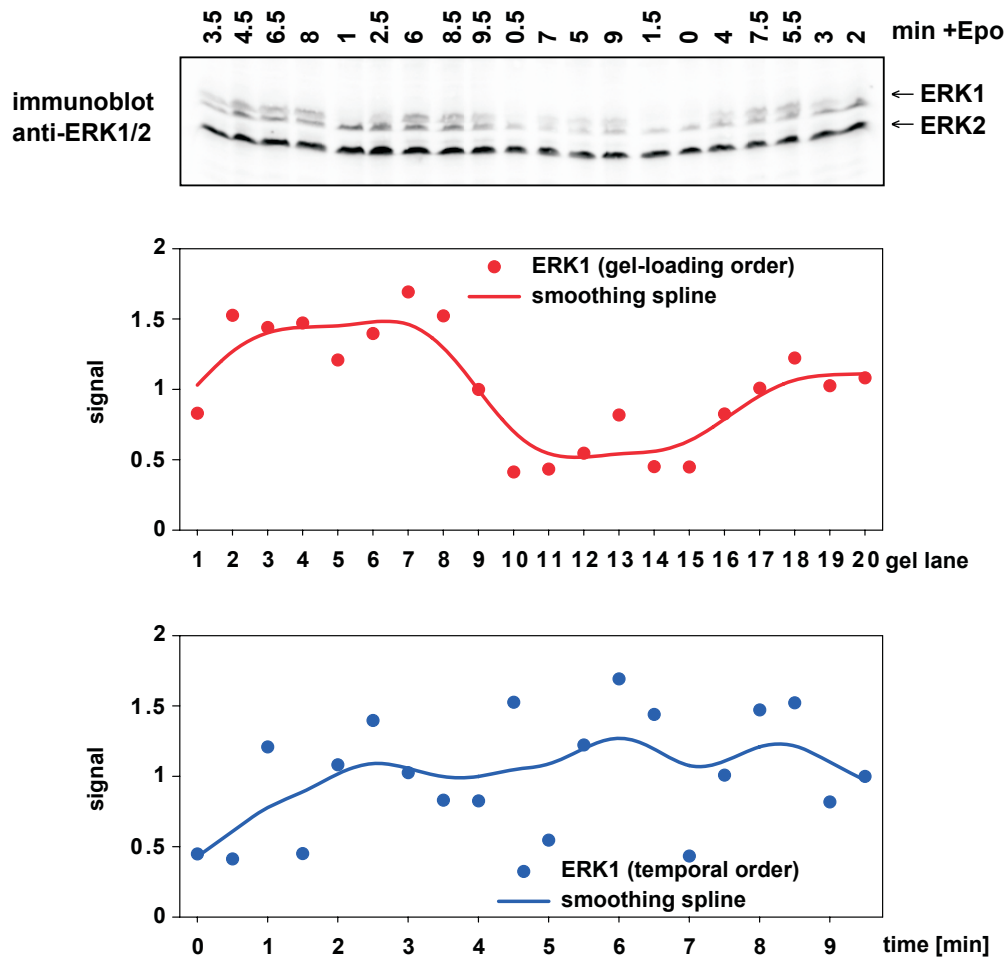


Figure 7: Randomized immunoblotting prevents error correlations. BaF3 cells expressing the erythropoietin receptor were stimulated for 10 min with erythropoietin (Epo). Every 30 s samples were taken and loaded on a SDS polyacrylamide gel in a randomized order, separated and transferred to a membrane. The quantified data of the protein ERK1 are depicted both in gel-loading and temporal order. Displaying the data in gel-loading order leads to a pseudo-dynamic that is caused by a blotting error. This can be eliminated with a smoothing spline if the data is displayed in temporal order.

The best way to classify the errors is using a smoothing spline, a mathematical function consisting of joined polynomials and representing the approximate characteristics of a data series. For error determination we use e.g. proteins of the cytoskeleton, assuming these proteins to be in constant concentration in the cell and not to be modified upon stimulation. We termed these proteins 'normalizers'. A spline is calculated, smoothing the data of the normalizer. Since the blotting error is not only apparent in horizontal, but also vertical direction on the membrane, it is crucial that the normalizer has similar molecular weight than the protein of interest. The protein data are corrected using this spline (Fig. 8).

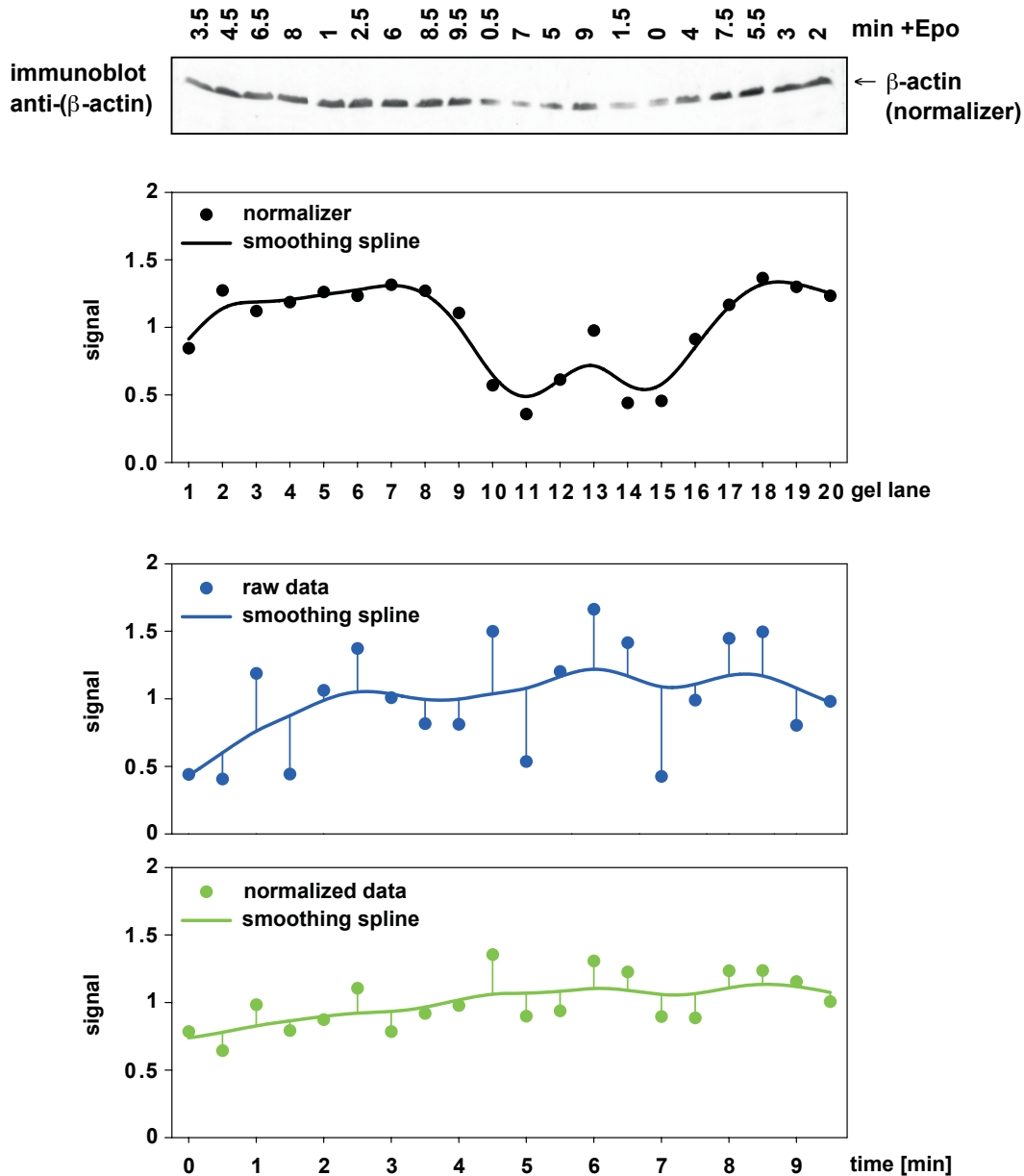


Figure 8: Normalization of immunoblotting data. The immunoblot depicted in Fig. 7 was incubated with an antibody against the cytoskeletal protein β -actin, serving as normalizer. A smoothing spline for the quantified data was calculated and used to normalize the ERK1 data. The distance of the data points to the spline is reduced and as expected a straight line for the time course arises.

The protein of interest frequently needs to be enriched prior to immunoblotting by immunoprecipitation (IP). The protein bound to an antibody is linked to a Protein A-sepharose matrix that can be separated from the cellular lysate by centrifugation. Indications for an IP include proteins of low cellular concentrations or the application of generic phospho-specific antibodies for detection. These antibodies recognize the phosphorylation of several proteins; IP allows the identification of a specific phosphorylated protein. No normalizers are present after IP. To circumvent this problem, we developed recombinant proteins harboring the same epitope as the precipitated protein, but differing slightly in molecular mass. We termed these proteins 'calibrators'. They can be added to the lysate prior to IP, therefore being

subjected to the same treatment as the protein of interest. On the immunoblot, the calibrator appears above or below the protein band and is used for data normalization. The concentration of the calibrator can be easily determined, allowing us to quantify the absolute amount of the protein of interest. Converted by the number of lysed cells, the number of proteins per cell and the stoichiometry of cellular components can be obtained, an important information for mathematical modeling and for the analysis of systems properties.

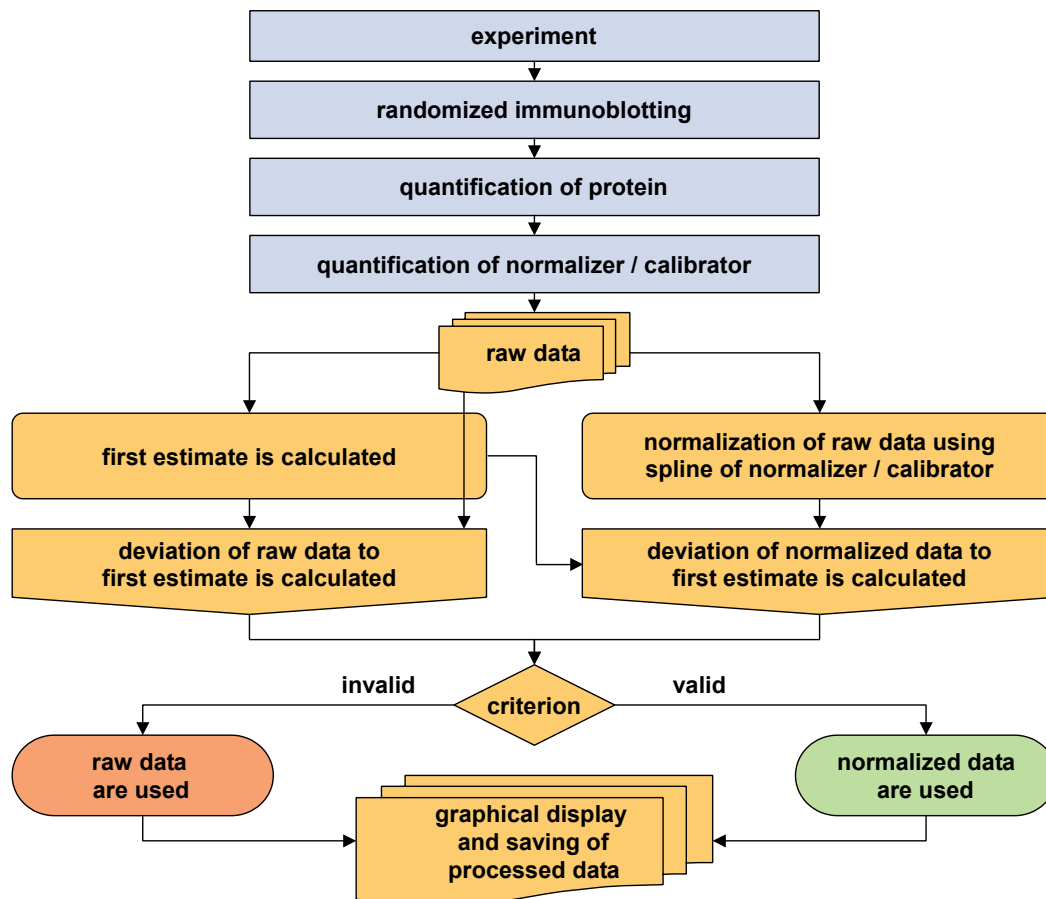


Figure 9: Automated data processing by GELINSPECTOR. The computer program imports quantified blotting data, performs data processing and verifies the normalization. Subsequently the processed data are saved.

To automate our methods we developed the computer software GELINSPECTOR (Schilling et al. 2005a). This program imports the quantified data of the measured proteins and normalizers / calibrators, calculates smoothing splines and performs data normalization. GELINSPECTOR is able to integrate data of several immunoblots. This is an important feature if the number of samples to be analyzed exceeds the number of lanes on the gel. For an unbiased data correction we introduced criteria. GELINSPECTOR calculates a first estimate of the data. For time series, the first estimate can be obtained by a smoothing spline. For samples analyzed in replicates, the mean of the data is used, while for dose response experiments a linear or sigmoidal relationship of stimulation dose and signal is calculated. Normalization is valid if the distance of the data to the first estimate is reduced after

normalization. Otherwise, the original data are retained. Thereby we can prevent deterioration of the data, as e.g. by an inaccurate measurement of the normalizer. The processed data are saved both as spreadsheet and diagram (Fig. 9).

Automatic data normalization and processing is crucial for quantitative data generation. Furthermore, to ensure transferability of data and to avoid problems due to heterogeneity of the biological background, standardization of the biological system under investigation and the methods applied is absolutely critical (Klingmüller et al. 2007). By developing quantitative immunoblotting as a robust and reliable technique for quantitative data acquisition under standardized conditions, we established an easy to handle and cost-effective method that permits the assembly of large data sets with high temporal resolution.

1.4.5 Modeling of the Epo-induced MAP-kinase network

To causatively understand Epo-induced MAP-kinase activation and to be able to predictably manipulate erythroid differentiation of progenitor cells, we performed mathematical modeling. A dynamic pathway model of the Epo-induced MAP-kinase signaling network was compiled comprising activation cascades and negative feedback loops (Schilling et al. 2007). A major challenge in establishing computational models is the frequent discrepancy between the number of parameters that have to be estimated and the amount of measured data points. To deal with this, we made use of the strategies discussed in the last chapter, measuring as many time points and proteins as feasible after continuous stimulation of primary murine erythroid progenitor cells. Furthermore, total concentrations of several proteins were measured. Using saturation binding assays with radiolabeled Epo as discussed in the next chapter, we determined the amount of cell surface EpoR. Additional concentrations were determined by quantitative immunoblotting using calibrators or standard proteins. Thus, the number of parameters to be estimated was decreased. Finally, we determined dependent parameters in our model by iterative rounds of parameter estimation and identifiability testing. Parameter dependencies detected include directly correlated and hyperbolically dependent parameters as well as parameter triplets describing a two-dimensional surface. We therefore could fix non-identifiable parameters to estimated values, enabling us to identify the remaining parameters with a standard deviation of less than 10%. Using our model, we could describe our immunoblotting data with sufficient accuracy. The analysis of the model variables revealed that the two negative feed-back loops delayed activation of the phosphatase SHP1 and inhibitory phosphorylation of SOS by activated ERK1/2 was responsible for fast deactivation of signaling despite continuous stimulation with Epo.

We performed sensitivity analysis to identify the parameters having the profoundest impact on signal propagation. Further simulations predicted that overexpression of a single ERK isoform would lead to feedback-mediated signal rerouting. This could be verified by isoform-specific overexpression, demonstrating that overexpression of an ERK isoform increases the negative feedback on SOS, which results in decreased activation of upstream molecules and the other ERK isoform. In contrast, overexpression of Raf was predicted to increase signaling of both ERK isoforms.

To investigate the impact of overexpressing MAP-kinase components on differentiation and proliferation of erythroid progenitor cells, we established protocols for preparation, retroviral transduction and sorting of primary murine CFU-E cells (Fig. 10).

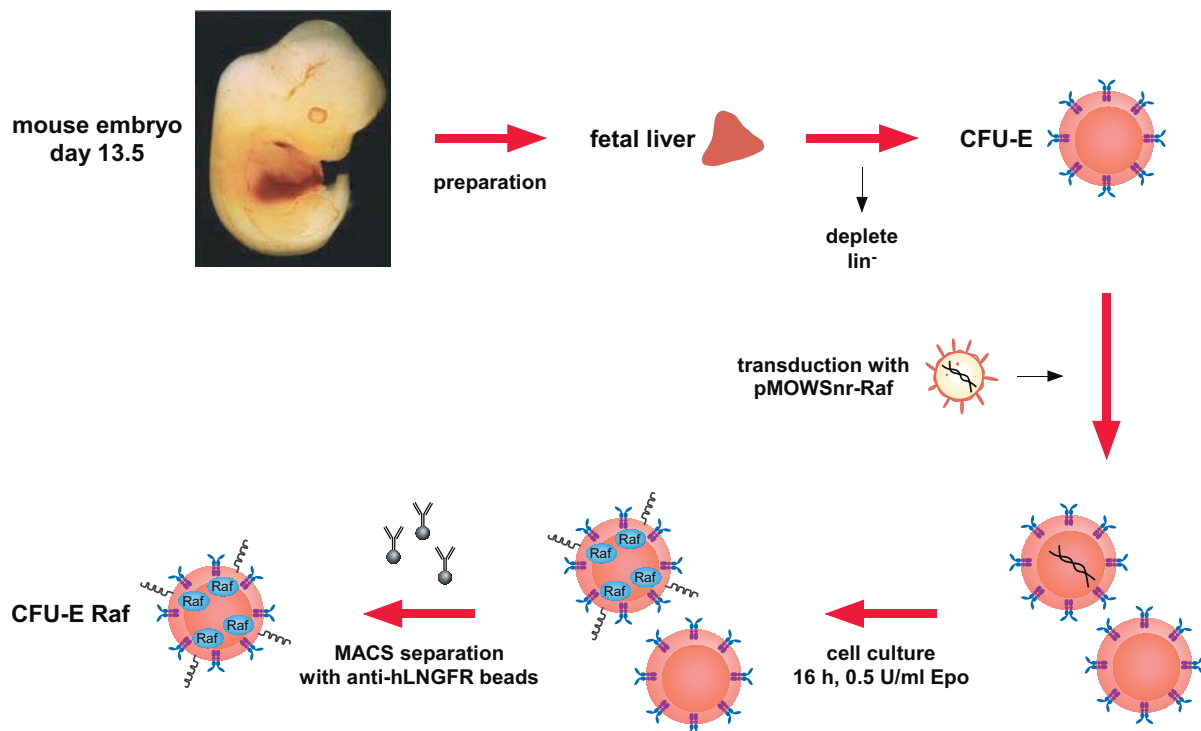


Figure 10: Preparation, retroviral transduction and sorting of erythroid progenitor cells. CFU-E cells are prepared from fetal livers and enriched by lineage depletion. After transduction with retroviral supernatants, cells are cultured overnight and positively selected by MACS sorting.

Primary CFU-E cells were prepared from fetal livers of Balb/c mouse embryos 13.5 days post conception as described (Ketteler et al. 2002a). CFU-E cells were retrovirally transduced with supernatants of Phoenix eco cells transfected with vectors harboring the gene of interest under control of the retroviral promoter and the cDNA coding for a truncated human low affinity nerve growth factor receptor (hLNGFR) under control of an SV40 promoter. Cells were cultivated overnight to allow for expansion and expression of the gene of interest. To enrich for transduced cells, positive selection was performed using magnetic cell sorting (MACS) with anti-hLNGFR beads.

To analyze the effect of overexpression of MAP-kinase components on proliferation of erythroid progenitor cells, CFU-E transduced with c-Raf, ERK1 or vector control were

cultivated for 16 h and positively sorted. Cells were incubated with Epo concentrations ranging from 10^{-3} to 10 U/ml and cultivated for 4 h. [^3H]-thymidine was added and cells were grown overnight. To measure proliferation, the amount of incorporated [^3H]-thymidine was determined (Fig. 11, left panel). As expected, vector control transduced cells showed a sigmoidal Epo-dependent proliferation curve as reported previously for the myeloid cell line 32D (Klingmüller et al. 1995). However, ERK1 and c-Raf transduced cells showed reduced Epo-dependent proliferation. In accordance with our model prediction, the effect of overexpressing c-Raf was stronger than overexpressing of ERK1. The phenotypes correlated with the simulated integrated response of ppERK1/2, defined as the combined area under the curve of activated ERK1 and ERK2. To investigate whether this effect is due to accelerated differentiation or increased apoptosis, we determined the apoptotic fraction of transduced and sorted erythroid progenitor cells using the TdT-mediated dUTP-FITC nick-end labeling (TUNEL) assay (Ketteler et al. 2003). It became apparent that a very low number of cells were apoptotic, independent of the vectors transduced (Fig. 11, right panel). Thus, hyperactivation of the MAP-kinase pathway reduces proliferation of erythroid progenitor cells without inducing apoptosis.

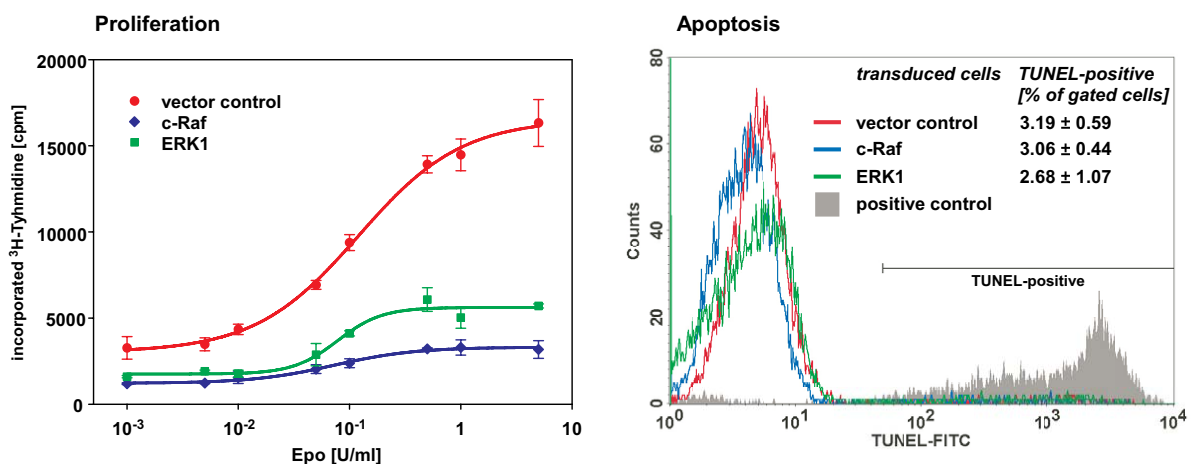


Figure 11: Overexpression of MAP-kinase components reduces proliferation of erythroid progenitor cells without inducing apoptosis. CFU-E were prepared and transduced with vector control, c-Raf or ERK1. Cells were cultivated overnight and proliferation was measured using thymidine incorporation, showing reduced proliferation for c-Raf and ERK1 transduced cells. Apoptosis was determined using TUNEL assay with low numbers of apoptotic cells for each transduction.

We extended the model to predict the effect of expressing kinase-defective ERK based on two hypotheses. In the first model, SOS activity is inhibited by complexation with ERK followed by phosphorylation; in the second model, phosphorylation is necessary to inhibit SOS signaling. If the first model was correct, overexpression of kinase-defective ERK would lead to reduced ERK signaling. Conversely, the second model predicted that overexpressing a kinase-defective ERK isoform would lead to reduced negative feedback signaling, resulting in similar phenotypes as overexpression of the wild-type isoform.

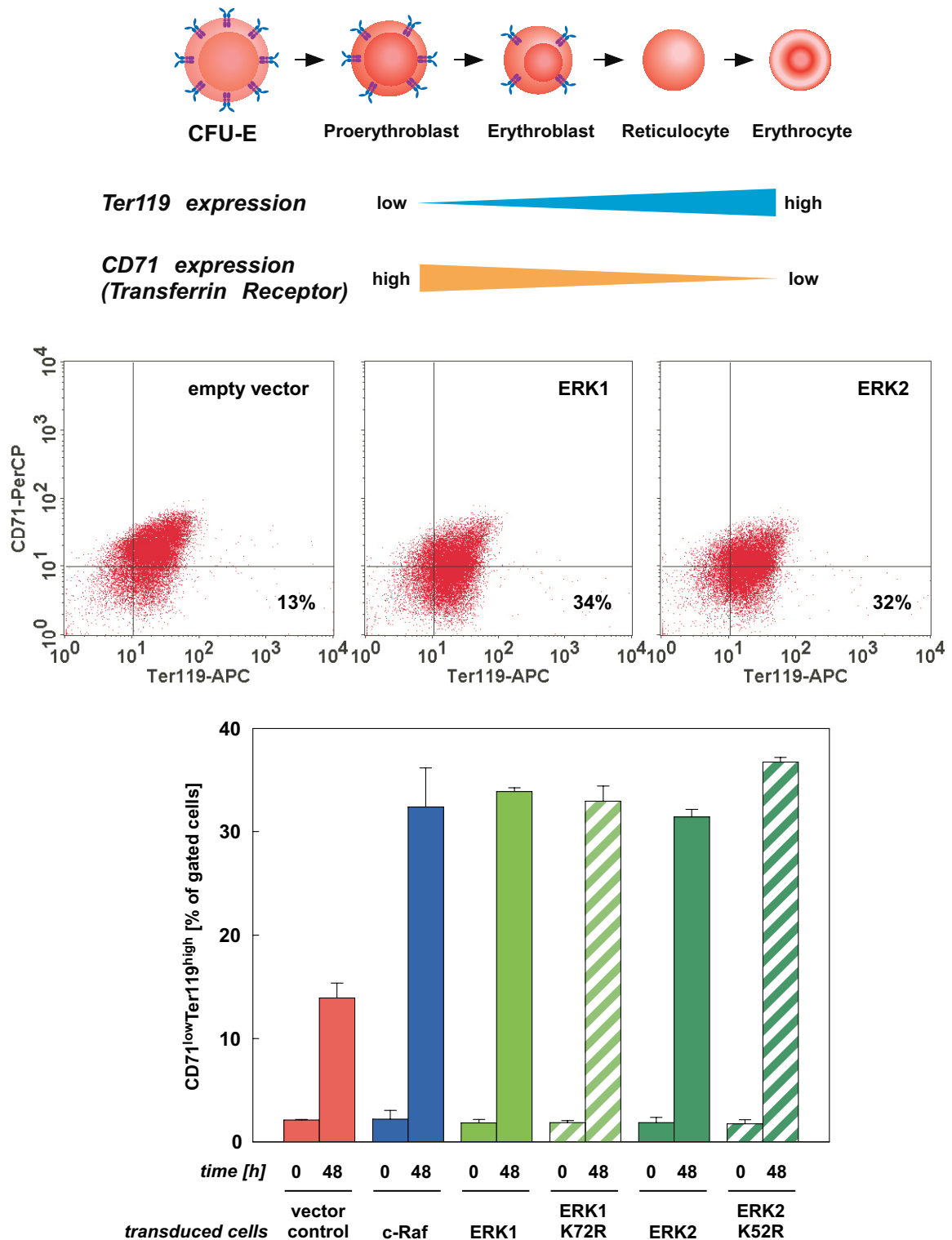


Figure 13: Overexpression of MAP-kinase components accelerates differentiation of erythroid progenitor cells. Differentiation of murine erythroid progenitor cells is characterized by increasing Ter119 and decreasing CD71 expression. CFU-E cells were prepared and transduced with vector control, c-Raf, wild type or kinase defective ERK. Differentiation was measured by FACS analysis directly after transduction and after two days, showing accelerated differentiation for c-Raf and ERK transduced cells.

Differentiation of murine erythroid progenitor cells can be analyzed using two cell surface markers. Ter119 is gradually upregulated, while CD71, the transferrin receptor, is downregulated during maturation from proerythroblasts to erythrocytes (Dumitriu et al. 2006).

To distinguish between the two models predicted, we transduced CFU-E with c-Raf, ERK1, ERK2, the kinase defective mutants ERK1 K71R and ERK2 K52R as well as vector control and measured differentiation by surface flow cytometry using antibodies against Ter119 and CD71. The fraction of cells expressing low levels of CD71 and high levels of Ter119 were counted directly after transduction and after culture in serum-free medium with 0.5 U/ml Epo for 48 h. Overexpression of MAP-kinase components resulted in larger fractions of differentiated cells, both for wild type and kinase defective ERK constructs (Fig 13). Thus, the second model is correct, demonstrating that phosphorylation, but not complexation of SOS inhibits signaling. Furthermore, we showed that hyperactivation of the MAP-kinase signaling network leads to accelerated partial differentiation in murine erythroid progenitor cells.

The function of erythrocytes critically depends on their hemoglobin concentration. Hemoglobin is accumulated during terminal maturation of erythroblasts. To investigate hemoglobinization in erythroid progenitor cells, we transduced CFU-E cells with c-Raf, ERK1, ERK2, the kinase defective mutants ERK1 K71R and ERK2 K52R as well as vector control. Hemoglobinization was measured by intracellular flow cytometry using antibodies against hemoglobin α . Cells were cultured in serum-free medium with 0.5 U/ml Epo for 48 h. Increased expression of hemoglobin was visible in all cells. However, overexpression of MAP-kinase components reduced hemoglobinization, both for wild type and kinase defective ERK constructs (Fig. 12). Thus, MAP-kinase hyperactivation leads to reduced hemoglobinization in primary erythroid cells.

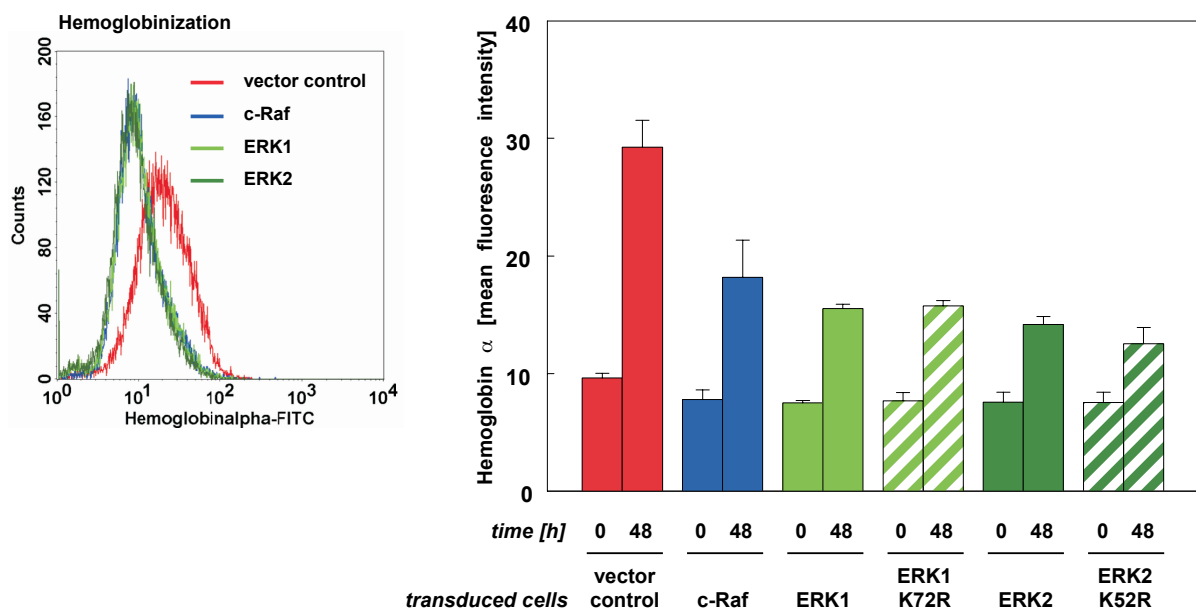


Figure 12: Overexpression of MAP-kinase components reduces hemoglobinization of erythroid progenitor cells. CFU-E were prepared and transduced with vector control, c-Raf, wild type or kinase-defective ERK. Hemoglobinization was measured by FACS analysis directly after transduction and after two days, showing reduced hemoglobinization for c-Raf and ERK transduced cells.

Therefore, hyperactivation of the Epo-induced MAP-kinase pathway in primary murine CFU-E by overexpression of signaling components accelerates differentiation and thereby reduces proliferation of erythroid progenitor cells. This does not induce apoptosis, but reduces hemoglobinization, possibly by restricting the time window cells can express the globin genes. Contrary results were obtained with gene knock-out studies, as c-Raf deficient erythroblasts were shown to differentiate much faster than their wild-type counterparts and cells expressing a kinase-dead c-Raf differentiated faster *in vitro* (Rubiolo et al. 2006). C-Raf deficient embryos are growth-retarded and anemic and die at midgestation. It was speculated that this anemic phenotype is due to premature erythroblast differentiation at the expense of renewal, depleting the fetal liver of erythroid precursors (Kolbus et al. 2002). However, these results were obtained with heterogeneous population of erythroid progenitor cells cultivated in presence of receptor tyrosine kinase ligands that strongly activate the MAP-kinase pathway. Our results were obtained using a pure population of CFU-E cells cultivated with Epo as the only growth factor, showing that hyperactivation of the Epo-induced MAP-kinase pathway in erythroid progenitor cells directly accelerates differentiation and reduces hemoglobinization

In conclusion, by combining mathematical modeling with the generation of a large set of quantitative data in primary erythroid progenitor cells, we provided a computational model of the Epo-induced MAP-kinase signaling network. The systems biology approach applied provided unexpected and counterintuitive results such as feedback-mediated signal rerouting and overexpression of kinase-defective and kinase-active isoforms leading to similar phenotypes. These predicted effects were validated by stimulation of transduced primary murine erythroid progenitor cells with Epo and by the analysis of the differentiation and hemoglobinization of these cells. Additionally, the sensitivity analysis provided novel targets for efficient and predictable interventions. These targets can be evaluated in the future using mathematical models, opening new possibilities for the treatment of anemia and leukemia.

1.4.6 Endocytosis of the Epo receptor

The first steps in activation of a signal transduction pathways such as the MAP-kinase signaling network depends on ligand binding to its cognate receptor. The responsiveness of a cell to a specific ligand is determined by the amount and specificity of receptors on the plasma membrane. Receptor surface expression is a function of maturation and internalization kinetics of the receptor and crucial for activation of signal transduction. Recruitment of inhibitory molecules to the receptor such as phosphatases and the activation of negative feed-back loops contribute to signal termination (Hilton 1999; Schlessinger 2000). In addition, ligand-induced receptor endocytosis of receptor tyrosine kinases has been

proposed to be involved in downregulation of cell surface receptors, thereby reducing the cellular responsiveness towards the ligand (Waterman and Yarden 2001). The erythropoietin receptor (EpoR) is produced in the endoplasmic reticulum, processed and matured in the Golgi complex as a preformed dimer and translocated to the plasma membrane, where it can bind to its ligand, erythropoietin (Epo). Both free EpoR and Epo-bound receptors can be internalized to early or recycling endosomes, the latter allowing return to the cell surface. After trafficking to late endosomes, ligand and receptor are degraded in lysosomes (Fig 14).

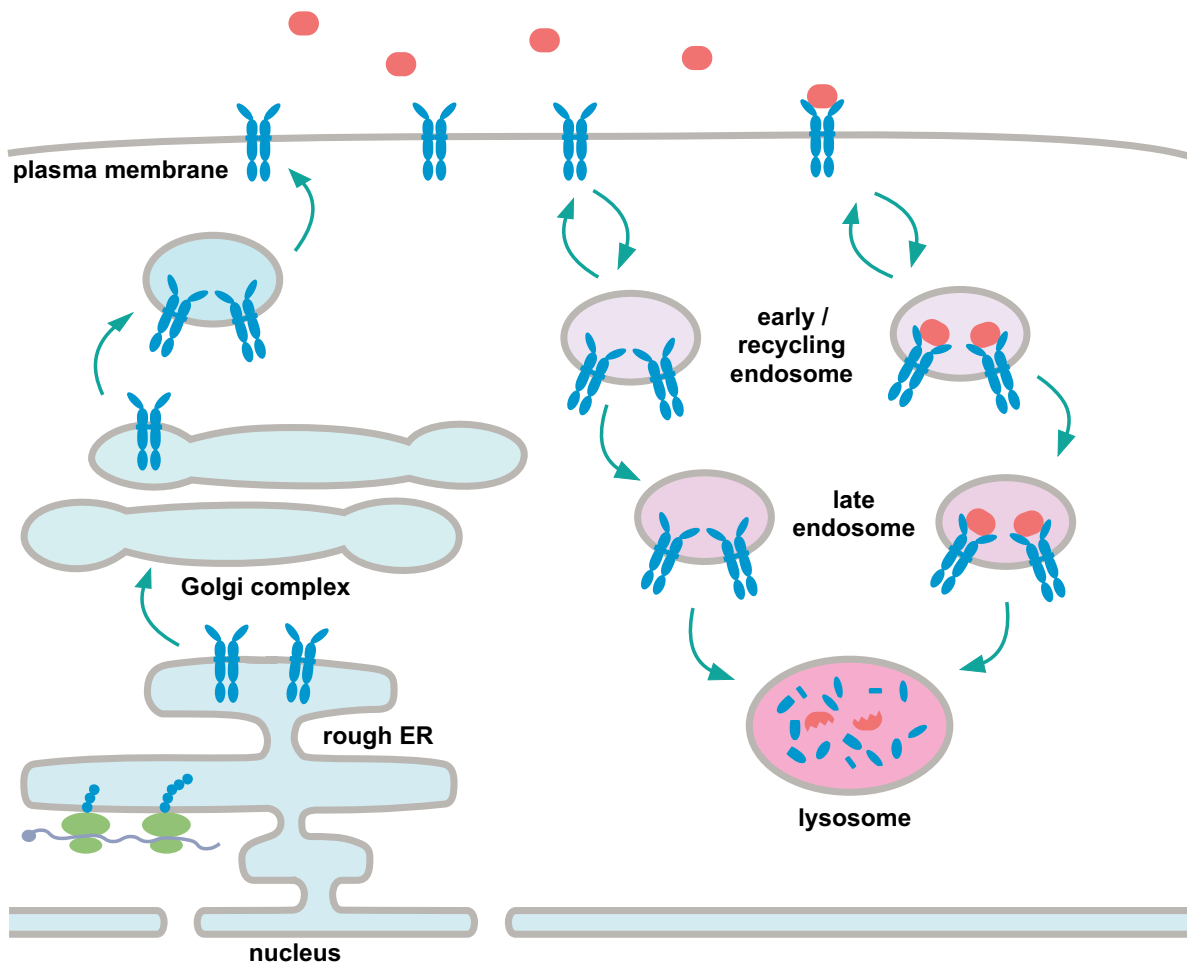


Figure 14: Schematic representation of maturation and internalization of the erythropoietin receptor (EpoR). Preformed receptor dimers mature in the endoplasmic reticulum (ER) and Golgi complex and traffic to the plasma membrane, where they can bind to their ligand, erythropoietin (Epo). Receptors on the plasma membrane are subject to constant turnover by constitutive and ligand-induced internalization to early and recycling endosomes. After internalization, both receptor and ligand can get processed to late endosomes, followed by degradation in lysosomes.

Ligand-induced EpoR internalization has been shown to depend on cytoplasmic residues of the EpoR (Levin et al. 1998). EpoR endocytosis independent of JAK2 activation or receptor tyrosine phosphorylation was documented (Beckman et al. 1999). Furthermore, the steady state of cell surface EpoR can be perturbed by inhibiting cysteine proteinases, resulting in reduced degradation and increased amounts of the receptor at the plasma membrane (Neumann et al. 1996). Internalization and degradation kinetics of recombinant

human Epo was determined by kinetic modeling, showing that about 40% of internalized Epo is degraded, while 60% is resecreted in a biologically active form (Gross and Lodish 2006).

Since it was proposed that internalization of plasma membrane receptors is responsible for adaptation to stimulation and thereby for termination of signaling, we performed data-based mathematical modeling of the dynamic behavior of EpoR activation and endocytosis. To first step in modeling of receptor activation was to determine the amount of binding sites on the plasma membrane and the affinity of the ligand to its receptor. Saturation binding assays with radiolabeled ligands can be used to determine the following two parameters: K_D , the affinity of the ligand to the receptor, which is the ratio of the off-rate k_{off} divided by the on-rate k_{on} , and B_{max} , the amount of maximally bound ligand. Knowing B_{max} , the amount of cell surface receptors can be calculated, as each Epo molecule is supposed to be bound by an EpoR dimer.

We performed saturation binding assays with three types of cells. First, cells of the murine pro B cell line BaF3 were retrovirally transduced with HA-tagged EpoR (Ketteler et al. 2002b) and selected with puromycin. Second, CFU-E cells were isolated from murine fetal livers using the methods described in the previous chapter. Third, CFU-E cells were cultivated in serum-free medium containing 0.5 U/ml Epo for 16 h. These cells resemble the early proerythroblast differentiation stage, as they have already slightly increased hemoglobin content without yet having reduced their size (see Fig. 2 for erythroid lineage in mice).

To determine maximal binding and binding affinity, cells were incubated with increasing concentration of [125 I]-Epo, free from cell-bound [125 I]-Epo was separated and radioactivity was measured using a gamma counter. Unspecific binding was determined by incubating the cells with excess unlabeled Epo. Specifically bound Epo was plotted against free Epo and a one-site regression analysis was performed. B_{max} and K_D were calculated, indicating maximal specific binding and free Epo concentration for half-maximal specific binding, respectively (Fig. 15, upper panel). To visualize the results in a linear way, data was displayed according to Scatchard (Scatchard 1949). The ratio of specifically bound to free Epo was plotted against specifically bound Epo and the regression curves were linearized, with B_{max} representing the intercept on the abscissa and K_D the negative reciprocal of the slope (Fig 4, lower panel). As expected, Scatchard plots showed a linear relationship, indicating a single Epo binding site.

Both BaF3 HA-EpoR and freshly isolated CFU-E cells express a receptor with an affinity of about 150 pM, consistent with published data. CFU-E cells cultivated for 16 h, however, express a receptor with much weaker affinity. According to the values for B_{max} , BaF3 HA-EpoR cells express about 15000 molecules of EpoR on their surface, freshly isolated CFU-E cells about 2900 molecules and CFU-E cells cultivated for 16 h about 18000

molecules per cell. Interestingly, CFU-E cells seem to decrease the affinity but increase the expression of the EpoR during cultivation in Epo.

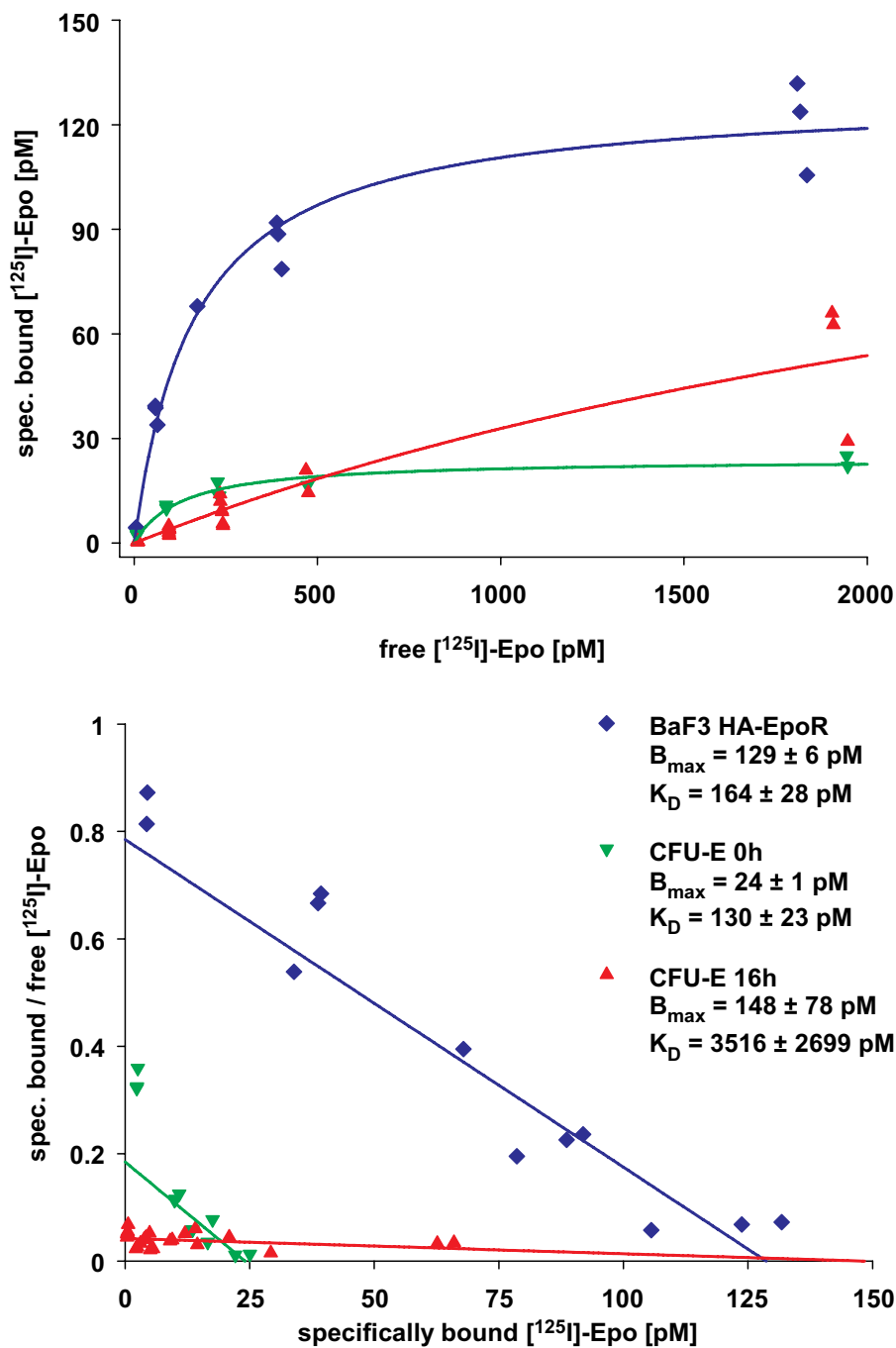


Figure 15: Saturation binding and Scatchard analysis for specifically bound $[^{125}\text{I}]\text{-Epo}$. BaF3 HA-EpoR cells, freshly isolated CFU-E cells and CFU-E cells cultivated for 16 h were incubated with increasing concentrations of $[^{125}\text{I}]\text{-Epo}$ and specifically bound Epo was plotted versus free Epo. B_{\max} and K_D were calculated using one-site saturation regressions. Data was plotted according to Scatchard with solid lines representing the B_{\max} and K_D values calculated by one-site saturation regression.

To analyze ligand-induced and constitutive EpoR endocytosis, we performed data-based mathematical modeling. We determined the kinetics of EpoR-bound $[^{125}\text{I}]\text{-Epo}$ in BaF3 HA-EpoR cells at 37°C (Becker et al. 2007). To measure the kinetics of constitutive EpoR

internalization, we expressed streptavidin binding peptide tagged (SBP)-EpoR in BaF3 cells. BaF3 SBP-EpoR cells were treated with [¹²⁵I]-streptavidin at 37°C and constitutive EpoR endocytosis was measured. Two ordinary differential equation (ODE)-based models were compiled and parameters were estimated. The parameters revealed that ligand binding to the EpoR increases receptor internalization by a factor of 3.5. Approximately 80% of internalized Epo are recycled to the medium intact, while only 20% are degraded. The trajectories of receptor-bound Epo, both on the surface and within the cell, describe a sharp peak in the first minutes followed by a rapid decrease with the concentration rising gradually again. We performed sensitivity analysis to examine *in silico* which parameters determine this intriguing kinetics. A larger on-rate k_{on} results in a higher, sharper and faster peak, while varying the affinity constant K_D had little effect. Therefore, it is more informative to determine the on-rate kinetics than to measure the affinity of the ligand to the receptor for designing efficient Epo derivatives. Furthermore, our simulations predicted that higher turnover of the receptor on the plasma membrane would result in a slightly higher, but shallower peak. On the other hand, faster internalization of ligand-bound receptor would lead to a smaller but sharper peak. The parameters values we determined in cells allow a steep rise with a rapid decline of Epo-bound receptor. As this represents the signaling-competent receptor complexes, we compared the kinetics of phosphorylated EpoR with the trajectory of the sum of plasma membrane and internalized Epo-bound EpoR of our model. Interestingly, the curves match in the first 45 min of simulation with Epo. At later times, EpoR signaling is terminated, probably by the activation of negative feedback loops. Thus, EpoR internalization is not responsible for long-term attenuation, but determines a sharp signaling peak by rapid endocytosis of ligand-bound receptors.

