Quantitative Analysis of the Cytokine-Mediated Apoptosis-Survival Cell Decision Process

by

Kevin A. Janes

B.S., Biomedical Engineering
B.A., Spanish
Johns Hopkins University, 1999

SUBMITTED TO THE BIOLOGICAL ENGINEERING DIVISION IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREEE OF

DOCTOR OF PHILOSOPHY IN BIOENGINEERING

AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2005

Signature of Author: _____
Biological Engineering Division
May 9, 2005

Certified by: _____
Douglas A. Lauffenburger
Whitaker Professor of Biologcial Engineering, Chemical Engineering, and Biology
Thesis Supervisor

Certified by: _____
Michael B. Yaffe
Howard S. and Linda Stern Associate Professor of Biology and Biological Engineering
Thesis Supervisor

Approved by: _____
K. Dane Wittrup
J.R. Mares Professor of Chemical Engineering and Biological Engineering
Chairman, Thesis Committee

Quantitative Analysis of the Cytokine-Mediated Apoptosis-Survival Cell Decision Process

by

Kevin A. Janes

Submitted to the Biological Engineering Division
on May 1, 2005 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in
Bioengineering

ABSTRACT

How do cells sense their environment and decide whether to live or to die? This question has drawn considerable interest since 1972, when it was first discovered that cells have an intrinsic ability to self-destruct through a process called apoptosis. Since then, apoptosis has been shown to play a critical role in both normal physiology and disease. In addition, many of the basic molecular mechanisms that control apoptosis have been revealed. Yet despite the known list of interactions and regulators, it remains difficult to inspect the network of apoptosis-related proteins and predict how cells will behave. The challenge is even greater when one considers interactions with other networks that are anti-apoptotic, such as growth-factor networks.

In this thesis, we develop an approach to measure, analyze, and predict how complex intracellular signaling networks transduce extracellular stimuli into cellular fates. This approach entails three interrelated aims: 1) to develop high-throughput, quantitative techniques that measure key nodes in the intracellular network; 2) to characterize the quantitative changes in network state and cell behavior by exposing cells to diverse fate-changing stimuli; and 3) to use data-driven modeling approaches that analyze large signaling-response datasets to suggest new biological hypotheses. These aims were focused on an apoptosis-survival cell-fate decision process controlled by one prodeath cytokine, tumor necrosis factor (TNF), and two prosurvival stimuli, epidermal growth factor (EGF) and insulin.

We first developed radioactive- and fluorescence-based high-throughput assays for quantifying activity changes in the kinases that catalyze key phosphorylation events downstream of TNF, EGF, and insulin. By combining these assays with techniques measuring other important posttranslational modifications, we then compiled over 7000 individual protein measurements of the cytokine-induced network. The signaling measurements were combined with over 1400 measurements of apoptotic responses by using partial least squares (PLS) regression approaches. These signaling-apoptosis regression models predicted apoptotic responses from cytokine-induced signaling patterns alone. Furthermore, the models helped to reveal the importance of previously unrecognized autocrine cytokines in controlling cell fate. This thesis has therefore shown how cell decisions, like apoptosis-versus-survival, can be understood and predicted from the quantitative information contained in the upstream signaling network.

Thesis Supervisor: Douglas A. Lauffenburger
Title: Whitaker Professor of Biologcial Engineering, Chemical Engineering, and Biology

Thesis Supervisor: Michael B. Yaffe
Title: Howard S. and Linda Stern Associate Professor of Biology and Biological Engineering

BIOGRAPHICAL NOTE

Kevin Janes received a bachelor of science in Biomedical Engineering and a bachelor of arts in Spanish from Johns Hopkins University in 1999. While at Johns Hopkins, Kevin worked in the laboratory of Dr. Kam Leong on nonviral delivery vehicles for gene therapy. By graduation, Kevin had received numerous research and academic achievement awards, and he became a member of Alpha Eta Mu Beta and Tau Beta Pi honor societies.