Thus, by applying a systems biology approach, we uncovered the parameters determining the kinetics of ligand-induced receptor complex formation and therefore signal activity of the receptor. We further determined that 80% of internalized Epo recycles to the medium. This allows EpoR activation without depleting Epo in extracellular space, being especially important for potentially low Epo levels in the hematopoietic stem cell niche (Noe et al. 1999). We additionally realized that the parameters controlling EpoR endocytosis permit fast activation as well as rapid deactivation of receptor complexes. In conclusion, by data-based mathematical modeling we determined receptor and ligand kinetics, uncovering unexpected systems properties such as the importance of both receptor turnover and recycling for plasma membrane prevalence. Furthermore, we identified the parameters determining early phase kinetics of EpoR activation. The data-based mathematical model provides an important basis for the targeted design of more potent Epo derivatives.

1.4.7 References

- Albeck, J. G., MacBeath, G., White, F. M., Sorger, P. K., Lauffenburger, D. A. and Gaudet, S. (2006). Collecting and organizing systematic sets of protein data. *Nat Rev Mol Cell Biol* **7**(11): 803-12.
- Alberghina, L. and Westerhoff, H. V. (2005). *Systems Biology: Definitions and Perspectives*. Berlin, Heidelberg, Springer.
- Barber, D. L., Corless, C. N., Xia, K., Roberts, T. M. and D'Andrea, A. D. (1997). Erythropoietin activates Raf1 by an Shc-independent pathway in CTLL-EPO-R cells. *Blood* **89**(1): 55-64.
- Becker, V., Schilling, M., Bachmann, J., Hengl, S., Maiwald, T., Timmer, J. and Klingmüller, U. (2007). Internalization controls early phase kinetics of Epo receptor activation. *Manuscript in preparation*.
- Beckman, D. L., Lin, L. L., Quinones, M. E. and Longmore, G. D. (1999). Activation of the erythropoietin receptor is not required for internalization of bound erythropoietin. *Blood* **94**(8): 2667-75.
- Bhalla, U. S., Ram, P. T. and Iyengar, R. (2002). MAP kinase phosphatase as a locus of flexibility in a mitogen-activated protein kinase signaling network. *Science* **297**(5583): 1018-23.
- Blüthgen, N. and Herzog, H. (2003). How robust are switches in intracellular signaling cascades? *J Theor Biol* **225**(3): 293-300.
- Bos, J. L. (1989). ras oncogenes in human cancer: a review. *Cancer Res* **49**(17): 4682-9.
- Brightman, F. A. and Fell, D. A. (2000). Differential feedback regulation of the MAPK cascade underlies the quantitative differences in EGF and NGF signalling in PC12 cells. *FEBS Lett* **482**(3): 169-74.
- Brummer, T., Naegele, H., Reth, M. and Misawa, Y. (2003). Identification of novel ERK-mediated feedback phosphorylation sites at the C-terminus of B-Raf. *Oncogene* **22**(55): 8823-34.
- Buday, L., Warne, P. H. and Downward, J. (1995). Downregulation of the Ras activation pathway by MAP kinase phosphorylation of Sos. *Oncogene* **11**(7): 1327-31.
- Camps, M., Chabert, C., Muda, M., Boschert, U., Gillieron, C. and Arkinstall, S. (1998). Induction of the mitogen-activated protein kinase phosphatase MKP3 by nerve growth factor in differentiating PC12. *FEBS Lett* **425**(2): 271-6.
- Chen, Z., Gibson, T. B., Robinson, F., Silvestro, L., Pearson, G., Xu, B., Wright, A., Vanderbilt, C. and Cobb, M. H. (2001). MAP kinases. *Chem Rev* **101**(8): 2449-76.

-
- Csete, M. E. and Doyle, J. C. (2002). Reverse engineering of biological complexity. *Science* **295**(5560): 1664-9.
- D'Andrea, A. D., Lodish, H. F. and Wong, G. G. (1989). Expression cloning of the murine erythropoietin receptor. *Cell* **57**(2): 277-85.
- Dard, N. and Peter, M. (2006). Scaffold proteins in MAP kinase signaling: more than simple passive activating platforms. *Bioessays* **28**(2): 146-56.
- Dumitriu, B., Patrick, M. R., Petschek, J. P., Cherukuri, S., Klingmüller, U., Fox, P. L. and Lefebvre, V. (2006). Sox6 cell-autonomously stimulates erythroid cell survival, proliferation, and terminal maturation and is thereby an important enhancer of definitive erythropoiesis during mouse development. *Blood* **108**(4): 1198-207.
- Elion, E. A. (1998). Routing MAP kinase cascades. *Science* **281**(5383): 1625-6.
- English, J. M. and Cobb, M. H. (2002). Pharmacological inhibitors of MAPK pathways. *Trends Pharmacol Sci* **23**(1): 40-5.
- Ferrell, J. E., Jr. (1996). Tripping the switch fantastic: how a protein kinase cascade can convert graded inputs into switch-like outputs. *Trends Biochem Sci* **21**(12): 460-6.
- Ferrell, J. E., Jr. and Machleder, E. M. (1998). The biochemical basis of an all-or-none cell fate switch in *Xenopus* oocytes. *Science* **280**(5365): 895-8.
- Gross, A. W. and Lodish, H. F. (2006). Cellular trafficking and degradation of erythropoietin and novel erythropoiesis stimulating protein (NESP). *J Biol Chem* **281**(4): 2024-32.
- He, T. C., Jiang, N., Zhuang, H. and Wojchowski, D. M. (1995). Erythropoietin-induced recruitment of Shc via a receptor phosphotyrosine-independent, Jak2-associated pathway. *J Biol Chem* **270**(19): 11055-61.
- Hilton, D. J. (1999). Negative regulators of cytokine signal transduction. *Cell Mol Life Sci* **55**(12): 1568-77.
- Hoffman, R., Benz, E., Shattil, S., Furie, B., Cohen, H. and Silberstein, L. (1995). Hematology, basic principles and practice. New York, Churchill Livingstone.
- Hornberg, J. J. and Westerhoff, H. V. (2006). Oncogenes are to lose control on signaling following mutation: should we aim off target? *Mol Biotechnol* **34**(2): 109-16.
- Hoshino, R., Chatani, Y., Yamori, T., Tsuruo, T., Oka, H., Yoshida, O., Shimada, Y., Ari-i, S., Wada, H., Fujimoto, J. and Kohno, M. (1999). Constitutive activation of the 41-/43-kDa mitogen-activated protein kinase signaling pathway in human tumors. *Oncogene* **18**(3): 813-22.
- Johnson, L., Greenbaum, D., Cichowski, K., Mercer, K., Murphy, E., Schmitt, E., Bronson, R. T., Umanoff, H., Edelman, W., Kucherlapati, R. and Jacks, T.

-
- (1997). K-ras is an essential gene in the mouse with partial functional overlap with N-ras. *Genes Dev* **11**(19): 2468-81.
- Karnitz, L. M., Burns, L. A., Sutor, S. L., Blenis, J. and Abraham, R. T. (1995). Interleukin-2 triggers a novel phosphatidylinositol 3-kinase-dependent MEK activation pathway. *Mol Cell Biol* **15**(6): 3049-57.
- Ketteler, R., Glaser, S., Sandra, O., Martens, U. M. and Klingmüller, U. (2002a). Enhanced transgene expression in primitive hematopoietic progenitor cells and embryonic stem cells efficiently transduced by optimized retroviral hybrid vectors. *Gene Ther* **9**(8): 477-87.
- Ketteler, R., Heinrich, A. C., Offe, J. K., Becker, V., Cohen, J., Neumann, D. and Klingmüller, U. (2002b). A functional green fluorescent protein-erythropoietin receptor despite physical separation of JAK2 binding site and tyrosine residues. *J Biol Chem* **277**(29): 26547-52.
- Ketteler, R., Moghraby, C. S., Hsiao, J. G., Sandra, O., Lodish, H. F. and Klingmüller, U. (2003). The cytokine-inducible Scr homology domain-containing protein negatively regulates signaling by promoting apoptosis in erythroid progenitor cells. *J Biol Chem* **278**(4): 2654-60.
- Kholodenko, B. N. (2000). Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. *Eur J Biochem* **267**(6): 1583-8.
- Kirschner, M. W. (2005). The meaning of systems biology. *Cell* **121**(4): 503-4.
- Kitano, H. (2001). *Foundations of Systems Biology*. Cambridge, MA, MIT Press.
- Kitano, H. (2002). Systems biology: a brief overview. *Science* **295**(5560): 1662-4.
- Klingmüller, U., Bergelson, S., Hsiao, J. G. and Lodish, H. F. (1996). Multiple tyrosine residues in the cytosolic domain of the erythropoietin receptor promote activation of STAT5. *Proc Natl Acad Sci U S A* **93**(16): 8324-8.
- Klingmüller, U., Lorenz, U., Cantley, L. C., Neel, B. G. and Lodish, H. F. (1995). Specific recruitment of SH-PTP1 to the erythropoietin receptor causes inactivation of JAK2 and termination of proliferative signals. *Cell* **80**(5): 729-38.
- Klingmüller, U., Schilling, M., Bohl, S. and Pfeifer, A. (2007). The need for standardization in systems biology. *Forward Look in Systems Biology in press*.
- Klingmüller, U., Wu, H., Hsiao, J. G., Toker, A., Duckworth, B. C., Cantley, L. C. and Lodish, H. F. (1997). Identification of a novel pathway important for proliferation and differentiation of primary erythroid progenitors. *Proc Natl Acad Sci U S A* **94**(7): 3016-21.
- Kolbus, A., Pilat, S., Husak, Z., Deiner, E. M., Stengl, G., Beug, H. and Baccarini, M. (2002). Raf-1 antagonizes erythroid differentiation by restraining caspase activation. *J Exp Med* **196**(10): 1347-53.

-
- Kollmann, M., Lovdok, L., Bartholome, K., Timmer, J. and Sourjik, V. (2005). Design principles of a bacterial signalling network. *Nature* **438**(7067): 504-7.
- Levin, I., Cohen, J., Supino-Rosin, L., Yoshimura, A., Watowich, S. S. and Neumann, D. (1998). Identification of a cytoplasmic motif in the erythropoietin receptor required for receptor internalization. *FEBS Lett* **427**(2): 164-70.
- Livnah, O., Stura, E. A., Middleton, S. A., Johnson, D. L., Jolliffe, L. K. and Wilson, I. A. (1999). Crystallographic evidence for preformed dimers of erythropoietin receptor before ligand activation. *Science* **283**(5404): 987-90.
- Marshall, C. J. (1995). Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation. *Cell* **80**(2): 179-85.
- Mason, J. M., Beattie, B. K., Liu, Q., Dumont, D. J. and Barber, D. L. (2000). The SH2 inositol 5-phosphatase Ship1 is recruited in an SH2-dependent manner to the erythropoietin receptor. *J Biol Chem* **275**(6): 4398-406.
- Matsumoto, A., Masuhara, M., Mitsui, K., Yokouchi, M., Ohtsubo, M., Misawa, H., Miyajima, A. and Yoshimura, A. (1997). CIS, a cytokine inducible SH2 protein, is a target of the JAK-STAT5 pathway and modulates STAT5 activation. *Blood* **89**(9): 3148-54.
- Murphy, L. O., MacKeigan, J. P. and Blenis, J. (2004). A network of immediate early gene products propagates subtle differences in mitogen-activated protein kinase signal amplitude and duration. *Mol Cell Biol* **24**(1): 144-53.
- Neubauer, H., Cumano, A., Muller, M., Wu, H., Huffstadt, U. and Pfeffer, K. (1998). Jak2 deficiency defines an essential developmental checkpoint in definitive hematopoiesis. *Cell* **93**(3): 397-409.
- Neumann, D., Yuk, M. H., Lodish, H. F. and Lederkremer, G. Z. (1996). Blocking intracellular degradation of the erythropoietin and asialoglycoprotein receptors by calpain inhibitors does not result in the same increase in the levels of their membrane and secreted forms. *Biochem J* **313** (Pt 2): 391-9.
- Noble, D. (2002). Modeling the heart--from genes to cells to the whole organ. *Science* **295**(5560): 1678-82.
- Noe, G., Riedel, W., Kubanek, B. and Rich, I. N. (1999). An ELISA specific for murine erythropoietin. *Br J Haematol* **104**(4): 838-40.
- Orton, R. J., Sturm, O. E., Vyshemirsky, V., Calder, M., Gilbert, D. R. and Kolch, W. (2005). Computational modelling of the receptor-tyrosine-kinase-activated MAPK pathway. *Biochem J* **392**(Pt 2): 249-61.
- Palis, J. and Yoder, M. C. (2001). Yolk-sac hematopoiesis: the first blood cells of mouse and man. *Exp Hematol* **29**(8): 927-36.
- Pouyssegur, J. and Lenormand, P. (2003). Fidelity and spatio-temporal control in MAP kinase (ERKs) signalling. *Eur J Biochem* **270**(16): 3291-9.

-
- Ravichandran, K. S., Lorenz, U., Shoelson, S. E. and Burakoff, S. J. (1995). Interaction of Shc with Grb2 regulates association of Grb2 with mSOS. *Mol Cell Biol* **15**(2): 593-600.
- Remy, I., Wilson, I. A. and Michnick, S. W. (1999). Erythropoietin receptor activation by a ligand-induced conformation change. *Science* **283**(5404): 990-3.
- Rubiolo, C., Piazzolla, D., Meissl, K., Beug, H., Huber, J. C., Kolbus, A. and Baccarini, M. (2006). A balance between Raf-1 and Fas expression sets the pace of erythroid differentiation. *Blood* **108**(1): 152-9.
- Sasagawa, S., Ozaki, Y., Fujita, K. and Kuroda, S. (2005). Prediction and validation of the distinct dynamics of transient and sustained ERK activation. *Nat Cell Biol* **7**(4): 365-73.
- Scatchard, G. (1949). The attractions of proteins for small molecules and ions. *Ann. N.Y. Acad. Sci.* **51**(4): 660-672.
- Schilling, M., Maiwald, T., Bohl, S., Kollmann, M., Kreutz, C., Timmer, J. and Klingmüller, U. (2005a). Computational processing and error reduction strategies for standardized quantitative data in biological networks. *Febs J* **272**(24): 6400-11.
- Schilling, M., Maiwald, T., Bohl, S., Kollmann, M., Kreutz, C., Timmer, J. and Klingmüller, U. (2005b). Quantitative data generation for systems biology: the impact of randomisation, calibrators and normalisers. *Syst Biol (Stevenage)* **152**(4): 193-200.
- Schilling, M., Maiwald, T., Hengl, S., Timmer, J. and Klingmüller, U. (2007). A data-based MAP-kinase model reveals systems properties and predicts cellular decisions. *Manuscript in preparation*.
- Schlessinger, J. (2000). Cell signaling by receptor tyrosine kinases. *Cell* **103**(2): 211-25.
- Schoeberl, B., Eichler-Jonsson, C., Gilles, E. D. and Müller, G. (2002). Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nat Biotechnol* **20**(4): 370-5.
- Socolovsky, M., Fallon, A. E., Wang, S., Brugnara, C. and Lodish, H. F. (1999). Fetal anemia and apoptosis of red cell progenitors in Stat5a^{-/-}5b^{-/-} mice: a direct role for Stat5 in Bcl-X(L) induction. *Cell* **98**(2): 181-91.
- Solit, D. B., Garraway, L. A., Pratilas, C. A., Sawai, A., Getz, G., Basso, A., Ye, Q., Lobo, J. M., She, Y., Osman, I., Golub, T. R., Sebolt-Leopold, J., Sellers, W. R. and Rosen, N. (2006). BRAF mutation predicts sensitivity to MEK inhibition. *Nature* **439**(7074): 358-62.
- Swameye, I., Müller, T. G., Timmer, J., Sandra, O. and Klingmüller, U. (2003). Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by databased modeling. *Proc Natl Acad Sci U S A* **100**(3): 1028-33.

-
- Tauchi, T., Damen, J. E., Toyama, K., Feng, G. S., Broxmeyer, H. E. and Krystal, G. (1996). Tyrosine 425 within the activated erythropoietin receptor binds Syp, reduces the erythropoietin required for Syp tyrosine phosphorylation, and promotes mitogenesis. *Blood* **87**(11): 4495-501.
- Troyanskaya, O. G., Dolinski, K., Owen, A. B., Altman, R. B. and Botstein, D. (2003). A Bayesian framework for combining heterogeneous data sources for gene function prediction (in *Saccharomyces cerevisiae*). *Proc Natl Acad Sci U S A* **100**(14): 8348-53.
- Tsuji, M., Inanami, O. and Kuwabara, M. (2001). Induction of neurite outgrowth in PC12 cells by alpha -phenyl-N-tert-butylintron through activation of protein kinase C and the Ras-extracellular signal-regulated kinase pathway. *J Biol Chem* **276**(35): 32779-85.
- von Bertalanffy, L. (1968). General system theory. New York, G. Braziller.
- Waterman, H. and Yarden, Y. (2001). Molecular mechanisms underlying endocytosis and sorting of ErbB receptor tyrosine kinases. *FEBS Lett* **490**(3): 142-52.
- Weissman, I. L., Anderson, D. J. and Gage, F. (2001). Stem and progenitor cells: origins, phenotypes, lineage commitments, and transdifferentiations. *Annu Rev Cell Dev Biol* **17**: 387-403.
- Wiener, N. (1948). Cybernetics or Control and Communication in the Animal and the Machine. Cambridge, MA, MIT Press.
- Wolkenhauer, O., Ullah, M., Wellstead, P. and Cho, K. H. (2005). The dynamic systems approach to control and regulation of intracellular networks. *FEBS Lett* **579**(8): 1846-53.
- Wu, H., Liu, X., Jaenisch, R. and Lodish, H. F. (1995). Generation of committed erythroid BFU-E and CFU-E progenitors does not require erythropoietin or the erythropoietin receptor. *Cell* **83**(1): 59-67.
- Yamada, S., Taketomi, T. and Yoshimura, A. (2004). Model analysis of difference between EGF pathway and FGF pathway. *Biochem Biophys Res Commun* **314**(4): 1113-20.
- Yeung, K., Seitz, T., Li, S., Janosch, P., McFerran, B., Kaiser, C., Fee, F., Katsanakis, K. D., Rose, D. W., Mischak, H., Sedivy, J. M. and Kolch, W. (1999). Suppression of Raf-1 kinase activity and MAP kinase signalling by RKIP. *Nature* **401**(6749): 173-7.
- Zhang, J. and Lodish, H. F. (2004). Constitutive activation of the MEK/ERK pathway mediates all effects of oncogenic H-ras expression in primary erythroid progenitors. *Blood* **104**(6): 1679-87.
- Zimmermann, S., Rommel, C., Ziogas, A., Lovric, J., Moelling, K. and Radziwill, G. (1997). MEK1 mediates a positive feedback on Raf-1 activity independently of Ras and Src. *Oncogene* **15**(13): 1503-11.

2. ORIGINAL PUBLICATIONS

2.1 Strategies for standardizing quantitative data

Schilling M, Maiwald T, Bohl S, Kollmann M, Kreutz C, Timmer J, Klingmüller U.

Computational processing and error reduction strategies for standardized quantitative data in biological networks.

FEBS J. 2005 Dec;272(24):6400-11.

This publication contains 6 figures. The experiments depicted in Figures 1, 2, 3, 4, and 5 were initiated, planned and performed by Marcel Schilling. The manuscript was largely written by Marcel Schilling (figures and text).

Computational processing and error reduction strategies for standardized quantitative data in biological networks

Marcel Schilling^{1,*}, Thomas Maiwald^{2,*}, Sebastian Bohl¹, Markus Kollmann², Clemens Kreutz², Jens Timmer² and Ursula Klingmüller¹

¹ German Cancer Research Center, Heidelberg, Germany

² Freiburg Center for Data Analysis and Modeling, University of Freiburg, Germany

Keywords

data processing; error reduction; normalization; quantitative immunoblotting; signaling pathways

Correspondence

U. Klingmüller, German Cancer Research Center, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany
Fax: +49 6221 424488
Tel: +49 6221 424481
E-mail: u.klingmueller@dkfz.de

*Authors who contributed equally to the work presented in this article.

(Received 8 September 2005, revised 25 October 2005, accepted 27 October 2005)

doi:10.1111/j.1742-4658.2005.05037.x

High-quality quantitative data generated under standardized conditions is critical for understanding dynamic cellular processes. We report strategies for error reduction, and algorithms for automated data processing and for establishing the widely used techniques of immunoprecipitation and immunoblotting as highly precise methods for the quantification of protein levels and modifications. To determine the stoichiometry of cellular components and to ensure comparability of experiments, relative signals are converted to absolute values. A major source for errors in blotting techniques are inhomogeneities of the gel and the transfer procedure leading to correlated errors. These correlations are prevented by randomized gel loading, which significantly reduces standard deviations. Further error reduction is achieved by using housekeeping proteins as normalizers or by adding purified proteins in immunoprecipitations as calibrators in combination with criteria-based normalization. Additionally, we developed a computational tool for automated normalization, validation and integration of data derived from multiple immunoblots. In this way, large sets of quantitative data for dynamic pathway modeling can be generated, enabling the identification of systems properties and the prediction of targets for efficient intervention.

Systems biology holds great promise for the targeted development of therapies and more cost-effective drug development. By combining experimental data with mathematical modeling of the dynamic behavior of complex biological networks [1,2], systems biology aims to identify systems properties and to predict perturbation-sensitive targets. However, the major limitation at present is the lack of reliable quantitative data. To determine, test and validate the quantitative accuracy of models, and to capture the characteristic dynamic behavior of systems, techniques that quantitatively and selectively measure biochemical reactions within the cell must be developed [3]. Additionally, a

comprehensive set of quantitative and time-resolved data is required to conduct a systems-level analysis [4]. Recent reports show that by analyzing quantitative data generated using fluorescence microscopy [5], electrophoretic mobility shift assays [6] or immunoblotting [7,8], new biological insights can be obtained. However, before this approach can be used for biomedical applications, standardized procedures for data acquisition, reliable normalization methods and generally applicable algorithms for data processing have to be developed.

Cellular responses are regulated by complex signaling networks, and subtle changes in protein concentration

Abbreviations

CCD, charge-coupled device; ECL, enhanced chemiluminescence; Epo, erythropoietin; EpoR, erythropoietin receptor; GST, glutathione S-transferase; HA, hemagglutinin-tagged; HRP, horseradish peroxidase; Hsc70, cellular heat shock cognate protein 70; IL-6, interleukin-6; IP, immunoprecipitation; MAP kinase, mitogen-activated protein kinase; PDI, protein disulfide isomerase; PVDF, poly(vinylidene difluoride); STAT, signal transducer and activator of transcription.

or protein modification can trigger the onset of diseases. For the analysis of proteins in complex mixtures, one of the most widely used techniques is immunoblotting, which is based on electrophoresis and transfer to a membrane. The presence of specific proteins on the membrane is detected by antibodies in combination with the utilization of chemiluminescent substrates and exposure to X-ray films. However, because the linear range of X-ray films is very limited, quantification by charged-coupled device (CCD) camera detection is preferable [9]. For rare proteins (such as certain signaling components), prepurification by immunoprecipitation (IP) is required prior to immunoblotting, potentially increasing the overall error owing to additional steps involved in the procedure. To date, only relative values that are difficult to compare between independent experiments have been generated by immunoblotting. Thus, reliable algorithms for error reduction and data processing are required to employ immunoblotting for the generation of high-quality quantitative data.

Another problem in normalization of data from different sources arises from the fact that signaling pathways have been primarily studied in the context of propagatable cell lines. However, as such cell lines have lost restrictive growth control mechanisms, it is of great importance to analyze the behavior of signaling pathways in primary cells. As material that can be isolated from animals or patients is very limited, it is of pressing importance that existing data be combined and compared. Mammalian cells grow either in suspension or attached to a support. Suspension cells are primarily cells of hematopoietic origin and are particularly suited for biochemical studies on cell populations with high temporal resolution because they permit bulk stimulation and rapid sampling. For biochemical studies in adherent cells, separate stimulations are required for each time-point, potentially resulting in a higher sample-to-sample variation. Even more difficult is the analysis of proteins in patient samples. To eliminate errors introduced by the measurement process and to ensure comparability of results, we have developed robust normalization procedures for biochemical data.

We use the erythropoietin receptor (EpoR)-induced activation of ERK1 in the hematopoietic suspension cell line, BaF3-hemagglutinin-tagged (HA)-EpoR, and the interleukin-6 (IL-6)-induced activation of the signal transducer and activator of transcription (STAT)3 in adherent primary hepatocytes, as model systems to establish a robust procedure for error reduction and to develop reliable algorithms for data processing, facilitating the generation of high-quality data by quantitative immunoblotting.

Results

Standardized generation of absolute values

The reliable generation of large data sets depends on the strategies used to achieve comparable results among individual experiments. To achieve this, we convert the relative signals, which are usually generated by immunoblotting, to absolute numbers, such as molecules per cell. As an example, the abundance of the mitogen-activated protein (MAP)-kinase family members, ERK1 and ERK2, in cytoplasmic lysates from BaF3-HA-EpoR cells, was determined by analyzing, in parallel, a serial dilution of purified recombinant ERK2 protein (Fig. 1A, upper panel). The CCD camera-based quantification of recombinant ERK2 was plotted against the number of molecules loaded on the gel. As demonstrated by a linear regression passing through the origin (Fig. 1A, lower panel) and extensive additional studies (see the Supplementary material) the detection was linearly proportional to protein concentration over at least two orders of magnitude. By using a linear regression model (detailed in the Supplementary material) relative signals of endogenous ERK1 and ERK2 were converted to molecules per cell, indicating that in the cytoplasm of an BaF3-HA-EpoR cell, 107 000 ERK1 molecules and 318 000 ERK2 molecules are present. This determination requires the recombinant and the endogenous proteins to be analyzed on the same immunoblot and to share the same antibody epitope. As the CCD camera-based detection is proportional to the number of epitopes, it can even be applied to proteins of different molecular mass, such as isoforms or partial fusion proteins, thus permitting the concomitant determination of multiple signaling components. In addition to ensuring comparability of independent experiments, absolute values can be used to determine the stoichiometry of cellular components, critical for obtaining insights into the quantitative behavior of biological networks.

Error determination of the measurement process

To estimate the inherent noise of data generated by the immunoblotting technique, error determinations were performed. A serial dilution of purified recombinant ERK2 protein was analyzed eight times by immunoblotting using an anti-ERK immunoglobulin (Fig. 1B, upper panel) and quantified by CCD camera-based detection. The estimated error was calculated as the standard deviation of the CCD camera-based measurements. Plotting signal strength vs. estimated error revealed that the expected error behavior of a

conventional CCD camera-based photon counting process cannot be recovered. The systematic error inherent in this technique can phenomenologically be described by a sublinear function. Within our measurement range, $\approx 20\%$ error for each data point is estimated, whereas for weaker signals this percentage is increased (Fig. 1B, lower panel). This noise consists of two different contributions: pipetting errors, which are constant within a lane but uncorrelated from lane to lane; and blotting errors, which are highly correlated from

lane to lane. Pipetting errors arise from differences in cell number, gel loading and antibody detection, while blotting errors are caused by inhomogeneities of the gel or the blot.

Eliminating correlated errors by randomized sample loading

To determine steps predominantly contributing to the error obtained by quantitative immunoblotting analysis, we monitored a time-course of erythropoietin (Epo)-induced activation of ERK1 in BaF3-HA-EpoR cells. Identical samples of cytoplasmic lysates were loaded, in a randomized manner, onto two gels, transferred to membranes (blot 1 and blot 2) and analyzed by three repetitive cycles of ERK immunoglobulin reprobings and application of the chemiluminescent substrate (Fig. 2A). Quantification of the signals (Fig. 2B, upper panel) showed that the data obtained by the two blots differed significantly. To reduce the effects of uncorrelated errors, we employed a cubic spline, the smoothness of which is determined by generalized cross-validation. It has been shown previously that time-course behavior can be estimated from noisy data by smoothing splines [10–12]. We emphasize that a sufficiently dense grid of time-points is necessary to keep the bias of this method small. Smoothing of the data is performed to average over the errors contributed by pipetting, electrophoresis and transfer, and other sources of noise.

Surprisingly, uncorrelated errors resulting from antibody detection and reprobings had little effect on the results, as the splines smoothing the data obtained by

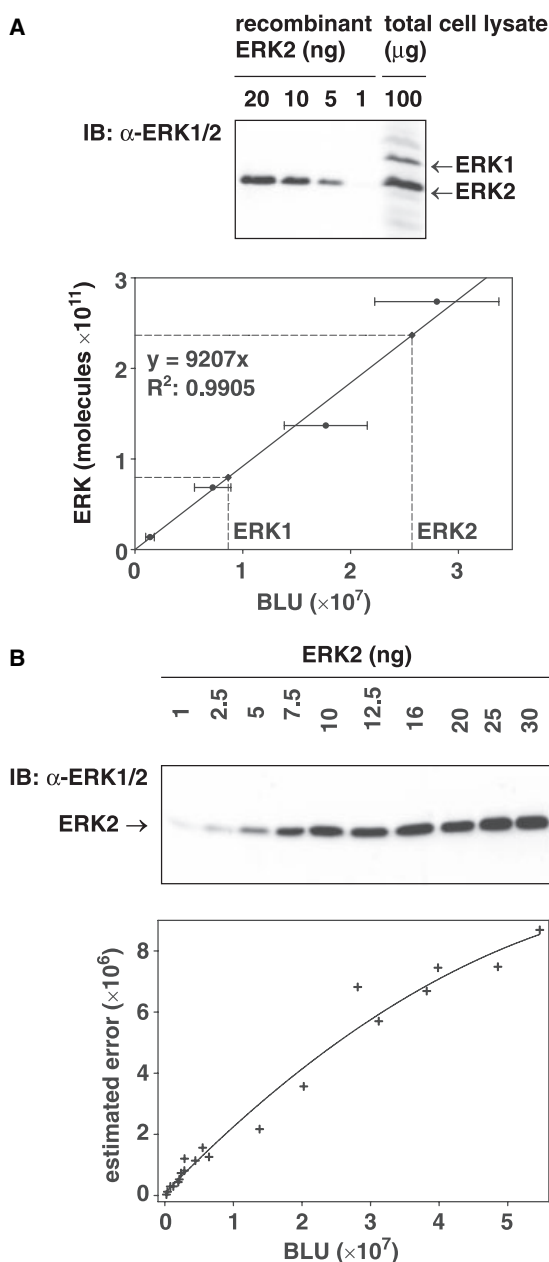


Fig. 1. Conversion of relative values to absolute protein concentrations and error estimation of quantitative immunoblotting. (A) A dilution series of recombinant ERK2 protein, as well as 100 μg of total cellular lysate prepared from BaF3-HA-EpoR cells, were analyzed by quantitative immunoblotting with anti-ERK immunoglobulin. The biomedical light unit (BLU) values of the dilution series were plotted against the number of molecules loaded onto the gel [amount (g)/ MW_{ERK2} ($\text{g}\cdot\text{mol}^{-1}$) $\times N_A$ ($\text{molecules}\cdot\text{mol}^{-1}$)] and a linear regression through the origin was applied. The slope was used for converting the signals of the total cellular lysate to molecules per cell. Error bars represent estimated errors of the total ERK2 dilution series, as determined in (B). (B) A dilution series of purified ERK2 was separated eight times by SDS/PAGE (10% acrylamide) and transferred to a membrane that was probed with anti-ERK immunoglobulin and subsequently developed with enhanced chemiluminescence (ECL) or ECL advance substrate. The estimated error of the quantified signals was calculated as the standard deviation of the data. To determine the noise inherent in this technique, the signal strength was plotted vs. estimated error and was described by a sublinear function showing a 20% error for each data point within our measurement range.

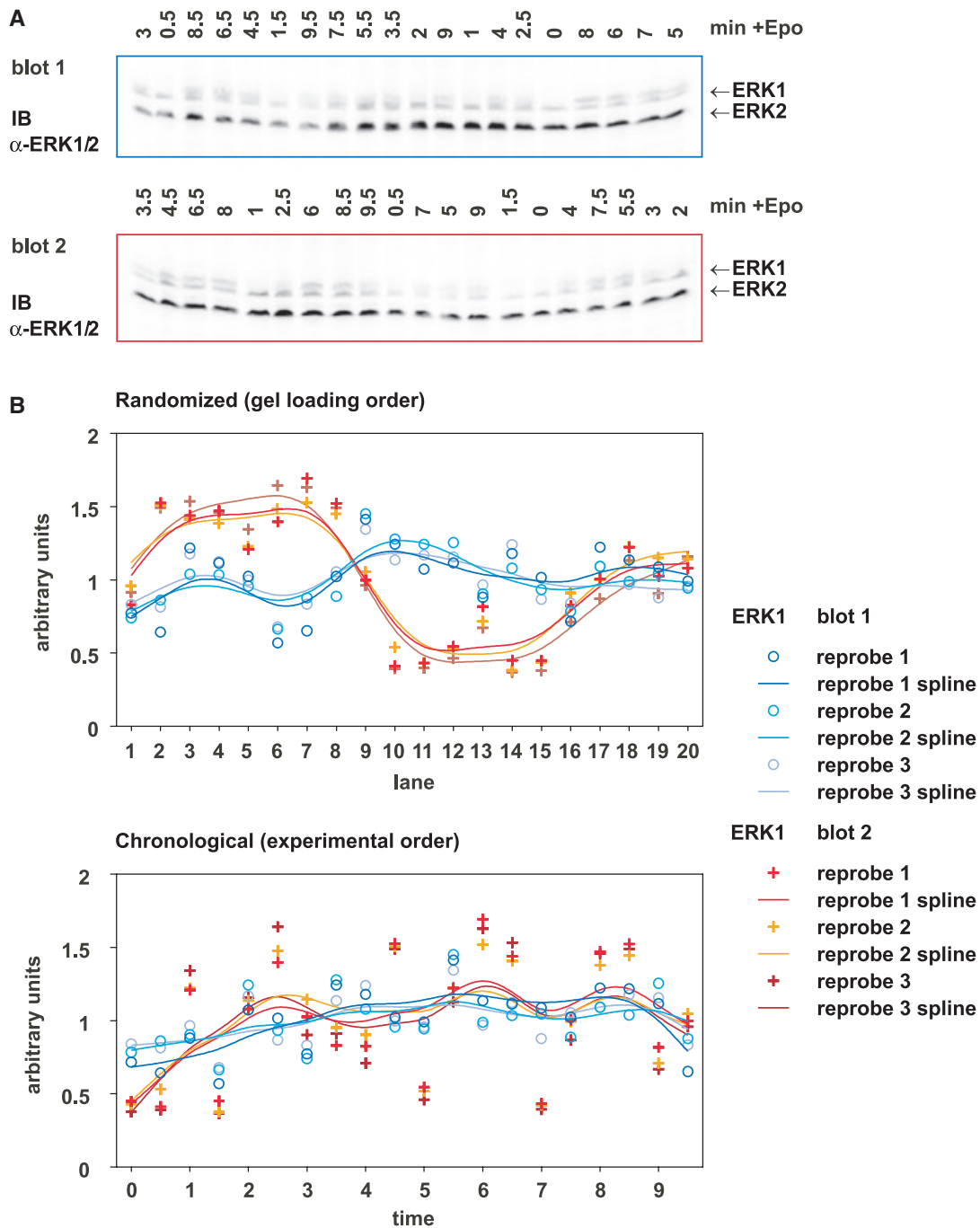


Fig. 2. Randomized sample loading ensures uncorrelated errors. (A) BaF3-HA-EpoR cells were starved and stimulated with 50 units·mL⁻¹ erythropoietin (Epo) for 9.5 min, with samples of 1×10^7 cells taken every 30 s. Cells were lysed, and 75 μ g of the total cellular lysate at each time-point was separated by two 17.5% SDS polyacrylamide gels using two distinct randomized sample loading orders. Each immunoblot was analyzed by three repetitive cycles of detection with anti-ERK immunoglobulin and subsequent removal of the antibodies by treatment with β -mercaptoethanol and SDS. The obtained signals for ERK1 were quantified by Lumilager analysis. (B) The data show strongly correlated errors when arranged in gel loading order, which are specific for a particular blot but are not affected by reprobing procedures. By arranging the data in chronological order, these correlations are eliminated and the data can be smoothed by spline approximations, as indicated by solid lines. Randomization reduced the standard deviation of the smoothing splines by a factor of 14.

successive reprobing of the same blot were nearly identical. However, the analysis revealed that the data obtained for neighboring lanes was strongly correlated. The apparently different results obtained for identical samples showed that the blotting error leads to aberrant dynamic behavior. Detailed analysis of large data sets revealed a strong correlation between neighboring lanes in immunoblotting analysis, resulting in substantial systematic errors. To separate this spatial correlation from true temporal dynamics in time-course data, we developed standard operating procedures for randomized sample loading, separating consecutive time-points by a minimum number of lanes. This loading scheme was varied from experiment to experiment to minimize gel border effects. The procedure thereby ensures uncorrelated errors (Fig. 2B, lower panel) and thus facilitates the detection of true dynamic behavior. In this case, randomization reduced the standard deviation of the smoothing splines from 18.6% to 1.4% and thus significantly improves the data quality.

Data correction using normalizers

To reduce the effect of the blotting error and improve the data quality, we used endogenous proteins as normalizers. The time-course of Epo-induced phosphorylation of ERK1 was detected by immunoblotting using a phosphospecific anti-pERK immunoglobulin (Fig. 3A). Subsequently, the antibody was removed and the blot was reprobed, first with an anti-ERK immunoglobulin to determine the total amount of ERK1 in the cytoplasmic lysates and, second, with a mixture of antibodies against endogenous proteins. These proteins, which we termed normalizers, are highly expressed, their levels are not changed during the course of the experiment and antibodies are available that permit efficient detection. As shown in Fig. 3A, the blotting error is strongly influenced by the position of a protein within a blot, as evidenced by the analysis of β Actin (42 kDa), protein disulfide isomerase (PDI; 58 kDa), and heat shock cognate protein 70 (Hsc70; 73 kDa) covering the entire separation range of the polyacrylamide gel. Therefore, the signal of a normalizer of similar molecular mass to the protein of interest has to be used to distinguish blotting error from the true protein concentration. The levels of pERK1 and ERK1 were normalized with a smoothing spline applied to the β Actin signal. As shown in Fig. 3B, this procedure enabled us to correct for blotting errors in our signals. As expected, the normalized data shows a constant concentration of ERK1 over the entire observation time. By employing purified ERK2 as standard, relative signals for ERK1 were

converted to molecules per cell and the proportion of phosphorylated ERK1 was determined by analyzing the fraction of protein that was detected by the anti-ERK immunoglobulin at a higher position in the blot. This ensures the comparability of normalized data derived from independent experiments.

Recombinant proteins as calibrators for IP

For certain proteins, immunoblotting is not capable of generating quantitative data. This problem can be caused by antibodies with weak affinity to the protein, cross-reaction with other proteins resulting in a high background, or by the use of generic phosphotyrosine antibodies. In such cases, the protein of interest has to be prepurified by IP, prior to electrophoresis.

As normalizers are not captured by the antibodies used for the IP, we have established a method to correct for blotting errors as well as inaccuracies in the multistep IP procedure, and to normalize the results obtained. We generated proteins (which we termed calibrators) that share the same epitope as the protein of interest, but differ in molecular mass. Adding a defined amount of calibrator to the lysate prior to IP permits normalization of the results obtained by CCD camera-based detection. We fused the protein domain containing the epitope of the antibody used for IP to a affinity tag for purification (Fig. 4A). Using only part of the protein, calibrators of large proteins or transmembrane proteins could easily be expressed in *Escherichia coli* and purified using affinity beads. We determined the concentration of the calibrators by analyzing a BSA dilution series and the calibrator in a Coomassie Blue-stained gel and quantifying the signals. To define the optimal amount of calibrator that should be added to the IP while still avoiding saturation of the antibodies, increasing concentrations of the calibrator, glutathione *S*-transferase-tagged (GST)-EpoR, were added to lysates of BaF3-HA-EpoR cells prior to IP (Fig. 4B). Plotting the concentration of calibrator added to the lysates vs. signals for HA-EpoR and GST-EpoR showed that the calibrator signal increased linearly in a range between 2.5 and 100 ng. This suggested that the use of a calibrator not only permits quantitative data generation, but also conversion of relative values to absolute protein concentrations. The addition of the calibrator had no effect on the signal for the HA-EpoR up to concentrations of 500 ng of GST-EpoR, indicating that the antibody was in large excess compared with HA-EpoR. Using this data, we calculated that 40 ng of GST-EpoR should be added to lysates to obtain comparable signals for HA-EpoR and the calibrator (Fig. 4C).

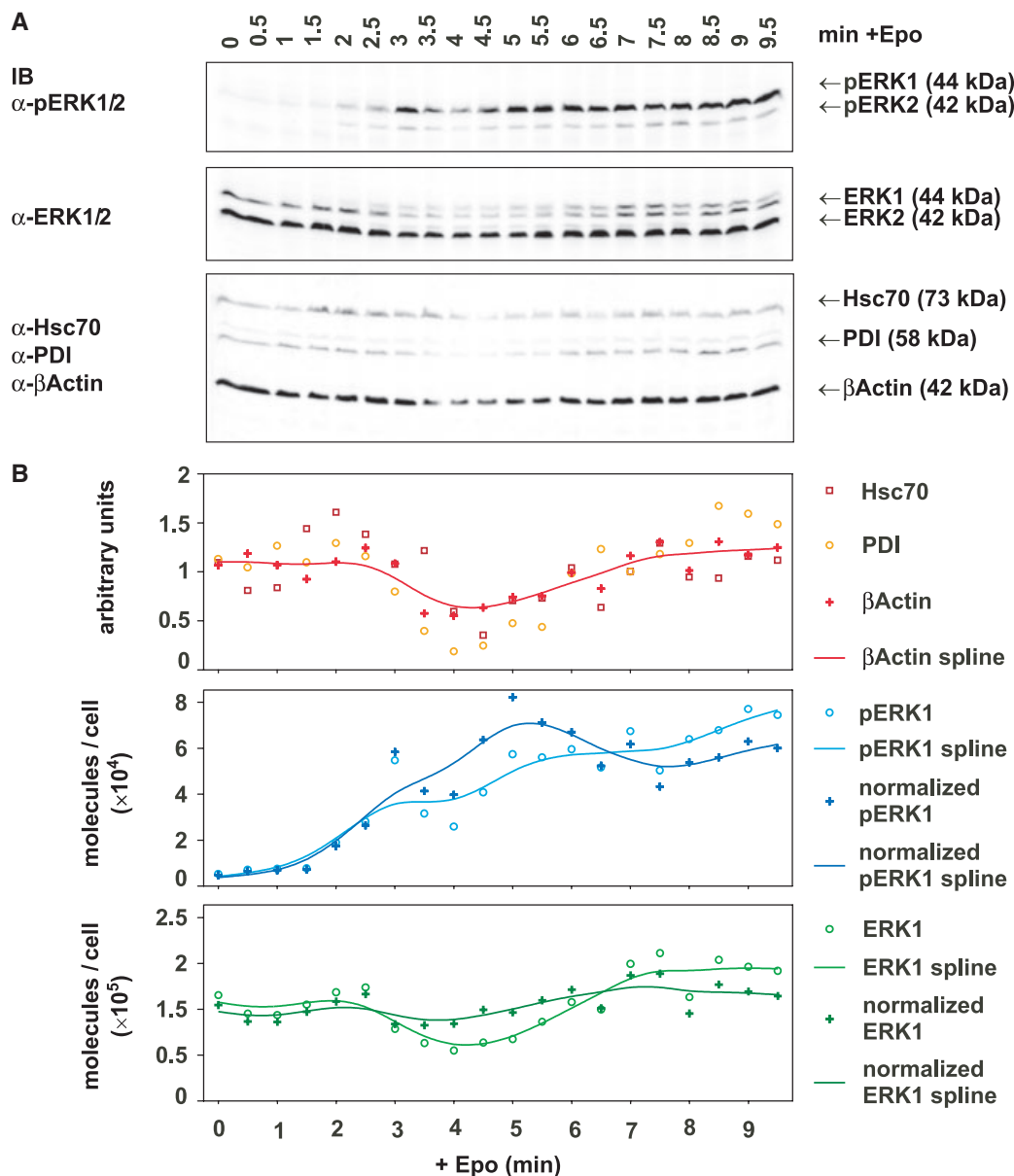


Fig. 3. Correction of phosphorylated and total ERK1 signals using normalizers. (A) BaF3-HA-EpoR cells were starved and stimulated with 50 units·mL⁻¹ erythropoietin (Epo) for 9.5 min, with samples of 1×10^7 cells taken every 30 s. Cells were lysed and 75 μg of total cellular lysate at each time-point was separated by electrophoresis on a 17.5% SDS polyacrylamide gel. The immunoblot was analyzed with anti-pERK immunoglobulin, and then reprobed, first with anti-ERK immunoglobulin and second with an anti-heat shock cognate protein 70 (Hsc70)/anti-protein disulfide isomerase (PDI)/anti-(βActin) immunoglobulin mixture. All signals were quantified by Lumimager analysis. (B) The βActin signal was spline-smoothed and used to normalize pERK1 and ERK1 signals, having similar molecular masses. pERK1 and ERK1 signals were converted to number of molecules per cell using the protein standard depicted in Fig. 1. Smoothing spline curves through original and normalized data are shown as solid lines.

Using calibrators for error reduction

The impact of calibrators on data quality is exemplified by an EpoR time-course experiment with

randomized gel loading. We stimulated BaF3-HA-EpoR cells with Epo for up to 10 min and added 40 ng of GST-EpoR to each cytoplasmic lysate to control for errors during the IP procedure (Fig. 5A). In

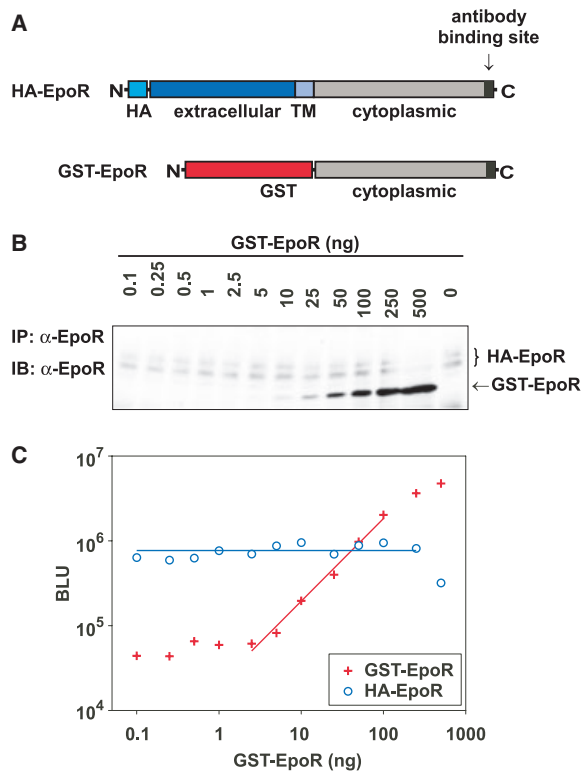


Fig. 4. Titration of the glutathione *S*-transferase tagged-erythropoietin receptor (GST-EpoR) calibrator in immunoprecipitation. (A) The domain structure of hemagglutinin-tagged HA-EpoR is schematically depicted and the binding epitope for the anti-EpoR immunoglobulin is indicated. The calibrator, GST-EpoR, consists of the protein domain containing the antibody-binding site fused to an affinity tag for purification. (B) BaF3-HA-EpoR cells were starved, stimulated with 50 units·mL⁻¹ erythropoietin (Epo) for 5 min and lysed. Increasing amounts of recombinant GST-EpoR were added to the lysates and both the GST-EpoR calibrator and the HA-EpoR were immunoprecipitated with anti-EpoR immunoglobulin. The samples were separated on a 10% SDS polyacrylamide gel. The immunoblot was analyzed with anti-EpoR immunoglobulin and quantified by Lumilimager analysis. (C) Concentrations of the calibrator were plotted vs. the signals obtained for the HA-EpoR and the GST-EpoR calibrator. A red line depicts the linear relationship between the calibrator concentration added to the lysate and the detected signal within a range of 2.5–100 ng of calibrator addition. The blue line depicting the average signal of the HA-EpoR intersects at 40 ng of GST-EpoR, indicating comparable signals for the calibrator and the HA-EpoR.

addition, the calibrator was used to correct for blotting errors, thereby significantly improving data quality. However, correction steps can be detrimental to the data if a calibrator yields noisy signals or is exposed to different gel/transfer inhomogeneities as the protein of interest owing to a large difference in molecular mass. We therefore developed criteria for automated data correction in IP experiments, as described in the

Supplementary material. One necessary condition for these criteria is randomized sample loading. As shown in the Supplementary material, by combining randomized sample loading with calibrators, the standard deviation of immunoblotting data can be improved by more than twofold. The corrected data (Fig. 5B) show the expected behavior of a continuous increase in phosphorylated HA-EpoR and a constant level of total HA-EpoR for 10 min after stimulation with Epo.

Computational data processing using GELINSPECTOR

For automated data processing and to permit data merging of samples analyzed on separate blots, we developed the computer algorithm GELINSPECTOR. This algorithm calculates smoothing splines for the normalizers or calibrators and normalizes blotting data using these splines. Furthermore, the program verifies the normalization, integrates multiple data sets and visualizes the results. To validate our approach, we investigated the effect of our algorithm on time-course data generated from primary hepatocytes. We combined sample randomization with criteria-mediated error reduction using Calnexin and Hsc70 as normalizers. By loading time-points alternating on two gels, the number of data points that could be analyzed together was increased beyond the capacity of a single gel (Fig. 6A). Applying GELINSPECTOR enabled us to normalize the signals and significantly decrease the standard deviation from a smoothing spline, resulting in time-course data with a high temporal resolution (Fig. 6B). The high reproducibility of the time-course dynamics for phosphorylated and total cytoplasmic STAT3 obtained by immunoblotting of cytoplasmic lysates, as well as immunoprecipitates (data not shown), demonstrated that our automated computational data processing is robust and reliably applicable for both methods. These tools facilitate the standardized and automated generation of quantitative data and permit the cost-effective assembly of large, high-quality data sets.

Discussion

Quantitative data generation is becoming increasingly important for obtaining insight into the dynamic behavior of complex biological networks, to elucidate systems properties and to predict targets for biomedical applications. We show that by randomized sample loading and computational data processing, including criteria-based normalization, high-quality quantitative data can reliably be generated by immunoblotting, a

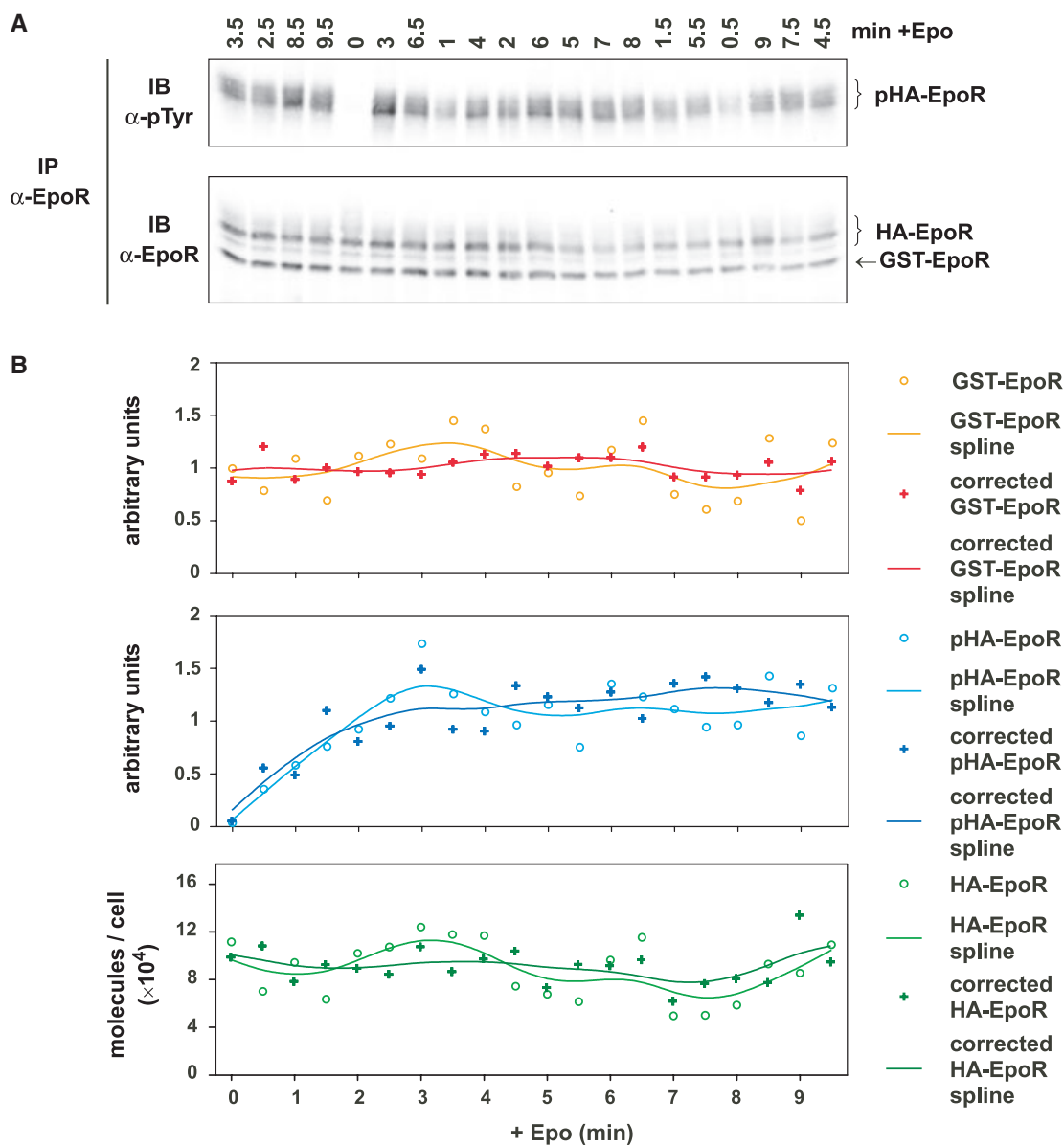


Fig. 5. Correction of hemagglutinin-tagged-erythropoietin receptor (HA-EpoR) signals with the glutathione *S*-transferase (GST)-EpoR calibrator. (A) BaF3-HA-EpoR cells were starved and stimulated with 50 units·mL⁻¹ erythropoietin (Epo) for the indicated time. A total of 1×10^7 cells was lysed and 40 ng of GST-EpoR was added to each lysate. Immunoprecipitation was performed using anti-EpoR immunoglobulin, followed by separation on a 10% SDS polyacrylamide gel with randomized sample loading. The immunoblot was analyzed with anti-pTyr and anti-EpoR immunoglobulin and quantified by Lumilmager analysis. (B) Time after Epo stimulation was plotted against the signals of HA-EpoR and the calibrator GST-EpoR. A spline smoothing the calibrator signal was used to correct pEpoR signals, whereas the EpoR signal was corrected and converted to molecules per cell. Splines are depicted as solid lines.

widely applied technique. By systematically determining steps contributing to the variability of the experimental data, we identified gel and transfer inhomogeneities as the major source for correlated errors. These correlations could be eliminated by

randomized sample loading, and error reduction was achieved by the use of normalizers or calibrators in combination with computational data processing. By converting relative signals to absolute values, comparable results can be obtained from independent

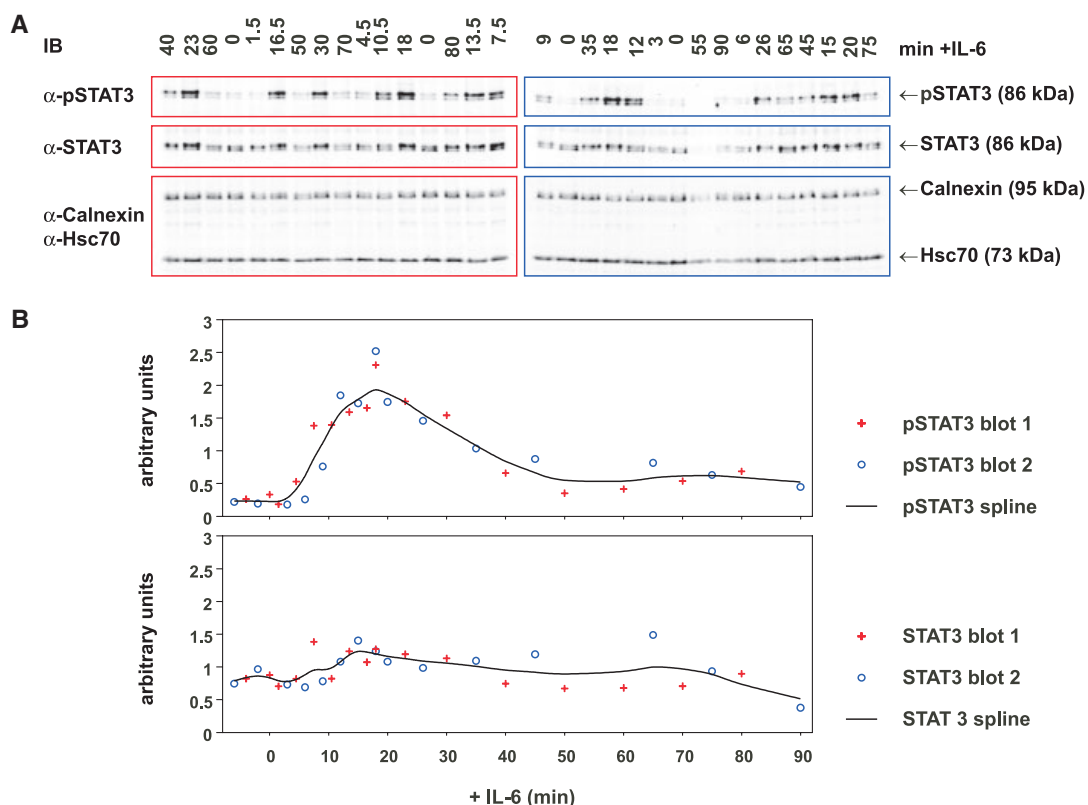


Fig. 6. Quantitative data generation of primary hepatocytes using the computer algorithm GELINSPECTOR. (A) Primary mouse hepatocytes were prepared from mouse livers. A total of 2×10^6 cells for each time-point was cultured on collagen-coated dishes and starved. Interleukin-6 (IL-6) was added ($40 \text{ ng}\cdot\text{mL}^{-1}$) and the cells were lysed at the indicated time-points. Cytoplasmic lysates were separated by two 10% SDS polyacrylamide gels. Sample loading was randomized with every second time-point on the second gel. Quantitative immunoblotting was performed with anti-phosphorylated signal transducer and activator of transcription 3 (pSTAT3), anti-signal transducer and activator of transcription (STAT3), and an anti-Calnexin/anti heat shock cognate protein 70 (Hsc70) mixture. (B) Immunoblotting data were automatically processed by GELINSPECTOR using Calnexin/Hsc70 signals as normalizers, and the data points were spline-smoothed, as indicated by solid lines.

experiments and used for the assembly of large sets of quantitative data.

Randomized sample analysis is a general strategy to prevent correlated errors, for example in double-blind comparative clinical studies [13] and in the design of DNA microarray experiments [14]. Here, we use this approach to separate spatial blotting effects from real changes in protein levels (i.e. their true dynamic behavior). By simulations of typical time-course experiments, we demonstrated that randomization reduces the standard deviation of immunoblotting data by more than twofold (see the Supplementary material for simulations). Sample randomization is thus a simple procedure that significantly improves data quality without increasing experimental efforts.

To reduce errors inherent in blotting techniques, such as inhomogeneities in the gel as well as transfer, normalizers are used that are present at a similar

position in the blot as the molecule of interest and which are detectable with a strong constant signal. We identified several housekeeping proteins of different molecular mass that can be reliably used as normalizers. The normalization procedure cannot be applied if a normalizer differs too much in molecular mass from the protein of interest because it is exposed to different gel/transfer inhomogeneities and therefore does not permit an adequate estimation to be made of the blotting error. To ensure accuracy of data normalization, we applied spline approximation and developed data processing criteria. The resulting computer algorithm, GELINSPECTOR, compares the standard deviation of both the normalized and the unprocessed data to a first estimate of the values. Only if the normalized values are closer to the estimate, is normalization by computational data processing accurate and results in significantly improved data quality.

In the case of grouped data, such as mutant to wild-type comparisons used in diagnostic approaches, the first estimate is the mean value of the same sample loaded in replicates. In other cases, where a continuous dependency exists (such as in time-course or in dose-response experiments), the first estimate is a regression curve if the functional relationship is known or a smoothing spline if unknown. These functions are implemented in GELINSPECTOR.

Our method is not only applicable to proteins concentrated by IP, but also, as we show in Fig. 6A, for the detection of proteins in total cellular lysates of primary hepatocytes. Furthermore, our data processing procedures permit the quantification of low abundance proteins, or modifications, as demonstrated for the Epo-induced phosphorylation of ERK1/2 (Fig. 3A). The EpoR, a member of the hematopoietic cytokine receptor family, activates the MAP kinase signaling cascade to a much lesser extent than receptor tyrosine kinases, such as the epidermal growth factor receptor or the platelet derived growth factor receptor.

Similarly to normalizers, calibrators added in IP experiments permit criteria-based data normalization. Importantly, calibrators, in addition, facilitate the conversion of relative signals to absolute values, such as molecules per cell. For the analysis of cellular lysates, this can be achieved by co-loading known amounts of recombinant proteins onto the gel, which are detected by the same antibody as the protein of interest. Using microscopic techniques, the volume of a cell can be estimated, allowing conversion of molecules per cell to protein concentrations. The generation of absolute values provides additional information regarding absolute protein concentrations that cannot only be used to compare signals derived from independent immunoblot experiments, but also to identify the amount of a given protein in a single cell and to determine the stoichiometry of cellular components [15].

The proposed methods can be applied to other blotting techniques, such as northern and Southern blotting analysis, as inhomogeneities in gel and transfer are likely to cause correlated errors in all blotting data. Similarly, correlations can be eliminated by randomization and the errors can be reduced by criteria-based normalization.

Recently developed strategies for quantitative determination of protein levels and modifications include mass spectrometry techniques based on isotope-coded affinity tags [16] and isotope-coded protein labels [17]. By labeling different samples with distinct isotopes, relative changes can be quantified using mass spectrometry. It is even possible to determine absolute values by the addition of synthesized peptides of known quanti-

ties as standards. However, these methods are still very expensive, technically demanding and have the disadvantages of requiring large amounts of cellular material.

By developing quantitative immunoblotting as a robust and reliable technique for quantitative data acquisition under standardized conditions, we establish an easy to handle and cost-effective alternative that permits the assembly of large data sets with high temporal resolution. This provides an important tool for diagnostic purposes and the targeted development of novel therapeutic applications.

Experimental procedures

Cell lines and primary cell cultures

The retroviral expression vector, pMOWS, containing HA-EpoR cDNA, was introduced into BaF3 cells by retroviral transduction. Cell lines stably expressing HA-EpoR (BaF3-HA-EpoR) were selected and maintained in RPMI 1640 (Invitrogen, Carlsbad, CA, USA) in the presence of puromycin.

Primary hepatocytes were isolated from male Black-6 mice (6–8 weeks old) (Charles River, Wilmington, MA, USA). Livers were perfused with Hanks buffer supplemented with collagenase II (Biochrom, Berlin, Germany). Experiments were carried out in accordance with the German Animal Welfare Act of 12 April 2002 and the European Council Directive of 24 November 1986. Intact liver capsules were transferred into Williams' medium (Biochrom) supplemented with fetal bovine serum, insulin, L-glutamine and dexamethasone. Hepatocytes were removed from the capsules, enriched by centrifugation and cultured on collagen I-coated dishes (BD Biosciences, Franklin Lakes, NJ, USA) in Williams' medium E (Biochrom) supplemented with L-glutamine and dexamethasone.

Expression, purification and quantification of recombinant proteins

Unphosphorylated purified ERK2 was purchased from Cell Signaling Technologies (Beverly, MA, USA). The cytoplasmic domain of the EpoR was cloned into pGEX-2T (Amersham Biosciences, Piscataway, NJ, USA) and expressed in *E. coli* BL21 CodonPlus-RIL bacteria (Stratagene, La Jolla, CA, USA). Proteins were extracted by lysozyme lysis and sonication. Glutathione agarose beads (Sigma-Aldrich, St Louis, MO, USA) were added to lysates and proteins were eluted by the addition of reduced glutathione (Sigma-Aldrich). For the quantification of purchased and purified proteins, dilution series of purified BSA (Sigma-Aldrich) and the recombinant proteins were separated by 10% SDS/PAGE and stained with Coomassie Brilliant Blue.

The gel was documented using the trans-illumination mode of a LumiImager (Roche Diagnostics, Mannheim, Germany). Proteins were quantified using LUMIANALYST software (Roche Diagnostics).

Time-course experiments

BaF3-HA-EpoR cells were starved for 5 h in RPMI 1640 (Invitrogen) supplemented with $1 \text{ mg}\cdot\text{mL}^{-1}$ BSA (Sigma-Aldrich) and then stimulated with $50 \text{ units}\cdot\text{mL}^{-1}$ Epo (Cilag-Jansen, Bad Homburg, Germany). For each time-point, 10^7 cells were taken from the pool of cells and lysed by the addition of $2 \times$ Nonidet P-40 lysis buffer, thereby terminating the reaction.

A total of 2×10^6 primary hepatocytes were cultured for 24 h after plating on collagen I-coated 60 mm dishes (BD Biosciences) in Williams' medium E (Biochrom) supplemented with L-glutamine and dexamethasone. Cells were starved for 5 h in Williams' medium E supplemented with L-glutamine. Each dish was stimulated with $40 \text{ ng}\cdot\text{mL}^{-1}$ IL-6 in Williams' medium E containing L-glutamine. The medium was removed from the cells, $1 \times$ Nonidet P-40 lysis buffer was added, and cells were collected using a cell scraper.

Immunoprecipitation and quantitative immunoblotting

For IP, cytosolic lysates were incubated with anti-EpoR immunoglobulin (Santa Cruz, La Jolla, CA, USA) or anti-STAT3 immunoglobulin (Cell Signaling Technologies). Immunoprecipitated proteins and total cellular lysates were separated by SDS/PAGE and transferred to poly(vinylidene difluoride) (PVDF) or nitrocellulose membranes. Proteins were immobilized with Ponceau S solution (Sigma-Aldrich) followed by immunoblotting analysis using the anti-phosphotyrosine mAb, 4G10 (Upstate Biotechnology, Lake Placid, NY, USA), the anti-(tyrosine phosphorylated STAT3) immunoglobulin or the anti-(double phosphorylated p44/42 MAP kinase) immunoglobulin (both Cell Signaling Technologies). Antibodies were removed by treating the blots with β -mercaptoethanol and SDS, as described previously [18]. Reprobes were performed using anti-EpoR (Santa Cruz), anti-STAT3 or anti-(p44/42 MAP kinase) (both Cell Signaling Technologies) immunoglobulins. For normalization, antibodies against β Actin (Sigma-Aldrich), PDI, Hsc70 and Calnexin (all Stressgen, Victoria, Canada) were used. Secondary horseradish peroxidase (HRP)-coupled antibodies (anti-rabbit HRP, anti-mouse HRP, protein A HRP) were purchased from Amersham Biosciences. Immunoblots against phosphorylated EpoR and total EpoR were incubated with enhanced chemiluminescence (ECL) substrate (Amersham Biosciences) for 1 min, and exposed for 10 min on a LumiImager (Roche Diagnostics). All other immunoblots were incubated with ECL Advance substrate (Amersham

Biosciences) for 2 min, and exposed for 1 min on a LumiImager (Roche Diagnostics). For quantifications, LUMIANALYST software (Roche Diagnostics) was used.

Spline approximation and signal normalization

Smoothing splines were applied to the noisy data to estimate the actual values. Their smoothness was determined by generalized cross-validation, minimizing the mean square error between the estimated time-course and the data [10,12]. Splines were used for criteria-mediated error reduction by GELINSPECTOR, as described in the Supplementary material.

Computational data processing by GELINSPECTOR

The computer algorithm GELINSPECTOR requires MATLAB 6.5 and the freely available statistics environment R1.9 or above. It visualizes the blotting error in a gel domain, rearranges randomized gel loadings into chronological order, applies the presented criteria-mediated normalization procedure, and merges data deriving from one experiment measured on several gels. Choices to calculate a first estimate include constant, sigmoidal and spline functions. Splines are calculated using the mgcv library [11] from R or the MATLAB spline toolbox. After specifying proteins of interest, normalizers, calibrators, the measurement files and some global options (such as the preferred figure format pdf, eps, jpg or png), GELINSPECTOR works completely automatically. Figures of all important steps, and files with the processed data, are created. GELINSPECTOR is available from the authors.

Acknowledgements

We thank Stefan Rose-John for the generous gift of IL-6, Sabine McNelly for technical help with hepatocyte preparation and Jan G. Hengstler for the development of standard operating procedures for the preparation of primary hepatocytes. We also thank Nils Blüthgen and Peter J. Nickel for helpful suggestions and Ute Baumann for excellent technical assistance. We are grateful to Jennifer Reed for critically reading the manuscript.

This work was supported by the funding priority 'Systems of Life – Systems Biology' of the German Federal Ministry of Education and Research (BMBF).

References

- 1 Kholodenko BN, Demin OV, Moehren G & Hoek JB (1999) Quantification of short term signaling by the epidermal growth factor receptor. *J Biol Chem* **274**, 30169–30181.

- 2 Schoeberl B, Eichler-Jonsson C, Gilles ED & Muller G (2002) Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nat Biotechnol* **20**, 370–375.
- 3 Bhalla US, Ram PT & Iyengar R (2002) MAP kinase phosphatase as a locus of flexibility in a mitogen-activated protein kinase signaling network. *Science* **297**, 1018–1023.
- 4 Kitano H (2002) Systems biology: a brief overview. *Science* **295**, 1662–1664.
- 5 Nelson DE, Ihekweaba AE, Elliott M, Johnson JR, Gibney CA, Foreman BE, Nelson G, See V, Horton CA, Spiller DG *et al.* (2004) Oscillations in NF-kappaB signaling control the dynamics of gene expression. *Science* **306**, 704–708.
- 6 Hoffmann A, Levchenko A, Scott ML & Baltimore D (2002) The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation. *Science* **298**, 1241–1245.
- 7 Swameye I, Muller TG, Timmer J, Sandra O & Klingmuller U (2003) Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by data-based modeling. *Proc Natl Acad Sci USA* **100**, 1028–1033.
- 8 Bentele M, Lavrik I, Ulrich M, Stosser S, Heermann DW, Kalthoff H, Krammer PH & Eils R (2004) Mathematical modeling reveals threshold mechanism in CD95-induced apoptosis. *J Cell Biol* **166**, 839–851.
- 9 Feather-Henigan K, Hersey S, Johnson A, Milosevich GM & Hines K (1999) Immunoblot imaging with a cooled CCD camera and chemiluminescent substrates. *Am Biotechnol Lab* **17**, 44–46.
- 10 Craven P & Wahba G (1979) Smoothing noisy data with spline functions. *Numer Math* **31**, 377.
- 11 Wood SN (2003) Thin plate regression splines. *J R Statist Soc B* **65**, 95–114.
- 12 Green P & Silverman B (1994) *Nonparametric Regression and Generalized Linear Models*. Chapman & Hall, London.
- 13 Hill AB (1951) The clinical trial. *Br Med Bull* **7**, 278–282.
- 14 Kerr MK (2003) Design considerations for efficient and effective microarray studies. *Biometrics* **59**, 822–828.
- 15 Li M & Hazelbauer GL (2004) Cellular stoichiometry of the components of the chemotaxis signaling complex. *J Bacteriol* **186**, 3687–3694.
- 16 Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH & Aebersold R (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* **17**, 994–999.
- 17 Schmidt A, Kellermann J & Lottspeich F (2005) A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* **5**, 4–15.
- 18 Klingmuller U, Lorenz U, Cantley LC, Neel BG & Lodish HF (1995) Specific recruitment of SH-PTP1 to the erythropoietin receptor causes inactivation of JAK2 and termination of proliferative signals. *Cell* **80**, 729–738.

Supplementary material

The following supplementary material is available for this article online:

Doc. S1. Computational processing and error reduction strategies for standardized quantitative data in biological networks.

This material is available as part of the online article from <http://www.blackwell-synergy.com>

**Supplementary Information:
Computational Processing and Error Reduction Strategies for
Standardized Quantitative Data in Biological Networks**

Marcel Schilling,^{1,2} Thomas Maiwald,^{3,2} Sebastian Bohl,¹ Markus
Kollmann,³ Clemens Kreutz,³ Jens Timmer,³ and Ursula Klingmüller¹

¹*German Cancer Research Center, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany.*

²*These authors contributed equally.*

³*FDM, Freiburg Center for Data Analysis and Modeling,
University of Freiburg, Eckerstr. 1, 79104 Freiburg, Germany.*

(Dated: October 25, 2005)

I. ERROR CLASSIFICATION

Immunoblotting as a technique for quantitative analysis of protein concentrations has several sources of noise, which can be divided into three classes:

- (i) pipetting errors $f(j)$ change the amount of each protein in the j -th lane by the same factor, reflecting, e.g., the different amount of lysate loaded on each gel lane
- (ii) blotting errors $g(j, m)$, observed as brighter and darker areas, arise from inhomogeneities of the gel and transfer to the membrane. They are highly correlated for neighboring lanes, j , and rows, m .
- (iii) contributions independent from the loaded protein concentrations. They are modeled as Gaussian noise, $\eta(j, m)$.

As shown in the main text, error contributions from pipetting are small compared to the highly correlated errors of the blotting technique (see Fig. 2b, main text). Figure 7 gives an example of a blotting error and Figure 1 displays the mean autocorrelation function (ACF) for 13 gels of the normalizer protein β Actin and for 8 gels of the calibrator GST-EpoR depending on the lane distance.

The effects on the concentration $x^*(t_j)$ of a given protein in the lysate at time t_j , resulting in the measured concentration $x(t_j)$, can be described by

$$x(t_j) = [1 + g(j, m)] [1 + f(j)] x^*(t_j) + \eta(j, m). \quad (1)$$

Here, the time point t_j after stimulation corresponds to lane j , the smooth systematic error g depends on the lane index j and on the molecular weight m , measured in kD. The pipetting error f depends only on the lane number j , since it changes the amount of all proteins in a lane by the same factor. Errors arising from (i) and (iii) are uncorrelated among different lanes, resulting in $\langle f(j)f(j') \rangle = 0$ and $\langle \eta(j, m)\eta(j', m) \rangle = 0$ for $j \neq j'$.

In the following we use normalizers and calibrators and a randomized, non-chronological loading of the

lanes, to identify and reduce the highly correlated blotting errors, $g(j, m)$.

II. ELIMINATION OF THE BLOTGING ERROR

The highly correlated errors arising from the blotting technique vary gradually over the lanes and make it difficult to extract the actual values. To eliminate the correlations among the lanes we employ non-chronological gel loading. Here, the subsequent time-points after stimulation are loaded randomized on the gel under the condition that consecutive time points are separated by minimum number of 4 lanes for 20 time points (compare loading example in Figure 7). By applying this method, the errors between consecutive time points are uncorrelated. We are able to estimate the true time-course from the data by rearranging the time points in chronological order. As shown in Figure 2C, we employ a cubic spline whose smoothness is determined by generalized cross-validation. This technique demands statistical independent errors as generated by the randomized gel loading. The estimation of a time-course from noisy data by smoothing splines has been worked out in detail in Refs. [1–4]. We emphasize that a sufficiently dense grid of time-points is necessary to keep the bias of this method small.

For the case that a normalizer protein, $x_n(j)$, can be measured with a similar molecular weight as the protein of interest, it is possible to estimate the blotting error $g(j, m)$ as $x_n^*(j) = \text{const}$ by definition. The true signal is then given by

$$\hat{x}^*(t_j) \approx \frac{x_i(t_j)}{\bar{x}_n(t_j)}. \quad (2)$$

Here, $\bar{x}_n(t)$, denotes the smoothing spline generated from the data set $\{x_n(t_j)\}$ by keeping the *lane ordering* of the randomly loaded gels. Smoothing of the data is performed in order to average over error contributions arising from pipetting, $f(j)$, and other sources of noise, $\eta(j, m)$. We further denote by $\tilde{x}^*(t)$ and $\tilde{x}(t)$ the time-courses of the smoothing splines generated from

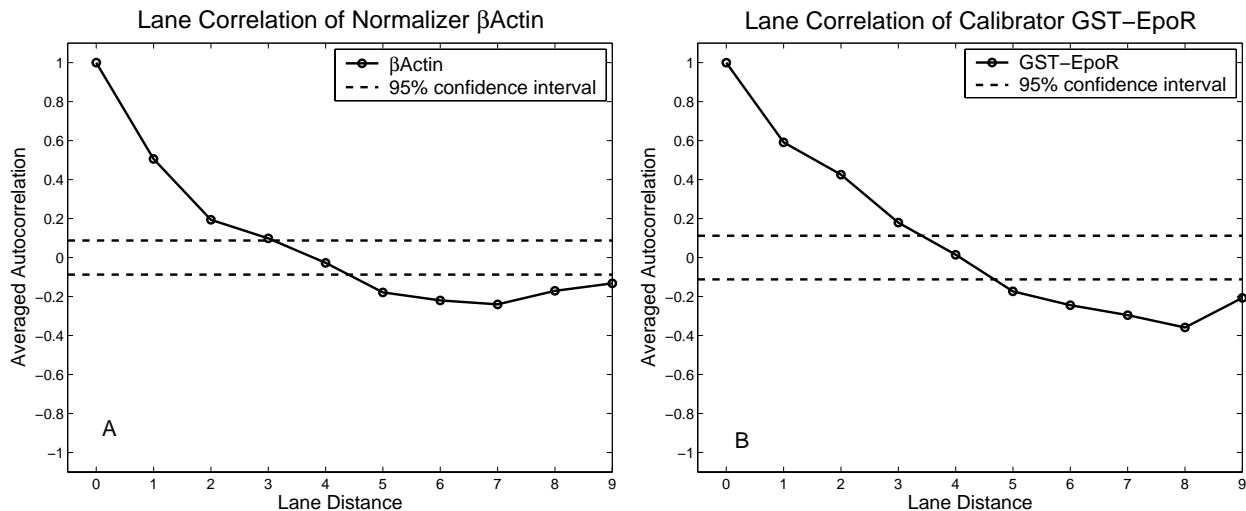


FIG. 1: Autocorrelation of the normalizer protein β Actin (A) and the calibrator protein GST-EpoR (B) in the gel domain. The two-sided 95% confidence interval for the averaged autocorrelation function of a purely random process is in both cases not preserved, indicating a strong correlation of neighbored gel lanes.

the *chronological* ordered data sets $\{\hat{x}^*(t_j)\}$, $\{x(t_j)\}$. The residuals $|\tilde{x}^*(t_j) - \hat{x}^*(t_j)|$, are expected to be significantly smaller than the residuals without employing a normalizer, $|\tilde{x}(t_j) - x(t_j)|$, as we have readily accounted for the blotting errors in the first case. This in turn gives us a reliable measure for the quality of the used normalizer protein.

III. ERROR REDUCTION VIA RANDOMIZATION AND NORMALIZERS - SIMULATION STUDY

Simulations of typical immunoblotting experiments were performed by generating a simulated signal with quadratic rise and exponential decay and a maximum at half lane number, equidistantly sampled (Figure 2B). This simulates a typical time-course experiment after stimulation with a hormone. The true signal $x^*(t_j)$ was processed with the two main sources of errors as described in the previous section, a pipetting and a blotting error. In detail:

1. A multiplicative, uncorrelated pipetting error was applied as shown in Figure 2A representing errors derived from unequal cell number or errors in pipetting the cellular lysates:

$$x'(t_j) = x^*(t_j) \cdot (1 + \sigma\varepsilon(j)) \quad \varepsilon(j) \in N(0, 1).$$

2. A multiplicative, strongly correlated blotting error was applied, representing errors from differences in migration in the SDS polyacrylamide gel or unequal

transfer to the membrane:

$$x(t_j) = x'(t_j) \cdot (1 + g(j)),$$

with the blotting error $g(j)$ represented by a sine function with mean zero and phase, amplitude and frequency consistent with experimental observations.

The processing was applied to a chronological and to a randomized true signal, x_{rand}^* and x_{chron}^* , respectively, leading to "measurements" like in Figure 2B. Note that the chronological signal is rather smooth but changes the characteristic of the true signal: The maximum occurs earlier and a new minimum is observed at $t = 15$. The randomized signal on the other hand is very noisy, but does not introduce systematic effects. The smoothed processed randomized signal \tilde{x}_{rand} is very close to the true time-course, whereas the smoothed processed chronological signal \tilde{x}_{chron} still keeps correlated deviations from the true signal (Figure 2C). The correlation structure of the deviations can be investigated via the autocorrelation function (Figure 2D). For uncorrelated errors, the autocorrelation function should drop from 1 at $\tau = 0$ into the 95% confidence interval for $\tau > 0$. This is not the case for the processed chronologically signal, which can lead to misleading conclusions if methods are applied which assume uncorrelated noise.

Besides visual inspection of the autocorrelation function, the improvement of data quality by means of a randomized gel loading can be quantified by the *error reduc-*

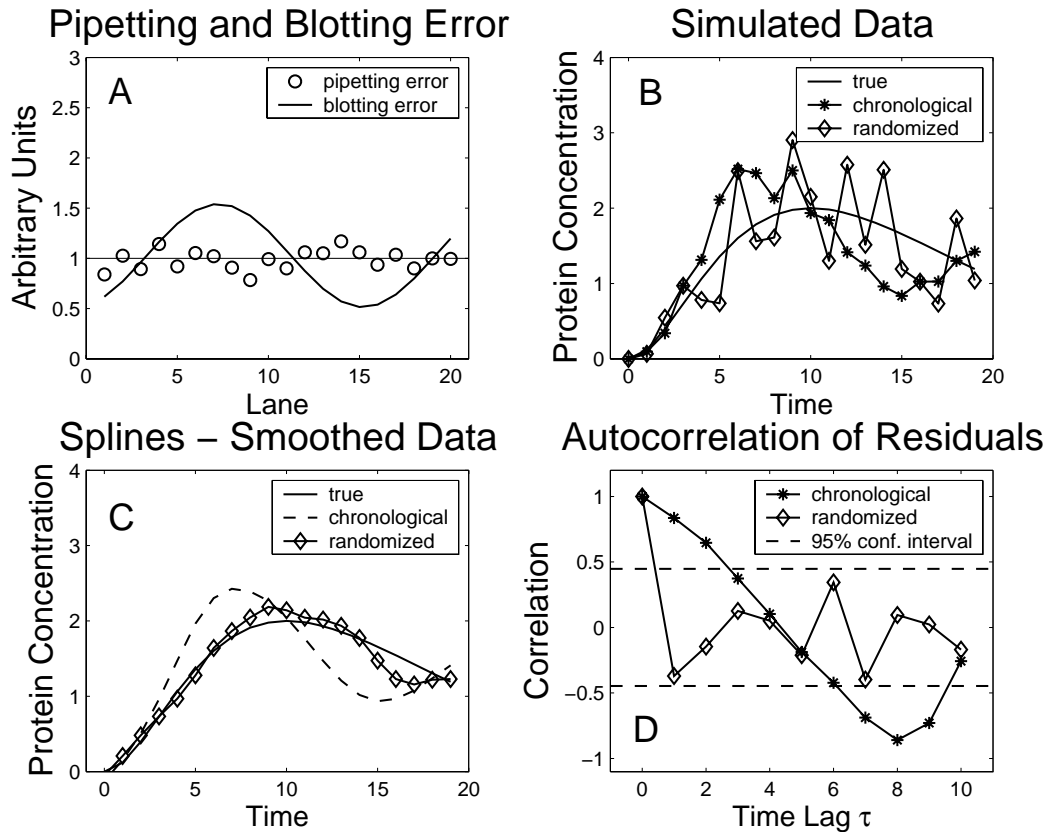


FIG. 2: Effect of randomization on immunoblotting data. (A) Simulated uncorrelated pipetting error and highly correlated, sine-like blotting error. (B) Simulated signal perturbed with the pipetting and blotting error in a chronological and randomized manner. Only the randomized procedure does not change the characteristics of the true signal, as the smoothed data show (C). (D) The residuals of the perturbed to the true signal exhibit a strong autocorrelation for the chronological procedure, which is not agreeable with white noise. This can be achieved by randomizing.

tion factor:

$$\alpha = \frac{\sqrt{\sum_j (\tilde{x}_{rand}(t_j) - x^*(t_j))^2}}{\sqrt{\sum_j (\tilde{x}_{chron}(t_j) - x^*(t_j))^2}}$$

For the illustrated data set the achieved reduction of the standard deviation was $\alpha \approx 0.4$. The reduction can only be quantified when the actual values are available, which is not the case in experimental measurements. Hence, the question arises whether a general error reduction factor can be established by randomizing or whether it depends on experimental parameters like the number of lanes, strength of signal maximum, blotting error or pipetting error. A simulation study showed that for small pipetting errors an error reduction factor of 0.45 ± 0.1 could be established independently from other parameters. Details of the study follow in the next section.

A. Quantifying the Error Reduction using Randomization and Calibrators

To determine the usefulness of randomization for the improvement of data quality, several parameters were varied including the number of lanes (10 to 100, Fig. 3A), the number of sine periods of the blotting error (0.8 to 2.2, Fig. 3B), the strength of the blotting error (ratio of smallest to largest value ranging from 1.5 to 10, Fig. 3C), the maximum signal strength (0.1 to 20, Fig. 3D), and the strength of the pipetting error (σ ranging from 0 to 1, Fig. 4). During the variation of one parameter, the other parameters were fixed:

- Number of lanes: 20
- Number of sine periods of the blotting error: 1
- Strength of the blotting error (max/min): 3
- Maximum signal strength: 2

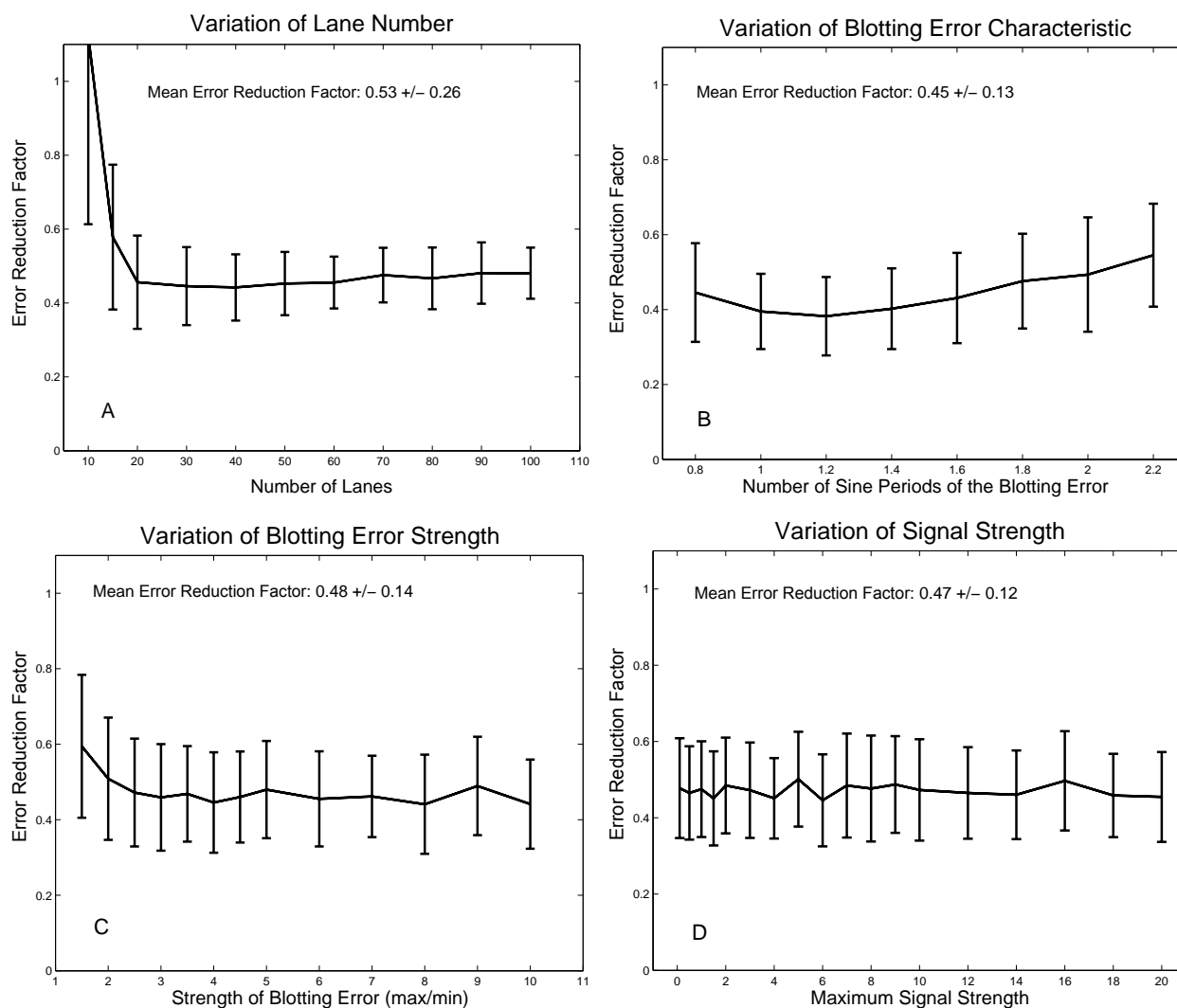


FIG. 3: Reduction of standard deviation of smoothed simulated blotting data by randomization. (A) For gels with more than 15 gel lanes randomization reduces robustly standard deviation by 50%. This holds also for variation of blotting error characteristics like number of periods (B) or error strength (C) and for a wide range of signal maximum (D).

- Standard deviation of the pipetting error: 0.1

Figures 3 and 4 display the error reduction factor for all parameter variations. At least 20 lanes should be used to achieve an optimal improvement. For the other investigated parameter ranges no strong effect is observed for all variations except for the strength of the pipetting error (Fig. 4). Since pipetting errors are uncorrelated, they cannot be reduced by randomization - if the fraction of the pipetting errors increases, the randomization takes less effects. In general, randomization decreases the standard deviation in quantitative immunoblotting to ca. 0.45 of the value without randomization, as long as the pipetting error is not too large. An approach to control the pipetting error in experiments is sampling the same

number of cells for each time point or measuring and adjusting total protein concentration.

B. Criteria for Employing Normalization with Normalizers and Calibrators

Calibrators and normalizers possess a constant concentration. Fluctuations occur only as measurement errors. Since the blotting error changes slowly from lane to lane and other errors like the pipetting error are rather uncorrelated, the blotting error can be estimated by smoothing the calibrator or normalizer signal, e.g. with a smoothing spline. Based on this blotting error estimate, the protein of interest can be normalized.

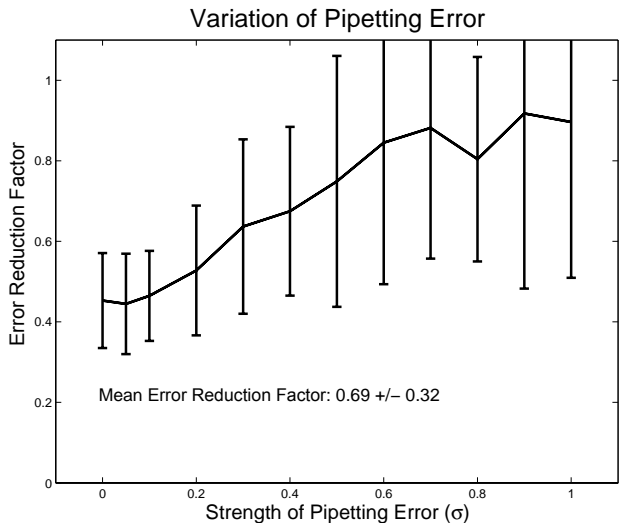


FIG. 4: Reduction of standard deviation of smoothed simulated blotting data by randomization for variation of the pipetting error strength. Increasing the pipetting error – blotting error ratio decreases the error reduction factor, since only blotting errors are tackled by randomization.

However, since the blotting error is a local property of the gel, normalizers and calibrators are required with a similar molecular weight as the protein of interest. If the molecular weight of the normalizer is different it does not reflect the blotting error for the protein of interest. We therefore developed criteria for employing normalizers and calibrators.

Figure 5A shows a simulated blotting error and a good estimation, corresponding to the smoothed signal of an appropriate normalizer in a real experiment. Smoothing the processed randomized signal leads to an acceptable estimation of the true signal. Smoothing the normalized signal yields virtually the true signal itself, as shown in Figure 5C. Even the correlation structure of the estimation error in the gel domain is improved.

The estimation of the blotting error displayed in Figure 6A is inaccurate: A strong phase shift can be observed, corresponding to a skewed gradient of the blotting error depending on the position on the blot. In this situation, normalizing the data increases the deviation of the estimated signal from the true signal. Hence, a criterion whether a normalization is applicable would be a decreased standard deviation of the estimated signal. This, though, requires knowledge of the true time course, which is not available. Instead, the smoothed curve of the randomized but not yet normalized signal is used as preliminary estimator of the true signal. If the normalizer is applicable, a new estimate can be calculated based on the randomized and

normalized data, otherwise the former estimate is kept.

The shown simulated data sets have the following standard deviations:

- Figure 5:
 - Randomized (true): 0.533
 - Randomized (estimation): 0.722
 - Randomized, normalized (true): 0.157
 - Randomized, normalized (estimation): 0.515
- Figure 6:
 - Randomized (true): 0.533
 - Randomized (estimation): 0.722
 - Randomized, normalized (true): 1.208
 - Randomized, normalized (estimation): 1.068

The estimated error decreases in case of Figure 5 and increases in case of Figure 6 if a normalizer is used. Hence, the normalization procedure is only applicable in the first case, reducing the true standard deviation from 0.533 to 0.157. In the other case it would increase the standard deviation from 0.533 to 1.208. This procedure works robustly for normalizers and calibrators, as long as randomized gel loading is applied.

C. Application to Stimulation Experiment

The randomizing and normalization procedure was applied on an erythropoietin (Epo)-induced time-course experiment resulting in phosphorylation of ERK1 and ERK2. Samples were loaded randomized and separated on 17.5% SDS polyacrylamide gel, and transferred to membranes that were developed with chemiluminescent substrates and quantified with the Lumi-Imager (Figure 7). We calculated the standard deviation of the signals to their spline approximation to 2.524 for pERK1 and 0.455 for pERK2. Normalization with β Actin reduced the standard deviation to the spline approximation to 1.878 for pERK1 and 0.262 for pERK2. The reduced lane-correlation for the normalized data confirms the quality of data processing. In this case the correlation structure of the systematic blotting error could be disrupted validating the normalization.

IV. CALCULATION OF MOLECULES PER CELL

A. Linearity of imaging unit

Quantification of a protein P measured by immunoblotting is performed via chemiluminescence detection yielding total intensities P_{blu} which are proportional

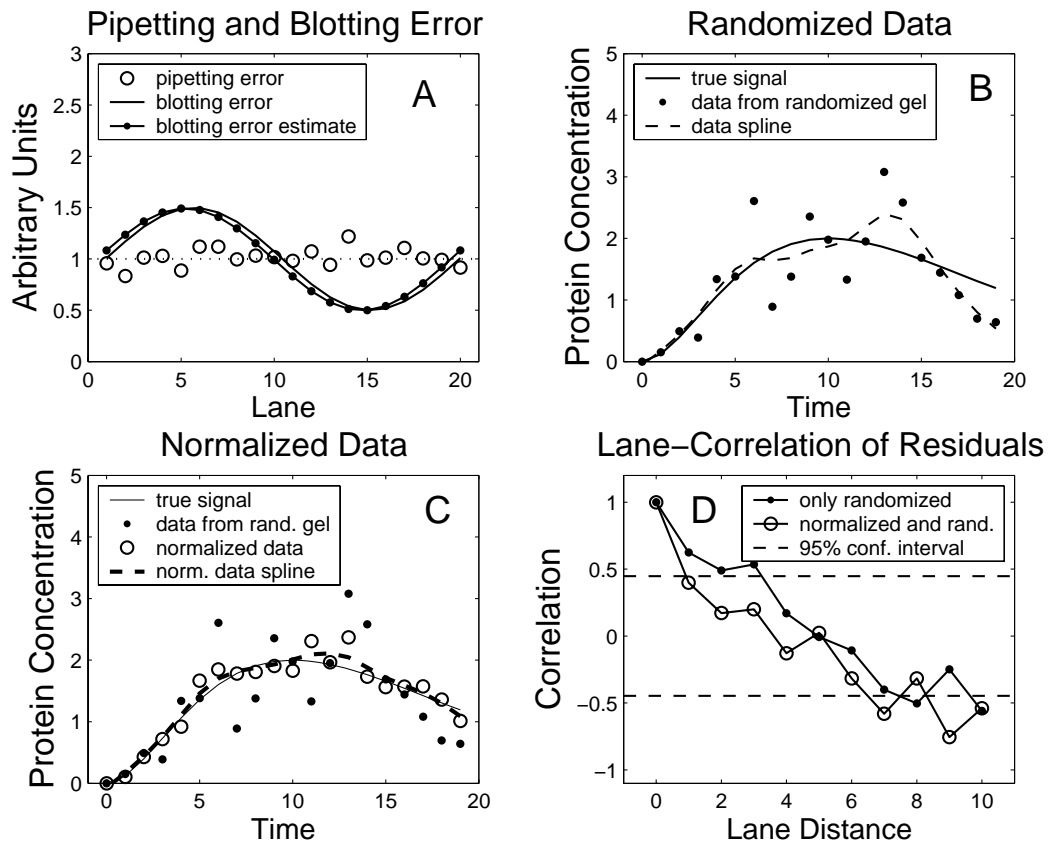


FIG. 5: Normalization of simulated time-course data depicting a valid procedure according to our criteria. (A) The blotting error is very well estimated corresponding to a suitable normalization protein. The perturbation of a simulated signal (B) is strongly reduced after normalization (C). Even the correlation in gel domain of the residuals is improved (D). The autocorrelation, i.e., the correlation in time domain, agrees for both randomized signals with white noise (not shown).

to the total number of molecules P_{tmlc} on the blot (Fig. 8). The linear relationship reads

$$P_{tmlc} = a P_{blu}$$

with a proportionality factor a and 0 y-axis interception. The factor a has to be determined for each protein species and for every blot, since the amount of antibody added varies for different blots and the antibody affinity differs for different proteins.