After graduation, Kevin received a Fulbright Award for his proposal entitled, "Biomedical Polymers for Mucosal Drug Delivery." He traveled to Santiago de Compostela, Spain to work with Dr. María-José Alonso in the Department of Pharmaceutical Technology at La Universidad de Santiago de Compostela. Kevin published several articles on his work involving modified chitosan polymers for the nanoencapsulation of proteins and chemotherapeutics.

Kevin started his graduate thesis in 2000 with the Biological Engineering Division at MIT. Working under the joint supervision of Drs. Douglas Lauffenburger and Michael Yaffe, Kevin pursued his thesis entitled, "Quantitative Analysis of the Cytokine-Mediated Apoptosis-Survival Cell Decision Process." While at MIT, Kevin was supported by a graduate fellowship from the Whitaker Foundation.

After graduation, Kevin will begin a postdoctoral fellowship at Harvard Medical School in the Department of Cell Biology with Dr. Joan Brugge. His work will focus on using an *in vitro* model of mammary gland development to understand the molecular mechanisms of tumor initiation and progression in the breast. Kevin recently received postdoctoral fellowship support from the American Cancer Society for his future project entitled, "Cell specification during mammary acinar morphogenesis *in vitro*."

ACKNOWLEDGMENTS

This thesis would not have been possible without the tremendous support from colleagues, advisors, family, and friends. Foremost, I must thank my two closest collaborators: John Albeck and Suzanne Gaudet. The contributions of John and Suzanne are distributed throughout this work and are too numerous to mention. In the same breath, I thank my three mentors: my coadvisors, Doug Lauffenburger and Mike Yaffe, and committee member, Peter Sorger. Their enthusiasm, flexibility, and support were invaluable during the course of this project.

Other collaborators provided key contributions to specific aspects of this thesis. I thank Melissa Shults and Barbara Imperiali for initiating the project described in Chapter 2.3, as well as Rebecca Fry and Manlin Luo for the emerging project described in Chapter 5.5. John Burke and Kate Leitermann provided some of the data presented in Chapter 1.3. In addition, I received critical contributions from three undergraduate researchers—Jason Kelly (Chapter 3.3–3.5), Lili Peng (Chapter 2.2), and Brian Chase (Chapter 2.4)—as well as assistance from three lab helpers: Andrea Urmanita, Stephanie Reed, and Jerry Adler.

On a personal note, I thank my parents for their unconditional support of my work and my continued education. Last, but certainly not least, I thank my fiancé Vivian Fan, whose love counterbalances for me the daily tribulations of scientific research.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 1

Introduction

1.1. Complexity in biological signal transduction networks

1.1.1. Introduction

Cytokines and their receptors regulate cell function and fate via a complex network of signaling cascades (here, "cytokine" refers generally to peptidyl ligands [20]). Cytokine-receptor interactions have been studied extensively, as have the intracellular kinases, adaptor proteins, GTPases, and transcription factors that constitute the signals [21]. The binding of a single cytokine to its cognate receptor typically induces multiple signals that vary with time. Despite the wealth of information about individual signaling proteins, systems-level understanding of the signal transduction network remains poor [22]. Consequently, many attempts to target signaling proteins with small-molecule and biological therapeutics have been disappointing, notwithstanding clear connections between signaling mutations and human diseases such as cancer [23]. In this Introduction, we discuss some of the challenges in studying biological signaling networks and in relating these networks to the control of cellular phenotypes.

1.1.2. Signaling specificity

One confounding feature of signaling networks is that multiple cytokines activate the same signaling cascades while still eliciting different physiological effects [24]. Epidermal growth factor (EGF) and insulin, for example, both activate extracellular-regulated kinase (ERK) and the kinase Akt, but EGF is primarily mitogenic whereas insulin regulates metabolism [25,

26]. For receptor tyrosine kinases (RTK's) such as the EGF and insulin receptors, there even exists some evidence that downstream signaling pathways might confer generic signals, which are subsequently tuned by cell-specific expression of transcription factors to activate target genes [27, 28]. Though provocative, these studies conflict with knockin models of pathway-deficient platelet-derived growth factor (PDGF) and c-Met receptors [29, 30]. In these studies, replacement of the endogenous receptor with one lacking the ability to signal through particular branches of the network conferred a subset of the defects present in the corresponding knockout. These results suggest that endogenously expressed receptors convey distinct information through different pathways in the downstream network.