B. Requirements for the standard/calibrator protein

The reference protein R realized by a standard or calibrator protein should

- contain the same epitope binding to the antibody as the protein of interest,
- have a known molecular weight R_{mw} ,
- be added to the lysate with a known amount R_g .

C. Calculation of the proportionality factor

The total number of reference proteins in the lysate is given as

$$R_{tmlc} = \frac{N_A R_g}{R_{mw}},$$

with Avogadro constant $N_A = 6.022 \cdot 10^{23}$. If the imaging unit measures the intensity R_{blu} , the proportionality factor can be calculated as

$$a = \frac{R_{tmlc}}{R_{blu}} = \frac{R_{tmlc} N_A R_g}{R_{blu} R_{mw}}.$$

If possible, one should measure the reference protein several times and estimate a by linear regression. This provides also a standard deviation for a .

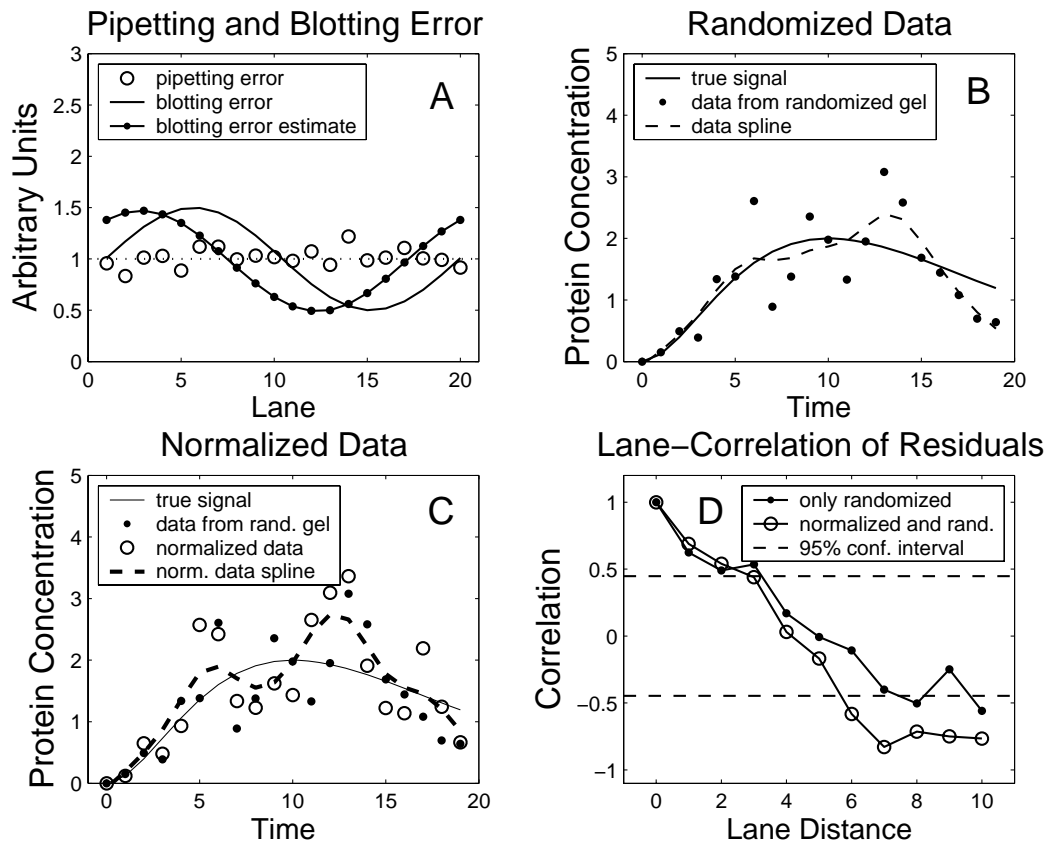


FIG. 6: Normalization of simulated time-course data depicting a rejected procedure according to our criteria. (A) The blotting error estimate is phase shifted, corresponding with a too distant normalization protein. The perturbation of a simulated signal (B) cannot be reduced with normalization (C) and the lane-to-lane correlation is not improved (D).

D. Calibrators

Calibrator proteins harbor the same antibody epitope as the protein of interest, P , yet possess a different molecular weight than P resulting in a distinct band in the immunoblot analysis. If analysis of total cellular lysates are performed, a few lanes of the immunoblot have to be used for the standard protein to facilitate parallel detection.

E. Calculation of molecules per cell

P_{tmlc} is the total number of molecules of the investigated lysate. If the number of cells in the lysate is

available, the molecule number per cell can be calculated as

$$P_{mlc} = \frac{P_{tmlc}}{\# \text{ cells}}.$$

- [1] P. Green and B. Silverman, *Nonparametric Regression and Generalized Linear Models* (Chapman and Hall, London, 1994).
 [2] A. Buja, T. Hastie, and R. Tibshirani, *Ann. Stat.* **17**, 453

- (1989).
 [3] P. Craven and G. Wahba, *Numer. Math.* **31**, 377 (1979).
 [4] S. Wood, *J. Royal Statistical Soc. B.* **65**(1), 95 (2003).

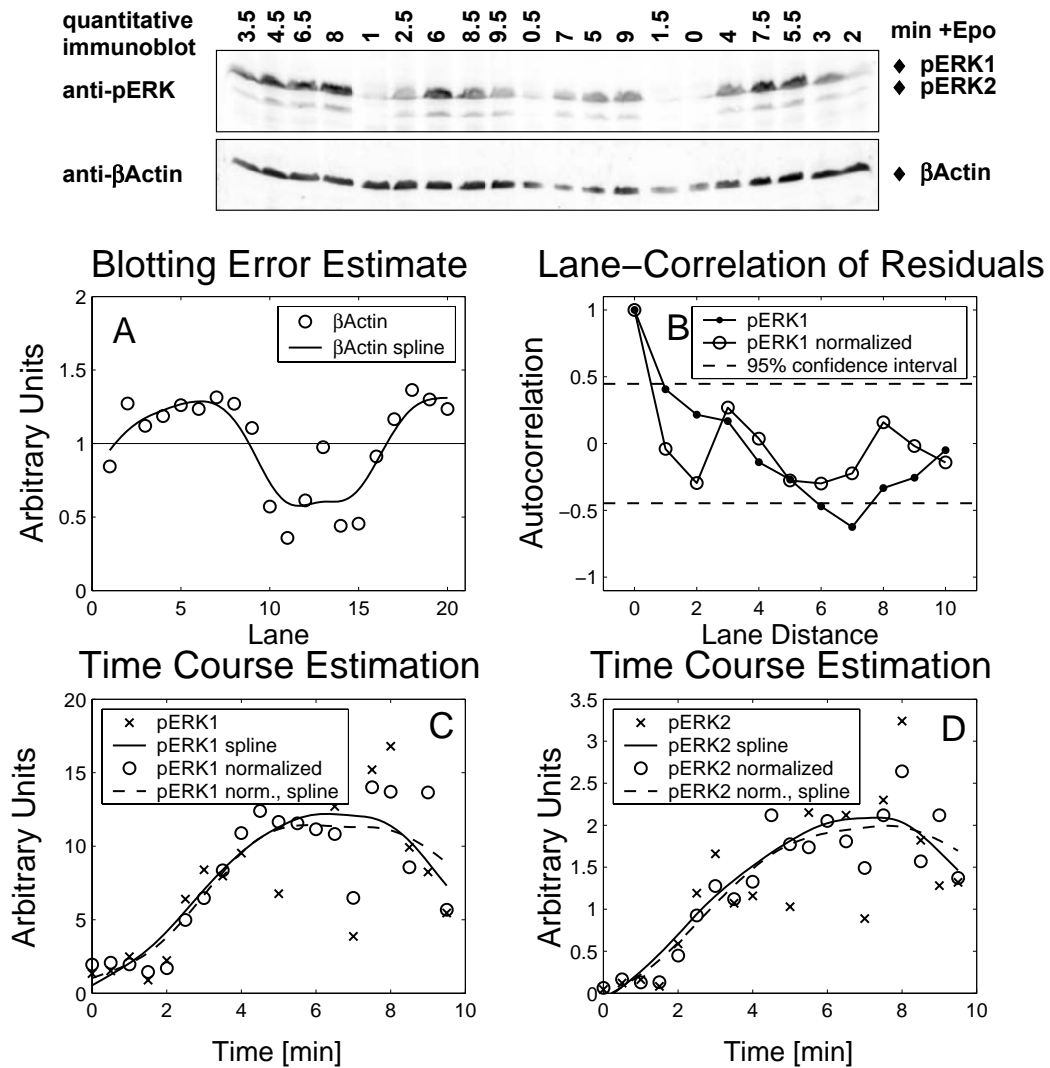


FIG. 7: Randomization and normalization of an Erythropoietin-induced time-course experiment. BaF3-EpoR cells are stimulated with 50 units/ml Epo resulting in ERK phosphorylation. Gel electrophoresis has been applied with a randomized, non-chronological gel loading with β Actin as normalizer protein (upper panel). (A) Smoothed measurements of β Actin serve as estimate of the strong, sine-like blotting error. (C, D) Normalization reduces significantly standard deviation of pERK1/2 measurements compared to a spline-smoothed pERK1/2 signal (cont. line), which serves as first estimate of the true signal.

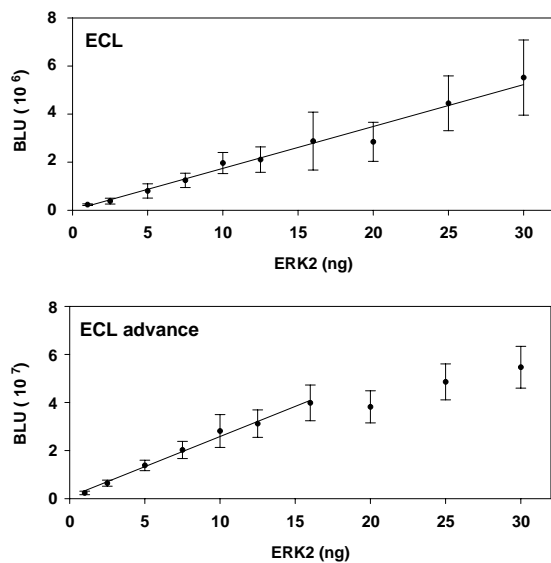


FIG. 8: A dilution series of purified ERK2 was separated eight times by a 10% SDS polyacrylamide gel and transferred to a membrane that was probed with anti-ERK antibody and subsequently developed with ECL or ECL advance. To determine linearity, the amount of ERK2 was plotted versus measured signal strength. Signals were linear up to 4×10^7 BLU.

2.2 Quantitative data generation for systems biology

Schilling M, Maiwald T, Bohl S, Kollmann M, Kreutz C, Timmer J, Klingmüller U.

Quantitative data generation for systems biology: the impact of randomisation, calibrators and normalisers.

IEE Proc Syst Biol. 2005 Dec;152(4):193-200.

This publication contains 6 figures. All experiments were initiated and planned by Marcel Schilling. The experiments depicted in Figures 4, 5, and 6 were performed by Marcel Schilling. The manuscript was largely written by Marcel Schilling (figures and text).

Quantitative data generation for systems biology: the impact of randomisation, calibrators and normalisers

M. Schilling, T. Maiwald, S. Bohl, M. Kollmann, C. Kreutz, J. Timmer and U. Klingmüller

Abstract: Systems biology is an approach to the analysis and prediction of the dynamic behaviour of biological networks through mathematical modelling based on experimental data. The current lack of reliable quantitative data, especially in the field of signal transduction, means that new methodologies in data acquisition and processing are needed. Here, we present methods to advance the established techniques of immunoprecipitation and immunoblotting to more accurate and quantitative procedures. We propose randomisation of sample loading to disrupt lane correlations and the use of normalisers and calibrators for data correction. To predict the impact of each method on improving the data quality we used simulations. These studies showed that randomisation reduces the standard deviation of a smoothed signal by $55\% \pm 10\%$, independently from most experimental settings. Normalisation with appropriate endogenous or external proteins further reduces the deviation from the true values. As the improvement strongly depends on the quality of the normaliser measurement, a criteria-based normalisation procedure was developed. Our approach was experimentally verified by application of the proposed methods to time course data obtained by the immunoblotting technique. This analysis showed that the procedure is robust and can significantly improve the quality of experimental data.

1 Introduction

Blotting techniques are widely used to analyse components in biological systems. They are based on the separation of components according to the molecular weight within a gel and transfer to a membrane followed by a detection process. The presence of proteins and/or their modifications in complex mixtures is examined by immunoblotting using specific antibodies in combination with chemiluminescence detection. So far, the data generated by immunoblotting have been primarily qualitative but the recent report of data-based mathematical modelling of the JAK-STAT signalling pathway [1] demonstrates the potential of using quantitative immunoblotting for systems biology approaches.

Here, we suggest new methodologies to improve data acquisition and data processing for quantitative immunoblotting. We propose randomised gel loading to transform correlated blotting errors into uncorrelated blotting errors by loading samples on the gel in a non-chronological order: Neighbouring lanes on the gel are used, not for consecutive time points, but in a randomised way. Furthermore, we suggest normalisation using data of calibrators (purified proteins of a different molecular weight from that of the protein of interest and the same antibody binding epitope added to cell lysates prior to immunoprecipitation) and

normalisers (endogenous proteins quantified by reprobing the immunoblot). Smoothing splines of the normaliser or calibrator signals were employed to correct immunoblotting data in an unbiased, criteria-mediated framework.

In addition, we present a quantitative analysis of the effects of these improvements by assessing the influence of each method on the standard deviation and correlation structure of simulated data and measurement processes. By applying our procedures to a data set comprising five independent experiments, each measuring eight protein species with 20 time points, we validated the performance of our approach method under experimental conditions.

2 Randomisation reduces standard deviation of immunoblotting data more than two-fold

Simulations of typical immunoblotting experiments were performed by generating a simulated signal x^* with quadratic rise and exponential decay and a maximum at half gel slot number, equidistantly sampled (Fig. 1b). This simulates a typical time course experiment after stimulation with a hormone. The true signal x^* was processed as follows:

First, a multiplicative, uncorrelated pipetting error of strength σ was applied, as shown in Fig. 1a, representing errors derived from unequal cell number or errors in pipetting the cellular lysates:

$$x' = x^* \cdot (1 + \varepsilon) \quad \varepsilon \sim N(0, \sigma)$$

Secondly, a multiplicative, strongly correlated blotting error was applied, representing errors from differences in migration in the SDS polyacrylamide gel or unequal transfer to the membrane, a common and probably underestimated problem in immunoblotting (compare the estimated blotting error in Figs. 4a and 6a and d)

$$x = x' \cdot g$$

© IEE, 2005

IEE Proceedings online no. 20050044

doi:10.1049/ip-syb:20050044

Paper first received 1st July and in revised form 7th September 2005

M. Schilling, S. Bohl and U. Klingmüller are with the German Cancer Research Center, Im Neuenheimer Feld 280, Heidelberg D-69120, Germany

T. Maiwald, M. Kollman, C. Kreutz and J. Timmer are with the Freiburg Centre for Data Analysis & Modelling, University of Freiburg, Eckerstr. 1, Freiburg D-79104, Germany

M. Schilling and T. Maiwald contributed equally to this work.

E-mail: m.schilling@dkfz.de

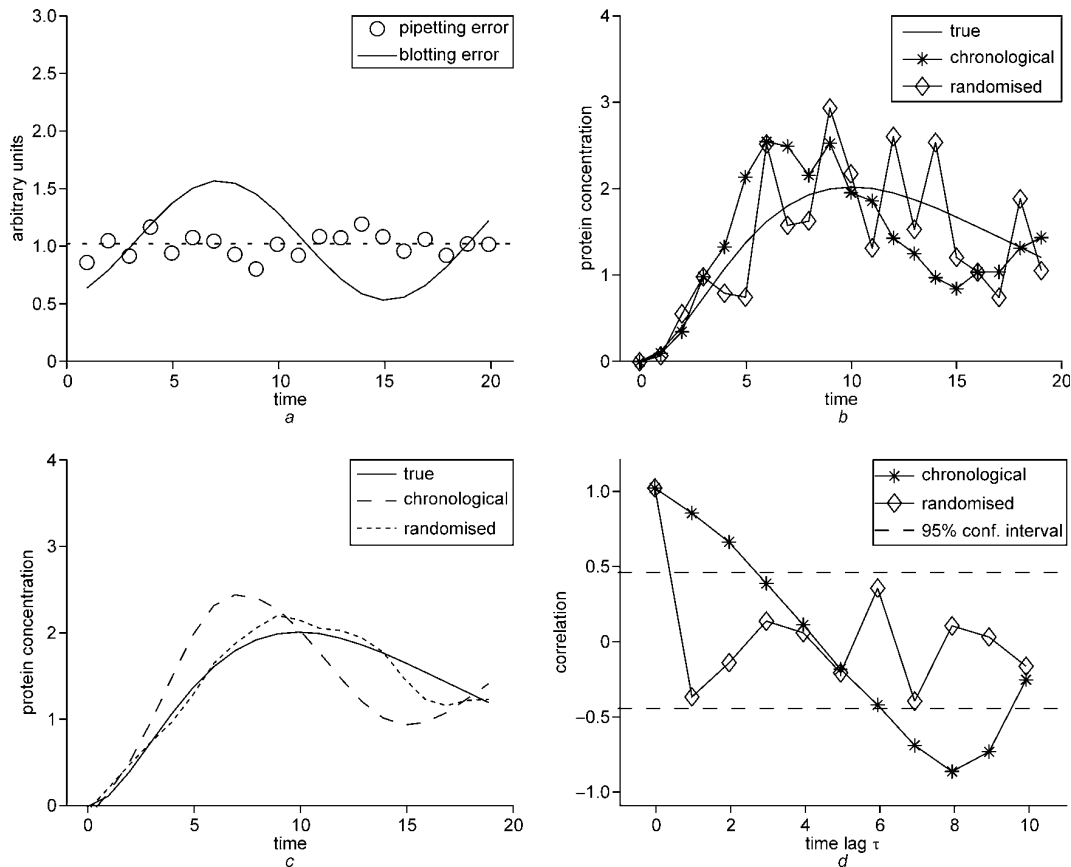


Fig. 1 Effect of randomisation on immunoblotting data

a Simulated uncorrelated pipetting error and highly correlated, sine-like blotting error

b Simulated signal perturbed with pipetting and blotting error in chronological and randomised manner

c Only randomised procedure does not change characteristics of true signal, as smoothed data show

d Residuals of perturbed to true signal exhibit strong autocorrelation for chronological procedure, which is not agreeable with white noise. Randomisation prevents this autocorrelation

with the blotting error g represented by a sine function with mean 1 and varied phase, amplitude and frequency.

The processing was applied to a chronological signal and to a randomised true signal, respectively, leading to simulated measurements such as in Fig. 1*b*. Note that the chronological signal is rather smooth but changes the characteristic of the true signal: the maximum occurs earlier, and a new minimum is observed at $t = 15$. The randomised signal, on the other hand, is very noisy, but does not introduce misleading effects. The smoothed, processed, randomised signal is very close to the true one, whereas the smoothed, processed, chronological signal conserves the correlated deviations from the true signal (Fig. 1*c*).

The correlation structure of the deviations can be investigated through the autocorrelation function (Fig. 1*d*). For uncorrelated errors, the autocorrelation function should drop from 1 at $\tau = 0$ into the 95% confidence interval for $\tau > 0$ [2]. This is not the case for the processed chronological signal, which can lead to incorrect conclusions if methods assuming uncorrelated noise are applied. Besides visual inspection of the autocorrelation function, the improvement of data quality can be quantified by the error reduction factor by randomisation, defined as the reduction of the standard deviation of the smoothed signal by means of randomisation

error reduction factor by randomisation

$$= \frac{\text{standard deviation (smoothed chronological data)}}{\text{standard deviation (smoothed randomised data)}}$$

For the illustrated data set, an error reduction factor of 0.4 was calculated, i.e. the standard deviation could be decreased by 60%. The reduction can only be quantified when the true data are available, which is certainly not the case in real measurements. Hence, the question arises whether a general reduction factor can be established by randomising or whether it depends on experimental parameters such as the number of lanes, maximum signal strength, blotting error or pipetting error. We performed a simulation study showing that, for small pipetting errors, a reduction to $45\% \pm 10\%$ could be established independently from other parameters.

3 Simulation study

Several parameters were varied quantitatively to assess the effect of randomisation. This included the number of lanes (10–100), the number of sine periods of the blotting error (0.8–2.2), the strength of the blotting error (ratio of smallest to largest value ranging from 1.5 to 10), the strength of the pipetting error (σ ranging from 0 to 1) and the maximum signal strength (0.1–20).

During the variation of one parameter, the other parameters were fixed:

- number of lanes: 20
- number of sine periods of the blotting error: 1
- strength of the blotting error (max/min): 3
- standard deviation of the pipetting error: 0.1
- maximum signal strength: 2.

Figure 2 displays the error reduction factor for all parameter variations. Variation of lane number (Fig. 2a) showed that randomisation is recommended for 15 or more lanes. For the other investigated parameter ranges, no strong effect could be observed for all variations (Figs. 2b–d), except for the strength of the pipetting error (Fig. 2e).

As pipetting errors are uncorrelated, they cannot be reduced by randomisation: if the fraction of the pipetting errors increases, the randomisation has less effect. In general, randomisation decreases the standard deviation in quantitative immunoblotting by $55\% \pm 10\%$ of the value without randomisation, as long as the standard deviation of the pipetting error does not exceed 20%. An approach

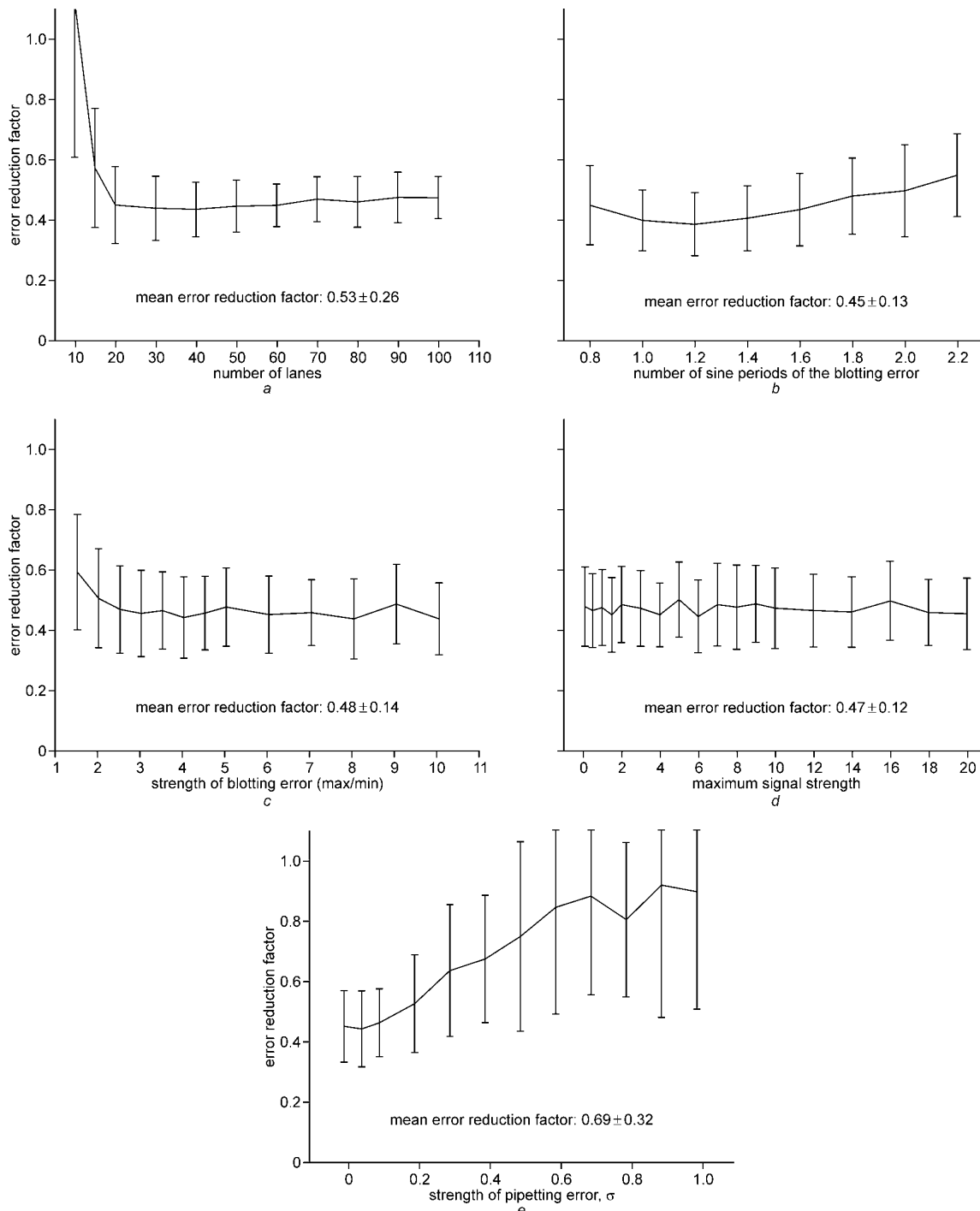


Fig. 2 For gels with more than 15 gel lanes, randomisation reduces standard deviation robustly by 50%

a Variation of lane number

b Variation of blotting error characteristics

c Variation of error strength

d Variation of signal strength

e Variation of pipetting error

Increasing pipetting error to blotting error ratio decreases error reduction factor, as only blotting errors are tackled by randomisation

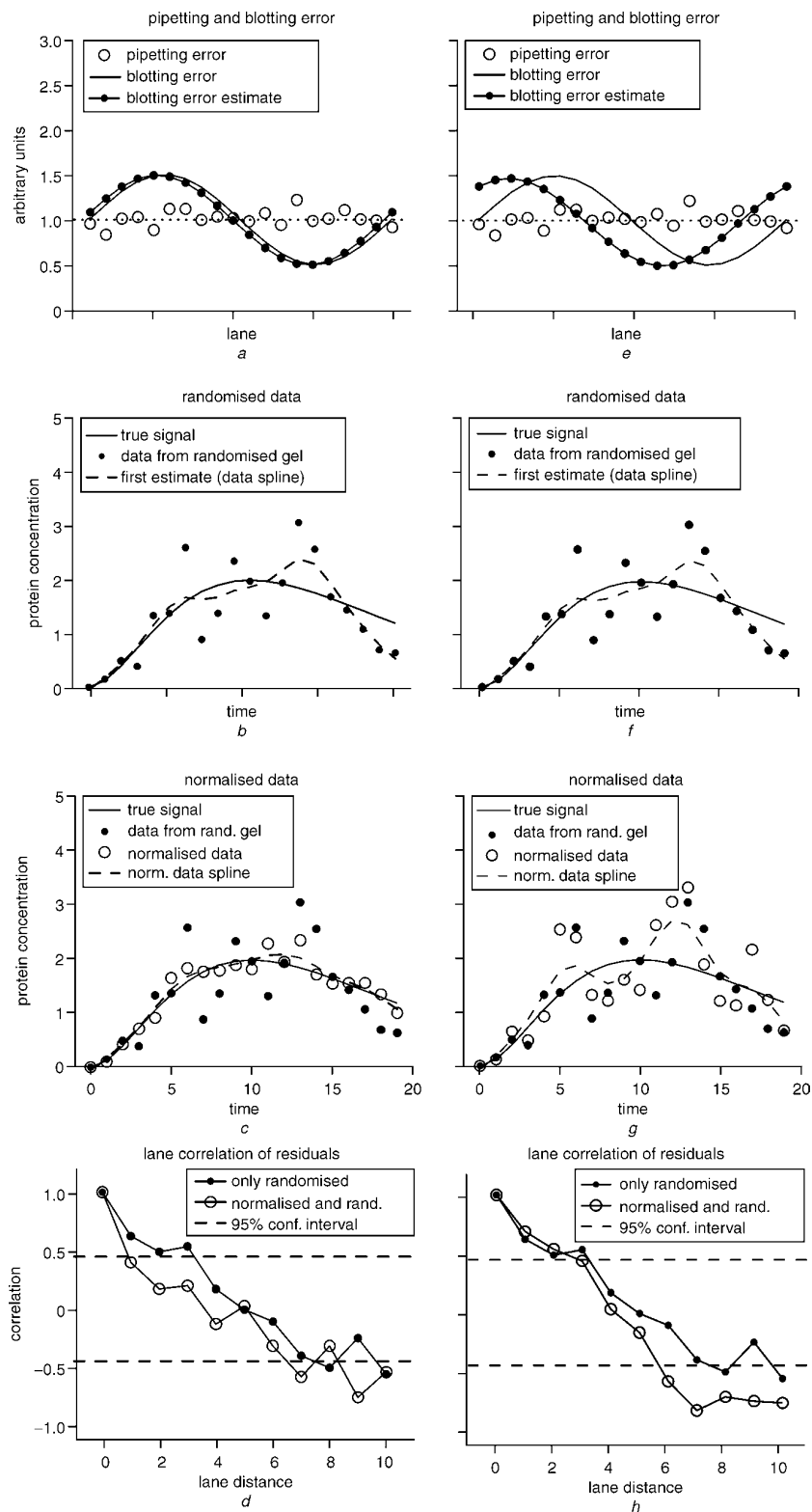


Fig. 3 Normalisation of simulated time course data

Left panel: valid procedure according to our criteria

a Blotting error is well estimated corresponding to suitable normalisation protein

b Perturbation of simulated signal

c Perturbation is strongly reduced after normalisation

d Correlation in gel domain of residuals is improved

Autocorrelation, i.e. correlation in time domain, agrees for both randomised signals with white noise (not shown)

Right panel: procedure rejected according to our criteria

e Blotting error estimate is phase shifted, corresponding to too distant normalisation protein

f Perturbation of simulated signal

g Perturbation cannot be reduced with normalisation

h Lane-to-lane correlation is not improved

to control the pipetting error in experiments is sampling the same number of cells for each time point or measuring and adjusting total protein concentration.

4 Criteria for employing calibrators and normalisers further to improve quantitative immunoblotting data

Calibrators and normalisers possess a constant concentration. Fluctuations occur only as measurement errors. As the blotting error changes gradually from lane to lane, and other errors such as the pipetting error are rather uncorrelated, the blotting error can be estimated by smoothing of the calibrator or normaliser signal, e.g. with a smoothing spline [3]. The smoothing is carried out with a cubic

spline approximation, the smoothness being determined by generalised cross-validation. Based on this blotting error estimate, the protein of interest can be normalised by division. However, as the blotting error is a local property of the gel, normalisers and calibrators with a similar molecular weight to that of the protein of interest are required. If the position of the normaliser is too distant on the blot, or its signal is too noisy, the normalisation procedure can even be detrimental to the data. We therefore developed criteria for employing normalisers and calibrators. The following discussion applies equally well to normalisers and calibrators.

Figure 3a shows a simulated blotting error and a good estimation, corresponding to the smoothed signal of an appropriate normaliser in a real experiment. Smoothing

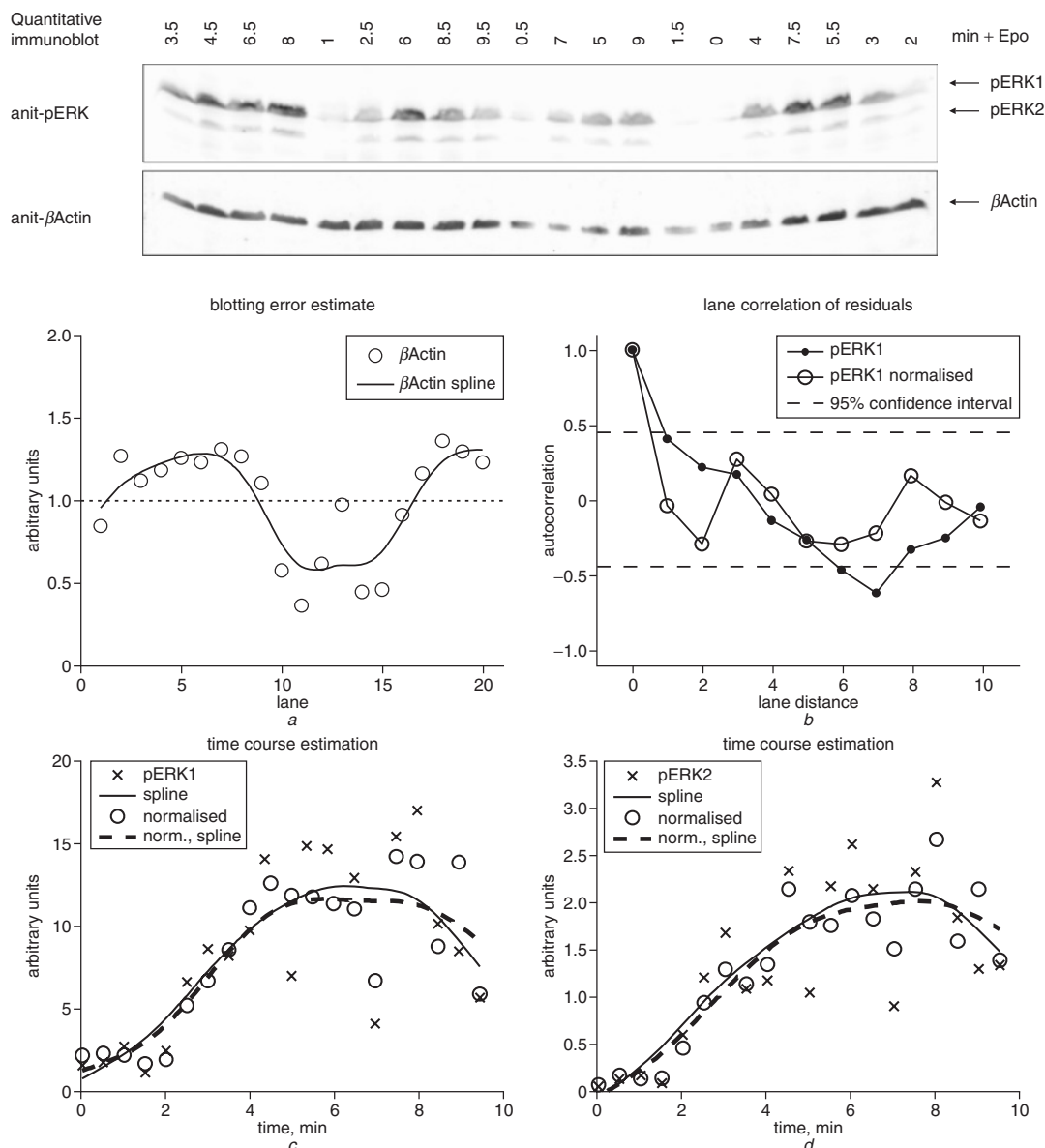


Fig. 4 Randomisation and normalisation of erythropoietin-induced time course experiment

BaF3-EpoR cells are stimulated with 50 units ml^{-1} Epo resulting in ERK phosphorylation. Gel electrophoresis has been performed with randomised, non-chronological gel loading with β Actin as normaliser protein (upper panel)

a Smoothed measurements of β Actin serve as estimate of strong, sine-like blotting error

b Normalisation destroys autocorrelation

Normalisation significantly reduces standard deviation of pERK1/2 measurements compared with spline-smoothed pERK1/2 signal (solid line), which serves as first estimate of true signal

c pERK1

d pERK2

the processed randomised signal leads to an acceptable estimation of the true signal (Fig. 3b). Smoothing the normalised signal yields virtually the true signal itself, as shown in Fig. 3c. Even the correlation structure of the estimation error in the gel domain is improved (Fig. 3d). The estimation of the blotting error displayed in Fig. 3e is inaccurate: a strong phase shift can be observed, corresponding to a normaliser measurement of a different molecular weight from the protein of interest. The blotting error can be simulated by a sine wave surface plot covering the entire blot area, resulting in a phase shift of the blotting error estimate when moving vertically away from the position of the protein of interest. In this situation, normalising the data increases the deviation of the estimated signal from the true one (Fig. 3g). Hence, a criterion for whether normalisation is applicable would be a decreased standard deviation of the estimated signal. Unfortunately, this requires knowledge of the true curve, which is not given. Instead, the smoothing spline curve of the randomised but not yet normalised signal is used as first estimate of the true signal. If the normalisation procedure is valid, a new spline curve can be calculated based on the randomised and normalised data; otherwise, the former values are kept. The error reduction factor by normalisation is defined as

$$\text{error reduction factor by normalisation} = \frac{\text{standard deviation of data from first estimate}}{\text{standard deviation of normalised data from first estimate}}$$

The shown simulated data sets possess the following standard deviations:

(a) Case 1 (Figs. 3a–d):

- standard deviation of the data from the true signal: 0.533
- standard deviation of the data from first estimate: 0.722
- standard deviation of the normalised data from the true signal: 0.157
- standard deviation of the normalised data from first estimate: 0.515
- error reduction factor: $0.515/0.722 = 0.7133$

(b) Case 2 (Figs. 3e–h):

- standard deviation of the data from the true signal: 0.533
- standard deviation of the data from first estimate: 0.722
- standard deviation of the normalised data from the true signal: 1.208
- standard deviation of the normalised data from first estimate: 1.068
- error reduction factor: $1.068/0.722 = 1.4792$

The estimated standard deviation improves in case 1 (error reduction factor <1) but becomes worse in case 2 (error reduction factor >1). Hence, the normalisation procedure is only applicable in the first case, reducing the true standard deviation from 0.533 to 0.157, i.e. to 30%. This procedure works robustly for normalisers and calibrators, as long as the immunoblot is randomised.

5 Application to experimental time course

The randomisation and normalisation procedure was applied on an erythropoietin-induced time course experiment resulting in phosphorylation of ERK1 and ERK2. Samples were loaded in a randomised order and separated on a 17.5% SDS polyacrylamide gel. Membranes were developed with chemiluminescent substrates and quantified with a CCD camera (Fig. 4). We calculated the standard

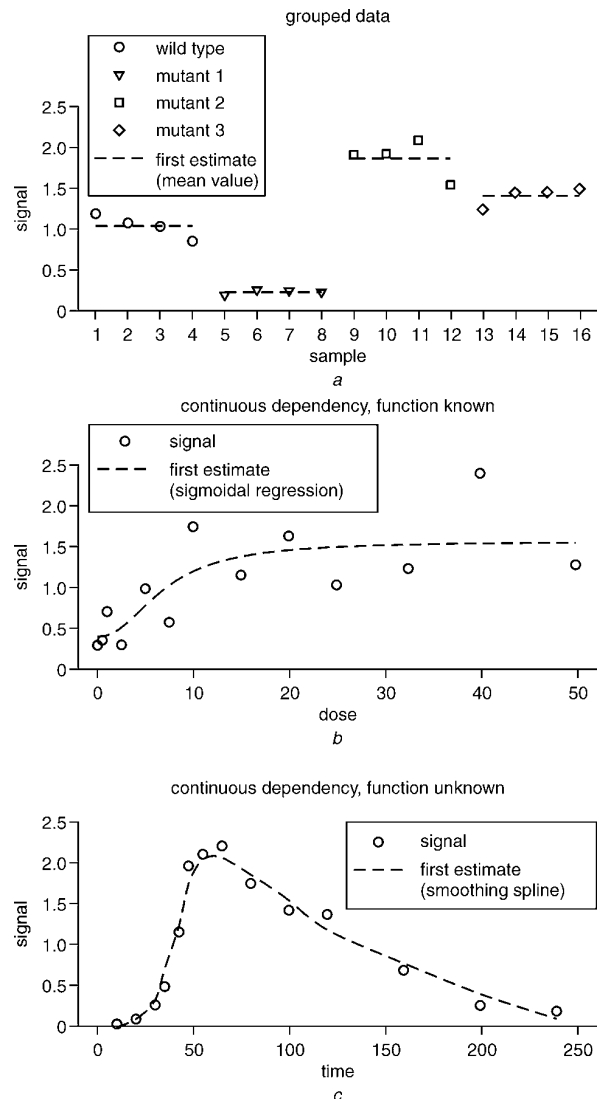


Fig. 5 First estimates used for criteria-mediated error reduction

a In grouped data, such as mutant to wild type comparisons, first estimate is calculated as mean value of samples loaded in replicates
b If known continuous dependency between data points exists, first estimate is calculated as regression function; for example, sigmoidal regression estimates dose-response experiment
c In cases where function is unknown, including time course experiments, first estimate consists of smoothing spline
 These are artificial data for illustration purposes

deviation of the signals from their spline approximation as 2.524 for pERK1 and 0.455 for pERK2. Normalisation with β Actin reduced the standard deviation to the spline approximation to 1.878 (74%) for pERK1 and 0.262 (58%) for pERK2. The reduced lane correlation for the normalised data confirms the quality of data processing. In this case, the correlation structure of the systematic blotting error could be disrupted, thus validating the normalisation.

6 First estimates used for criteria-mediated error reduction

The proposed criterion for error reduction needs a first estimate, which is compared with the measured and the normalised data, respectively. Above, we discussed time course data, where a smoothing spline adequately describes

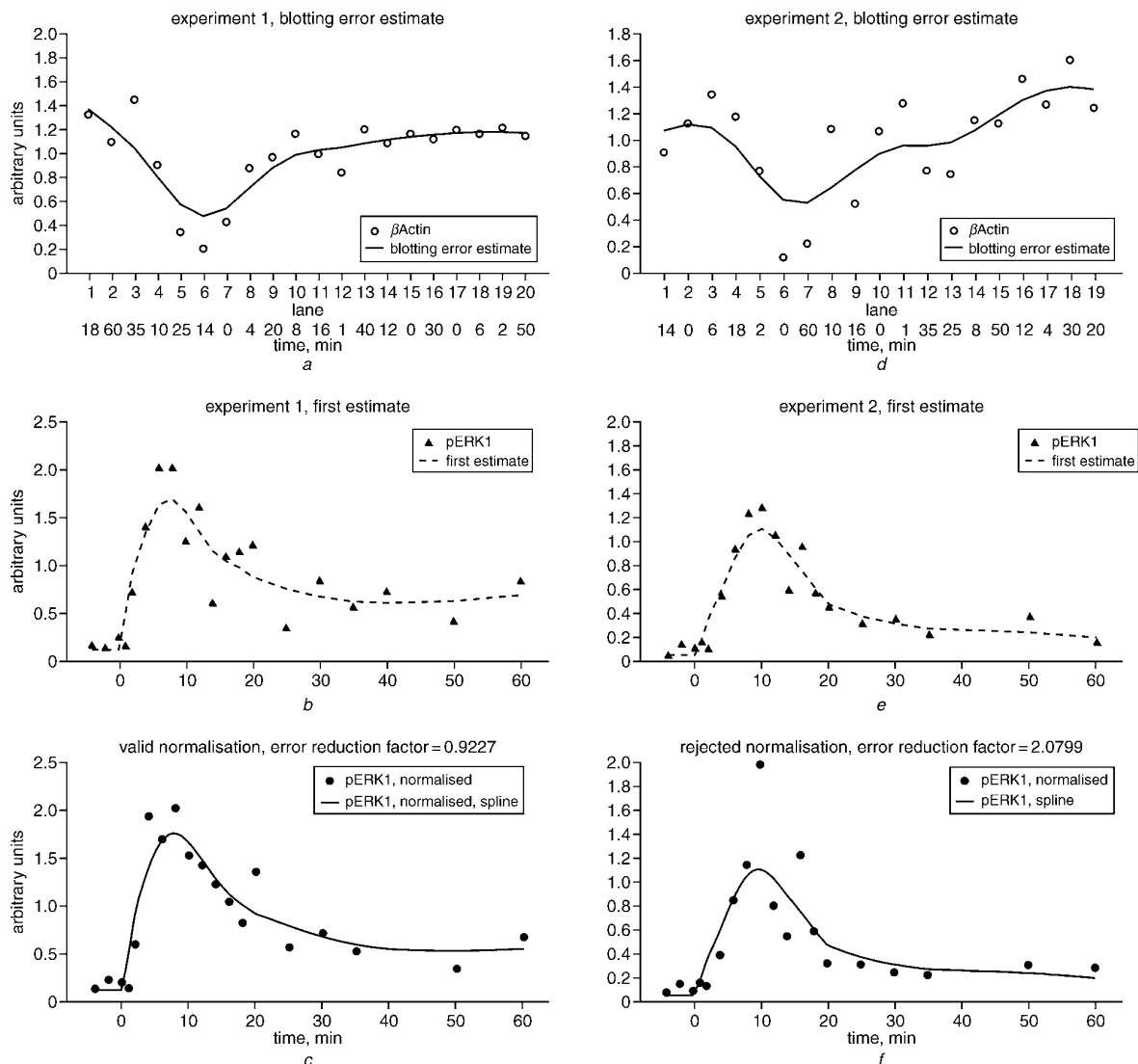


Fig. 6 Normalisation of experimental quantitative immunoblotting data

Left panel: valid procedure according to our criteria

a Data points of β Actin and smoothing spline serving as blotting error estimate

b Smoothing spline (dashed line) serves as first estimate for data points of pERK1, which are normalised using blotting error estimate shown in *a*

c As normalised values are closer to first estimate (error reduction factor = 0.9227), normalised values are kept, and new smoothing spline is calculated (solid line)

Right panel: procedure rejected according to our criteria

d Data points of β Actin and smoothing spline serving as blotting error estimate

e Smoothing spline (solid line) serves as first estimate for data points of pERK2, which are normalised using blotting error estimate shown in *d*

f As normalised values are not closer to first estimate (error reduction factor = 2.0799), normalisation is rejected, and original data are retained

the unknown functional dependency between samples. However, our procedures can also be applied to other experimental settings. Therefore we developed three categories for first estimates, as shown in Fig. 5

(i) In grouped data, as in mutant-to-wild type comparisons, the first estimate is the mean value of replicates.

(ii) In experiments with a known continuous functional dependency between time points and known function, such as dose response assays, the first estimate is calculated by a regression function.

(iii) For experiments with an unknown continuous functional dependency between time points, including time course analysis, the first estimate is represented by a smoothing spline.

Using this approach, we are able to process data derived from any immunoblotting experiment robustly. Care has

to be taken if a regression is used as a first estimate. It is important to show in advance that the signal behaves as expected to prevent incorrect use of the criterion. If there is uncertainty, a smoothing spline might be more appropriate, the only prerequisite being a smooth signal behaviour.

7 Data processing of time course experiments in murine cell lines

To analyse the robustness of our data processing methods, we applied the standard operating procedures and data normalisation criteria to a large, yet noisy, data set obtained by quantitative immunoblotting of mouse cellular lysates. The murine BaF3 cell line was transfected with five different HA-tagged EpoR variants and stimulated for 1 h with 50 units ml^{-1} Epo. A total of 20 samples were taken at regular intervals so that the changes in phosphorylation and

total amount of four different proteins (the Epo receptor (EpoR), Janus kinase 2 (JAK2) and extracellular regulated kinase 1 (ERK1) and 2 (ERK2)) could be followed. EpoR and JAK2 were enriched by immunoprecipitation, and ERK1 and ERK2 were analysed from total cellular lysates. SDS gel electrophoresis was performed with randomised sample loading, and phosphorylation levels were measured using quantitative immunoblotting with phospho-specific antibodies. Furthermore, total amounts of these proteins were determined by reprobings of the membranes with the respective antibodies. Calibrators of EpoR and JAK2 added to the lysates prior to immunoprecipitation were quantified in parallel. The membranes used for measuring ERK1 and ERK2 were reprobated once more, and β Actin was measured, serving as normaliser for these proteins.

After quantification of all proteins with the LumiImager system, we applied criteria-mediated normalisation procedures to the data. As the data generated were very noisy, we tested whether our methods could improve their quality. We calculated the error reduction factor of the standard deviation of the values compared with a smoothing spline, which served as first estimate. In 33% of all cases, the error reduction factor was smaller than 1, resulting in a valid normalisation procedure. In all other instances, the error reduction factor was larger than 1. In these cases, normalisation was rejected and the original data were retained.

Figure 6 shows two examples of criteria-mediated error reduction. In the left panel, the normalisation was valid, with the corrected values closely following the smoothing spline (Fig. 6c). In the right panel, normalisation was rejected, and the original data were kept. The most common reason for rejected normalisation was poor quality of the normaliser or calibrator measurement, as evidenced by the deviation of data points of the normaliser β Actin from the blotting error estimate (Fig. 6d). However, as immunoblotting was performed by randomised gel loading, no misleading effects occurred that could lead to false interpretation. For visual investigation, the numbers were approximated with a smoothing spline, displaying convincing time course dynamics (Fig. 6f). By comparing the data obtained with previous experiments, we could show that our data processing methods substantially improved the reliability of measurements in an unbiased manner.

8 Conclusions

We show that, by the suggested data processing procedures, the standard deviations of data generated by quantitative immunoblotting can be decreased to approximately 55%, thereby increasing data quality substantially. This is largely independent of the experimental setting and quality of the immunoblotting procedure. In contrast to traditional chronological sample loading, we demonstrate the benefit of randomised immunoblots. Randomisation is most useful if the signal and blotting error are in the same frequency range. As the number of measurements is often highly

limited in biochemical experiments, the sampling of time courses is rather coarse. This leads to similar frequencies of the signal and the blotting error. By quantification and plotting of the signal intensity against time, the same information is obtained as with chronological immunoblots; however, as the correlation between neighbouring lanes is disrupted, the standard deviation of the smoothed signal is reduced more than two-fold, leading to data of higher quality and therefore to fewer experiments being necessary for novel biological insight.

Furthermore we demonstrate that calibrators and normalisers allow for correction of immunoblot data, provided there is a good estimate of the blotting error. By generating criteria for data correction, we developed a robust method to enhance data quality further. The application to a large set of experimental data validated our approach. The normalisation criterion not only validates normalisation procedures, but also assesses the quality of the immunoblotting experiment. Valid normalisation criteria and decreased standard deviation to the first estimate are indicators of adequate measurements. If our standard operating procedures for experimental design and data processing are employed, many aspects of quantitative immunoblotting can be automated.

Subtle parameter changes in biological systems can change the state of a cell and trigger the onset of diseases. Therefore quantitative measurements with the highest resolution possible are necessary so that we can understand, predict and interfere with these networks. The limitation of current systems biology is often the lack of data to test the quantitative accuracy of mathematical models, requiring new measurement techniques [4]. With the presented procedure, the established technique of quantitative immunoblotting is developed into a robust and reliable method to generate high-quality quantitative data for systems biology.

9 Acknowledgments

The authors would like to thank Nils Blüthgen for stimulating discussions and Peter J. Nickel and Verena Becker for reading the manuscript critically. We would also like to thank the nine undergraduate students participating in the practical course ‘Components and mechanisms of signal transduction’ for the data analysed in Section 7.

10 References

- 1 Swameye, I., Muller, T.G., Timmer, J., Sandra, O., and Klingmuller, U.: ‘Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by data-based modeling’, *Proc. Natl. Acad. Sci. USA*, 2003, **100**, (3), pp. 1028–1033
- 2 Priestley, M.: ‘Spectral analysis and time series’ (Academic Press, London, 1989)
- 3 Wood, S.N.: ‘Thin plate regression splines’, *J. Royal Statistical Soc. B*, 2003, **65**, (1), pp. 95–114
- 4 Bhalla, U.S., Ram, P.T., and Iyengar, R.: ‘MAP kinase phosphatase as a locus of flexibility in a mitogen-activated protein kinase signaling network’, *Science*, 2002, **297**, (5583), pp. 1018–1023

2.3 Cellular decisions predicted by MAPK modeling

Schilling M, Maiwald T, Hengl S, Timmer J, Klingmüller U.

A data-based MAP-kinase model reveals systems properties and predicts cellular decisions

To be submitted

This publication contains 4 figures. All experiments were initiated and planned by Marcel Schilling. The experiments depicted in Figures 1, 2b, 2c, 3 and 4 were performed by Marcel Schilling. The manuscript was largely written by Marcel Schilling (figures and text).

A data-based MAP-kinase model reveals systems properties and predicts cellular decisions

Marcel Schilling^{1*}, Thomas Maiwald^{2*}, Stefan Hengl², Jens Timmer², and Ursula Klingmüller¹

¹ German Cancer Research Center, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

² Freiburg Center for Data Analysis and Modeling, Eckerstr. 1, University of Freiburg, 79104 Freiburg, Germany

* These authors contributed equally

Running title: Cellular decisions predicted by MAPK modeling

The cells of the hematopoietic system are continuously renewed in a tightly regulated proliferation and differentiation process. Key regulator of the erythroid lineage is the hormone erythropoietin (Epo) that binds to the erythropoietin receptor (EpoR) triggering activation of several signal transduction pathways including the MAP-kinase signaling network. To analyze this network by a systems biology approach we monitored pathway components after stimulation of primary murine erythroid progenitor cells by quantitative immunoblotting. An isoform-specific mathematical model of the signaling network was established and kinetic parameters were estimated by parameter fitting approaches. Sensitivity analysis revealed parameters having the profoundest impact on signal propagation and simulations predicted that overexpression of a single ERK isoform would trigger feedback-mediated rerouting of signaling, which was confirmed by isoform-specific overexpression. To predict the effect of expressing kinase-defective ERK isoforms, the model was extended based on feedback inhibition of SOS by complexation or by phosphorylation. The latter model predicted that expressing kinase-defective ERK isoforms would lead to reduced negative feedback signaling, resulting in similar phenotypes as overexpression of the wild-type isoform. We verified the predictions experimentally by demonstrating that the integrated response of activated ERK directly correlates with accelerated differentiation of primary erythroid progenitor cells. Thus, the developed mathematical model allows for a quantitative analysis of this important signaling network and predicts targets for modulation and intervention.

Introduction

Erythrocytes are continuously renewed from a small number of pluripotent stem cells in a stringently controlled process of differentiation and proliferation. Erythropoietin (Epo) is the key regulator of red blood cell production (Krantz 1991). The importance of Epo signaling has been shown by targeted inactivation of Epo and its cognate receptor (EpoR), both being necessary and irreplaceable for definitive erythropoiesis *in vivo* (Wu et al. 1995; Lin et al. 1996). Erythropoiesis is the process by which multipotent hematopoietic stem cells differentiate into mature, non-nucleated erythrocytes. Hematopoietic stem cells differentiate first into myeloid stem cells and then into CFU-GEMM cells (colony forming unit-granulocyte, erythrocyte, monocyte, megakaryocyte). The cytokines interleukin (IL)-3 and GM-CSF trigger development into BFU-E cells (burst forming unit-erythroid) and eventually CFU-E cells (colony forming unit-erythroid) (Richmond et al. 2005). These cells express the EpoR and depend on Epo for survival and proliferation as well as terminal differentiation (Wu et al. 1995; Lin et al. 1996). Beyond the late basophilic erythroblast stage, the level of EpoR expression declines and the cells differentiate independently of Epo (Koury and Bondurant 1988). The cells then eject their nuclei, entering the reticulocyte stage and terminally differentiate into erythrocytes.

The EpoR is expressed on the plasma membrane as a preformed dimer (Livnah et al. 1999) with its associated tyrosine kinase Janus kinase 2 (JAK2). Upon Epo-binding, JAK2 trans-phosphorylates and activates itself (Witthuhn et al. 1993). Subsequently, JAK2 phosphorylates tyrosine residues in the cytoplasmic tail of the EpoR. These phosphotyrosine residues serve as docking sites for several Src-homology 2 (SH2) domain-containing proteins. Hereby, several important signal transduction networks are activated, triggering survival, proliferation and differentiation of erythroid progenitor cells. The signaling networks responding to Epo include signal transducer and activator of transcription (STAT) 5a/b (Klingmüller et al. 1996), phosphatidylinositol (PI) 3-kinase activating protein kinase B (PKB)/AKT (Klingmüller et al. 1997) and protein kinase C (PKC) ϵ , as well as the mitogen activated protein (MAP)-kinase cascade.

The MAP-kinase pathway is activated upon Epo-stimulation by recruitment of the complex of growth factor receptor-bound protein (Grb) 2 and son of sevenless (SOS) to the receptor. Grb2 can interact with the EpoR directly by binding to Y464 (Barber et al. 1997) or indirectly by binding either to SHIP1 (Mason et al. 2000) or SHP2 (Tauchi et al. 1996). Alternatively, SHC can interact with phosphorylated JAK2 (He et al. 1995), serving as an adaptor protein for Grb2. Furthermore, it was suggested that activation of PKC ϵ by PI 3-

kinase can activate Raf and MEK (Klingmüller et al. 1997) independently of SHC/Grb2. In each case, EpoR activation leads to the recruitment of the guanine-nucleotide exchange factor SOS to the membrane, triggering activation of Ras and Raf, culminating in the activation of the MAP-kinase cascade (Chen et al. 2001). The dual-specific MAP kinase kinases MEK1 and MEK2 are serine phosphorylated by Raf, in turn threonine / tyrosine phosphorylating ERK1 and ERK2. ERK1/2 can dimerize (Khokhlatchev et al. 1998), translocate into the nucleus and activate transcription factors. The role of MAP-kinase activation in primary erythroid cells remains to be identified. Studies analyzing the blood of c-Raf knock-out mice or c-Raf-deficient cells cultivated in serum-free medium supplemented with Epo, stem cell factor, dexamethasone and insulin-like growth factor (Kolbus et al. 2002) suggested that Raf delays differentiation. Furthermore, overexpression of constitutively active Ras blocked terminal erythroid differentiation in erythroid progenitor cells cultivated in medium supplemented with fetal calf serum (Zhang and Lodish 2004). These experiments were performed in the presence of receptor tyrosine kinase (RTK) ligands that strongly activate the MAP-kinase pathway. It is therefore uncertain whether the observed phenotypes are a consequence of EpoR- or RTK-signaling

A pathway is weakly activated if its component kinases are phosphorylated to a low degree, while strong activation occurs if one or more kinases are converted essentially completely to the phosphorylated state. It was shown that weakly activated pathways display higher amplification potential and shorter signal duration (Heinrich et al. 2002). Mathematical modeling of MAP-kinase pathways has provided insight into various aspects of this important signaling network (reviewed in Orton et al. 2005). However, so far only strongly activated receptor tyrosine kinase-dependent MAP-kinase cascades were modeled. Cytokine receptors such as the EpoR are weak activators of the MAP-kinase pathway.

Here, we present the first data-based mathematical model of a weakly activated MAP-kinase signaling network and demonstrate the inherent high amplification and short signal duration properties. We identified the parameters in control of the amplitude and duration of signaling and show how they can be modified. We classified the MAP-kinase network as an independent module and demonstrated the function of the negative feedbacks in the system. And finally we determined whether the amplitude of signaling, the duration or the combination of both is responsible for cellular decisions such as differentiation or proliferation.

Results

Mathematical modeling

We developed a mathematical model of the Epo-induced MAP-kinase signaling network (Fig. 1a). The Epo concentration is used as input function, leading to phosphorylation of JAK2, which in turn phosphorylates cell surface EpoR. Receptor activation leads to recruitment of inactive SHP1 to the membrane. After a delay, membrane-associated SHP1 is activated and dephosphorylates the EpoR and JAK2. We modeled the recruitment of SOS to the membrane to be only dependent on phosphorylated EpoR, treating the receptor complex with SHC and GRB2 as a single compound. Membrane-associated SOS can be released from the membrane or can trigger GDP to GTP exchange of Ras, which subsequently leads to phosphorylation of Raf. Again, we treated this as a single step, represented by Raf (see Supplementary Fig. S3). Subsequently, phosphorylated Raf activates the MAP-kinase phosphorylation cascade. We used a two step phosphorylation and dissociation model for the isoforms MEK2, MEK1, ERK1 and ERK2. Phosphorylation parameters were chosen as specific for the substrate isoform. Phosphatases were assumed to be constant in the time window analyzed (70 min) and unspecific for both isoforms. Additionally, we incorporated a negative feedback loop from activated ERK1/2 to SOS. Several negative feedbacks are proposed in the MAP-kinase signaling network, including inhibitory phosphorylation of Raf or SOS (Buday et al. 1995). In our system the main negative feedback is phosphorylation of SOS, increasing the apparent molecular weight of SOS, which can be seen as a shift in the immunoblot (see Supplementary Fig. S2). Phosphorylation of SOS leads to dissociations of the GRB2-SOS complex from the activated receptor. A constant phosphatase can remove the phosphorylation. Our mathematical model consists of a set of 32 ordinary differential equations (ODE) with 24 parameters and 9 species (Supplementary Fig. S1).

Data acquisition and parameter estimation

Data were acquired using primary murine erythroid cells continuously stimulated with Epo for 70 min and 30 time points were subjected to quantitative immunoblotting (Supplementary Fig. S2). Data for phosphorylated and total EpoR, JAK2, MEK2, MEK1, ERK1, ERK2 and SOS were quantified, normalized using calibrator or normalizer proteins and merged as described (Schilling et al. 2005a). Total levels of the proteins were roughly constant during the time investigated, thus enabling us to use mass conservation in our model. The antibodies used recognize only double-phosphorylated MEK and ERK. We therefore could define 8 observables, notably pEpoR, pJAK2, ppMEK2, ppMEK1, ppERK1,

ppERK2, pSOS and SOS + mSOS, the latter being the fraction of SOS that did not shift in the SDS-PAGE. As the results for the phosphorylated proteins were only relative values, we needed to include a scaling factor for each species, increasing the number of parameters by 5. Total protein levels could be determined for several components. We calculated molecules per cell for surface EpoR by saturation binding assay with [¹²⁵I]-labeled Epo and for the total amount of JAK2 as well as the MEK and ERK isoforms using quantitative immunoblotting (Schilling et al. 2005a and data not shown). Furthermore, the fraction of double-phosphorylated ERK could be determined by analyzing the fraction with an increased apparent molecular weight in the immunoblot. This reduced the parameter space by 7 parameters. The remaining 31 parameters were estimated using PottersWheel (www.potterswheel.de, Maiwald et al., submitted). To reduce the model to an identifiable one, we performed iterative cycles of parameter estimation and detection of dependent parameters (Supplementary Fig. S3). We identified 10 groups of parameter doublets or triplets that were dependent (Supplementary Fig. S4) and fixed the corresponding parameters accordingly. Finally, we succeeded in identifying 21 parameters with a standard deviation of less than 10%. Fig. 1b depicts the immunoblotting data with the corresponding fit trajectory for the 8 observables. χ^2 -values indicate good agreement of the model with the data, as they are comparable to the number of data points. Indeed, we could calculate a p-value of 0.96 that our model is sufficient for describing the data.

Mechanism for fast adaptation to a change in ligand concentrations

The trajectories show that the proteins observed are rapidly activated after stimulation with Epo, reaching basal levels again after 30 – 40 min. Thus, signaling is only transient and the cells show a fast adaptation to a change in ligand concentrations. Interestingly, the activation and deactivation of MEK and ERK describe a much sharper peak than phosphorylation and dephosphorylation at the receptor level. Plotting the trajectories for the variables (Supplementary Fig. S5a) indicates that this effect is caused by the negative feedback of ppERK1/2 to SOS, leading to a very narrow time window for activation of Raf, namely the first 10 min. On the other hand, it takes 20 min before SHP1 is completely activated. This negative feedback reduces phosphorylation levels of EpoR and JAK2. Signal amplification is important to transform small changes at the receptor level into cellular decisions (Sourjik and Berg 2002). We therefore calculated signal amplification as a function of total activated molecules at signal maximum (Supplementary Fig. S5b). As expected, the cytokine receptor EpoR activates the MAP-kinase pathway only weakly, leading to a substantial signal reduction between EpoR and the membrane associated

factors SOS and Raf. The cytoplasmic proteins MEK and ERK amplify this weak signal massively by virtue of the two-step phosphorylation and dephosphorylation mechanisms. Interestingly, total protein concentrations differ only by a factor of 3.5, while the concentrations of activated proteins vary by a factor of 5000. Signal amplification, which is a key feature of most signal transduction systems (Heinrich et al. 2002), is thus achieved at the level of activated molecules.

Identification and modification of parameters in control of signal amplitude and duration

To analyze which reactions in the pathway control activation of the MAP-kinases, we performed a sensitivity analysis. We determined control coefficients for peak amplitude, integrated response (area under the curve), peak time and duration of the activation of double-phosphorylated ERK1 and ERK2. We calculated the control coefficients for the initial protein concentrations (Fig. 2a and Supplementary information, Fig S6) of our model. The initial concentrations were either determined experimentally or by parameter estimation and the values are in physiologically meaningful ranges that are consistent with published concentrations. Initial concentrations of EpoR, SOS and Raf control peak amplitude for ppERK1/2. The level of SHP1 had virtually no control over ppERK1/2. Interestingly, ERK1 had major positive control over the peak amplitude of ppERK1 and minor negative control over ppERK2 while the control coefficients of ERK2 were reciprocal. Control over peak time and duration was small compared to amplitude and integrated response, indicating that the negative feedback wiring of the Epo-induced MAP-kinase pathway prevents sustained signaling. We also calculated control coefficients for the parameters of our model (Supplementary Fig. S6). As some parameters were non-identifiable, we concentrated our analysis on the identifiable ones. Positive control over the peak amplitude and integrated response of ppERK1 and ppERK2 was distributed among the parameters associated with activation of SOS, Raf and ERK. Highest negative control was associated with dephosphorylation of Raf, followed by the first dephosphorylation step of MEK and ERK. Dephosphorylation of Raf and MEK were also the only parameters significantly controlling peak time and duration of ppERK1/2, supporting the theory that phosphatases have the strongest influence on duration of signaling. As can be shown by summation theorems, the sum of all control coefficients for the parameters is equal to 0 for peak amplitude and to -1 for the other quantities analyzed (Hornberg et al. 2005b). The large control coefficients of parameters associated with Raf

phosphorylation and dephosphorylation are consistent with the oncogenic properties of Raf.

The MAP-kinase network acts as an independent module

As the control coefficients were calculated for infinitesimal changes, we tested whether the results are also applicable for larger variations of initial parameter values. Therefore, we performed overexpression simulations using our mathematical model. The effect on the time course of double-phosphorylated ERK1 and ERK2 was analyzed for three-fold elevated concentrations of SHP1, Raf and ERK1 (Fig 2b). As predicted by the sensitivity analysis, increasing the concentration of Raf led to higher amplitudes of both double-phosphorylated ERK1 and ERK2, while elevating the level of SHP1 had virtually no effect on these molecules, indicating that the MAP-kinase module is independently regulated by the negative feedback from ppERK1/2 to SOS. Thus, prolonged receptor activation does not lead to sustained ERK signaling. Furthermore, overexpression of ERK1 resulted in an increase in activated ERK1, but a decrease in activated ERK2. To determine the cause of this observation, we analyzed the change in activation of the signaling proteins (Supplementary Fig. S7a). Overexpression of ERK1 increases the negative feedback on SOS, slightly reducing the activation of SOS and Raf. This change in amplitude is increased by the signal amplification mechanisms and reduces the activation of the downstream molecules MEK and ERK. Thus overexpression of an ERK isoform leads to diversion of the signal towards this isoform by feedback-mediated rerouting. To experimentally validate this finding, we overexpressed c-Raf and ERK1 by retrovirally transducing erythroid progenitor cells. After positive sorting of the cells, overexpression levels were in the same range as the concentrations assumed in the simulations. We analyzed phosphorylation levels after stimulation with Epo for 7 min by quantitative immunoblotting (Supplementary Fig. S7b). Activated molecules per cell at peak time are depicted in Fig. 2c. The levels agree with our model predictions, showing that indeed overexpression of Raf enhances signaling, while overexpression of an ERK isoform leads to feedback-mediated rerouting of signaling.

Function of the negative feedback

As signal rerouting by ERK depends on the negative feedback to SOS, we speculated on the effect of expressing a kinase-defective isoform of ERK. Thus, we extended our model to accommodate expression of the kinase-defective isoforms ERK1 K71R or ERK2 K52R (Robbins et al. 1993). As it is unclear whether complexation of activated ERK1/2 is sufficient for inhibiting SOS activity or if the subsequent phosphorylation is necessary, we

extended our model in two different ways. In the first model (Fig 3a), activated ERK1/2 or kinase-defective ERK1 can bind to membrane-associated SOS, inhibiting signaling. Only kinase-functional ERK1/2 can subsequently catalyze phosphorylation of SOS. The additional parameters k_{on} , k_{off} and k_{cat} were estimated while fixing the remaining parameters of the model to the determined values, resulting in an affinity constant in the micromolar range for the association (Supplementary Fig. S7c). We simulated the trajectories of ppERK1/2 for control cells, cells additionally expressing ERK1 twice the endogenous level of ERK1, cells additionally expressing ERK1 K71R twice the endogenous level of ERK1 as well as cells additionally expressing Raf twice the endogenous level of Raf. Using this information, we calculated the integrated response for ppERK1 and ppERK2. In the first model, the combined integrated response is increased by overexpression of ERK1 but decreased by expression of ERK1 K71R. Overexpression of Raf increased the integrated response even more.

In the second model (Fig 3b), activated ERK1/2 phosphorylates membrane-associated SOS, thereby inhibiting signaling. Kinase-defective ERK1 can associate to and dissociate from membrane-associated SOS without inhibiting its function. Both reactions are dependent on the same kinetic parameter that was used for the phosphorylation reaction. We simulated the trajectories of ppERK1/2 for control cells, cells additionally expressing ERK1 twice the endogenous level of ERK1, cells additionally expressing ERK1 K71R twice the endogenous level of ERK1 as well as cells additionally expressing Raf twice the endogenous level of Raf and calculated the integrated response for ppERK1/2. The combined integrated response is increased both by overexpression of ERK1 and expression of ERK1 K71R. This is due to the reduction of free membrane-associated SOS by activated kinase-defective ERK1 that can be phosphorylated by activated kinase-functional ERK1/2. Overexpression of Raf increased the integrated response to the same extent as in the first model.

Thus, we can distinguish between the two models by analyzing the effect of expressing kinase-defective ERK isoforms. If the first model was correct, expression of kinase-functional or kinase-defective ERK1/2 would result in opposite effects. On the other hand, if the second model was correct, expression of kinase-functional or kinase-defective ERK1/2 would result in similar phenotypes.

The integrated signal response is responsible for cellular decisions

To analyze the effect of hyperactivating the MAP-kinase pathway on proliferation and differentiation of erythroid progenitor cells, we retrovirally transduced primary erythroid

progenitor cells with c-Raf, ERK1 or vector control and measured Epo-dependent proliferation by thymidine incorporation (Fig. 4a). Hyperactivation of the MAP-kinase pathway resulted in strongly reduced proliferation. In accordance with our model prediction, overexpression of Raf had a stronger effect than overexpression of ERK1. To distinguish between the two models suggested, we retrovirally transduced erythroid progenitor cells with c-Raf, ERK1, ERK2, the kinase-defective mutants ERK1 K71R and ERK2 K52R as well as vector control. Cells were cultivated for 48 h in serum free medium supplemented with Epo and differentiation was determined by surface flow cytometry analyzing the fraction of cells expressing low levels of CD71 (transferrin receptor) and high levels of the erythropoiesis marker Ter119 (Fig 3b). Remarkably, overexpression of c-Raf and both wild type and kinase-defective ERK1/2 accelerated differentiation substantially. We next examined hemoglobinization in these cells by intracellular flow cytometry using antibodies against hemoglobin α (Dumitriu et al. 2006). In cells transduced with vector control, hemoglobin expression was greatly increased after cultivation in serum free medium supplemented with Epo (Fig. 3b). However, overexpression of c-Raf or both wild type and kinase-defective ERK constructs significantly reduced hemoglobinization in erythroid progenitor cells treated with Epo. As predicted by the second model, expression of kinase-defective ERK isoforms resulted in similar phenotypes as overexpressing wild-type ERK isoforms. Thus, we can rule out the first model, demonstrating that membrane-associated SOS needs to be phosphorylated to prevent signaling. The phenotypes of differentiation of primary erythroid cells directly correlates with the predicted integrated response of double-phosphorylated ERK1/2, demonstrating the predictive power of the computational model. Thus, hyperactivation of the MAP-kinase pathway leads to accelerated partial differentiation and reduced hemoglobinization of erythroid progenitor cells.

Discussion

We compiled an isoform-specific dynamic model of the Epo-induced signaling network and discovered novel systems properties. A major challenge in mathematical modeling is the frequent discrepancy between the number of parameters and experimentally measured data points. We addressed this issue in several ways. We measured as many data points and proteins as possible during a single stimulation experiment by quantitative immunoblotting, making use of recently developed methodologies for data normalization and integration (Schilling et al. 2005a). By determining initial concentrations of several proteins the parameter space could be decreased. The stoichiometry of signaling components is a vital information that has often been neglected. Finally, by iterative rounds of parameter estimation and identifiability testing, we could determine dependent, non-identifiable parameters. This allowed us to fix non-identifiable parameters and also enabled us to recognize over-parameterization of one step facilitating model reduction. Thus, we were able to create a mathematical model with identifiable parameters describing Epo-induced MAP-kinase activation and feedback-mediated deactivation.

By sensitivity analysis, we discovered that overexpression of an upstream kinase, i.e. Raf, leads to increased signaling, while overexpression of a single isoform, i.e. ERK1, leads to diversion of signaling towards this isoform by feedback-mediated rerouting. When qualitatively analyzing these results, one might speculate that the reason for this phenomenon might be competition of the ERK isoforms for the same upstream kinase, i.e. activated MEK. If this was the case, the sum of activated ERK would stay constant when one isoform was elevated. However, the phosphorylation levels of ERK1 and ERK2 demonstrate that this is not the case, as the total amount of activated ERK1 and ERK2 is increased if ERK1 is overexpressed. We further identified the phosphorylation level of the MAP-kinases as being independent of the expression level of SHP1. The MAP-kinase signaling network therefore acts as an autonomous module with a feedback regulating signal termination even in the case of prolonged receptor activation. Thus, SHP1-dependent phenotypes are most likely independent of Epo-induced MAP-kinase signaling. To qualitatively predict the effect of expressing kinase-defective ERK mutants, we extended the model based on two different hypotheses. Feedback inhibition of SOS was accomplished either by complexation and subsequent phosphorylation or by phosphorylation only. The two models predicted the same outcome for overexpressing a wild-type protein, but divergent effects for expressing a kinase-defective mutant. We therefore hyperactivated the MAP-kinase pathway in primary murine erythroid progenitor

cells by overexpression of signaling components. Expression of kinase-defective ERK resulted in similar phenotypes as overexpression of wild-type ERK, supporting our model of inhibition of membrane-associated SOS by phosphorylation, demonstrating that complexation is insufficient for inhibiting SOS. The notion that expression of a kinase-defective protein would lead to a similar phenotype as overexpression of the wild-type protein is a very counterintuitive result, but can qualitatively and quantitatively be explained by the negative feedback. Several studies have been performed expressing both kinase-defective ERK and constitutively active Ras or Raf. In that case, the negative feedback loop on SOS cannot regulate signaling. This explains the transformation potential of constitutively active Ras and Raf. In healthy cells, hyperactivation of ERK1/2 is directly counteracted by the negative feedback on SOS, reducing signaling. If Ras or Raf are constitutively active, the negative feedback is futile.

Hyperactivation of the MAP-kinase pathway reduced proliferation, accelerated differentiation and reduced hemoglobinization of erythroid progenitor cells, possibly by restricting the time window cells can express the globin genes. Contrary results were obtained with gene knock-out studies, as c-Raf deficient erythroblasts were shown to differentiate faster than their wild-type counterparts and cells expressing a kinase-defective c-Raf differentiated faster in vitro (Kolbus et al. 2002). c-Raf deficient embryos are growth-retarded and anemic and die at midgestation. It was speculated that this anemic phenotype is due to premature erythroblast differentiation at the expense of renewal, depleting the fetal liver of erythroid precursors (Rubiolo et al. 2006). However, in these studies, the differentiation of a heterogeneous population of erythroid progenitor cells was analyzed in the presence of receptor tyrosine kinase ligands that strongly activate the MAP-kinase pathway. In contrast, we analyzed a pure population of CFU-E cells stimulated with Epo as the only growth factor. Using this approach, we could identify the effect of altered Epo-dependent MAP-kinase signaling, demonstrating that hyperactivation of the Epo-induced MAP-kinase pathway in erythroid progenitor cells directly reduces hemoglobinization and accelerates differentiation.

In conclusion, we combined a large set of quantitative data derived from primary murine cells with mathematical modeling, providing a computational model of Epo-induced MAP-kinase signaling. Our systems biology approach provided counterintuitive results that could not be obtained by conventional methods, such as the isoform-specific effects of ERK overexpression, the lack of control of SHP1 on MAP-kinase activation, and the realization that expression of a kinase-defective ERK isoform hyperactivates MAP-kinase signaling by

reducing the negative feedback on SOS. Furthermore, sensitivity analysis provided targets for efficient interventions. We aim to use our computational model in the future for systems-oriented drug design (Kitano 2007), opening new possibilities for treatments of anemia and leukemia.

Materials and methods

Primary Cell Cultures

Colony-forming units erythroid (CFU-E) were prepared from fetal livers of 13.5-day-old embryos from Balb/c mice as described (Ketteler et al. 2002). Briefly, isolated fetal livers were resuspended in ice-cold phosphate-buffered saline (PBS) supplemented with 0.3% bovine serum albumin (BSA) and passed through a 40 μm cell strainer (BD Biosciences, Franklin Lakes, NJ), treated with Red Blood Cell Lysing Buffer (Sigma-Aldrich, St. Louis, MO) to remove erythrocytes and washed by centrifugation through 0.3% BSA/PBS. For negative depletion, fetal liver cells were incubated with rat antibodies against the following surface markers: GR1, CD41, CD11b, CD14, CD45, CD45R/B220, CD4, CD8 (all purchased from BD Pharmingen), Ter119 (gift from Dr Albrecht Müller, Julius-Maximilians-University, Würzburg, Germany) and with the rat monoclonal antibody YBM/42 (gift from Dr S. Watt) for 30 minutes at 4°C. Cells were washed 3 times in 0.3% BSA/PBS and were incubated for 30 minutes at 4°C with anti-rat antibody-coupled magnetic beads and negative sorted with MACS columns according to the manufacturer's instructions (MACSbeads; Miltenyi Biotech, Bergisch-Gladbach, Germany).

CFU-E cells were cultivated for 16 h in Iscove's Modified Dulbecco's Medium, 30% fetal calf serum, 50 μM β -mercaptoethanol supplemented with 0.5 unit/ml Epo.

Time-course Experiments

CFU-E cells cultivated for 16 h were starved in Panserin 401 (PAN Biotech, Aidenbach, Germany) supplemented with 1 mg/ml BSA (Sigma-Aldrich, St. Louis, MO) for 1 h and were stimulated with 50 U/ml Epo (Cilag-Jansen, Bad Homburg, Germany). For each time point, 8×10^6 cells were taken from the pool of cells and lysed by the addition of 2 \times Nonidet P-40 lysis buffer, thereby terminating the reaction.

Immunoprecipitation and Quantitative Immunoblotting

For immunoprecipitation, 40 ng of GST-EpoR and 50 ng of GST-JAK2 were added as calibrator to each cytosolic lysate. The lysates were incubated with anti-EpoR antibodies (Santa Cruz, La Jolla, CA) and anti-JAK2 serum (Upstate Millipore, Billerica, MA). For cellular lysates, protein concentrations were measured using BCA assay (Pierce, Rockford, IL). Immunoprecipitated proteins and 50 μg of cellular lysates were loaded in a randomized fashion on a SDS polyacrylamide gel as described (Schilling et al. 2005b), separated by electrophoresis and transferred to PVDF or nitrocellulose membranes.

Proteins were immobilized with Ponceau S solution (Sigma-Aldrich, St. Louis, MO) followed by immunoblotting analysis using the anti-phosphotyrosine monoclonal antibody 4G10 (Upstate Biotechnology, Lake Placid, NY), the anti-doublephosphorylated MEK1/2 antibody (Cell Signaling Technologies, Beverly, MA) or the anti-SOS-1 antibody (Santa Cruz). Antibodies were removed by treating the blots with β -mercaptoethanol and SDS as described (Klingmüller et al. 1995). Reprobes were performed using anti-EpoR antibody (Santa Cruz), anti-JAK2 serum (Upstate), anti-doublephosphorylated p44/42 MAPK, anti-p44/42 MAPK, and anti-MEK1/2 antibodies (all Cell Signaling). For normalization, antibodies against β -actin (Sigma-Aldrich, St. Louis, MO) and Clathrin HC (Santa Cruz) were used. Secondary horseradish peroxidase coupled antibodies (anti-rabbit HRP, anti-goat HRP, protein A HRP) were purchased from Amersham Biosciences, Piscataway, NJ. Immunoblots were incubated with ECL substrate (Amersham) for 1 min and exposed for 10 min on a Lumilmager (Roche Diagnostics, Mannheim, Germany) or with ECL Advance substrate (Amersham) for 2 min and exposed for 1 min on a Lumilmager. For quantifications, LumiAnalyst software (Roche) was used.

Computational Data Processing

Quantitative immunoblotting data was processed using GellInspector software (Schilling et al. 2005a). The following normalizers were used: GST-JAK2 for pJAK2 and JAK2, GST-EpoR for pEpoR and EpoR, β -actin for ppERK1, ppERK2, ERK1 and ERK2, and Clathrin for pSOS and SOS. For first estimates, csaps – splines were used with a smoothness of 0.2 for pJAK2, JAK2, pEpoR and EpoR, 0.5 for ppERK1, ppERK2, ERK1 and ERK2, and 0.3 for pSOS and SOS. As the values for ppMEK1, ppMEK2 and pSOS had a rather high background, the lowest value of the timecourse was subtracted from all data points.

For the calculation of phosphorylation levels after overexpression of c-Raf and ERK1, we scaled the raw values of ppERK1 and ppERK2 to the values of the shifted fraction of ERK1 and ERK2. Data were normalized with the sum of phosphorylated and unphosphorylated ERK2 and scaled to the number of phosphorylated molecules per cell as calculated for control cells.

Mathematical modeling, parameter estimation and simulations

Modeling was performed using the Matlab toolbox PottersWheel (www.potterswheel.de, Maiwald et al., submitted). Parameters were estimated in logarithmic parameter space using a trust region optimization approach. For each fit, 100 iterations were performed with a χ^2 tolerance of 10^{-7} and fit parameters tolerance of 10^{-7} . Using the best fit as starting

value, 1000 fits were performed, each time varying all parameters with a disturbance strength s of 0.4 corresponding to $p_{\text{new}} = p_{\text{old}} \times 10^{(s \times e)}$, with e being normally distributed with variance 1 and mean 0. Initial values for EpoR, JAK2, MEK1, MEK2 were fixed to the values determined by saturation binding or quantitative immunoblotting divided by 10^4 . To determine the scaling factor of ppERK, we scaled the raw values of the 9 strongest ppERK1 and ppERK2 levels to the values of the shifted fraction of ERK1 and ERK2. Then, we calculated the fraction of phosphorylated to total protein for these 9 time points. These values were multiplied with the total concentration of ERK1 and ERK2. The scaling factor was defined as the mean of these values divided by the values of the 9 strongest ppERK1 and ppERK2 signals. Simulations of different input scenarios were performed by temporarily overwriting the driving functions of the model and calculating the derived variables using the driving input designer of PottersWheel. Similarly, the initial values of SHP1, Raf and ERK1 were multiplied three-fold and variables of ppERK1 and ppERK2 were obtained.

Non-identifiability and sensitivity analysis

Sensitivity analysis was applied to investigate relative changes of derived system quantities K as a result of relative changes in parameter values p_i

$$S_{p_i}^K = \frac{p_i}{K} \cdot \frac{\partial K}{\partial p_i} \quad (1)$$

We analyzed the following quantities: Peak amplitude is defined as the concentration at the maximum, integrated response is the area under the curve from stimulation start to the time when the curve drops to 10% of its maximum, peak time is the time at the maximum and duration is the distance between the time when the curve first reaches 10% of its maximum and the time it drops to 10% of its maximum.

Hornberg et al. (2005a) derived summation laws for sensitivities of derived system quantities like signal amplitude, signal duration, and integrated response. The proofs for the summation laws (Hornberg et al. 2005b) can easily be extended to show the existence of summation laws for the system quantities investigated in our approach.

$$\sum_i S_{p_i}^{\text{peak amplitude}} = 0 \quad (2)$$

$$\sum_i S_{p_i}^{\text{integrated response}} = -1 \quad (3)$$

$$\sum_i S_{p_i}^{\text{peak time}} = -1 \quad (4)$$

$$\sum_i S_{p_i}^{\text{duration}} = -1 \quad (5)$$

Most models are non-identifiable, i.e., there exist model parameters that cannot be determined unambiguously. Often, non-identifiability manifests itself in functionally related parameters (see linear relationships, hyperbolas and two-dimensional surfaces in Supplementary Fig. S4). However, sensitivity analysis is a local approach, because derivations are evaluated at a certain point in parameter space (local sensitivity analysis). Thus, without prior knowledge, it can in principle not be determined statistically at which point in parameter space sensitivity analysis has to be performed. To deal with this problem, we take the following approach: The model is fitted N-times to data (N=1000). Each fit yields different estimates for the non-identifiable parameters. Non-identifiabilities are detected by non-parametric bootstrap-based identifiability testing with the mean optimal transformation approach (MOTA, Hengl et al., submitted) and sensitivities (1) are calculated at the actual point along the non-identifiability, here the linear relationships, the hyperbolas and the two-dimensional surfaces. As the derived system variables are invariant to changes along the non-identifiabilities, the results of the sensitivity analysis do not depend on the values we chose for the non-identifiable parameters.

Plasmids and retroviral transduction

To generate retroviral expression vectors, a multiple cloning site with the restriction sites BamHI, PstI, PmlI, EcoRI, NdeI and BclI was introduced into the BamHI EcoRI locus of pMOWS (Ketteler et al. 2002). Furthermore, the puromycin resistance cassette was replaced by HindIII AfeI digestion with the LNGFR cDNA (Miltenyi Biotech, Bergisch-Gladbach, Germany), resulting in the vector pMOWSnrMCS.

The cDNA of ERK1 was cloned into pMOWSnrMCS by digesting pGEX-ERK1 (Klingmüller et al. 1997) with BamHI and EcoRI resulting in pMOWSnr-ERK1. pMOWSnr-ERK1 K71R was created by PCR mutagenesis, changing position 71 from AAG to AGG. The cDNA of ERK2 and ERK2 K52R was cloned into pMOWSnrMCS by PCR amplification of pcDNA3-HA-ERK2 (kind gift of John Blenis, Harvard Medical School) and p3XFLAG-CMV7-ERK2 K52R (kind gift of Melanie Cobb, University of Texas) and digestion with BamHI and NdeI resulting in pMOWSnr-ERK2 and pMOWSnr-ERK2 K52R, respectively. ERK cDNAs are from rat origin, with 100% amino acid identity to mouse. pMOWSnr-cRaf was created by digesting pCMV-cRaf (kind gift of Michael Reth, University of Freiburg) with BamHI EcoRI and subcloning into pMOWSnrMCS digested with BamHI and BclI. c-Raf cDNA is from human origin, with 97% identity to mouse.

Retroviral expression vectors were transiently transfected into Phoenix-eco cells using the Calcium-phosphate method (Ketteler et al. 2002). 6 hours after transfection of Phoenix cells, the medium was changed to Iscove's Modified Dulbecco's Medium containing 50 μ M β -mercaptoethanol and 30% fetal calf serum (IMDM). Twenty-four hours after transfection, the virus-containing supernatant was harvested and filtered through a 0.45- μ m filter. For spin infection, 4.5 ml of supernatant containing 8 μ g/ml Polybren (Sigma-Aldrich) was added to 5×10^6 CFU-E in 500 μ l Panserin 401 (PAN Biotech, Aidenbach, Germany) and centrifuged for 2 h in an Heraeus centrifuge with 2500 rpm at room temperature. The cells were washed three times with Panserin 401 and seeded in Panserin 401 supplemented with 0.5 U/ml Epo and cultured for 16 or 24 h.

MACS selection, thymidine incorporation and FACS analysis

For proliferation assays, transduced CFU-E cells cultivated for 16 h were positive selected using MACSelect LNGFR selection kit (Miltenyi Biotech, Bergisch-Gladbach, Germany) according to the manufacturer's instructions. 1.5×10^4 LNGFR positive cells were cultivated in Panserin 401 supplemented with 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1 and 5 U/ml Epo for 4 h. [3 H]-thymidine was added and cells were cultivated for 20 h. Cells were harvested and incorporated radioactivity was measured using a scintillation counter.

For FACS analysis, transduced CFU-E cells cultivated for 24 h were positive selected using MACSelect LNGFR selection kit (Miltenyi Biotech, Bergisch-Gladbach, Germany) according to the manufacturer's instructions. LNGFR positive cells were cultivated in Panserin 401 for another 24 h. Cells were stained with APC-conjugated anti-Ter119 antibody (eBioscience, San Diego, CA) and a Biotin-conjugated anti-CD71 antibody (BD Pharmingen, Franklin Lakes, NJ) in combination with PerCP-conjugated streptavidin (BD Pharmingen). For hemoglobin content, cells were permeabilized and stained with an anti-hemoglobin α antibody (Santa Cruz, La Jolla, CA) in combination with a FITC-conjugated anti-goat secondary antibody (DakoCytomation, Hamburg, Germany). Gated cells were analyzed for hemoglobin α and CD71/Ter119 surface marker expression by a FACSCalibur (BD Biosciences, Franklin Lakes, NJ) and analyzed with CellQuest Software (BD Biosciences).

Acknowledgements

We thank John Blenis and Melanie Cobb for the generous gifts of ERK2 and ERK2 K52R cDNA. We are grateful to Véronique Lefebvre for providing the protocols for hemoglobinization assays. We also thank Ute Baumann for excellent technical assistance.

References

- Barber, D. L., Corless, C. N., Xia, K., Roberts, T. M. and D'Andrea, A. D. (1997). Erythropoietin activates Raf1 by an Shc-independent pathway in CTLL-EPO-R cells. *Blood* **89**(1): 55-64.
- Buday, L., Warne, P. H. and Downward, J. (1995). Downregulation of the Ras activation pathway by MAP kinase phosphorylation of Sos. *Oncogene* **11**(7): 1327-31.
- Chen, Z., Gibson, T. B., Robinson, F., Silvestro, L., Pearson, G., Xu, B., Wright, A., Vanderbilt, C. and Cobb, M. H. (2001). MAP kinases. *Chem Rev* **101**(8): 2449-76.
- Dumitriu, B., Patrick, M. R., Petschek, J. P., Cherukuri, S., Klingmüller, U., Fox, P. L. and Lefebvre, V. (2006). Sox6 cell-autonomously stimulates erythroid cell survival, proliferation, and terminal maturation and is thereby an important enhancer of definitive erythropoiesis during mouse development. *Blood* **108**(4): 1198-207.
- He, T. C., Jiang, N., Zhuang, H. and Wojchowski, D. M. (1995). Erythropoietin-induced recruitment of Shc via a receptor phosphotyrosine-independent, Jak2-associated pathway. *J Biol Chem* **270**(19): 11055-61.
- Heinrich, R., Neel, B. G. and Rapoport, T. A. (2002). Mathematical models of protein kinase signal transduction. *Mol Cell* **9**(5): 957-70.
- Hornberg, J. J., Binder, B., Bruggeman, F. J., Schoeberl, B., Heinrich, R. and Westerhoff, H. V. (2005a). Control of MAPK signalling: from complexity to what really matters. *Oncogene* **24**(36): 5533-42.
- Hornberg, J. J., Bruggeman, F. J., Binder, B., Geest, C. R., de Vaate, A. J., Lankelma, J., Heinrich, R. and Westerhoff, H. V. (2005b). Principles behind the multifarious control of signal transduction. ERK phosphorylation and kinase/phosphatase control. *Febs J* **272**(1): 244-58.
- Ketteler, R., Glaser, S., Sandra, O., Martens, U. M. and Klingmüller, U. (2002). Enhanced transgene expression in primitive hematopoietic progenitor cells and embryonic stem cells efficiently transduced by optimized retroviral hybrid vectors. *Gene Ther* **9**(8): 477-87.
- Khokhlatchev, A. V., Canagarajah, B., Wilsbacher, J., Robinson, M., Atkinson, M., Goldsmith, E. and Cobb, M. H. (1998). Phosphorylation of the MAP kinase ERK2 promotes its homodimerization and nuclear translocation. *Cell* **93**(4): 605-15.
- Kitano, H. (2007). A robustness-based approach to systems-oriented drug design. *Nat Rev Drug Discov* **6**(3): 202-210.
- Klingmüller, U., Bergelson, S., Hsiao, J. G. and Lodish, H. F. (1996). Multiple tyrosine residues in the cytosolic domain of the erythropoietin receptor promote activation of STAT5. *Proc Natl Acad Sci U S A* **93**(16): 8324-8.

- Klingmüller, U., Lorenz, U., Cantley, L. C., Neel, B. G. and Lodish, H. F. (1995). Specific recruitment of SH-PTP1 to the erythropoietin receptor causes inactivation of JAK2 and termination of proliferative signals. *Cell* **80**(5): 729-38.
- Klingmüller, U., Wu, H., Hsiao, J. G., Toker, A., Duckworth, B. C., Cantley, L. C. and Lodish, H. F. (1997). Identification of a novel pathway important for proliferation and differentiation of primary erythroid progenitors. *Proc Natl Acad Sci U S A* **94**(7): 3016-21.
- Kolbus, A., Pilat, S., Husak, Z., Deiner, E. M., Stengl, G., Beug, H. and Baccarini, M. (2002). Raf-1 antagonizes erythroid differentiation by restraining caspase activation. *J Exp Med* **196**(10): 1347-53.
- Koury, M. J. and Bondurant, M. C. (1988). Maintenance by erythropoietin of viability and maturation of murine erythroid precursor cells. *J Cell Physiol* **137**(1): 65-74.
- Krantz, S. B. (1991). Erythropoietin. *Blood* **77**(3): 419-34.
- Lin, C. S., Lim, S. K., D'Agati, V. and Costantini, F. (1996). Differential effects of an erythropoietin receptor gene disruption on primitive and definitive erythropoiesis. *Genes Dev* **10**(2): 154-64.
- Livnah, O., Stura, E. A., Middleton, S. A., Johnson, D. L., Jolliffe, L. K. and Wilson, I. A. (1999). Crystallographic evidence for preformed dimers of erythropoietin receptor before ligand activation. *Science* **283**(5404): 987-90.
- Mason, J. M., Beattie, B. K., Liu, Q., Dumont, D. J. and Barber, D. L. (2000). The SH2 inositol 5-phosphatase Ship1 is recruited in an SH2-dependent manner to the erythropoietin receptor. *J Biol Chem* **275**(6): 4398-406.
- Orton, R. J., Sturm, O. E., Vyshemirsky, V., Calder, M., Gilbert, D. R. and Kolch, W. (2005). Computational modelling of the receptor-tyrosine-kinase-activated MAPK pathway. *Biochem J* **392**(Pt 2): 249-61.
- Richmond, T. D., Chohan, M. and Barber, D. L. (2005). Turning cells red: signal transduction mediated by erythropoietin. *Trends Cell Biol* **15**(3): 146-55.
- Robbins, D. J., Zhen, E., Owaki, H., Vanderbilt, C. A., Ebert, D., Geppert, T. D. and Cobb, M. H. (1993). Regulation and properties of extracellular signal-regulated protein kinases 1 and 2 in vitro. *J Biol Chem* **268**(7): 5097-106.
- Rubiolo, C., Piazzolla, D., Meissl, K., Beug, H., Huber, J. C., Kolbus, A. and Baccarini, M. (2006). A balance between Raf-1 and Fas expression sets the pace of erythroid differentiation. *Blood* **108**(1): 152-9.
- Schilling, M., Maiwald, T., Bohl, S., Kollmann, M., Kreutz, C., Timmer, J. and Klingmüller, U. (2005a). Computational processing and error reduction strategies for standardized quantitative data in biological networks. *Febs J* **272**(24): 6400-11.
- Schilling, M., Maiwald, T., Bohl, S., Kollmann, M., Kreutz, C., Timmer, J. and Klingmüller, U. (2005b). Quantitative data generation for systems biology: the impact of randomisation, calibrators and normalisers. *Syst Biol (Stevenage)* **152**(4): 193-200.