How then is this signaling specificity achieved? In mammalian cells, the predominant mechanism is thought to be a combination of selective posttranslational modifications and protein-protein associations [31]. For RTK's, ligand binding causes the phosphorylation of key tyrosine (Tyr) residues, which serve as phospho-dependent docking sites for Src-homology 2 (SH2) or phosphotyrosine binding (PTB) domain-containing adaptor proteins. These adaptors recruit cytoplasmic kinase and other proteins to the membrane for activation, which subsequently leads to phosphorylation cascades mediated by serine-threonine (Ser-Thr) and dual-specificity protein kinases [21]. Serine-threonine phosphorylation by these transducer and effector kinases modulates transcription-factor activity and thus gene expression [32]. In addition, these phosphorylation events initiate a second class of phospho(Ser-Thr)-dependent interactions through WW and Forkhead-associated (FHA) domain-containing proteins, as well as through proteins like 14-3-3 [33].

For each event in this cascade of signals, specificity is conferred by both sequence and structure. Protein kinases preferentially phosphorylate Ser, Thr, and Tyr residues based on the flanking amino acids surrounding the site [34] as well as more distant docking sequences [35]. Phospho-dependent binding proteins are likewise preferential towards certain neighboring residues [36], and structural studies have revealed that specificities can be highly complex at the molecular level [37]. Interestingly, protein kinase and phospho-dependent binding proteins only partially overlap in their sequence preference, and neither interaction is "hard-wired" to be 100% specific [33]. Together, this suggests that signaling proteins, as well as their upstream activators and downstream targets, are specific and yet also multifunctional.