- Sourjik, V. and Berg, H. C. (2002). Receptor sensitivity in bacterial chemotaxis. *Proc Natl Acad Sci U S A* **99**(1): 123-7.
- Tauchi, T., Damen, J. E., Toyama, K., Feng, G. S., Broxmeyer, H. E. and Krystal, G. (1996). Tyrosine 425 within the activated erythropoietin receptor binds Syp, reduces the erythropoietin required for Syp tyrosine phosphorylation, and promotes mitogenesis. *Blood* **87**(11): 4495-501.
- Witthuhn, B. A., Quelle, F. W., Silvennoinen, O., Yi, T., Tang, B., Miura, O. and Ihle, J. N. (1993). JAK2 associates with the erythropoietin receptor and is tyrosine phosphorylated and activated following stimulation with erythropoietin. *Cell* **74**(2): 227-36.
- Wu, H., Liu, X., Jaenisch, R. and Lodish, H. F. (1995). Generation of committed erythroid BFU-E and CFU-E progenitors does not require erythropoietin or the erythropoietin receptor. *Cell* **83**(1): 59-67.
- Zhang, J. and Lodish, H. F. (2004). Constitutive activation of the MEK/ERK pathway mediates all effects of oncogenic H-ras expression in primary erythroid progenitors. *Blood* **104**(6): 1679-87.

Figure legends

Figure 1: Mathematical modeling of the Erythropoietin-induced MAP-kinase signaling network.

(a) The dynamic signaling network model consists of 34 reactions (solid arrows) with enzymatic (E), mass action (MA) or delay (D) kinetics. Dashed arrows indicate enzymes catalyzing the particular reaction. Erythropoietin (Epo) is used as input function (U1) of the system. (b) Experimental data are depicted with circles and error bars indicating standard deviations calculated with smoothing splines. Trajectories of the fitted model variables are indicated by solid lines. χ^2 -values comparable to the number of data points indicate good agreement of the model with the data. The rapid adaptation of the system to the continuous stimulus is seen at the receptor level and, with an even faster kinetics, at the level of the MAP-kinase cascade.

Figure 2: Prediction and validation of MAP-kinase signaling control by protein levels.

(a) Control coefficients of the model initial concentrations for peak amplitude of ppERK1 and ppERK2 were calculated. While ERK1 has positive control over the peak amplitude of its own double-phosphorylation, it negatively controls double-phosphorylation of ERK2. The reciprocal case applies to ERK2. Raf has positive effect on both isoforms while SHP1 has none.

(b) The effect of three-fold elevated concentrations of SHP1, Raf and ERK1 were simulated *in silico* and trajectories of double-phosphorylated ERK1 and ERK2 were plotted. Overexpression of SHP1 was predicted to have virtually no effect, Raf increased phosphorylation of both isoforms, whereas overexpression of ERK1 increased ERK1 phosphorylation but decreased ERK2 phosphorylation. The peak amplitude at 7 min is indicated by a vertical dashed line. (c) Murine erythroid progenitor cells retrovirally transduced with c-Raf, ERK1 or vector control were stimulated with Epo or left untreated and subjected to quantitative immunoblotting with antibodies specific for double-phosphorylated and total ERK1/2. The data was quantified and converted to phosphorylated molecules per cell. The experimental results agree with the predictions, showing that overexpression of Raf triggers enhanced signaling, while overexpression of an ERK isoform leads to feedback-mediated signal rerouting.

Figure 3: Model extensions predicting the effect of expressing kinase-defective ERK isoforms. (a) To model SOS inhibition by complexation, the reaction mechanism of SOS phosphorylation was extended to contain a binding step and a reaction step. Activated kinase-defective ERK (ppERK1 K71R) can bind to membrane-associated SOS, inhibiting activation of Raf. The trajectories of ppERK1/2 are predicted for control cells as well as cells overexpressing Raf, ERK1 or ERK1 K71R. The combined integrated response was calculated, which was increased for overexpressed ERK1 and reduced for overexpressed ERK1 K71R. (b) To model SOS inhibition by phosphorylation, binding of activated kinase-defective ERK (ppERK1 K71R) to membrane-associated SOS was included in the model. The complex is still able to activate Raf and reduces the level of free membrane-associated SOS that can be phosphorylated. The trajectories of ppERK1/2 are predicted for control cells as well as cells overexpressing Raf, ERK1 or ERK1 K71R. The combined integrated response was calculated, which was increased for both overexpressed ERK1 and overexpressed ERK1 K71R.

Figure 4: Hyperactivation of the MAP-kinase pathway in erythroid progenitor cells leads to accelerated partial differentiation and reduced hemoglobinization. (a) Erythroid progenitor cells were retrovirally transduced, cultivated for 16 h and sorted to achieve a homogeneous population. Cells were incubated with increasing Epo concentrations for 24 h and proliferation was measured by thymidine incorporation. Overexpression of c-Raf and ERK resulted in reduced Epo-dependent proliferation. Error bars represent standard deviations of triplicates. Lines depict sigmoidal regression curves. (b) Erythroid progenitor cells were retrovirally transduced and cultivated for 48 h in serum-free medium supplemented with Epo. Differentiation was determined by analyzing the fraction of erythroid progenitor cells expressing low levels of CD71 and high levels of Ter119 surface marker using flow cytometry. Error bars represent standard deviations of triplicates. Differentiation was strongly accelerated for cells transduced with c-Raf, ERK1/2 as well as kinase-defective ERK1/2 as predicted by the inhibition of SOS by phosphorylation model. (a) Hemoglobinization was determined by analyzing the amount of intercellular hemoglobin α using flow cytometry. Error bars represent standard deviations of triplicates. Cells transduced with c-Raf, ERK1/2 as well as kinase-defective ERK1/2 showed reduced levels of hemoglobinization. Hyperactivation of the MAP-kinase pathway therefore leads to accelerated partial differentiation and reduced hemoglobinization of erythroid progenitor cells.

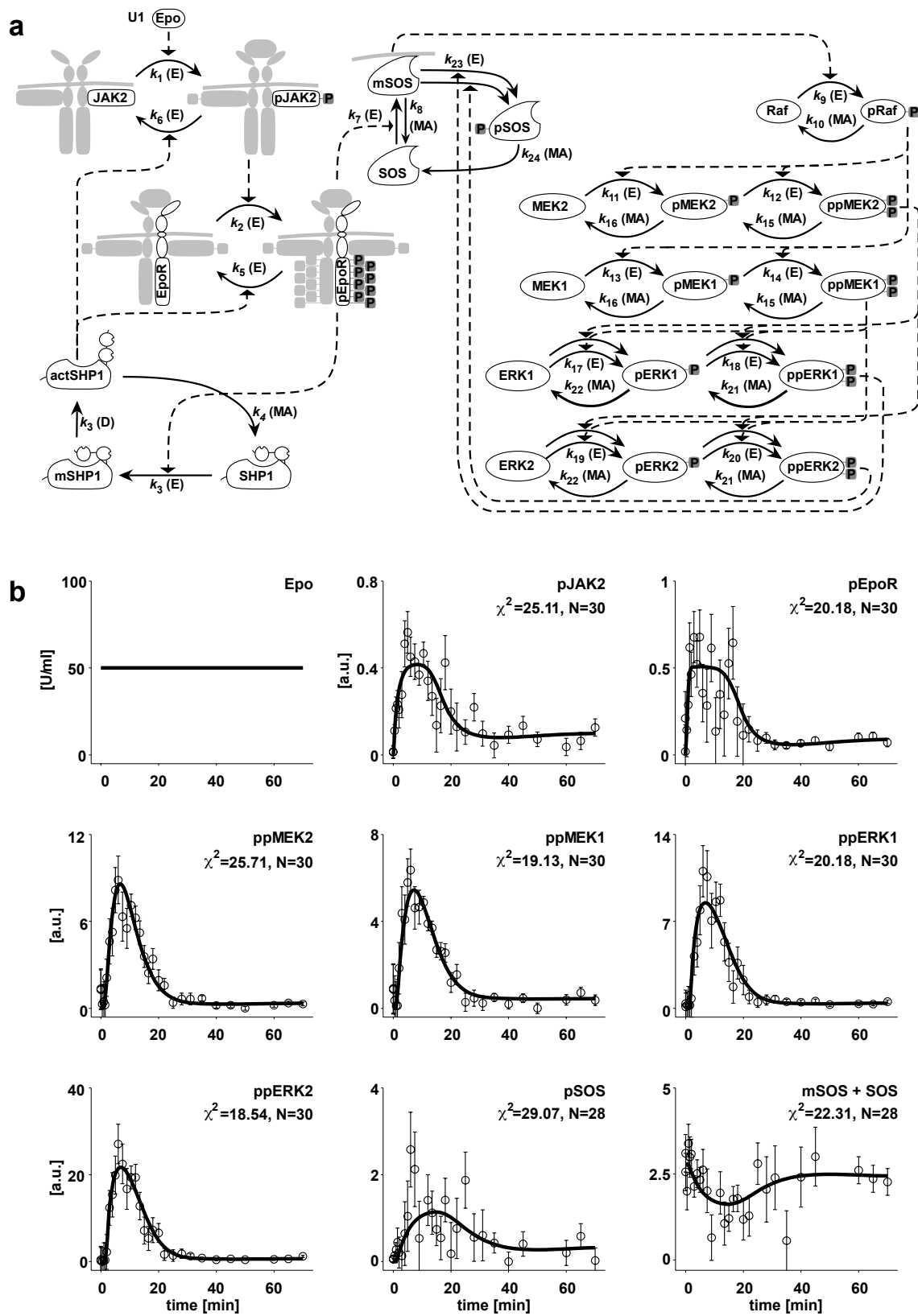


Figure 1, Schilling et al., 2007a

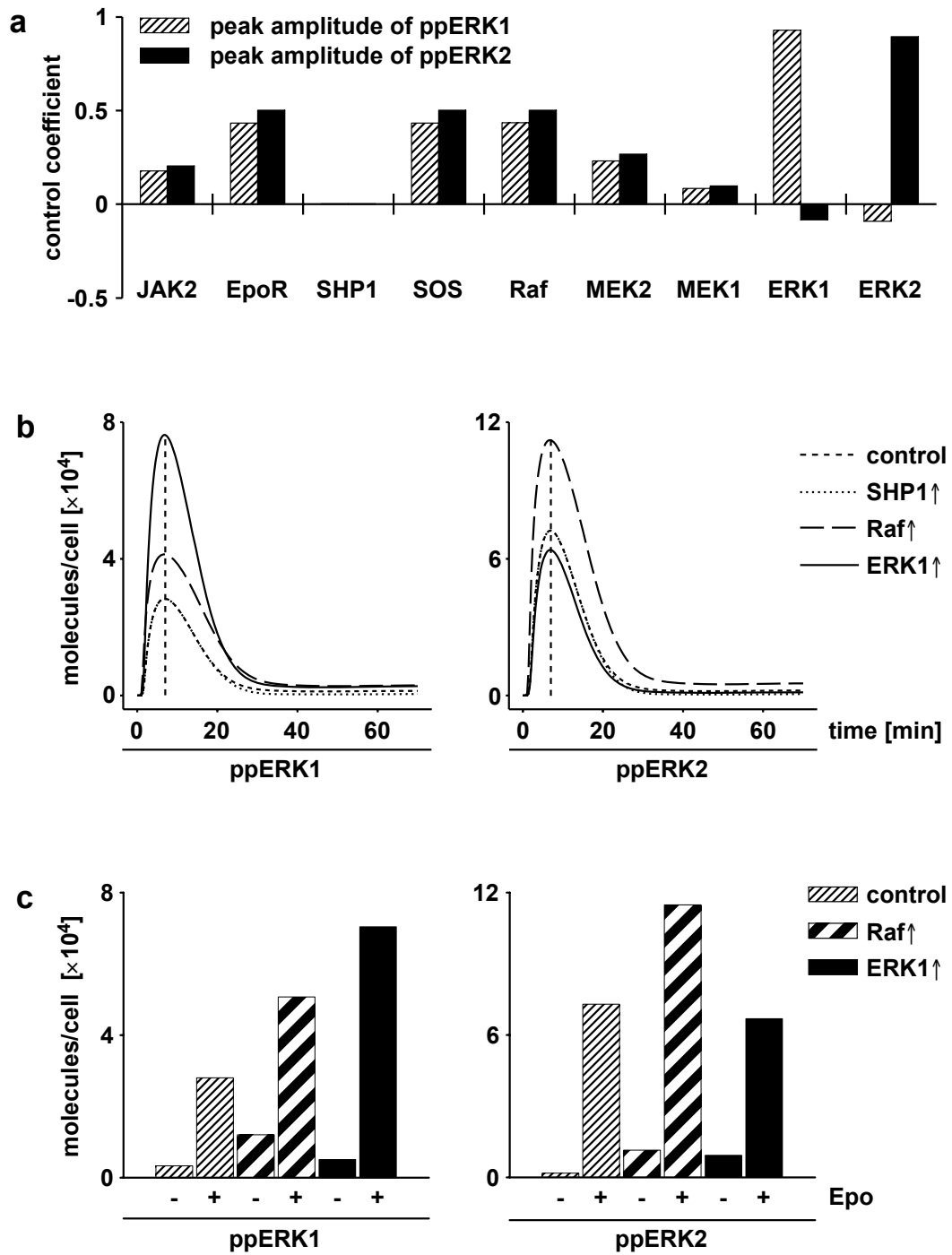


Figure 2, Schilling et al., 2007a

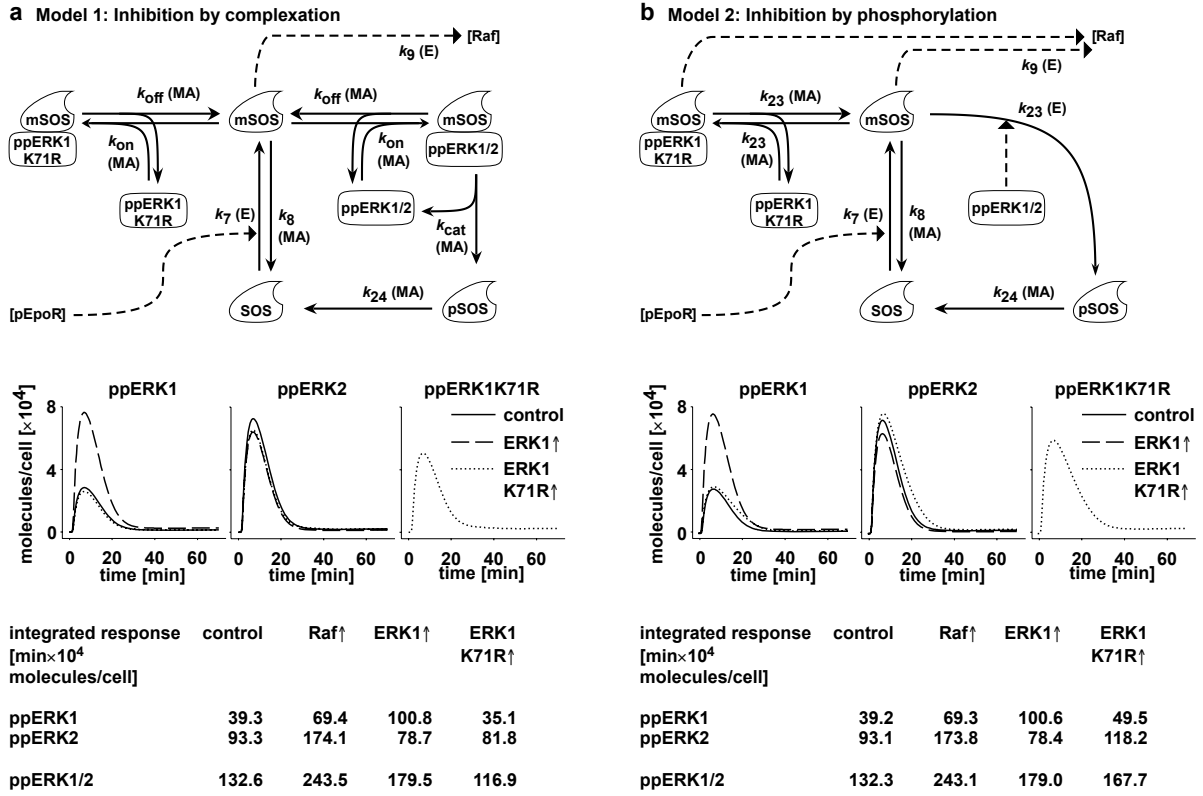


Figure 3, Schilling et al., 2007a

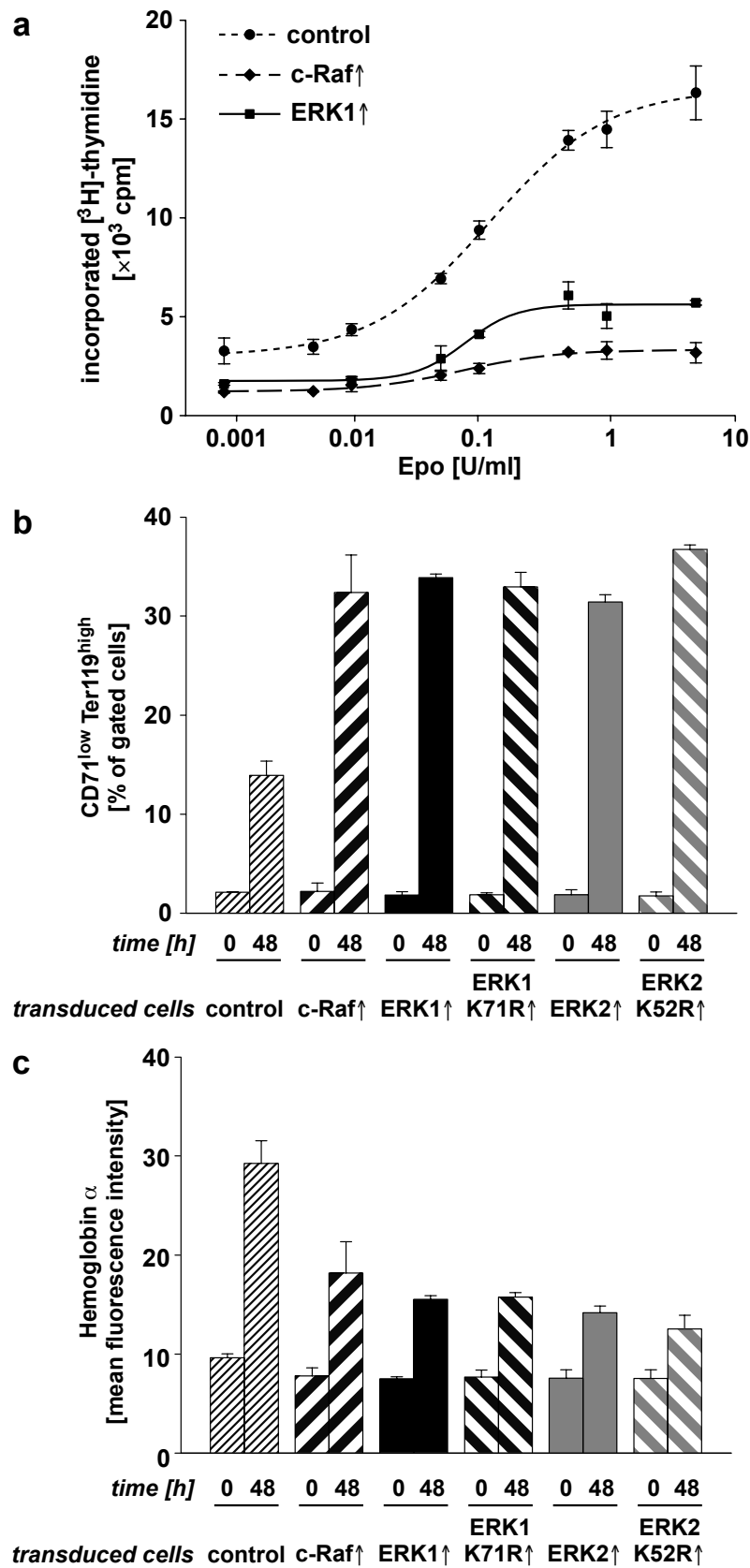


Figure 4, Schilling et al., 2007a

Supplementary figure legends

Figure S1: Ordinary differential equations, parameters, observables and scaling factors for the mathematical model. Ordinary differential equations (ODE) are shown for the mathematical model of the Epo-induced MAP-kinase pathway. Delayed activation of membrane-associated SHP1 was modeled using a 9-step compartmentalization reaction (linear chain trick). Parameters, observables and scaling factors are shown with the respective descriptions.

Figure S2: Quantitative immunoblotting data of primary erythroid progenitor cells stimulated with erythropoietin. Murine erythroid progenitor cells of the CFU-E stage were stimulated with 50 U/ml Epo and samples were taken up to 70 min after addition of Epo. Cellular lysates were either first subjected to immunoprecipitation (IP) or separated directly (TCL) in a randomized order by SDS-PAGE followed by quantitative immunoblotting (IB). GST-JAK2 and GST-EpoR were added prior to IP acting as calibrators. These calibrators as well as the normalizers β -actin and Clathrin HC were used for normalization of data.

Figure S3: Iterative rounds of parameter estimation and identifiability testing result in accurate determination of identifiable parameters. Parameters, initial concentrations and scaling factors are displayed in red for standard deviations larger than 10% and in green for standard deviations smaller than 10%. For each iterative round, 1000 fits were performed and the mean values and standard deviation of each parameter was calculated on the basis of the best 50% of fits. Non-parametric bootstrap-based identifiability testing with the mean optimal transformation approach (MOTA) revealed dependent parameter doublets or triplets. One parameter of each doublet and two parameters of each triplet were fixed to the value of the best fit and parameter estimation was performed again. After restraining the parameter space to 25 parameters, 6 parameters related to Ras and Raf were unidentifiable. Therefore, we condensed the reactions, treating the activation of Raf by membrane-associated SOS as a single step. With two additional round of parameter testing and MOTA, we could identify all of the remaining 21 parameters.

Figure S4: Parameter doublets and triplets analytically dependent as identified with the mean optimal transformation approach (MOTA). Analytically dependent parameters identified by MOTA are shown, with lower case letters corresponding to the dependent parameter doublets or triplets in Supplementary Fig. S3. Data points (black diamonds) represent the estimated parameter values of the best 500 fits of the specific parameter estimation round. The data points describe either a straight line (**c, g**), a hyperbola (**b, f, j**) or a two-dimensional surface (**a, d, e, h, i**). In the first two cases, fixing one parameter identifies the second parameter, in the latter case, two parameters have to be fixed.

Figure S5: Kinetics of the mathematical model variables. (a) Trajectories are depicted for all protein states of the dynamic pathway model. Data are shown for the observation period of 70 min after stimulation with Epo. As the number of molecules per cell were determined for several molecules, parameter estimation revealed the absolute concentration of the rest of the proteins. Thus, absolute phosphorylation levels of all proteins are determined. The activation of the negative feedback proteins activated SHP1 and phosphorylated SOS critically determines the kinetics of the other signaling molecules. (b) The time and the number of activated molecule at peak maximum for pathway components are shown. As expected, the cytokine receptor EpoR activates the MAP-kinase pathway only weakly. Thus, very few molecules at the plasma membrane are involved in signaling. However, the cytoplasmic proteins MEK and ERK amplify the signal impressively.

Figure S6: Sensitivity analysis demonstrates that most parameters control the amplitude of the Epo-induced MAP-kinase pathway. Control coefficients of the model parameters for peak amplitude, integrated response, peak time and duration of ppERK1 and ppERK2 are depicted. Positive control coefficients indicate higher values for the derived systems quantities for increasing parameters, while negative control coefficients indicate decreasing values for the quantities for increasing parameter values. Whereas control of peak amplitude and integrated response is shared by activation and deactivation parameters, peak time and duration is virtually only controlled by dephosphorylation steps. Note the large control of the steps associated with the oncogene Raf. As expected by the summation theorems, the sum for parametric control coefficients equals to 0 for peak amplitude and -1 for the other quantities. Control coefficients for initial concentrations

shows that control is distributed among EpoR, SOS, Raf and ERK1/2 and that the integrated response is mostly determined by the amplitude of signaling.

Figure S7: Feedback-mediated signal rerouting. (a) The first 7 min of the activated molecules are depicted for the dynamic pathway model (red dashed lines) and for three-fold overexpressed ERK1 (green solid line). Elevated ERK1 concentrations enhances the negative feed-back on membrane-associated SOS, but leaves pJAK2 and pEpoR unaffected. The slight reduction in activation of SOS and Raf is propagated and amplified to MEK and ERK, thus reducing ppERK2 levels. (b) Murine erythroid progenitor cells retrovirally transduced with c-Raf, ERK1 or vector control were stimulated with Epo for 7 min or left untreated and subjected to quantitative immunoblotting with antibodies specific for double-phosphorylated and total ERK1/2. Compared to total concentrations of ERK2, overexpression of Raf increased levels of both double-phosphorylated ERK1 and ERK2, while overexpression of ERK1 increased the levels of ppERK1 but reduced the levels of ppERK2. (c) Kinetic parameters determined for the SOS inhibition by complexation model. Rate constants for association, dissociation and phosphorylation reaction were estimated based on quantitative immunoblotting data. The dissociation constant was in the micromolar range.

Figure S8: Flow cytometric analysis of differentiation and hemoglobinization of primary erythroid progenitor cells. Erythroid progenitor cells were retrovirally transduced with ERK1 (green) or empty vector control and cultivated for 48 h in serum-free medium supplemented with Epo. (a) Differentiation was determined by analyzing the fraction of erythroid progenitor cells expressing low levels of CD71 and high levels of Ter119 surface marker using flow cytometry. An overlay of the measurements is shown. (b) Hemoglobinization was determined by analyzing the amount of intercellular hemoglobin α using flow cytometry. An overlay of the measurements is shown.

Ordinary differential equations

$$\begin{aligned}
\text{JAK2:} & \dot{x}_1 = -k_1 \cdot x_1 \cdot u1 + k_6 \cdot x_{10} \cdot x_{13} \\
\text{EpoR:} & \dot{x}_2 = -k_2 \cdot x_2 \cdot x_{10} + k_5 \cdot x_{11} \cdot x_{13} \\
\text{SHP1:} & \dot{x}_3 = -k_3 \cdot x_3 \cdot x_{11} + k_4 \cdot x_{13} \\
\text{SOS:} & \dot{x}_4 = -k_7 \cdot x_4 \cdot x_{11} + k_8 \cdot x_{14} \\
& + k_{24} \cdot x_{20} \\
\text{Raf:} & \dot{x}_5 = -k_9 \cdot x_5 \cdot x_{14} + k_{10} \cdot x_{15} \\
\text{MEK2:} & \dot{x}_6 = -k_{11} \cdot x_6 \cdot x_{15} + k_{16} \cdot x_{21} \\
\text{MEK1:} & \dot{x}_7 = -k_{13} \cdot x_7 \cdot x_{15} + k_{16} \cdot x_{22} \\
\text{ERK1:} & \dot{x}_8 = -k_{17} \cdot x_8 \cdot x_{16} - k_{17} \cdot x_8 \cdot x_{17} \\
& + k_{22} \cdot x_{23} \\
\text{ERK2:} & \dot{x}_9 = -k_{19} \cdot x_9 \cdot x_{16} - k_{19} \cdot x_9 \cdot x_{17} \\
& + k_{22} \cdot x_{24} \\
\text{pJAK2:} & \dot{x}_{10} = +k_1 \cdot x_1 \cdot u1 - k_6 \cdot x_{10} \cdot x_{13} \\
\text{pEpoR:} & \dot{x}_{11} = +k_2 \cdot x_2 \cdot x_{10} - k_5 \cdot x_{11} \cdot x_{13} \\
\text{mSHP1:} & \dot{x}_{12} = +k_3 \cdot x_3 \cdot x_{11} - k_3 \cdot x_{12} \\
\text{actSHP1:} & \dot{x}_{13} = +k_3 \cdot x_{32} - k_4 \cdot x_{13} \\
\text{mSOS:} & \dot{x}_{14} = +k_7 \cdot x_4 \cdot x_{11} - k_8 \cdot x_{14} \\
& - k_{23} \cdot x_{14} \cdot x_{18} - k_{23} \cdot x_{14} \cdot x_{19} \\
\text{pRaf:} & \dot{x}_{15} = +k_9 \cdot x_5 \cdot x_{14} - k_{10} \cdot x_{15} \\
\text{ppMEK2:} & \dot{x}_{16} = +k_{12} \cdot x_{21} \cdot x_{15} - k_{15} \cdot x_{16} \\
\text{ppMEK1:} & \dot{x}_{17} = +k_{14} \cdot x_{22} \cdot x_{15} - k_{15} \cdot x_{17} \\
\text{ppERK1:} & \dot{x}_{18} = +k_{18} \cdot x_{23} \cdot x_{16} + k_{18} \cdot x_{23} \cdot x_{17} \\
& - k_{21} \cdot x_{18} \\
\text{ppERK2:} & \dot{x}_{19} = +k_{20} \cdot x_{24} \cdot x_{16} + k_{20} \cdot x_{24} \cdot x_{17} \\
& - k_{21} \cdot x_{19} \\
\text{pSOS:} & \dot{x}_{20} = +k_{23} \cdot x_{14} \cdot x_{18} + k_{23} \cdot x_{14} \cdot x_{19} \\
& - k_{24} \cdot x_{20} \\
\text{pMEK2:} & \dot{x}_{21} = +k_{11} \cdot x_6 \cdot x_{15} - k_{12} \cdot x_{21} \cdot x_{15} \\
& + k_{15} \cdot x_{16} - k_{16} \cdot x_{21} \\
\text{pMEK1:} & \dot{x}_{22} = +k_{13} \cdot x_7 \cdot x_{15} - k_{14} \cdot x_{22} \cdot x_{15} \\
& + k_{15} \cdot x_{17} - k_{16} \cdot x_{22} \\
\text{pERK1:} & \dot{x}_{23} = +k_{17} \cdot x_8 \cdot x_{16} + k_{17} \cdot x_8 \cdot x_{17} \\
& - k_{18} \cdot x_{23} \cdot x_{16} - k_{18} \cdot x_{23} \cdot x_{17} \\
& + k_{21} \cdot x_{18} - k_{22} \cdot x_{23} \\
\text{pERK2:} & \dot{x}_{24} = +k_{19} \cdot x_9 \cdot x_{16} + k_{19} \cdot x_9 \cdot x_{17} \\
& - k_{20} \cdot x_{24} \cdot x_{16} - k_{20} \cdot x_{24} \cdot x_{17} \\
& + k_{21} \cdot x_{19} - k_{22} \cdot x_{24} \\
\text{mSHP1}_{\text{delay1}}: & \dot{x}_{25} = +k_3 \cdot x_{12} - k_3 \cdot x_{25} \\
\text{mSHP1}_{\text{delay2}}: & \dot{x}_{26} = +k_3 \cdot x_{25} - k_3 \cdot x_{26} \\
\text{mSHP1}_{\text{delay3}}: & \dot{x}_{27} = +k_3 \cdot x_{26} - k_3 \cdot x_{27} \\
\text{mSHP1}_{\text{delay4}}: & \dot{x}_{28} = +k_3 \cdot x_{27} - k_3 \cdot x_{28} \\
\text{mSHP1}_{\text{delay5}}: & \dot{x}_{29} = +k_3 \cdot x_{28} - k_3 \cdot x_{29} \\
\text{mSHP1}_{\text{delay6}}: & \dot{x}_{30} = +k_3 \cdot x_{29} - k_3 \cdot x_{30} \\
\text{mSHP1}_{\text{delay7}}: & \dot{x}_{31} = +k_3 \cdot x_{30} - k_3 \cdot x_{31} \\
\text{mSHP1}_{\text{delay8}}: & \dot{x}_{32} = +k_3 \cdot x_{31} - k_3 \cdot x_{32}
\end{aligned}$$

Parameters

JAK2 phosphorylation by Epo:	k_1
EpoR phosphorylation by pJAK2:	k_2
SHP1 activation by pEpoR:	k_3
actSHP1 deactivation:	k_4
pEpoR dephosphorylation by actSHP1:	k_5
pJAK2 dephosphorylation by actSHP1:	k_6
SOS recruitment by pEpoR:	k_7
mSOS release from membrane:	k_8
mSOS induced Raf phosphorylation:	k_9
pRaf dephosphorylation:	k_{10}
1st MEK2 phosphorylation by pRaf:	k_{11}
2nd MEK2 phosphorylation by pRaf:	k_{12}
1st MEK1 phosphorylation by pRaf:	k_{13}
2nd MEK1 phosphorylation by pRaf:	k_{14}
1st MEK dephosphorylation:	k_{15}
2nd MEK dephosphorylation:	k_{16}
1st ERK1 phosphorylation by ppMEK:	k_{17}
2nd ERK1 phosphorylation by ppMEK:	k_{18}
1st ERK2 phosphorylation by ppMEK:	k_{19}
2nd ERK2 phosphorylation by ppMEK:	k_{20}
1st ERK dephosphorylation:	k_{21}
2nd ERK dephosphorylation:	k_{22}
ppERK neg feedback on mSOS:	k_{23}
pSOS dephosphorylation:	k_{24}

Observables

pEpoR:	$y_1 = s_1 \cdot x_{11}$
pJAK2:	$y_2 = s_2 \cdot x_{10}$
ppMEK2:	$y_3 = s_3 \cdot x_{16}$
ppMEK1:	$y_4 = s_3 \cdot x_{17}$
ppERK1:	$y_5 = s_4 \cdot x_{18}$
ppERK2:	$y_6 = s_4 \cdot x_{19}$
pSOS:	$y_7 = s_5 \cdot x_{20}$
SOS + mSOS:	$y_8 = s_5 \cdot (x_4 + x_{14})$

Scaling factors

scale pEpoR:	s_1
scale pJAK2:	s_2
scale ppMEK:	s_3
scale ppERK:	s_4
scale SOS:	s_5

Figure S1, Schilling et al., 2007a

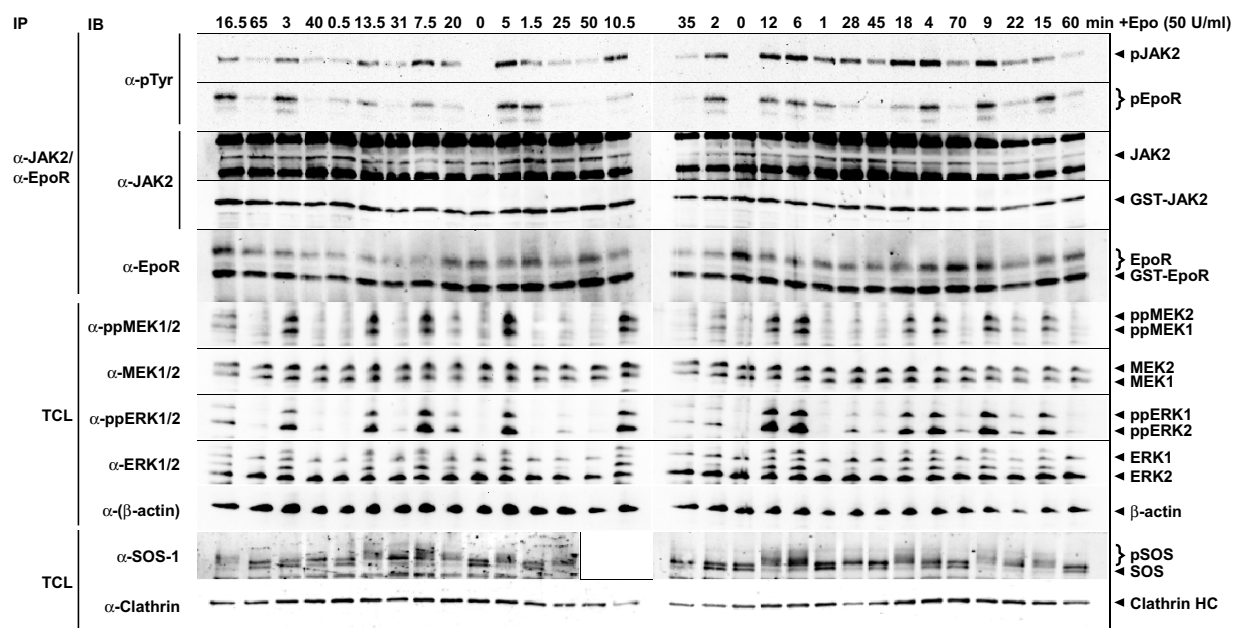


Figure S2, Schilling et al., 2007a

	34 parameters estimated	MOTA	28 parameters estimated	MOTA	25 parameters estimated	22 parameters estimated	MOTA	21 parameters estimated	Parameters
JAK2 phosphorylation by Epo [min ⁻¹ (U/ml)]	0.012 ± 0.000 (3%)		0.012 ± 0.000 (0%)		0.012 ± 0.000 (0%)	0.012 ± 0.000 (1%)		0.012 ± 0.000 (1%)	0.0122229
EpoR phosphorylation by pJAK2 [min ⁻¹ (10 ⁶ molecules ⁻¹)]	3.474 ± 1.655 (48%)	a	2.625 ± 0.015 (1%)		2.632 ± 0.020 (1%)	2.598 ± 0.025 (1%)		2.601 ± 0.026 (1%)	2.57635
SHP1 activation by pEpoR [min ⁻¹ (10 ⁶ molecules ⁻¹)]	0.411 ± 0.011 (3%)		0.415 ± 0.002 (0%)		0.415 ± 0.001 (0%)	0.416 ± 0.001 (0%)		0.416 ± 0.001 (0%)	0.415126
actSHP1 deactivation [min ⁻¹ (10 ⁶ molecules ⁻¹)]	0.020 ± 0.003 (12%)		0.021 ± 0.000 (1%)		0.021 ± 0.000 (1%)	0.020 ± 0.000 (2%)		0.020 ± 0.000 (2%)	0.0202363
pEpoR dephosphorylation by actSHP1 [min ⁻¹ (10 ⁶ molecules ⁻¹)]	1.256 ± 1.589 (127%)	a	fixed		fixed	fixed		fixed	1.30893
pJAK2 dephosphorylation by actSHP1 [min ⁻¹ (10 ⁶ molecules ⁻¹)]	0.351 ± 0.440 (125%)	b	0.470 ± 0.004 (1%)		0.471 ± 0.003 (1%)	0.458 ± 0.004 (1%)		0.457 ± 0.004 (1%)	0.460926
SOS recruitment by pEpoR [min ⁻¹ (10 ⁶ molecules ⁻¹)]	0.100 ± 0.004 (4%)		0.102 ± 0.001 (1%)		0.103 ± 0.001 (1%)	0.099 ± 0.001 (1%)		0.100 ± 0.002 (2%)	0.09871
mSOS release from membrane [min ⁻¹]	18.284 ± 13.500 (74%)		14.379 ± 8.389 (58%)	g	fixed	fixed		fixed	27.6573
Ras GDP exchange by mSOS [min ⁻¹]	25.761 ± 29.323 (114%)		22.374 ± 14.837 (66%)		30.824 ± 22.057 (72%)	removed		removed	-
Ras_GTP hydrolysis [min ⁻¹]	13.765 ± 13.355 (97%)		8.003 ± 2.888 (36%)		5.665 ± 0.736 (13%)	removed		removed	-
GTP_Ras / mSOS induced Raf phosphorylation [min ⁻¹ (10 ⁶ molecules ⁻¹)]	0.032 ± 0.104 (326%)		0.008 ± 0.008 (93%)		0.014 ± 0.010 (70%)	0.653 ± 0.416 (64%)	j	0.172 ± 0.006 (4%)	0.176785
pRaf dephosphorylation [min ⁻¹]	0.274 ± 0.037 (13%)		0.302 ± 0.272 (90%)		0.320 ± 0.371 (116%)	0.254 ± 0.002 (1%)		0.254 ± 0.002 (1%)	0.254708
1st MEK2 phosphorylation by pRaf [min ⁻¹ (10 ⁶ molecules ⁻¹)]	29.464 ± 41.491 (141%)		13.769 ± 9.164 (67%)	h	7.174 ± 0.360 (5%)	10.639 ± 0.697 (7%)		10.643 ± 0.711 (7%)	10.6134
2nd MEK2 phosphorylation by pRaf [min ⁻¹ (10 ⁶ molecules ⁻¹)]	240.093 ± 177.937 (74%)		283.626 ± 195.209 (69%)	i	175.909 ± 10.424 (6%)	114.769 ± 8.573 (7%)		114.743 ± 8.769 (8%)	114.297
1st MEK1 phosphorylation by pRaf [min ⁻¹ (10 ⁶ molecules ⁻¹)]	3.055 ± 3.898 (128%)		1.631 ± 0.925 (57%)	h, i	fixed	fixed		fixed	0.96932
2nd MEK1 phosphorylation by pRaf [min ⁻¹ (10 ⁶ molecules ⁻¹)]	967.853 ± 824.087 (85%)		1135.051 ± 639.626 (56%)	h, i	fixed	fixed		fixed	691.72
1st MEK dephosphorylation [min ⁻¹]	0.265 ± 0.129 (49%)		0.231 ± 0.008 (3%)		0.231 ± 0.003 (2%)	0.250 ± 0.003 (1%)		0.250 ± 0.004 (1%)	0.249181
2nd MEK dephosphorylation [min ⁻¹]	0.068 ± 0.019 (27%)		0.070 ± 0.004 (6%)		0.071 ± 0.001 (1%)	0.068 ± 0.001 (1%)		0.068 ± 0.001 (2%)	0.0674336
1st ERK1 phosphorylation by ppMEK [min ⁻¹ (10 ⁶ molecules ⁻¹)]	141.518 ± 112.735 (80%)		93.226 ± 3.317 (4%)		94.146 ± 3.543 (4%)	62.159 ± 1.337 (2%)		61.993 ± 1.510 (2%)	62.8869
2nd ERK1 phosphorylation by ppMEK [min ⁻¹ (10 ⁶ molecules ⁻¹)]	66.706 ± 58.235 (87%)	c	60.324 ± 0.116 (0%)		60.353 ± 0.136 (0%)	59.677 ± 0.170 (0%)		59.669 ± 0.151 (0%)	59.6234
1st ERK2 phosphorylation by ppMEK [min ⁻¹ (10 ⁶ molecules ⁻¹)]	56.724 ± 45.363 (80%)		41.077 ± 0.990 (2%)	e	41.410 ± 1.088 (3%)	30.403 ± 0.467 (2%)		30.330 ± 0.520 (2%)	30.6693
2nd ERK2 phosphorylation by ppMEK [min ⁻¹ (10 ⁶ molecules ⁻¹)]	54.240 ± 47.600 (88%)	c, d	fixed		fixed	fixed		fixed	48.7469
1st ERK dephosphorylation [min ⁻¹]	39.721 ± 38.468 (97%)	d	24.593 ± 0.080 (0%)		24.622 ± 0.109 (0%)	23.233 ± 0.082 (0%)		23.217 ± 0.093 (0%)	23.2891
2nd ERK dephosphorylation [min ⁻¹]	2.017 ± 2.349 (116%)	e	fixed		fixed	fixed		fixed	1.00149
ppERK neg feedback on mSOS [min ⁻¹ (10 ⁶ molecules ⁻¹)]	1886.884 ± 1626.068 (86%)		1275.571 ± 685.335 (54%)	g	2355.684 ± 85.459 (4%)	3043.851 ± 94.007 (3%)		3024.169 ± 104.049 (3%)	3116.04
pSOS dephosphorylation [min ⁻¹]	0.125 ± 0.006 (5%)		0.126 ± 0.001 (1%)		0.127 ± 0.001 (1%)	0.122 ± 0.001 (1%)		0.122 ± 0.001 (1%)	0.121478
JAK2 [×10 ⁶ molecules]	determined								2
EpoR [×10 ⁶ molecules]	determined								1
SHP1 [×10 ⁶ molecules]	22.805 ± 20.568 (90%)	a, b	fixed		fixed	fixed		fixed	7.8493
SOS [×10 ⁶ molecules]	2.545 ± 2.544 (100%)	f	2.615 ± 0.001 (0%)		2.614 ± 0.001 (0%)	2.619 ± 0.002 (0%)		2.619 ± 0.003 (0%)	2.61825
GDP_Ras [×10 ⁶ molecules]	69.043 ± 30.690 (44%)		67.056 ± 24.072 (36%)		65.126 ± 23.940 (37%)	removed		removed	-
Raf [×10 ⁶ molecules]	0.756 ± 1.061 (140%)		0.220 ± 0.148 (67%)		0.222 ± 0.162 (73%)	1.529 ± 1.075 (70%)	j	fixed	3.98858
MEK2 [×10 ⁶ molecules]	determined								11
MEK1 [×10 ⁶ molecules]	determined								24
ERK1 [×10 ⁶ molecules]	determined								7
ERK2 [×10 ⁶ molecules]	determined								21
scale pEpoR [×10 ⁶ molecules/a.u.]	0.498 ± 0.009 (2%)		0.500 ± 0.001 (0%)		0.500 ± 0.001 (0%)	0.505 ± 0.001 (0%)		0.505 ± 0.001 (0%)	0.504791
scale pJAK2 [×10 ⁶ molecules/a.u.]	0.210 ± 0.002 (1%)		0.210 ± 0.000 (0%)		0.210 ± 0.000 (0%)	0.210 ± 0.000 (0%)		0.210 ± 0.000 (0%)	0.209815
scale ppMEK [×10 ⁶ molecules/a.u.]	38.975 ± 12.672 (33%)	d, e	fixed		fixed	fixed		fixed	49.9407
scale ppERK [×10 ⁶ molecules/a.u.]	determined								3
scale SOS [×10 ⁶ molecules/a.u.]	2.307 ± 2.327 (101%)	f	fixed		fixed	fixed		fixed	1.05574
Best total χ^2 - value	178.329		178.312		178.308	180.229		180.230	
p-value for 'model is sufficient' based on degrees of freedom	0.884		0.933		0.950	0.955		0.960	

Figure S3, Schilling et al., 2007a

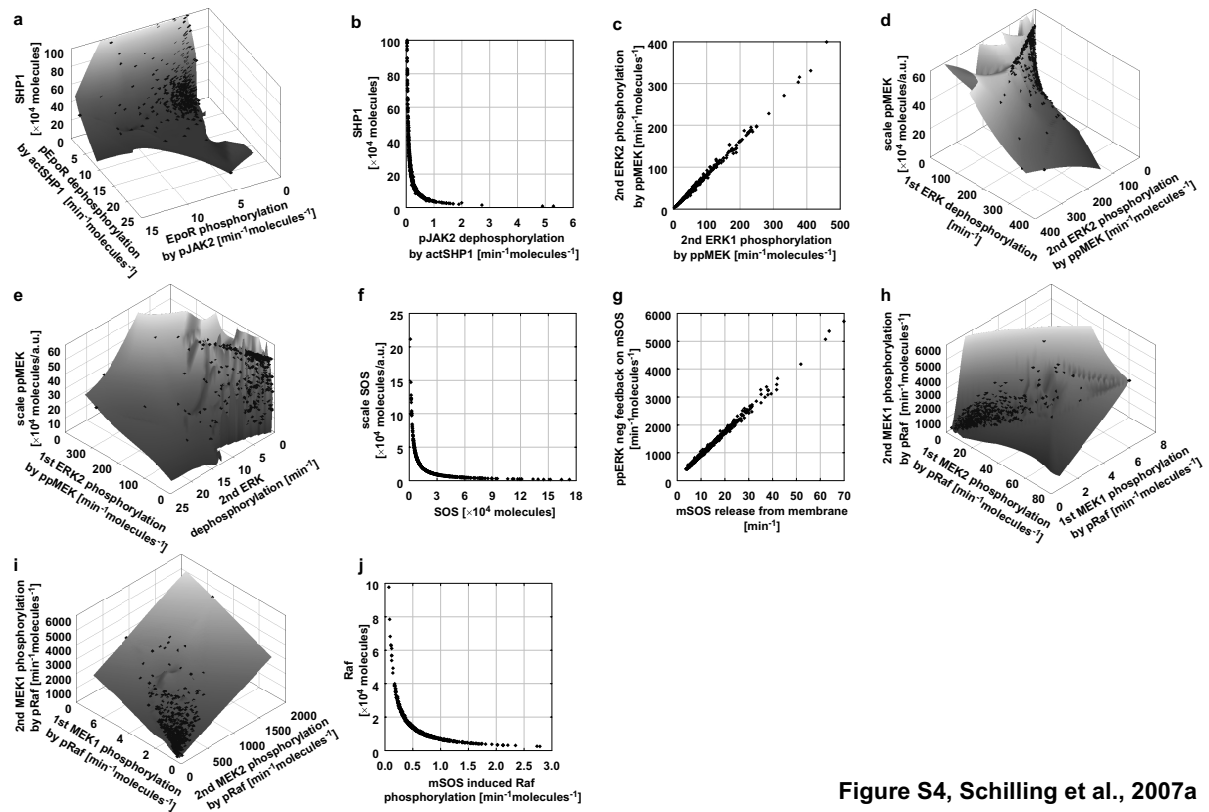
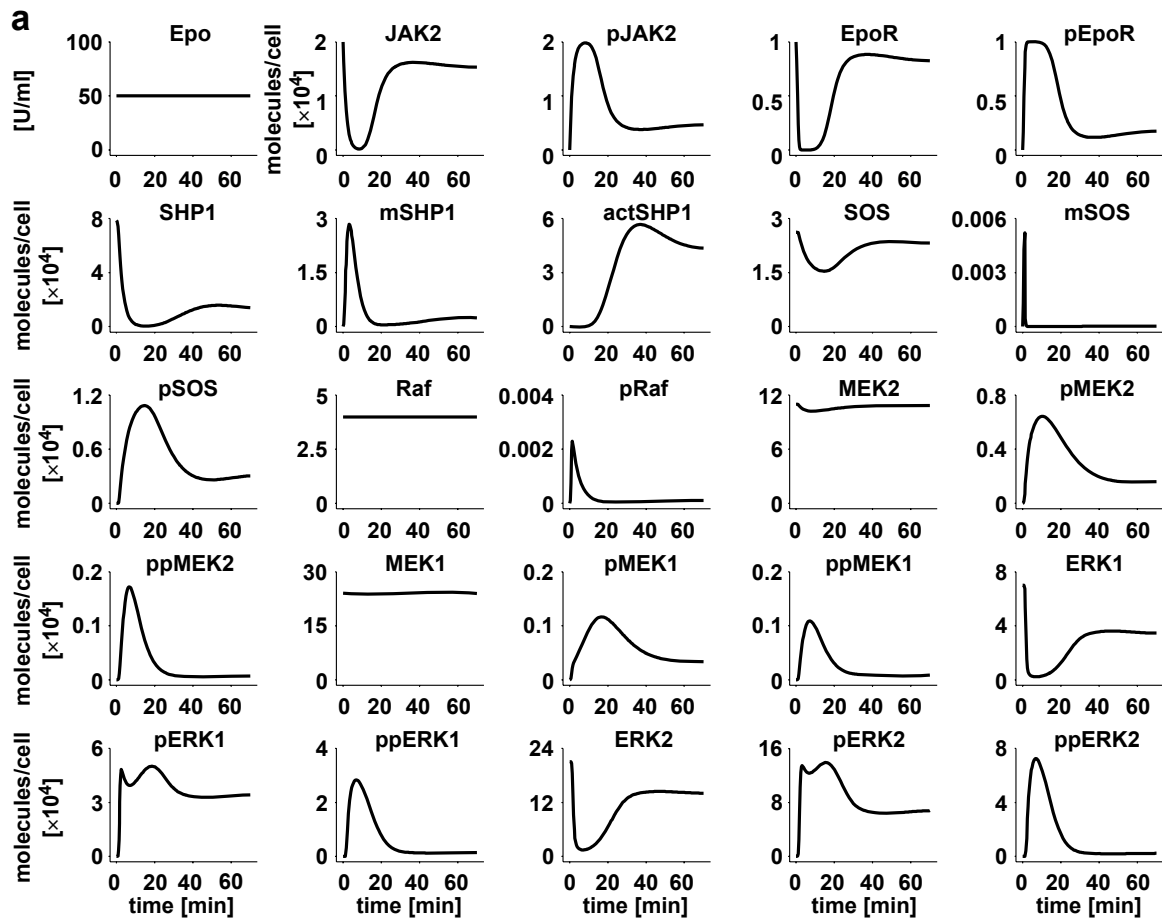


Figure S4, Schilling et al., 2007a



b

	<i>time at maximum</i>	<i>total molecules</i>	<i>activated molecules at maximum</i>	<i>signal amplification</i>
pJAK2	8:27 min	20000	19805	
pEpoR	4:13 min	10000	10000	0.5
mSOS	1:13 min	26183	52	0.005
pRaf	1:24 min	39886	23	0.44
ppMEK1/2	6:52 min	350000	2809	122
ppERK1/2	7:02 min	280000	100703	36

Figure S5, Schilling et al., 2007a

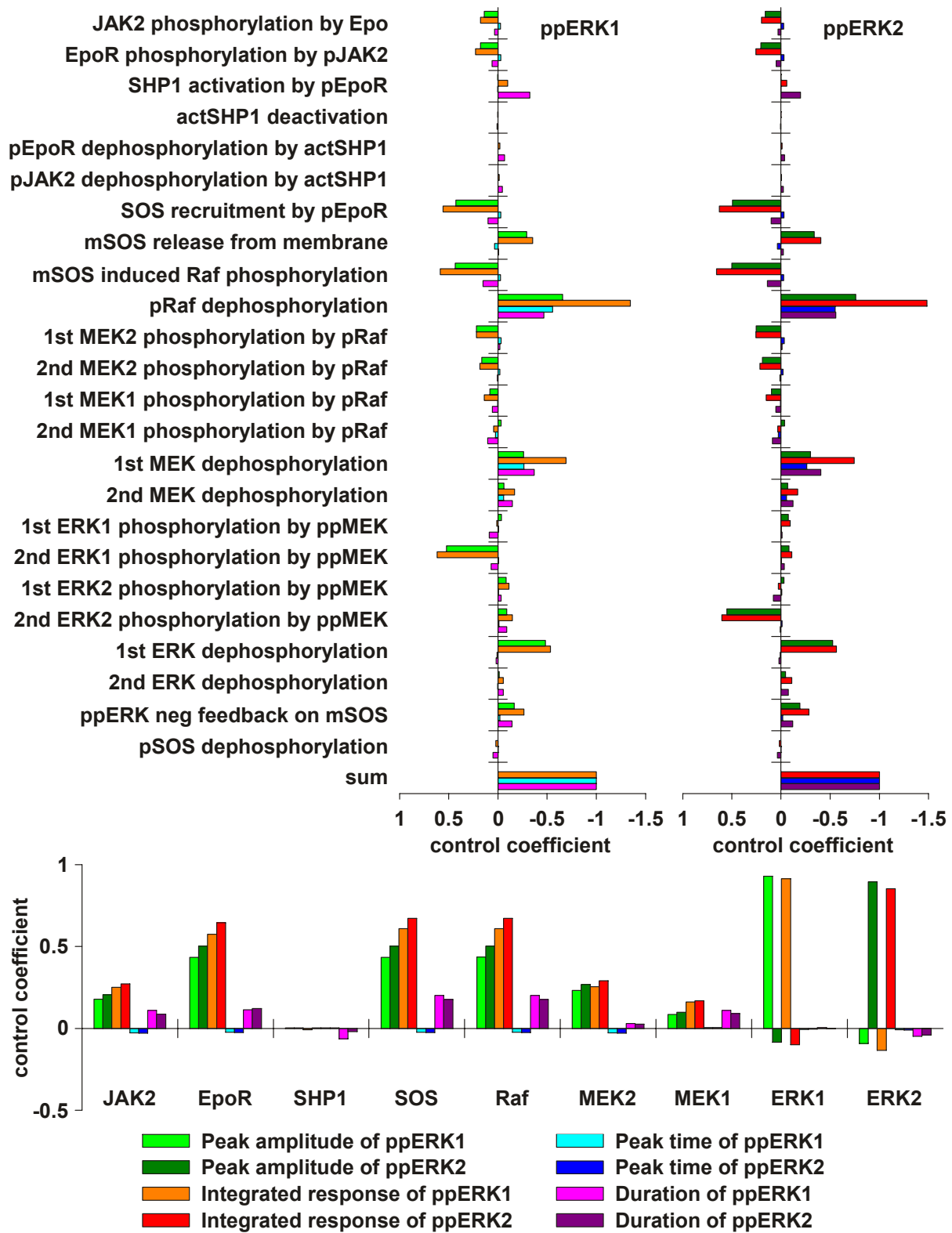


Figure S6, Schilling et al., 2007a

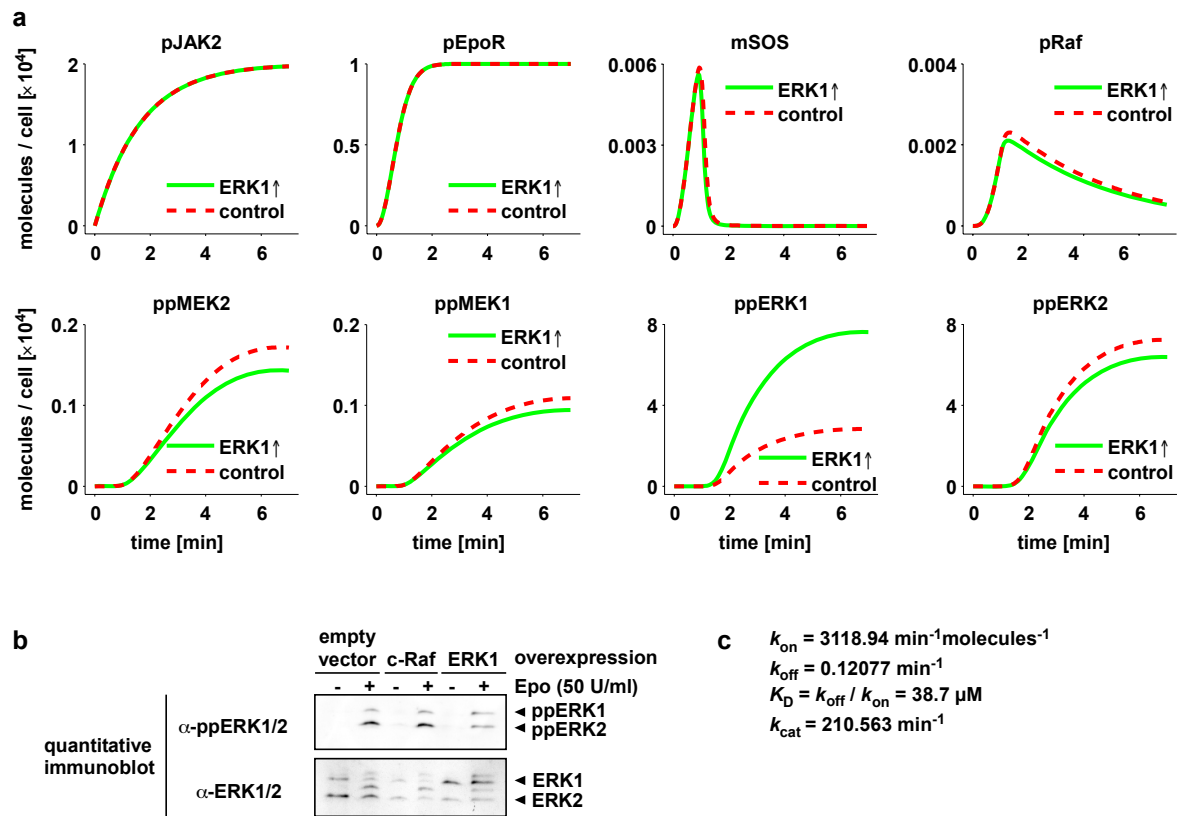


Figure S7, Schilling et al., 2007a

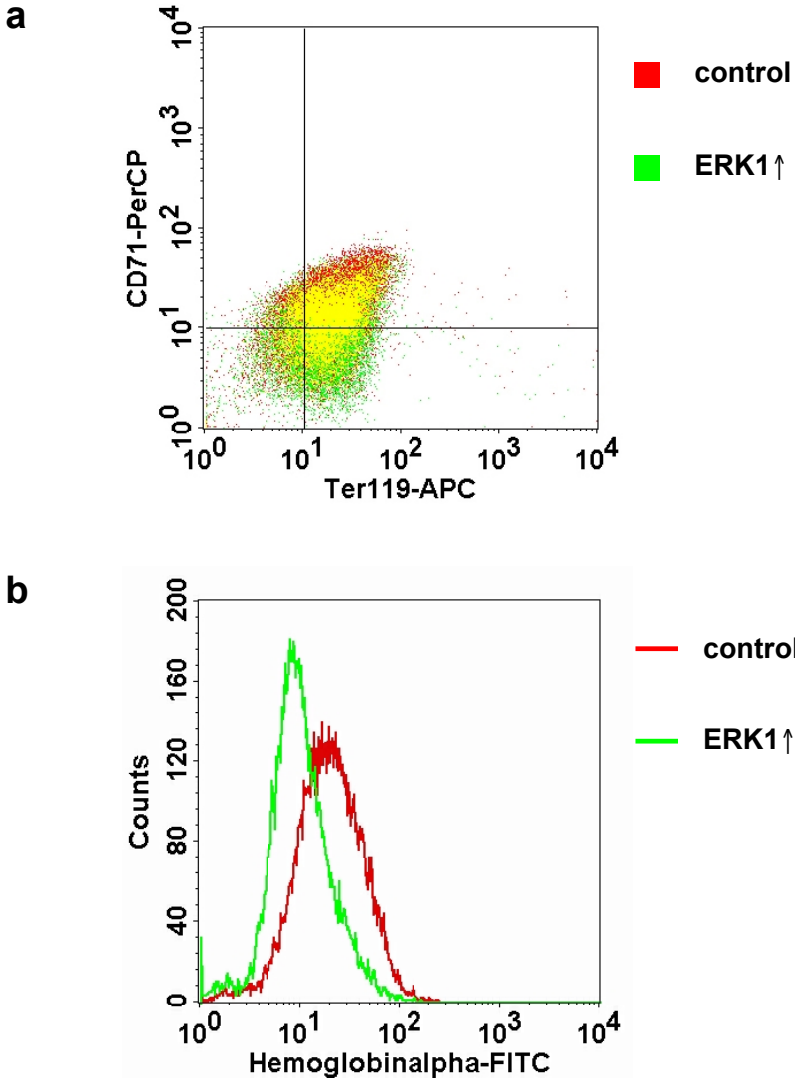


Figure S8, Schilling et al., 2007a

2.4 Internalization determining Epo receptor activation kinetics

Becker V, Schilling M, Bachmann J, Hengl S, Maiwald T, Timmer J, Klingmüller U

Internalization controls early phase kinetics of Epo receptor activation

To be submitted.

This manuscript contains 5 figures. The experiments depicted in Figures 1a, 1b and 3 were initiated, planned and performed by Marcel Schilling. The experiments depicted in Figure 4 were initiated and planned by Marcel Schilling. The manuscript was in part written by Marcel Schilling (figures and text).

Internalization controls early phase kinetics of Epo receptor activation

Verena Becker^{1,3}, Marcel Schilling^{1,3}, Julie Bachmann¹, Stefan Hengl², Thomas Maiwald², Jens Timmer², and Ursula Klingmüller¹

¹Systems Biology of Signal Transduction, German Cancer Research Center, 69120 Heidelberg, Germany

²Freiburg Center for Data Analysis and Modeling, University of Freiburg, 79104 Freiburg, Germany

³These authors contributed equally.

Correspondence should be addressed to U. K. (e-mail: u.klingmueller@dkfz-heidelberg.de)

Running head: Internalization determining Epo receptor activation kinetics

Initiation of signal transduction is critically determined by the amount of ligand-accessible receptor on the cell surface. We analyzed ligand-induced and constitutive endocytosis of the erythropoietin receptor (EpoR) by mathematical modelling based on quantitative data. The rapid decline of initial EpoR signalling is mediated by receptor internalization, while the amount of ligand-receptor complexes at the plasma membrane recovers after prolonged stimulation. Modelling revealed that ligand binding to the EpoR substantially increases receptor internalization. Sensitivity analysis demonstrated that the receptor turnover, the association rate of ligand binding as well as internalization critically shape the early phase kinetic behaviour of signalling-competent ligand-receptor complexes. We conclude that EpoR internalization is not responsible for long-term receptor attenuation, but induces rapid deactivation of EpoR signalling by accelerated endocytosis of ligand-bound receptors. Furthermore, to determine the potency of Epo derivatives and mimetics, it is more informative to measure the association kinetics rather than the ligand affinity. Thus, the data-based model provides an important basis for the targeted design of more potent Epo derivatives for the use in clinical applications.

INTRODUCTION

The amplitude of signal transduction initiation is critically determined by the amount of specific receptor on the plasma membrane. The duration of signalling is limited by recruitment of phosphatases and the activation of negative feedback loops^{1,2}. Furthermore, ligand-mediated internalization of receptors is proposed to control the signal duration by lysosomal degradation and downregulation of cell surface receptors, thus providing a mechanism for long term attenuation of signals emanating from the cell surface³. As cell surface receptors are present at low concentrations, the concept of signalling endosomes has recently been discussed to mediate amplification of signal transduction cascades through receptor tyrosine kinases (RTK)⁴. Thus, both the temporal as well as the spatial organization of signal transduction play a key role in shaping cellular responses to external stimuli.

Cytokines regulate the fine-tuned balance of self-renewal and fast adaptation in the hematopoietic system. Erythropoietin (Epo) and its cognate receptor (EpoR) are the key regulators of definitive erythropoiesis by promoting proliferation, survival, and terminal differentiation of erythroid progenitor cells at the colony forming unit-erythroid (CFU-E) stage⁵. The EpoR is expressed at low levels on the cell surface⁶. Upon ligand binding, the preformed homodimeric EpoR⁷ undergoes a conformational switch leading to activation of the associated Janus kinase (JAK) 2⁸, which phosphorylates cytoplasmic tyrosine residues of the EpoR. These phosphotyrosines serve as docking sites for signal-promoting molecules such as the signal transducer and activator of transcription (STAT) 5 as well as for negative regulators including the protein tyrosine phosphatase SHP-1 and members of the suppressor of cytokine signalling (SOCS) family^{9,10}.

Additionally, Epo binding induces internalization of the EpoR. Ligand-mediated EpoR internalization depends on motifs within the extracellular domain¹¹ and cytoplasmic domain of the receptor^{12,13} but does not strictly depend on receptor tyrosine phosphorylation or JAK2 activity¹⁴. Studies applying inhibitors of protein synthesis have suggested that Epo binding accelerates receptor internalization¹⁵.

The cellular function of EpoR internalization remains unclear. Endocytosis of the EpoR has been proposed to terminate signalling by lysosomal degradation of activated receptor complexes in UT-7 cells¹⁵. Alternatively, receptor endocytosis can target activated ligand-receptor complexes to sites of phosphatase activity as demonstrated for the interaction between internalized epidermal growth factor (EGF) receptor and the protein tyrosine phosphatase PTP-1B on the surface of the endoplasmic reticulum (ER)¹⁶.

Recently, internalization and degradation kinetics of recombinant human Epo (rhEpo) as well as of the hyperglycosylated darbepoetin alfa bound to the human EpoR have been determined by kinetic modelling, showing that about 40% of ligand is degraded after uptake,

while 60% of internalized Epo is resecreted intact¹⁷. However, so far little is known about how ligand-induced internalization influences cellular responsiveness and signal transduction through the EpoR.

To address this, we applied a systems biology approach by employing dynamic mathematical modelling of ligand-induced and constitutive EpoR internalization based on quantitative data. Unexpectedly, the amount of ligand-receptor complexes at the plasma membrane recovered after prolonged stimulation, although EpoR activation was not further detectable. Sensitivity analysis of Epo-EpoR internalization modelling identified the parameters receptor turnover and internalization as critical for the early kinetics of signalling-competent ligand-receptor complexes. Furthermore, we uncovered that the amount of bound Epo is largely determined by association kinetics rather than ligand affinity.

RESULTS

Dynamic mathematical model of ligand-mediated EpoR internalization

To gain insight into the kinetics of Epo-stimulated EpoR internalization, we reviewed the principle steps of ligand-mediated receptor endocytosis and translated them into a dynamic ordinary differential equations (ODE)-based mathematical model comprising 10 reactions and 8 species (Fig. 1a). Free cell surface EpoR is subjected to constant turnover, with production and plasma membrane transport being dependent on the *turnover* rate and the maximal amount of receptor on the cell surface B_{max} . Constitutive endocytosis and degradation of the EpoR depend on the amount of receptor on the cell surface and on the *turnover* rate. Therefore, in absence of ligand, the amount of cell surface EpoR is constant and equal to B_{max} . Free functional Epo in the medium binds to the EpoR with the on-rate k_{on} and dissociation occurs with the off-rate k_{off} . As the dissociation constant K_D can be determined experimentally, k_{off} is expressed as $k_{on} \times K_D$. Epo bound to its receptor is internalized and finally dissociates from the EpoR due to the pH-change during trafficking through endosomal and lysosomal compartments. Internalized receptor can recycle to the surface with a delay. Internalized Epo can proceed by two routes. It is either recycled back to the medium or degraded in the cell and released to the medium afterwards, unable to re-bind to the receptor. This is the minimal model that can both describe our data and is consistent with biological knowledge. The delay for EpoR recycling is absolutely necessary, while delays at other steps in the model including Epo binding and internalization as well as Epo recycling and release proved to be unnecessary to describe our data. Additionally, degradation and release of Epo are inevitably serial, not parallel steps.

To determine the parameters B_{max} and K_D , BaF3-HA-EpoR cells were treated with increasing concentrations of [125 I]-Epo and the free as well as the specifically bound [125 I]-Epo was measured (Fig. 1b). We calculated parameters by a one-site saturation regression analysis, revealing a dissociation constant $K_D = 164 \pm 28$ pM that is comparable to published data¹⁸. A kinetic measurement of Epo-induced EpoR internalization was performed for BaF3-HA-EpoR cells at 37°C. Cells were stimulated with [125 I]-Epo for up to 240 min. Unbound [125 I]-Epo in the medium was removed and surface-bound [125 I]-Epo was separated from internalized [125 I]-Epo by acid stripping (Fig. 1c). Ligand-receptor complexes were rapidly internalized, while [125 I]-Epo in the medium was not depleted within the time observed. Remarkably, after a first decline, the extent of cell surface-bound [125 I]-Epo recovered at 240 min, indicating that the amount of EpoR at the plasma membrane is not depleted for prolonged ligand stimulation. Measurements were performed in triplicates, the standard deviations being small enough for parameter estimation techniques.

Dynamic mathematical model of constitutive EpoR internalization

To determine Epo-independent internalization, we established BaF3 cells expressing a streptavidin-binding peptide tagged (SBP)-EpoR. Since streptavidin does not induce EpoR phosphorylation (Supplementary Information, Fig S1), this approach allowed us to follow constitutive EpoR endocytosis by incubation with radiolabelled streptavidin. The model for constitutive receptor internalization is analogous to the Epo-induced receptor internalization model (Fig. 2a), the main difference being that the receptor is internalized with a *turnover* rate that is independent of bound streptavidin. As the amount of cell surface EpoR is constant, it is not important whether the receptor is recycled to the surface or newly produced. Internalized streptavidin undergoes the same reactions as internalized Epo. The parameter *release* was assumed to be the same for both Epo and streptavidin, as parameter estimation resulted in very similar values and one can assume that degraded peptides are processed with similar reactions within intracellular compartments. To determine the rates for B_{max_strep} and K_D_strep , BaF3-SBP-EpoR cells were treated with increasing concentrations of [¹²⁵I]-streptavidin. [¹²⁵I]-Streptavidin did not efficiently bind to SBP-EpoR, therefore the dissociation constant $K_D = 2964 \pm 1189$ pM showed a high standard deviation (Fig. 2b). However, the binding affinity of [¹²⁵I]-streptavidin was within the reported nanomolar range¹⁹. A kinetic measurement of constitutive EpoR internalization was performed for BaF3-SBP-EpoR cells. Cells were stimulated with [¹²⁵I]-streptavidin at 37°C for up to 60 min. Free [¹²⁵I]-streptavidin in the medium was measured, while [¹²⁵I]-streptavidin on the surface was separated by acid stripping from intracellular [¹²⁵I]-streptavidin (Fig. 2c). The amount of cell surface-bound [¹²⁵I]-streptavidin increased within the observed time, indicative of the inefficient binding to the receptor. Measurements were performed in triplicates, showing standard deviations being small enough for parameter estimation techniques.

Parameter estimation of combined EpoR internalization models

We performed simultaneous parameter estimation using the two ODE-based models (see Supplementary Information, Fig. S2). EpoR delay was modelled using a 10-step compartmentalization reaction. All initial concentrations were set to 0, except for EpoR set to B_{max} or B_{max_strep} , as well as Epo and strep, which were set to the initial concentrations ± 100 pM used for the experiments. B_{max} or B_{max_strep} were fixed to the determined values. Parameters for both models were estimated simultaneously, with two shared parameters *turnover* and *release*. Parameter estimation was performed 500 times with different starting conditions. The trajectories for the best fit are depicted in Fig. 3. The model captures the dynamics of the system very well, including the steep initial decrease of a minor fraction of Epo in the medium, the rapid peak of Epo bound on the cell surface followed by a decrease before rising again and the saturation kinetics-like accumulation of internalized Epo. The data

for constitutive internalization are equally well represented. Low χ^2 -values compared to the number of data points indicate accurate explanation of the data by the dynamic mathematical models. To gain insight into the systems properties, we additionally plotted the trajectories of all variables of the Epo-induced EpoR internalization model (see Supplementary Information, Fig. S3) and uncovered interesting properties. Free Epo is in large excess, however, free EpoR on the surface is not completely occupied by Epo. Rather, it is rapidly reduced to about one third of the initial receptor concentration and then slowly replenished by receptor recycling. Receptor-bound Epo, both on the cell surface and within the cell, displays a remarkable kinetics. A sharp peak in the first minutes is followed by a rapid decline and the receptor concentration rises again after two hours. Finally, a rather substantial part of internalized Epo is degraded and released into the medium.

To investigate the identifiability of the estimated parameter values, we calculated the standard deviation of the parameters for the best 65% of the 500 fits (see Supplementary Information, Fig. S4). Five of 16 parameters exhibited standard deviations larger than 25%. As these parameters are divided in two groups of highly dependent parameters, we could fix two of these parameters and then re-estimate the parameters. By this procedure, all parameter exhibited standard deviations smaller than 25%. Comparing the parameters *turnover* and *internalization*, it became evident that Epo-binding increases EpoR internalization by a factor of 3.5. Moreover, approximately 80% of Epo are recycled to the medium intact, while only 20% are degraded.

Sensitivity analysis of ligand-receptor complex formation

To analyze parameters determining the intriguing kinetics of receptor-bound Epo, we performed a sensitivity analysis. The sum of surface and internalized ligand-receptor complexes that are capable of initiating signal transduction was investigated. Three quantities were defined for the analysis. The peak amplitude is the concentration of ligand receptor complexes at the first maximum, the peak time is the time at the first maximum, and the extrema amplitude is the concentration at the first maximum minus the concentration at the first minimum (Fig. 4a). The control coefficients of the parameters of the Epo-induced EpoR internalization model for these quantities were determined. Positive control coefficients indicate that values for peak amplitude, extrema amplitude and peak time increase with higher parameter values, while negative control coefficients indicate that higher parameters lead to lower values for these quantities. Higher absolute values of control coefficients thereby represent a larger influence of the parameter (Fig. 4b). The *turnover* rate has a minor positive impact on peak amplitude and peak time, but a massive negative effect on extrema amplitude. The effect of k_{on} is positive for both peak and extrema amplitude and negative for peak time, indicating that faster binding leads to faster and thus probably more pronounced

signalling. On the other hand, K_D has little control on all parameters, demonstrating that the binding kinetics is more important than the affinity of Epo binding to its receptor. *Internalization* possesses negative control on peak amplitude, minor positive control on extrema amplitude, and strong negative control on peak time. *Dissociation* has a minor negative impact on all quantities. The rest of the parameters have essentially no control on the kinetics of EpoR-bound Epo. As can be shown by summation theorems, the sum of all control coefficients is 0 for peak and extrema amplitude and -1 for peak time²⁰. As the control coefficients were calculated for infinitesimal changes, we tested whether the results are also applicable for larger parameter variations. Therefore, trajectories for the sum of surface and internalized EpoR-bound Epo were plotted depending on changes in the parameters *turnover* and *internalization* (Fig. 4c). As predicted by the sensitivity analysis, increased *turnover* led to a slightly higher peak but a shallower kinetic, while increased *internalization* caused a sharper but lower peak. Thus, sensitivity analysis allowed identification of the factors that cause the steep rise and rapid decline of EpoR-bound Epo in cells.

Time-course analysis of long-term EpoR activation

The kinetic measurements showed that the cell surface population of ligand-bound EpoR recovers after approximately 240 min. Therefore, unless negative regulators prevent signal initiation one would assume that EpoR and JAK2 activation would follow this kinetic behaviour and phosphorylated proteins would be detectable upon prolonged ligand stimulation. To examine the kinetics of receptor activation, BaF3 cells stably expressing HA-EpoR were stimulated with 5 U/ml Epo for up to 240 min and the amount of activated receptor and JAK2 were quantified (Fig. 5a). Phosphorylated EpoR showed peak intensity at 10 min, followed by a rapid decrease that reflects the kinetics of ligand-bound receptor complexes. EpoR activation was reduced to basal levels between 60 and 120 min (Fig. 5a). Despite recovery of the plasma membrane pool of receptor-ligand complexes at 240 min of stimulation, phosphorylated EpoR could not be detected at later time points. Analysis of JAK2 revealed that the activated kinase exhibits similar kinetics to phosphorylated EpoR (Fig. 5b). Thus, ligand-induced receptor internalization does not mediate long-term attenuation of EpoR activation but rather shapes the initial kinetics of ligand-receptor complex formation. Probably, negative regulators are responsible for termination of EpoR signalling upon prolonged stimulation.

DISCUSSION

The extent of cell surface expression regulates the sensitivity of receptors to external ligands. By quantitative measurement we determined that long-term attenuation of receptor activation is not mediated by downregulation of the receptor pool at the plasma membrane. By applying a systems biology approach, we uncovered the parameters that determine the kinetics of ligand-receptor complex formation and thus activation of EpoR signalling.

Receptor endocytosis has been proposed to control downregulation of activated receptor complexes and thus terminate signals emanating from the plasma membrane. For example, while the EGFR is rapidly internalized depending on autophosphorylation of the cytoplasmic domain and subsequently downregulated by lysosomal degradation²¹, ErbB3 endocytosis exhibits a significantly slower kinetics and the receptor recycles to the plasma membrane²². These studies indicate the existence of different strategies for receptor trafficking, even within one family of structurally related receptors. For the human myeloid cell line UT-7 expressing high levels of EpoR, downregulation of the activated receptor is proposed to be mediated by both proteasomal and lysosomal degradation¹⁵. However, our observations revealed that despite recovery of ligand-receptor complexes at the cell surface, EpoR activation is restrained upon prolonged stimulation, arguing against the assumption that long-term attenuation of EpoR signalling is mediated by internalization and downregulation of the receptor plasma membrane pool.

Therefore, other mechanisms terminate EpoR signalling, probably including the recruitment of phosphatases to the EpoR, the induction of negative feedback proteins, as well as degradation of the receptor-associated JAK2. A prominent phosphatase involved in negative regulation of EpoR signalling is SHP-1 that is recruited to the phosphorylated EpoR and subsequently inhibits JAK2 activation^{23,24}. Since the induction of negative feedback proteins requires gene transcription, recruitment of the SOCS family members CIS²⁵ and SOCS-3²⁶ to tyrosine-phosphorylated EpoR is likely to occur at later time points after EpoR activation. An alternative mechanism proposed suggested that upon Epo stimulation, JAK2 is autophosphorylated on a negatively regulating tyrosine within the receptor-binding FERM domain, which leads to dissociation from the EpoR and subsequent kinase degradation²⁷. This indicates that JAK2 activity may be the limiting factor for sustained EpoR phosphorylation. As the amount of SHP-1 remains constant after Epo stimulation (unpublished data), we speculate that the decreased ratio of JAK2 to SHP-1 favours dephosphorylation over phosphorylation of the receptor.

We identified recycling of the ligand and the receptor as key systems properties of EpoR trafficking. Approximately 80% of Epo recycles to the medium and therefore allows EpoR activation without depleting the ligand in the extracellular space, which is important in particular in the hematopoietic stem cell niche showing potentially low Epo concentrations²⁸.

Thus, Epo recycling enables erythroid progenitor cells to respond to Epo without affecting ligand concentration in the extracellular environment for adjacent cells. Recovery of the amount of ligand-receptor complexes at the plasma membrane upon prolonged stimulation critically depends on EpoR recycling. Association of JAK2 with the EpoR in the ER was reported to be necessary for receptor maturation by a chaperone-like mechanism and cell surface expression of newly synthesized proteins²⁹. As the overall intracellular receptor pool is depleted upon Epo stimulation (data not shown), we speculate that the decline in total receptor expression during ligand stimulation might correlate with decreased JAK2 levels, possibly resulting in an accelerated degradation of unprocessed EpoR protein.

Our modelling approach revealed that three parameters receptor *turnover*, the association rate k_{on} for ligand binding to the receptor, as well as receptor *internalization* determined the sharp peak of ligand-receptor complex formation. We suggest that these parameters values are a prerequisite to allow both rapid formation of signalling-competent receptor complexes and an immediate decline after peak levels. Analysis of EpoR and JAK2 phosphorylation levels indicated that rapid receptor internalization restricts the initial kinetics of signalling at the receptor level.

The majority of the EpoR resides within the ER and the Golgi complex, while only a minor part is transported to the cell surface⁶. In contrast to a recent study modelling Epo trafficking and degradation for human EpoR¹⁷, our modelling approach includes the turnover rate of the receptor, revealing that receptor turnover is critical for the early phase of ligand-receptor complex formation. This observation indicates that the continuous exchange of the EpoR plasma membrane pool and sampling of the extracellular environment is essential for signal transduction.

Sensitivity analysis of parameters describing the ligand-stimulated EpoR internalization demonstrated that k_{on} but not K_D is critical for both the extent as well as the rate to form signalling-competent ligand-receptor complexes. Thus, for designing efficient Epo derivatives such as darbepoetin alfa³⁰ or continuous erythropoietin receptor activator (CERA)³¹ used in clinical applications it would be more informative to determine the kinetics rather than the affinity of ligand binding.

In conclusion, by data-based mathematical modelling we determined receptor and ligand kinetics, uncovering unexpected systems properties such as the importance of both receptor turnover and recycling for plasma membrane prevalence. Furthermore, we identified the parameters determining early phase kinetics of EpoR activation. Thus, insights gained by the data-based mathematical model provides an important new basis for the targeted design of more potent Epo derivatives.

METHODS

Cells, plasmids, and antibodies

If not stated otherwise, chemicals were obtained from Sigma and cell culture media were purchased from Gibco. For streptavidin assays, stably transduced BaF3 cells were washed six times with biotin-free RPMI 1640 (PAN Biotech) and the experiment was performed in streptavidin binding medium (biotin-free RPMI 1640 supplemented with 2 mM L-Glutamine and 25 mM HEPES pH 7.4).

Phoenix eco and BaF3 cells were cultured as described³². Retroviral expression vectors were pMOWS puro³². To yield pMOWS-HA-EpoR, a *EcoRI* and *BamHI* fragment from the vector pMX-EpoR-HA-IRES-GFP (provided by S. Constantinescu, Ludwig Institute for Cancer Research, Brussels, Belgium) was inserted into the *PacI* and *BamHI* restriction sites of pMOWS-EpoR³³ and a consensus Kozak sequence was introduced 5' of the EpoR cDNA. To generate pMOWS-SBP-EpoR, the HA-tag from pMOWS-HA-EpoR was exchanged with an SBP-tag¹⁹. All clones were verified by sequence analysis.

The following antibodies were used: anti-EpoR (M-20) and anti-JAK2 (HR-758) (for immunoprecipitation) were purchased from Santa Cruz, anti-JAK2 (24B11) (for immunoblotting) was purchased from Cell Signaling, and anti-phosphotyrosine (clone 4G10) was purchased from UBI.

Transfection and transduction

Transfections were performed by calcium phosphate precipitation as described³². Transducing supernatants were generated 24h after transfection and supplemented with 8 µg/ml polybrene. BaF3 cells stably expressing EpoR variants were selected in 1.5 µg/ml puromycin 48 hours after transduction.

Binding assays

BaF3 cells (1×10^6) stably expressing HA-EpoR were incubated with 10 pM, 100 pM, 250 pM, 500 pM, or 2000 pM [¹²⁵I]-Epo (GE Healthcare) in 100 µl RPMI 1640 supplemented with 10% FCS for 4 h at RT. To separate free [¹²⁵I]-Epo, cells were centrifuged through a layer of FCS and cell-bound as well as free [¹²⁵I]-Epo was measured in a Cobra gamma counter (Packard). Specific binding was determined by subtracting the mean value of radioactivity of cells incubated with both [¹²⁵I]-Epo and 250 U/ml unlabelled Epo (Janssen-Cilag) (n=3) from the radioactivity of cells incubated in the absence of unlabelled Epo. Results are presented either with fitting a one-site saturation curve or as a Scatchard plot.

For determining the affinity of streptavidin to SBP-EpoR, BaF3 cells stably expressing SBP-EpoR (1×10^6) were incubated with 100 pM, 250 pM, 1000 pM, 2500 pM, or 5000 pM [¹²⁵I]-streptavidin (GE Healthcare) in 100 µl streptavidin binding medium for 4 h at RT. Cells were

washed three times with 500 μ l biotin-free RPMI 1640 and pooled supernatants as well as the cells were measured in a gamma counter. To determine specific binding of [125 I]-streptavidin, control cells were additionally incubated with 2.5 μ M unlabelled streptavidin.

Internalization assays

BaF3 cells (4×10^6) expressing HA-EpoR were washed and starved in RPMI 1640 supplemented with 1 mg/ml BSA for 3 h and stimulated with 2100 pM [125 I]-Epo (approximately equal to 5 U/ml) (GE Healthcare) in 100 μ l RPMI 1640 for the indicated time at 37°C. After stimulation, cells were immediately transferred to ice and free [125 I]-Epo was separated from the cells by centrifugation through a layer of FCS. Cell surface bound [125 I]-Epo was stripped by incubation with 4% acetic acid for 5 min on ice and centrifugation through FCS. The efficiency of acid stripping was about 95% and membrane integrity after acid stripping was confirmed by Trypan blue exclusion. Cell surface bound as well as internalized [125 I]-Epo was measured in a Cobra gamma counter (Packard). Specific binding was determined by subtracting values obtained with 500 U/ml unlabeled Epo (Janssen-Cilag).

To access constitutive receptor internalization, stably transduced BaF3-SBP-EpoR cells (1×10^6) were washed and starved for 1 h in streptavidin binding medium supplemented with 1 mg/ml BSA. Cells were subsequently incubated with 1000 pM [125 I]-streptavidin (GE Healthcare) in 100 μ l streptavidin binding medium for the indicated time at 37°C. After stimulation, cells were immediately transferred to ice and washed three times with 500 μ l biotin-free RPMI 1640 to separate unbound [125 I]-streptavidin from the cells. Cell surface bound [125 I]-streptavidin was stripped by incubation with 4% acetic acid for 5 min on ice and subsequent washing. Cell surface bound as well as internalized [125 I]-streptavidin was measured in a gamma counter. Specific binding was determined by subtracting values obtained with an excess of unlabeled streptavidin.

Immunoprecipitation and immunoblotting

Stably transduced BaF3 cells (1×10^7) were washed and starved in RPMI 1640 (Gibco) supplemented with 1 mg/ml BSA (Sigma) for 3 h. Cells were stimulated with 5 U/ml Epo (Janssen-Cilag) for the times indicated at 37°C and lysed with NP-40 lysis buffer (1% NP-40, 150 mM NaCl, 20 mM Tris pH 7.4, 10 mM NaF, 1 mM EDTA pH 8.0, 1 mM ZnCl₂ pH 4.0, 1 mM MgCl₂, 1 mM Na₃VO₄, 10% glycerol) supplemented with aprotinin and AEBSF (Sigma). Immunoprecipitates were eluted, separated by 10% SDS-PAGE, and transferred to a nitrocellulose membrane. Detection was performed using enhanced chemiluminescence (GE Healthcare). Quantitative immunoblotting data was processed using GellInspector software³⁴.

The following normalizers were used: GST-JH2JH1 for pJAK2 and JAK2 and GST-EpoR for pEpoR and EpoR. For first estimates, csaps – splines were used with a smoothness of 0.3.

Mathematical modeling

Modeling was performed using PottersWheel (Maiwald et al., unpublished). Parameter estimation was performed in logarithmized parameter space using trust region optimization method. For each fit, 100 iterations were performed with χ^2 tolerance of 10^{-7} and fit parameters tolerance of 10^{-7} . Using the best fit as starting value, 500 fits were performed, each time varying all parameters with a disturbance strength of 0.4. The boundaries for K_D and K_{D_strep} were confined by the standard deviations of the measurements, and B_{max} and B_{max_strep} were fixed to the measured values. All other parameters were estimated with boundaries between 10^{-7} and 10^3 , and none of the estimated parameters lay on these boundaries. The initial values for EpoR were set to the experimentally determined B_{max} or B_{max_strep} , while the initial values for Epo and $strep$ were set to the concentrations ± 100 pM used for experimental setup.

Sensitivity Analysis

Sensitivity analysis was applied to investigate relative changes of derived system quantities K (see Fig. 3) as a result of relative changes in parameter values p_i

$$S_{p_i}^K = \frac{p_i}{K} \cdot \frac{\partial K}{\partial p_i} \quad (1)$$

Hornberg et al.³⁵ derived summation laws for sensitivities of derived system quantities like *signal amplitude*, *signal duration*, and *area under curve*. The proofs for the summation laws²⁰ can easily be extended to show the existence of summation laws for the system quantities investigated in our approach.

$$\sum_i S_{p_i}^{\text{peak amplitude}} = 0 \quad (2)$$

$$\sum_i S_{p_i}^{\text{peak time}} = -1 \quad (3)$$

$$\sum_i S_{p_i}^{\text{extrema amplitude}} = 0 \quad (4)$$

Sensitivity analysis is a local approach, because derivations are evaluated at a certain point in parameter space (local sensitivity analysis).

Most models, however, are non-identifiable, i.e., there exist model parameters that cannot be determined unambiguously. Often, non-identifiability manifests itself in functionally related parameters (see linear relationship and hyperbola in Supplementary information, Fig. S3b,c). For example, the output functions of the constitutive EpoR internalization model are invariant under parameter variations along the hyperbola. Thus, without prior knowledge, it can in

principle not be determined statistically at which point in parameter space sensitivity analysis has to be performed. To deal with this problem, we take the following approach: The model is fitted N -times to data ($N=500$). Each fit yields different estimates for the non-identifiable parameters. Non-identifiabilities are detected with NBI and sensitivities (1) are calculated at the actual point along the non-identifiability, here the linear relationship and the hyperbola, respectively. As the derived system quantities (peak amplitude, peak time and extrema amplitude) are invariant to changes along the non-identifiabilities (linear and hyperbolic relationship), the results of the sensitivity analysis do not depend on the values we chose for the non-identifiable parameters. This is always the case for analytical non-identifiabilities. Sensitivities for the non-identifiable parameters are not determinable, yet in our model, the non-identifiable parameters had essentially no control over the derived system variables independent of their values. Therefore, the control coefficients we determined varied by less than 10^{-4} , which is due to numerical reasons.

REFERENCES

1. Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* **103**, 211-25 (2000).
2. Hilton, D. J. Negative regulators of cytokine signal transduction. *Cell Mol Life Sci* **55**, 1568-77 (1999).
3. Waterman, H. & Yarden, Y. Molecular mechanisms underlying endocytosis and sorting of ErbB receptor tyrosine kinases. *FEBS Lett* **490**, 142-52 (2001).
4. Miaczynska, M., Pelkmans, L. & Zerial, M. Not just a sink: endosomes in control of signal transduction. *Curr Opin Cell Biol* **16**, 400-6 (2004).
5. Wu, H., Liu, X., Jaenisch, R. & Lodish, H. F. Generation of committed erythroid BFU-E and CFU-E progenitors does not require erythropoietin or the erythropoietin receptor. *Cell* **83**, 59-67 (1995).
6. Yoshimura, A., D'Andrea, A. D. & Lodish, H. F. Friend spleen focus-forming virus glycoprotein gp55 interacts with the erythropoietin receptor in the endoplasmic reticulum and affects receptor metabolism. *Proc Natl Acad Sci U S A* **87**, 4139-43 (1990).
7. Livnah, O. et al. Crystallographic evidence for preformed dimers of erythropoietin receptor before ligand activation. *Science* **283**, 987-90 (1999).
8. Remy, I., Wilson, I. A. & Michnick, S. W. Erythropoietin receptor activation by a ligand-induced conformation change. *Science* **283**, 990-3 (1999).
9. Klingmüller, U. The role of tyrosine phosphorylation in proliferation and maturation of erythroid progenitor cells--signals emanating from the erythropoietin receptor. *Eur J Biochem* **249**, 637-47 (1997).
10. Richmond, T. D., Chohan, M. & Barber, D. L. Turning cells red: signal transduction mediated by erythropoietin. *Trends Cell Biol* **15**, 146-55 (2005).
11. Quelle, D. E., Quelle, F. W. & Wojchowski, D. M. Mutations in the WSAWSE and cytosolic domains of the erythropoietin receptor affect signal transduction and ligand binding and internalization. *Mol Cell Biol* **12**, 4553-61 (1992).
12. Levin, I. et al. Identification of a cytoplasmic motif in the erythropoietin receptor required for receptor internalization. *FEBS Lett* **427**, 164-70 (1998).
13. Flint-Ashtamker, G., Eisen-Lev, R., Cohen, J., Jun-shen Huang, L. & Neumann, D. Amino acid residues 268-276 of the erythropoietin receptor contain an endocytosis motif and are required for erythropoietin-mediated proliferation. *FEBS Lett* **518**, 189-94 (2002).
14. Beckman, D. L., Lin, L. L., Quinones, M. E. & Longmore, G. D. Activation of the erythropoietin receptor is not required for internalization of bound erythropoietin. *Blood* **94**, 2667-75 (1999).

15. Walrafen, P. et al. Both proteasomes and lysosomes degrade the activated erythropoietin receptor. *Blood* **105**, 600-8 (2005).
16. Haj, F. G., Verveer, P. J., Squire, A., Neel, B. G. & Bastiaens, P. I. Imaging sites of receptor dephosphorylation by PTP1B on the surface of the endoplasmic reticulum. *Science* **295**, 1708-11 (2002).
17. Gross, A. W. & Lodish, H. F. Cellular trafficking and degradation of erythropoietin and novel erythropoiesis stimulating protein (NESP). *J Biol Chem* **281**, 2024-32 (2006).
18. D'Andrea, A. D. et al. The cytoplasmic region of the erythropoietin receptor contains nonoverlapping positive and negative growth-regulatory domains. *Mol Cell Biol* **11**, 1980-7 (1991).
19. Keefe, A. D., Wilson, D. S., Seelig, B. & Szostak, J. W. One-step purification of recombinant proteins using a nanomolar-affinity streptavidin-binding peptide, the SBP-Tag. *Protein Expr Purif* **23**, 440-6 (2001).
20. Hornberg, J. J. et al. Principles behind the multifarious control of signal transduction. ERK phosphorylation and kinase/phosphatase control. *Febs J* **272**, 244-58 (2005).
21. Sorkin, A. & Waters, C. M. Endocytosis of growth factor receptors. *Bioessays* **15**, 375-82 (1993).
22. Waterman, H., Sabanai, I., Geiger, B. & Yarden, Y. Alternative intracellular routing of ErbB receptors may determine signaling potency. *J Biol Chem* **273**, 13819-27 (1998).
23. Klingmüller, U., Lorenz, U., Cantley, L. C., Neel, B. G. & Lodish, H. F. Specific recruitment of SH-PTP1 to the erythropoietin receptor causes inactivation of JAK2 and termination of proliferative signals. *Cell* **80**, 729-38 (1995).
24. Yi, T., Zhang, J., Miura, O. & Ihle, J. N. Hematopoietic cell phosphatase associates with erythropoietin (Epo) receptor after Epo-induced receptor tyrosine phosphorylation: identification of potential binding sites. *Blood* **85**, 87-95 (1995).
25. Ketteler, R. et al. The cytokine-inducible Scr homology domain-containing protein negatively regulates signaling by promoting apoptosis in erythroid progenitor cells. *J Biol Chem* **278**, 2654-60 (2003).
26. Hortner, M., Nielsch, U., Mayr, L. M., Heinrich, P. C. & Haan, S. A new high affinity binding site for suppressor of cytokine signaling-3 on the erythropoietin receptor. *Eur J Biochem* **269**, 2516-26 (2002).
27. Funakoshi-Tago, M., Pelletier, S., Matsuda, T., Parganas, E. & Ihle, J. N. Receptor specific downregulation of cytokine signaling by autophosphorylation in the FERM domain of Jak2. *Embo J* **25**, 4763-72 (2006).
28. Noe, G., Riedel, W., Kubanek, B. & Rich, I. N. An ELISA specific for murine erythropoietin. *Br J Haematol* **104**, 838-40 (1999).