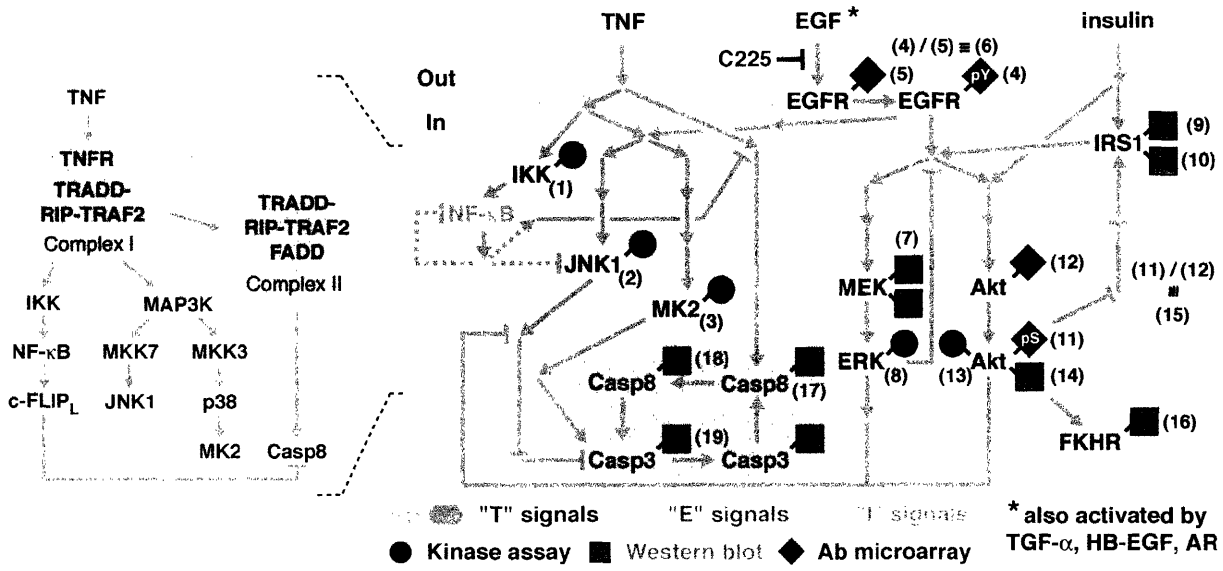

Figure 1-1. Schematic of the recognized intracellular signaling network shared by TNF, EGF, and insulin. Signals were defined as "T" signals (red or black), "E" signals (green), or "I" signals (blue) as described in Chapter 2.4. Dashed arrows indicated transcriptional pathways. Nineteen signals (parentheses) were selected throughout the network and measured by either high-throughput kinase assay [5], Western blotting, or Ab microarray [16]. P/total EGFR (signal 6) and P/total Akt (signal 15) signals were defined by taking the ratio of the phospho and total signals measured by antibody microarray.

1.1.3. Signaling antagonism and crosstalk

An additional complication in most signaling networks is that a single cytokine can activate multiple opposing signals. One clear example of this is the proinflammatory cytokine, tumor necrosis factor (TNF) [38-40]. TNF binding trimerizes the TNF receptor (TNFR) to recruit death domain (DD)-containing proteins, like TRADD and TRAF2, into the so-called death-inducing signaling complex (DISC) (Figure 1-1) [39]. A membrane-bound DISC complex (complex I) activates nuclear factor-κB (NF-κB) via a TRAF2/5–RIP–IκB kinase (IKK) pathway and the MKK7–c-jun N-terminal kinase (JNK) and MKK3–p38 stress kinase pathways via TRAF2–MAP3K [38, 40, 41]. Subsequently, a cytoplasmic DISC complex (complex II) forms that cleaves the executioner cysteine protease, caspase-3, via the binding of the initiator caspase-8 to FADD (Figure 1-1). Thus, TNF promotes prodeath responses by activating caspase-8 and caspase-3 [42], but TNF also activates the nuclear factor-κB (NF-κB) transcription factor, which is typically prosurvival [43]. TNF signaling is further antagonized transcriptionally, in that

complex I signals are downregulated by complex II signals by NF-κB mediated expression of c-$FLIP_L$ [44], which binds to caspase-8 and prevents FADD-induced activation [41].

In tissues, cells are exposed to multiple cytokines that combine synergistically and antagonistically to control cell fate. In colonic epithelia, for example, TNF is a key mediator of inflammatory bowel disease [45] whereas EGF and insulin-like growth factor (IGF) stimulate growth and repair of the mucosa [46, 47]. In many cell types, the proapoptotic functions of TNF are antagonized by EGF [48-50]. Likewise, insulin and the related insulin-like growth factors (IGFs) have also been shown to inhibit many TNF-induced responses [51-55]. For these types of antagonistic stimuli, communication between signaling pathways through crosstalk is critical in determining outcome. Attempts to understand the responses of cells to combinations of cytokines started several decades ago but were hindered by a lack of molecular information on signaling [56, 57].

Most work on the topic of conflicting stimuli has focused on crosstalk among intracellular signaling proteins with pro- and antiapoptotic functions [58]. For example, Akt has been shown to phosphorylate Bad and caspase-9 in response to prosurvival cytokines, which downregulate critical prodeath responses [59, 60]. Likewise, extracellular-regulated kinase (ERK) activity has been shown to prevent the activation of caspase-8 [61]. Conversely, it has been shown that prosurvival EGF-family receptors are targeted for degradation by caspases [62]. These interactions add new stimulus-specific edges to the signaling network and together shift the balance of pro- and antiapoptotic signals toward one fate or another [63]. A similar mechanism has been proposed for cytokines, like TNF, that activate opposing intracellular signals. Within the TNF-induced signaling network, there are several examples of pathway crosstalk, including degradation of prosurvival IKK by caspases [64], NF-κB-mediated downregulation of JNK1 [65], and the aforementioned c-$FLIP_L$ inhibition of complex II [41]. It is not clear why a single cytokine like TNF is endowed with so many overlapping positive and negative feedbacks for controlling cell behavior. However, control theory suggests that overlapping feedback optimizes the sensitivity and stability of systems for a wide ranges of inputs [66].