-
29. Huang, L. J., Constantinescu, S. N. & Lodish, H. F. The N-terminal domain of Janus kinase 2 is required for Golgi processing and cell surface expression of erythropoietin receptor. *Mol Cell* **8**, 1327-38 (2001).
 30. Macdougall, I. C. et al. Pharmacokinetics of novel erythropoiesis stimulating protein compared with epoetin alfa in dialysis patients. *J Am Soc Nephrol* **10**, 2392-5 (1999).
 31. Brandt, M., Lanzendörfer, M., Frische, J., Haselbeck, A. & Jarsch, M. Different receptor binding activity of C.E.R.A. (Continuous Erythropoietin Receptor Activator) compared with epoetin beta determined by surface plasmon resonance and competition assay on UT-7 cells. *Nephrol Dial Transplant* **21 (Suppl 4): iv9**, (Abstract SO018) (2006).
 32. Ketteler, R., Glaser, S., Sandra, O., Martens, U. M. & Klingmüller, U. Enhanced transgene expression in primitive hematopoietic progenitor cells and embryonic stem cells efficiently transduced by optimized retroviral hybrid vectors. *Gene Ther* **9**, 477-87 (2002).
 33. Swameye, I., Muller, T. G., Timmer, J., Sandra, O. & Klingmuller, U. Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by databased modeling. *Proc Natl Acad Sci U S A* **100**, 1028-33 (2003).
 34. Schilling, M. et al. Computational processing and error reduction strategies for standardized quantitative data in biological networks. *Febs J* **272**, 6400-11 (2005).
 35. Hornberg, J. J. et al. Control of MAPK signalling: from complexity to what really matters. *Oncogene* **24**, 5533-42 (2005).

FIGURE LEGENDS

Figure 1 Dynamic mathematical model describing ligand-induced EpoR internalization. **(a)** Graphical representation of the Epo-induced EpoR internalization model. Measured observables are represented by orange shading (Epo in medium), blue shading (Epo on surface), and green shading (Epo in cells). **(b)** BaF3-HA-EpoR cells (1×10^6) were incubated with increasing concentrations of [125 I]-Epo and specifically bound [125 I]-Epo was plotted versus free [125 I]-Epo (solid triangles). A one-site saturation regression was fitted to the data set (solid line) to determine B_{max} (maximal binding, long dash) and K_D (free [125 I]-Epo concentration for half-maximal binding, short dash). Scatchard analysis shows a linear relationship (inset). **(c)** BaF3-HA-EpoR cells (4×10^6) were incubated with 2100 pM Epo (approximately equal to 5 U/ml) [125 I]-Epo at 37°C for the times indicated. Unbound [125 I]-Epo was measured (Epo in medium, orange). Cells were acid-stripped to remove cell surface-bound [125 I]-Epo and supernatants (Epo on surface, blue) as well as cell pellets (Epo in cells, green) were measured. Error bars represent standard deviations of triplicates.

Figure 2 Dynamic mathematical model of constitutive EpoR internalization. **(a)** Graphical representation of constitutive EpoR internalization. Measured observables are represented by orange shading (streptavidin in medium), blue shading (streptavidin on surface), and green shading (streptavidin in cells). **(b)** BaF3-SBP-EpoR cells (1×10^6) were incubated with increasing concentrations of [125 I]-streptavidin and specifically bound [125 I]-streptavidin was plotted versus free [125 I]-streptavidin (solid triangles). A one-site saturation regression was fitted to the data set (solid line) to calculate B_{max_strep} (maximal binding) and K_{D_strep} (free [125 I]-streptavidin concentration for half-maximal binding). **(c)** BaF3-SBP-EpoR cells (1×10^6) were incubated with 1000 pM [125 I]-streptavidin at 37°C for the times indicated. Unbound [125 I]-streptavidin was measured (streptavidin in medium, orange). Cells were acid-stripped to remove surface-bound [125 I]-streptavidin and supernatants (streptavidin on surface, blue) and cell pellets (streptavidin in cells, green) were measured. Error bars represent standard deviations of triplicates.

Figure 3 Parameter estimation for mathematical models of Epo-induced and constitutive EpoR internalization. Experimental data are represented (hollow symbols) with standard deviations and data trajectories of the best fit are shown (solid lines). Parameter estimation was performed simultaneously for both ligand-induced and constitutive EpoR internalization models, with common parameters having the same value. χ^2 -values lower than the number of data points indicate good agreement of the model with the experimental values for both ligand-induced **(a, b, c)** and constitutive EpoR internalization **(d, e, f)**.

Figure 4 Sensitivity analysis reveals impact of *turnover* and *internalization* on kinetics of ligand-receptor complex formation. **(a)** The trajectory for the sum of cell surface and internalized EpoR-bound Epo is represented by a solid line. The three quantities analyzed are peak amplitude (concentration at first maximum, orange), extrema amplitude (concentration at first maximum minus concentration at first minimum, green) and peak time (time at first maximum, blue). **(b)** Control coefficients for the parameters of the Epo-induced internalization model on the sum of cell surface and internalized EpoR-bound Epo. Positive control coefficients indicate higher values for peak amplitude (orange), extrema amplitude (green) and peak time (blue) for increasing parameters, while negative control coefficients indicate decreasing values for the quantities for increasing parameter values. Higher absolute values of control coefficients represent larger control. **(c)** Time course simulations for the sum of cell surface and internalized EpoR-bound Epo are shown in green for increased parameters, in black for unchanged parameters and in orange for decreased parameters.

Figure 5 Long-term EpoR activation is restrained despite ligand-receptor complex prevalence. BaF3-HA-EpoR cells (1×10^7) were starved and stimulated with 5 U/ml Epo for up to 240 min. Immunoprecipitates with anti-EpoR antibodies **(a)** or anti-JAK2 antibodies **(b)** were separated by 10% SDS-PAGE and analysis of activated receptor was performed by immunoblotting with anti-phosphotyrosine antibodies.

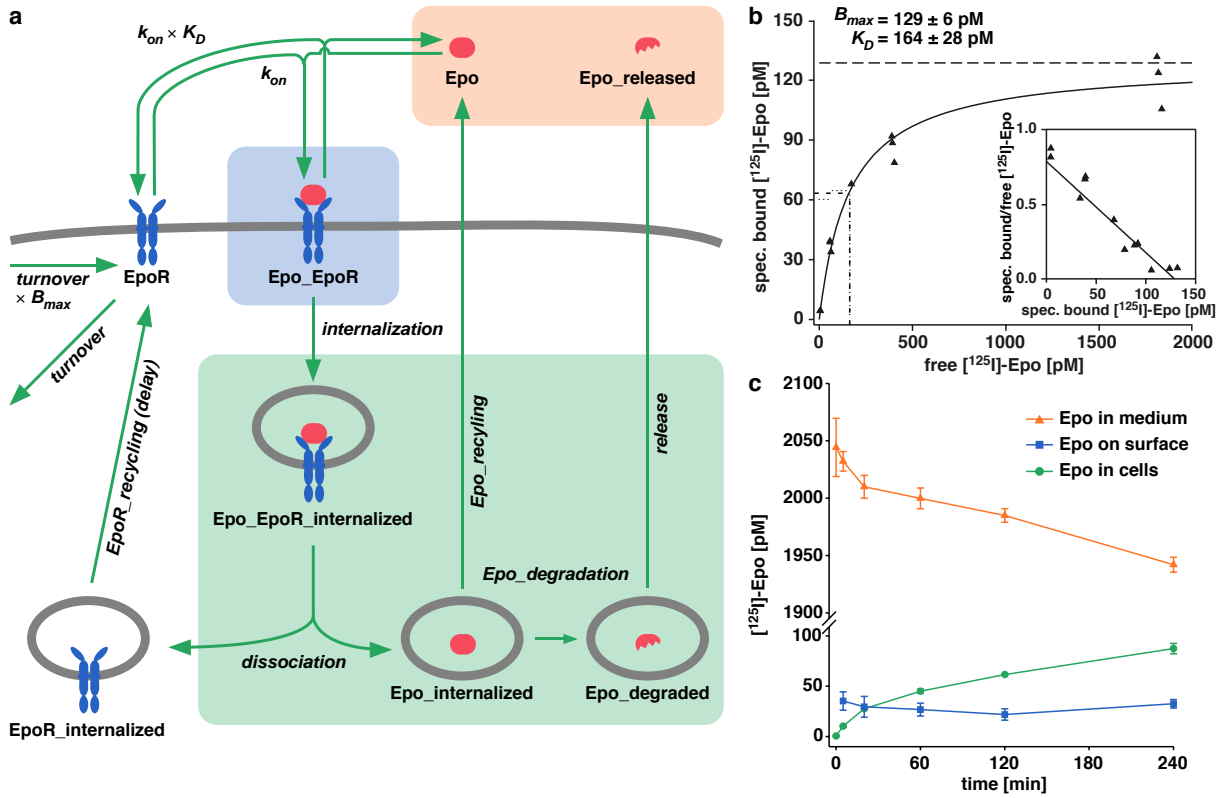


Figure 1, Becker, Schilling et al., 2007

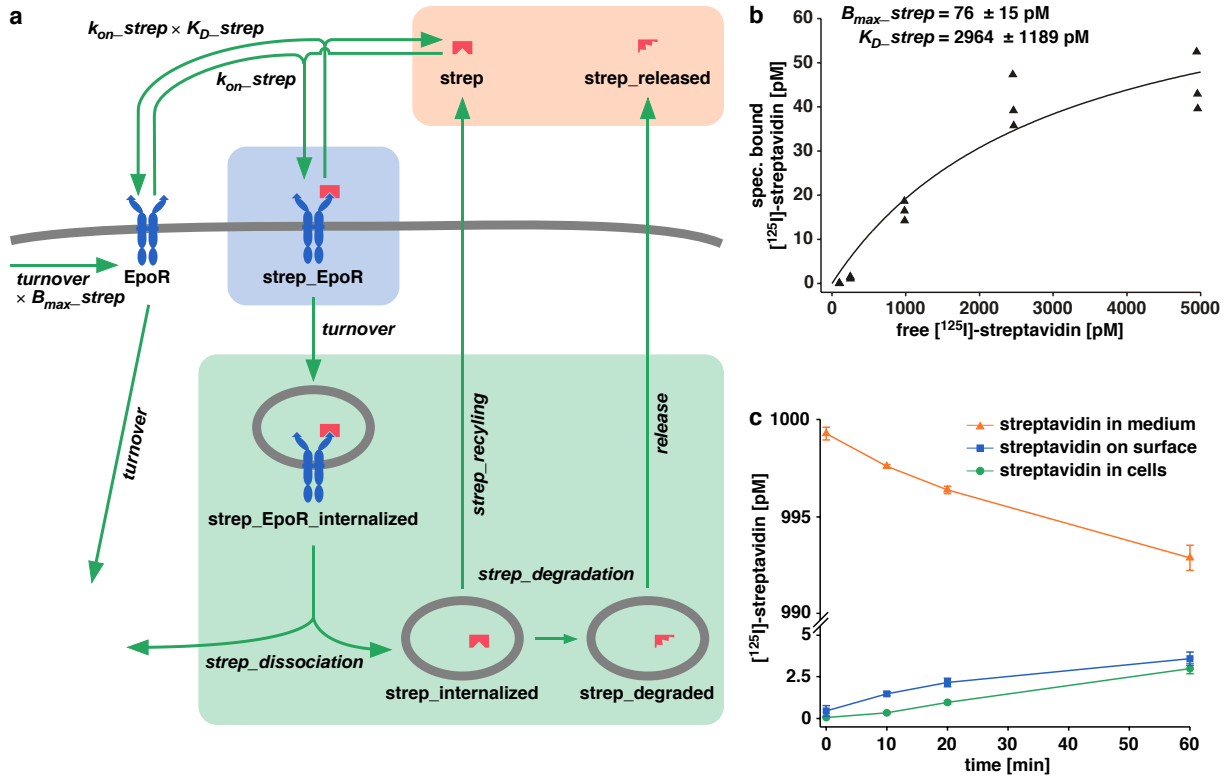


Figure 2, Becker, Schilling et al., 2007

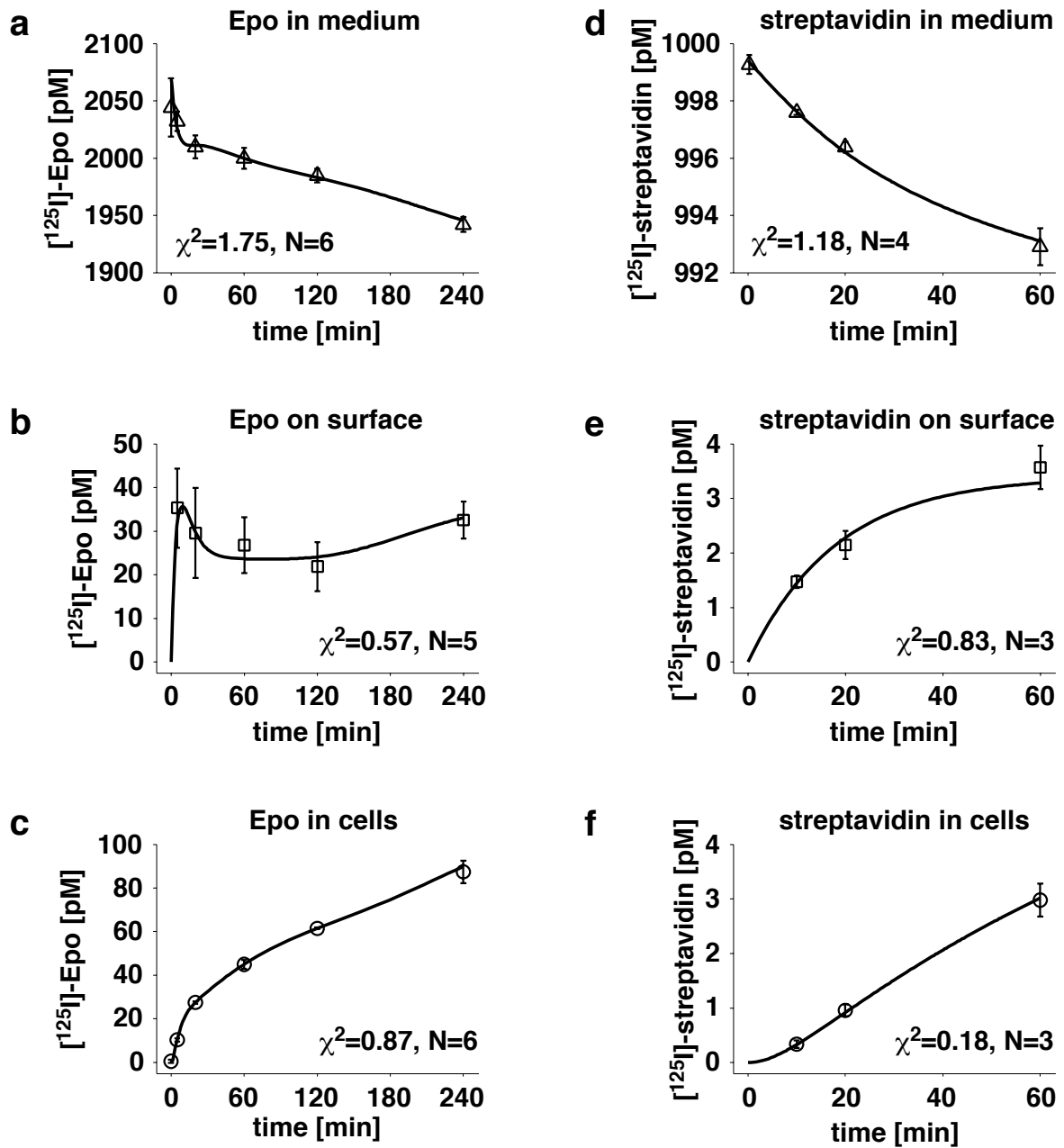


Figure 3, Becker, Schilling et al., 2007

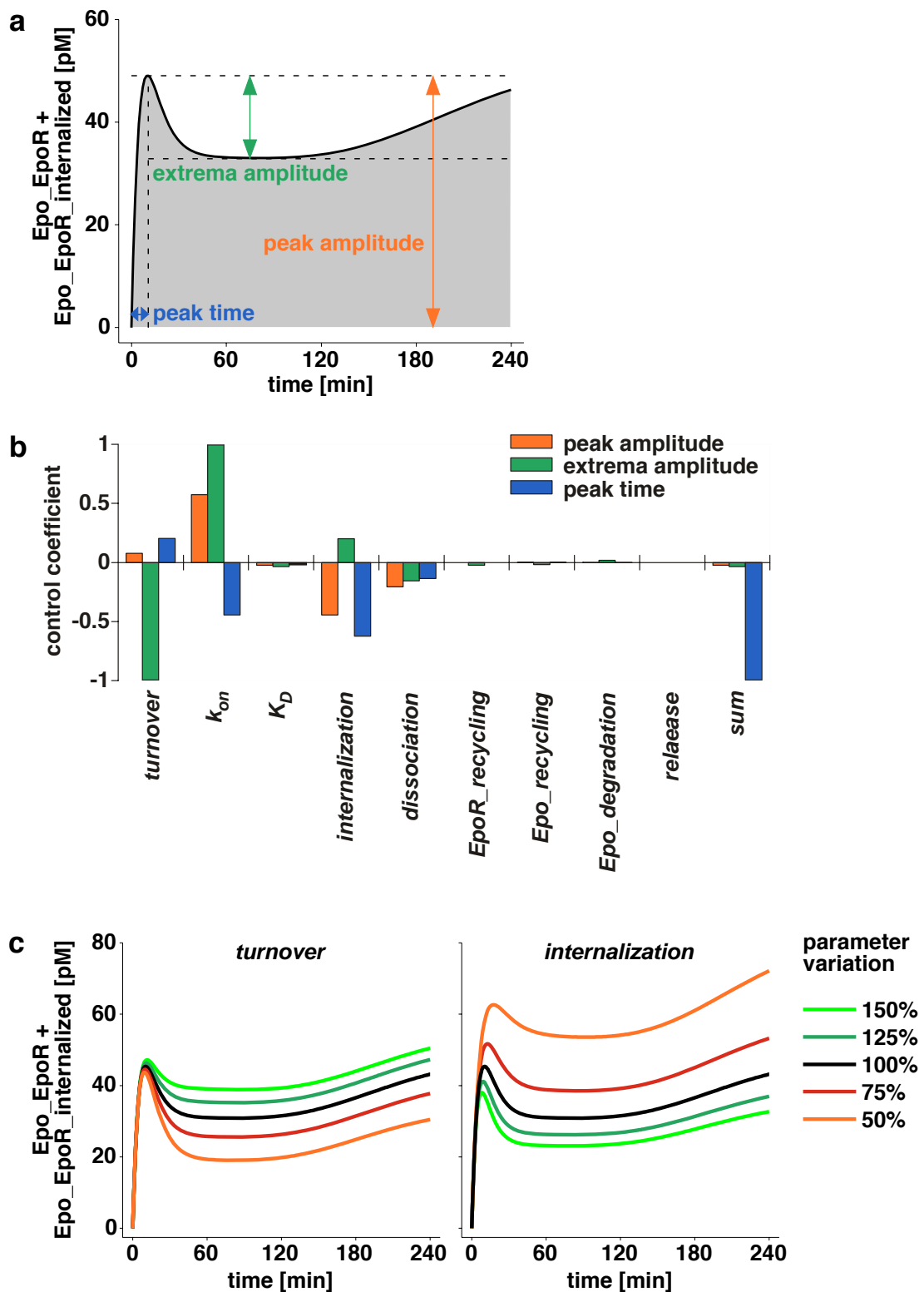


Figure 4, Becker, Schilling et al., 2007

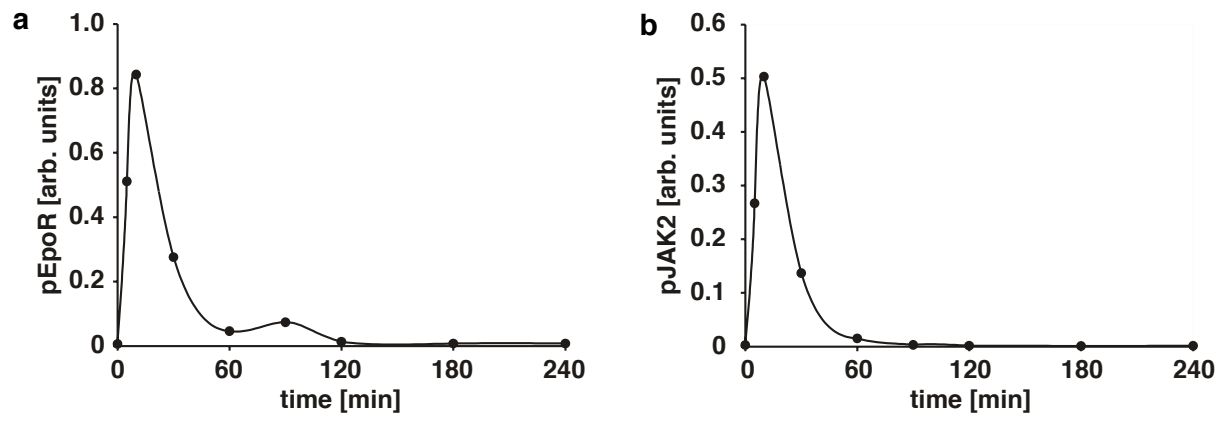


Figure 5, Becker, Schilling et al., 2007

Internalization controls early phase kinetics of Epo receptor activation

Verena Becker^{1,3}, Marcel Schilling^{1,3}, Julie Bachmann¹, Stefan Hengl², Thomas Maiwald², Jens Timmer², and Ursula Klingmüller¹

SUPPLEMENTARY INFORMATION

SUPPLEMENTARY RESULTS

Streptavidin does not induce EpoR phosphorylation

To determine the turnover rate of the EpoR in unstimulated cells, we made use of a streptavidin-binding peptide-tagged EpoR (SBP-EpoR) and incubation with [¹²⁵I]-streptavidin in medium devoid of biotin. To test whether the biotin-free medium in the experimental setup is comparable to the standard medium, we assessed the amount of receptor at the plasma membrane of BaF3 cells expressing HA-EpoR using anti-HA-antibodies. Flow cytometry analysis revealed no difference in viability and EpoR cell surface expression of cells washed and starved either in standard cultivation medium (+ biotin) or in biotin-free medium (- biotin) (Fig. S1a). Furthermore, while SBP-EpoR was activated upon Epo stimulation, streptavidin was binding to the receptor without inducing EpoR phosphorylation even at high concentrations (Fig. S1b). Therefore, the use of streptavidin allows determination of constitutive receptor internalization and turnover.

SUPPLEMENTARY METHODS**Flow cytometry**

To analyze cell surface expression, BaF3 cells expressing HA-EpoR were stained as described¹ and analyzed by flow cytometry using a FACSCalibur (Becton Dickinson). Live cells were gated by forward and side scatter.

Immunoprecipitation and immunoblotting

Stably transduced BaF3 cells (1×10^7) were washed and starved in streptavidin binding medium supplemented with 1 mg/ml BSA (Sigma) for 3 h. After stimulation with either 50 U/ml Epo (Janssen-Cilag) or rising concentrations of streptavidin for 20 min at 37°C, cells were lysed with NP-40 lysis buffer (1% NP-40, 150 mM NaCl, 20 mM Tris pH 7.4, 10 mM NaF, 1 mM EDTA pH 8.0, 1 mM ZnCl₂ pH 4.0, 1 mM MgCl₂, 1 mM Na₃VO₄, 10% glycerol) supplemented with aprotinin and AEBSF (Sigma). Immunoprecipitates were eluted, separated on 10% SDS-PAGE, and transferred to a nitrocellulose membrane. Detection was performed using enhanced chemiluminescence (GE Healthcare).

SUPPLEMENTARY REFERENCES

1. Ketteler, R. et al. A functional green fluorescent protein-tagged erythropoietin receptor despite physical separation of JAK2 binding site and tyrosine residues. *J Biol Chem* **277**, 26547-52 (2002).

SUPPLEMENTARY FIGURE LEGENDS

Figure S1 Biotin-free medium does not change EpoR cell surface expression and streptavidin does not induce EpoR phosphorylation. **(a)** BaF3-HA-EpoR cells were washed and starved for 1 h in standard RPMI 1640 (+ biotin; blue shading) and biotin-free medium (- biotin; red line). HA-EpoR cell surface expression was determined by FACS analysis using anti-HA antibodies as first antibody and Cy5-coupled secondary antibodies. Live cells were gated by forward and side scatter and a representative overlay is shown. Analysis revealed that cell surface expression of the EpoR is independent of biotin in the medium. **(b)** Mock-transduced BaF3 cells (ctrl) or BaF3-SBP-EpoR cells (1×10^7) were stimulated with 50 U/ml Epo or with increasing concentrations of streptavidin or left unstimulated. Immunoprecipitations were performed with anti-EpoR antibodies. Samples were separated by 10% SDS-PAGE and activated receptor analyzed by immunoblotting with anti-phosphotyrosine antibodies revealing phosphorylated EpoR after Epo stimulation, while streptavidin had no effect. Blots were reprobbed with anti-EpoR antibodies.

Figure S2 Ordinary differential equations (ODE), parameter values, initial values and observables for the mathematical models. **(a)** ODE are shown for the ligand-induced EpoR internalization model. EpoR recycling is modeled with a 10-step compartmentalization reaction. **(b)** ODE are shown for the constitutive EpoR internalization model.

Figure S3 Unobserved variables of the ligand-induced EpoR internalization model.

Figure S4 Non-parametric bootstrap-based identifiability testing (NBI) reveals dependent parameters. **(a)** Box plots of all estimated parameters are shown for the best 65% of 500 fits. Five parameters show a standard deviation larger than 25%. **(b,c)** NBI revealed parameter dependencies for a set of two and a set of three parameters. *Epo_recycling* and *Epo_degradation* are directly correlated (b), while *strep_dissociation*, *strep_recycling*, and *strep_degradation* are described by a skew hyperbola in space (c). **(d)** The parameters *Epo_degradation* and *strep_dissociation* were fixed and parameter estimation was repeated. Box plots of all estimated parameters are shown for the best 65% of 500 fits. All parameters show a standard deviation smaller than 25%.

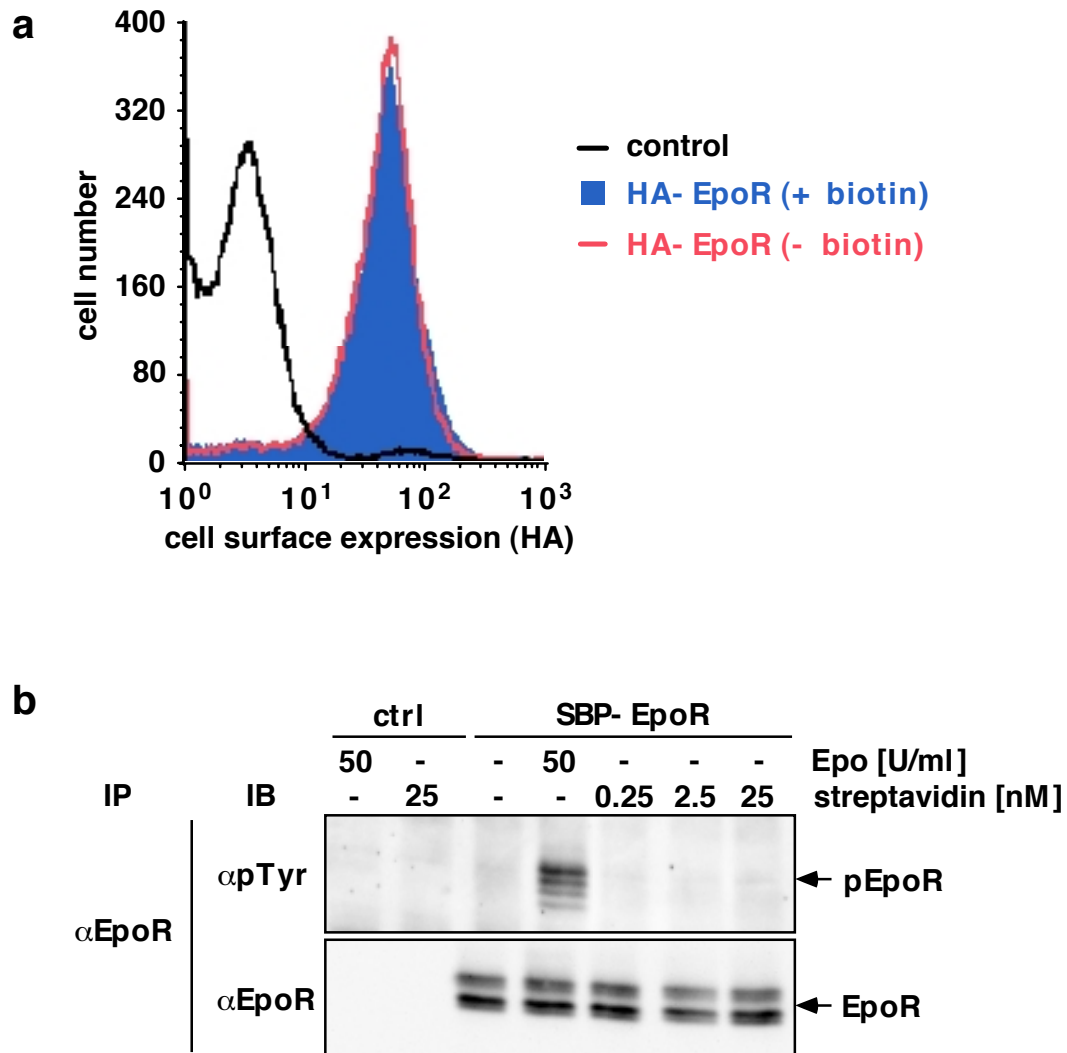


Figure S1, Becker, Schilling et al., 2007

a Ordinary differential equations

EpoR:	$\dot{x}_1 = k_1 \cdot k_2 - k_1 \cdot x_1 - k_3 \cdot x_2 \cdot x_1 + k_3 \cdot k_4 \cdot x_3 + k_7 \cdot x_{17}$
Epo:	$\dot{x}_2 = -k_3 \cdot x_2 \cdot x_1 + k_3 \cdot k_4 \cdot x_3 + k_6 \cdot x_5$
Epo_EpoR:	$\dot{x}_3 = k_3 \cdot x_2 \cdot x_1 - k_3 \cdot k_4 \cdot x_3 - k_5 \cdot x_3$
Epo_EpoR_internalized:	$\dot{x}_4 = k_5 \cdot x_3 - k_6 \cdot x_4$
Epo_internalized:	$\dot{x}_5 = k_6 \cdot x_4 - k_6 \cdot x_5 - k_7 \cdot x_5$
EpoR_internalized:	$\dot{x}_6 = k_6 \cdot x_4 - k_7 \cdot x_6$
Epo_degraded:	$\dot{x}_7 = k_9 \cdot x_5 - k_{10} \cdot x_7$
Epo_released:	$\dot{x}_8 = k_{10} \cdot x_7$
EpoR_recycling _{delay1} :	$\dot{x}_9 = k_7 \cdot x_6 - k_7 \cdot x_9$
EpoR_recycling _{delay2} :	$\dot{x}_{10} = k_7 \cdot x_9 - k_7 \cdot x_{10}$
EpoR_recycling _{delay3} :	$\dot{x}_{11} = k_7 \cdot x_{10} - k_7 \cdot x_{11}$
EpoR_recycling _{delay4} :	$\dot{x}_{12} = k_7 \cdot x_{11} - k_7 \cdot x_{12}$
EpoR_recycling _{delay5} :	$\dot{x}_{13} = k_7 \cdot x_{12} - k_7 \cdot x_{13}$
EpoR_recycling _{delay6} :	$\dot{x}_{14} = k_7 \cdot x_{13} - k_7 \cdot x_{14}$
EpoR_recycling _{delay7} :	$\dot{x}_{15} = k_7 \cdot x_{14} - k_7 \cdot x_{15}$
EpoR_recycling _{delay8} :	$\dot{x}_{16} = k_7 \cdot x_{15} - k_7 \cdot x_{16}$
EpoR_recycling _{delay9} :	$\dot{x}_{17} = k_7 \cdot x_{16} - k_7 \cdot x_{17}$

Parameters

turnover:	k_1
B_{max} :	k_2
K_{on} :	k_3
K_D :	k_4
internalization:	k_5
dissociation:	k_6
EpoR_recycling:	k_7
Epo_recycling:	k_8
Epo_degradation:	k_9
release:	k_{10}

Initial values

EpoR:	x_1
Epo:	x_2

Observables

Epo in medium:	$y_1 = x_2 + x_8$
Epo on surface:	$y_2 = x_3$
Epo in cells:	$y_3 = x_4 + x_5 + x_7$

b Ordinary differential equations

EpoR:	$\dot{x}_1 = k_1 \cdot k_2 - k_1 \cdot x_1 - k_3 \cdot x_2 \cdot x_1 + k_3 \cdot k_4 \cdot x_3$
strep:	$\dot{x}_2 = -k_3 \cdot x_2 \cdot x_1 + k_3 \cdot k_4 \cdot x_3 + k_6 \cdot x_5$
strep_EpoR:	$\dot{x}_3 = k_3 \cdot x_2 \cdot x_1 - k_3 \cdot k_4 \cdot x_3 - k_1 \cdot x_3$
strep_EpoR_internalized:	$\dot{x}_4 = k_1 \cdot x_3 - k_5 \cdot x_4$
strep_internalized:	$\dot{x}_5 = k_5 \cdot x_4 - k_6 \cdot x_5 - k_7 \cdot x_5$
strep_degraded:	$\dot{x}_6 = k_7 \cdot x_5 - k_8 \cdot x_6$
strep_released:	$\dot{x}_7 = k_8 \cdot x_6$

Parameters

turnover:	k_1
B_{max_strep} :	k_2
k_{on_strep} :	k_3
K_D_strep :	k_4
strep_dissociation:	k_5
strep_recycling:	k_6
strep_degradation:	k_7
release:	k_8

Initial values

EpoR:	x_1
strep:	x_2

Observables

streptavidin in medium:	$y_1 = x_2 + x_7$
streptavidin on surface:	$y_2 = x_3$
streptavidin in cells:	$y_3 = x_4 + x_5 + x_6$

Figure S2, Becker, Schilling et al., 2007

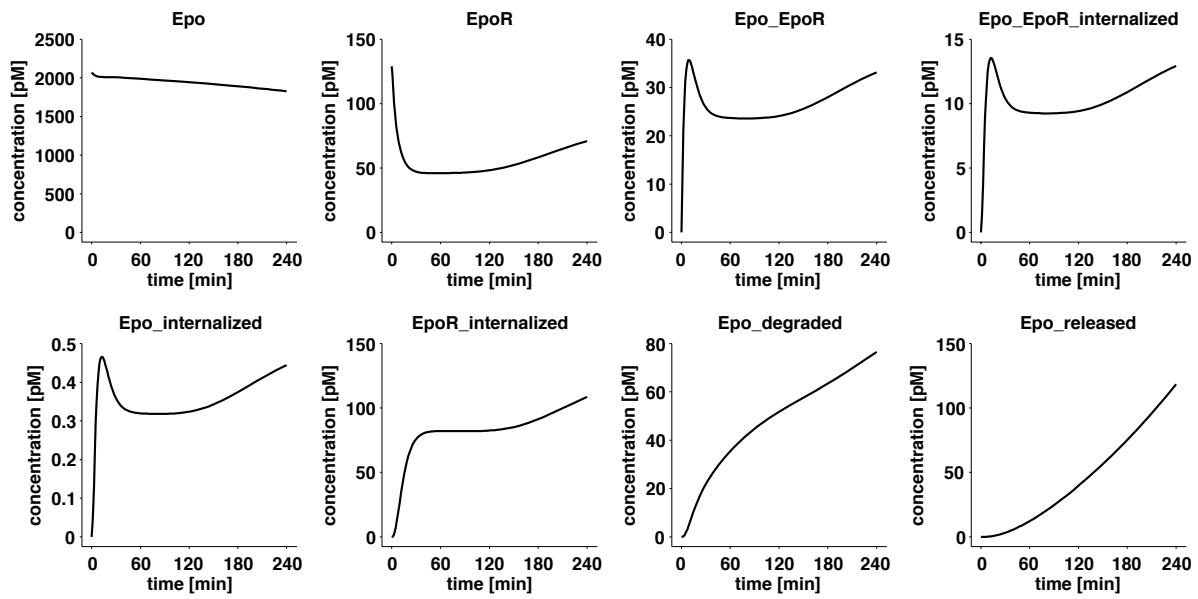


Figure S3, Becker, Schilling et al., 2007

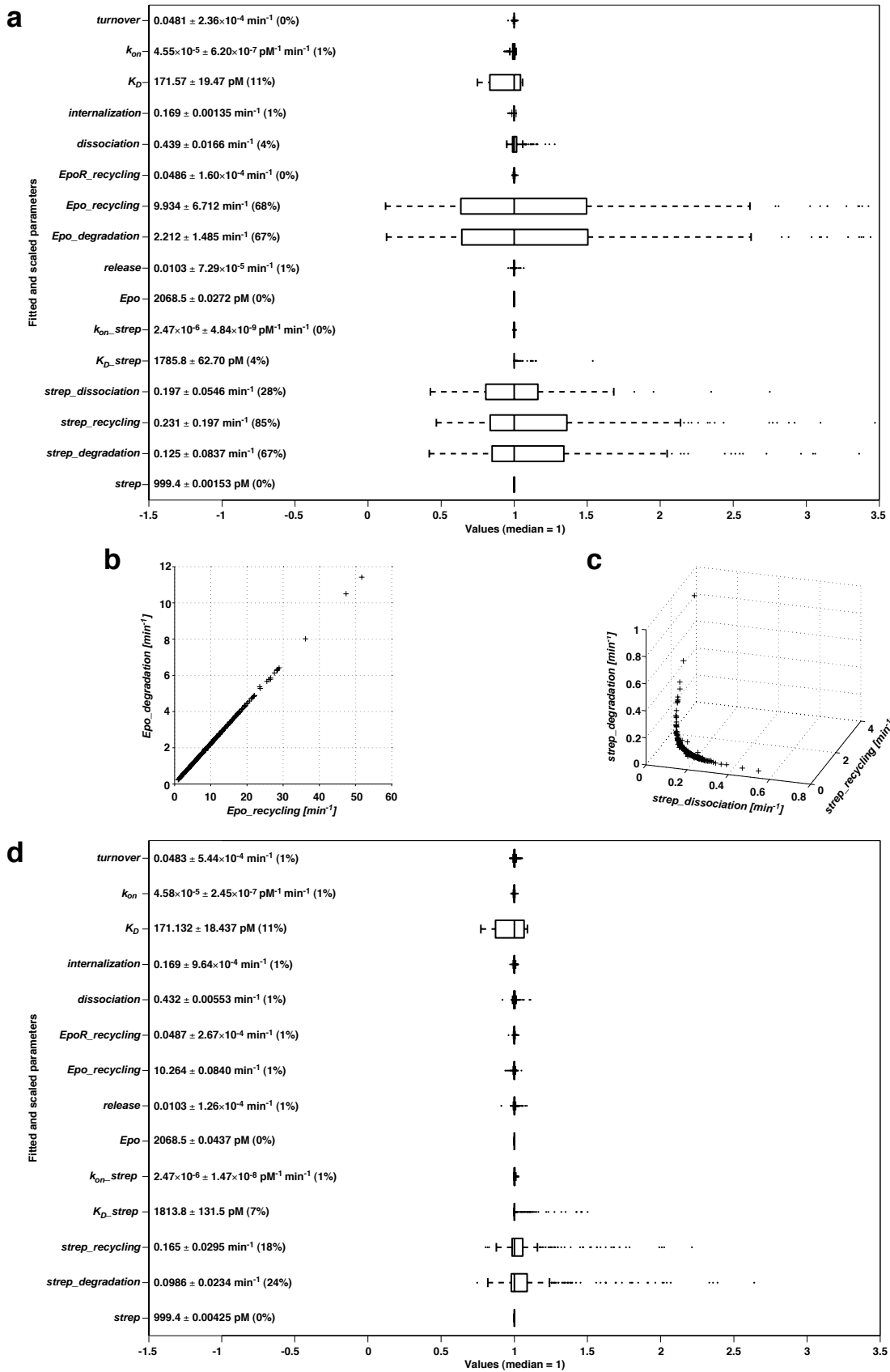


Figure S4, Becker, Schilling et al., 2007

3. APPENDIX

3.1 Abbreviations

ACF	autocorrelation function
ASK	apoptosis signal regulating kinase
ATF	activating transcription factor
Bcl	B-cell leukemia/lymphoma
BFU-E	burst forming unit erythrocyte
CCD	charge-coupled device
CFU-E	colony forming unit erythrocyte
CFU-GEMM	colony forming unit granulocyte erythrocyte monocyte macrophage
CIS	cytokine-inducible SH2-domain containing
DEF	docking site for ERK, FXFP
DNA	deoxyribonucleic acid
DUSP	dual specificity phosphatase
ECL	enhanced chemiluminescence
EGF	epidermal growth factor
Epo	erythropoietin
EpoR	erythropoietin receptor
ER	endoplasmic reticulum
ERK	extracellular signal-regulated kinase
FGF	fibroblast growth factor
FITC	fluorescein-5-isothiocyanat
Gab	GRB2 associated binding protein
GDP	guanine diphosphate
GRB	growth factor receptor-bound protein
GST	glutathione-S-transferase
GTP	guanine triphosphate
HA	hemagglutinin-tagged
hLNGFR	human low affinity nerve growth factor receptor
HOG	high osmolarity/glycerol
HRP	horseradish peroxidase
Hsc	cellular heat shock cognate protein
IB	immunoblot
IL	interleukin
IP	immunoprecipitation
IRS	insulin receptor substrate
JAK	Janus kinase
JIP	JNK interacting protein

JNK	c-Jun N-terminal kinase
KSR	kinase suppressor of Ras
MACS	magnetic cell sorting
MAP	mitogen-activated protein
MAPK	MAP kinase
MAPKAP	MAP kinase-activated protein kinase
MAPKK	MAP kinase kinase
MAPKKK	MAP kinase kinase kinase
MCA	metabolic control analysis
MEF	myocyte-enhancing factor
MEK	MAPK/ERK kinase
MEKK	MEK kinase
MKP	MAP-kinase phosphatase
Mnk	MAP kinase-interacting kinase
Mos	Moloney sarcoma oncogene
MP	MEK-partner
Msk	mitogen- and stress-activated protein
NGF	nerve growth factor
ODE	ordinary differential equation
OSM	osmosensing scaffold for MEKK3
PDI	protein disulfide isomerase
PerCP	peridinin chlorophyll protein
PI3K	phosphoinositide 3-kinase
PKC	protein kinase C
PVDF	polyvinylidene difluoride
Rac	Ras-related C3 botulinum substrate
Rap	Ras-related protein
Ras	rous avian sarcoma homologue
RKIP	Raf-kinase inhibitor protein
Rsk	ribosomal S6 kinase
SAPK	stress-activated protein kinase
SBP	streptavidin binding peptide
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SH2	Src homology 2
SHIP	SH2 inositol 5-phosphatase
SHP	SH2-containing phosphatase
SOS	son of sevenless
STAT	signal transducer and activator of transcription
STE	sterile yeast mutants
TAK	TGF β -activated kinase
TAO	thousand and one amino acid kinase
TCF	ternary complex factor
TPA	12-O-tetradecanoylphorbol-13-acetate

TUNEL	terminal deoxynucleotidyl transferase-mediated deoxyuridine 5'-triphosphate-FITC nick-end labeling
UV	ultraviolet radiation
Y	tyrosine

3.2 Curriculum vitae

Marcel Schilling

personal information

private address: Schulzengasse 20, 69120 Heidelberg, Germany

work address: Systems Biology of Signal Transduction (A150)
 German Cancer Research Center
 Im Neuenheimer Feld 280
 69120 Heidelberg
 Germany
 tel: +49 6221 42 4483
 fax: +49 6221 42 4488
 mail: m.schilling(at)dkfz.de

date of birth: January 8, 1977

place of birth: Binningen, Switzerland

citizenship: Swiss

education

primary school	1984 – 1989	Bottmingen, Switzerland
junior high school	1989 – 1993	Binningen, Switzerland
senior high school	1993 – 1997	Oberwil, Switzerland
Matura examination	December 1997	Oberwil, Switzerland
study of biology II	1998 – 2002	University of Basel
1 st intermediate diploma	October 1999	Biozentrum, University of Basel
2 nd intermediate diploma	October 2000	Biozentrum, University of Basel
diploma examination	October 2002	Biozentrum, University of Basel

research projects

diploma thesis

2001 – 2002

Friedrich Miescher Institute, Basel

"Cloning and characterization of a novel AU-rich
element-binding RNA helicase"

Supervisor: Dr. Yoshikuni Nagamine

PhD thesis

November 2002 – June 2003

Max Planck Institute for Immunobiology, Freiburg

July 2003 – March 2007

German Cancer Research Center, Heidelberg

"Strategies for High Quality Quantitative Data
Generation and Dynamic Modeling of the MAP-
Kinase Signaling Cascade"

Supervisor: PD Dr. Ursula Klingmüller

Participation in the International PhD Study Program
of the German Cancer Research Center

teaching experiences

- March 18 – 22, 2005 Assistant for practical course “Components and Mechanisms of Signal Transduction” (HP-F05; Master of Cell Biology, University of Heidelberg)
- May 1 – 5, 2006 Assistant for practical course “Components and Mechanisms of Signal Transduction” (HP-F06; Master of Cell Biology, University of Heidelberg)
- January – June 2006 Training and supervision of a student intern with the project: “Quantitative Measurement of MAPK Signaling in the Hematopoietic System”

scientific meetings

- October 2004 5th International Conference on Systems Biology (ICSB), Heidelberg (Poster Presentation)
- August 2005 Foundations of Systems Biology and Engineering (FOSBE), Santa Barbara (Poster Presentation)
- July 2006 Systems Biology of Mammalian Cells (SBMC), Heidelberg (Short Talk)
- October 2006 7th International Conference on Systems Biology (ICSB), Yokohama (Poster Presentation)

scientific workshops

October 2003	3 rd Workshop on Computation of Biochemical Pathways and Genetic Networks, Heidelberg
June 2004	Summer School in Systems Biology, Göteborg
June 2006	International Workshop Molecular Systems Biology, Bielefeld
October 2006	Receptor Tyrosine Kinase Training Course, Tokyo

press and public relations

German Cancer Research Center press release (13/12/2005): "The GELINSPECTOR: A Friend and Helper for Systems Biologists"

Dipl.-Biol. M. Schilling et al., "Quantitative Daten für die Systembiologie", BIOforum 04/2006, S. 33-35

patents

PCT/EP2006/069274	M. Schilling, S. Bohl, U. Klingmüller, T. Maiwald, J. Timmer "Method for Processing and Error Reduction for Standardized Quantitative Data in Biological Networks" European and international patent applied
-------------------	--

publications

Tran H, Schilling M, Wirbelauer C, Hess D, Nagamine Y. (2004) Facilitation of mRNA deadenylation and decay by the exosome-bound, DExH protein RHAU. *Mol Cell*. 13(1):101-11.

Schilling M, Maiwald T, Bohl S, Kollmann M, Kreutz C, Timmer J, Klingmüller U. (2005) Quantitative data generation for systems biology: the impact of randomisation, calibrators and normalisers. *IEE Proc Syst Biol*. 152(4):193-200.

Schilling M, Maiwald T, Bohl S, Kollmann M, Kreutz C, Timmer J, Klingmüller U. (2005) Computational processing and error reduction strategies for standardized quantitative data in biological networks. *FEBS J*. 272(24):6400-11.

Klingmüller U, Schilling M, Bohl S, Pfeifer AC. (2007) The need for standardization in systems biology. *Forward Look in Systems Biology*: in press.

Schilling M, Maiwald T, Hengl S, Timmer J, Klingmüller U. A data-based MAP-kinase model reveals systems properties and predicts cellular decisions. *Manuscript in preparation*.

Becker V, Schilling M, Bachmann J, Hengl S, Maiwald T, Timmer J, Klingmüller U. Internalization controls early phase kinetics of Epo receptor activation. *Manuscript in preparation*.

3.3 Erklärung

Ich erkläre hiermit, dass ich die vorgelegte Dissertation selbst verfasst und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfen bedient habe.

Weiterhin erkläre ich hiermit, dass ich an keiner anderen Stelle ein Prüfungsverfahren beantragt bzw. die Dissertation in dieser oder anderer Form bereits anderweitig als Prüfungsarbeit verwendet oder einer anderen Fakultät als Dissertation vorgelegt habe.

Datum:

Unterschrift:

Marcel Schilling