1.1.4. Signaling dynamics

Controlling cellular behaviors requires the correcting timing of signaling as much as the signal itself. For transcriptional networks in prokaryotes, it has been shown that a "just-in-time" principle applies to key regulators of metabolism and cell-cycle progression [67, 68]. Although less well understood in mammalian cells, the temporal aspects of signaling are a recognized means for controlling distinct cellular processes [69].

The mitogen-activated protein kinases (MAPK's) are perhaps the simplest example of time-dependent signals controlling distinct cellular outputs [70]. One of the earliest observations was in PC12 pheochromocytoma cells, where EGF and nerve growth factor (NGF) both activate ERK with different kinetics. EGF induces transient ERK activity and causes proliferation, whereas NGF induces sustained ERK activity and causes differentiation and neurite outgrowth [71]. A possible mechanism for interpreting transient and sustained ERK signals was discovered using quiescent Swiss 3T3 fibroblasts [72]. In this system, platelet-derived growth factor (PDGF) induces sustained ERK activity and promotes S phase entry, whereas EGF induces transient ERK activity and does not promote entry. It was shown that immediate-early gene (IEG) products induced by the initial transient spike of ERK were themselves ERK substrates. IEG phosphorylation by sustained ERK activity stabilizes the protein products from degradation and triggers S phase entry. Many immediate-early "ERK sensors" have now been implicated in the control of the G_1-S transition [73].

The stress-activated JNK MAPK has also been implicated to exert time-dependent control of cell behavior, in particular, toward apoptosis. The involvement of JNK signaling in apoptosis is controversial [74, 75], but transient JNK activation frequently correlates with survival and sustained JNK with apoptosis [76]. Mechanisms for both prosurvival JNK signaling through JunD phosphorylation [77] and prodeath signaling through Bim phosphorylation [78] have been described. More recently, the kinetics of JNK activity have been linked to reactive oxygen species (ROS), whereby ROS deactivate JNK phosphatases to cause sustained JNK activation and cell death by necrosis [79, 80]. Nevertheless, other JNK-dependent proapoptotic pathways have been suggested [81], and the molecular mechanisms appear to depend heavily on the experimental system [80]. Unlike for ERK [72], it has yet to be determined in a single system how different classes of JNK activation might exert opposing control on apoptosis.

The lack of real-time or high-throughput techniques to quantify MAPK signaling has prevented more detailed studies of dynamics. Although there is a sense that "transient" and

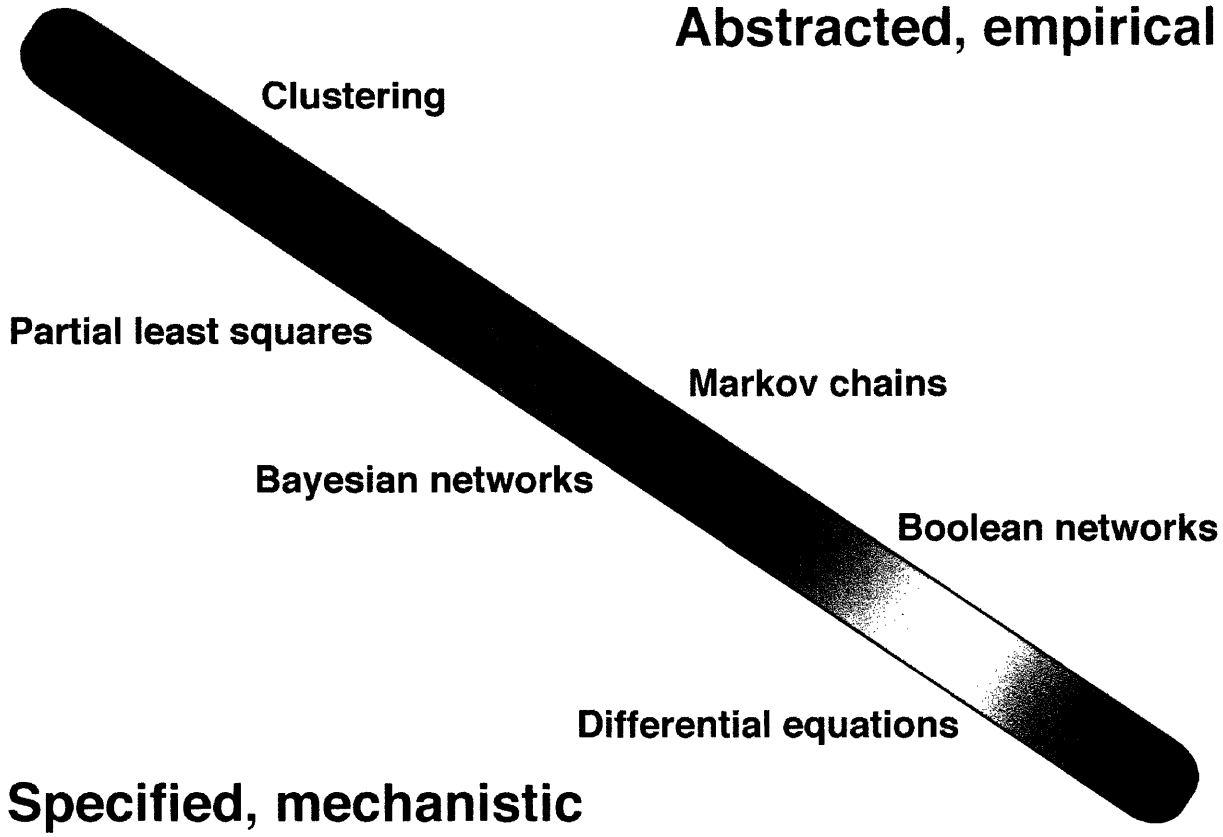

Figure 1-2. Spectrum of computational approaches for modeling protein networks: clustering [7], partial least squares [10], Markov models [15], Bayesian networks [17], Boolean networks [18], and differential equations [19]. The spectrum is not implied to be comprehensive.

"sustained" signals are categorically different (see Chapter 3.2.1), there is no strict definition of what these terms actually mean. This starkly contrasts other signaling fields, such as Ca^{2+} signaling, where real-time single-cell reporters have allowed oscillation frequencies and amplitudes to be quantified from "puffs" and "waves" of Ca^{2+} [82, 83]. Equivalent tools for enzyme-catalyzed cascades could investigate in greater detail the temporal fidelity of these pathways for changing gene expression and cellular phenotypes.

1.2. Clarity through quantitative modeling: the MAPK pathway

Increasingly, systematic methods are being to applied to the interpretation and computational analysis of cell signaling [84, 85]. These numerical methods lie on a spectrum of

approaches that vary in their level of abstraction and specificity (Figure 1-2) [86]. However, by far the most common approach for analyzing intracellular signaling networks involves modeling the fundamental biochemical reactions with ordinary differential equations (ODE's). These reaction-based, "mechanistic" models are meant to formalize the basic processes that underlie signal transduction. Although hurdles exist for constructing mechanistic models (see Chapter 1.3), these approaches are a valuable quantitative test of the aggregate biological understanding of a system [87]. By codifying the existing knowledge, one can then examine if this information is sufficient to explain and, more importantly, predict experimental results. There are several comprehensive reviews on mechanistic modeling applied toward signaling pathways and networks [88, 89]. Here, we restrict our discussion to MAPK signaling [90], where quantitative models have had a frequent scientific impact.

The MAPK signaling cascade is a module of three kinases [91, 92]. First, a MAP3K kinase activates the dual-specificity MAP2K by phosphorylation on two Ser-Thr residues. Then the MAP2K activates a MAPK Ser-Thr kinase by phosphoryating its Thr-X-Tyr motif, where X denotes Glu for ERK, Pro for JNK, and Gly for p38 MAPK. Many theoretical studies have focused on this cascade because (1) the basic biochemical events are agreed upon, and (2) many of the rate parameters have been determined by experiments. Perhaps most importantly, the MAPK cascade appears to be "insulated" from the rest of the signaling network, in that MAP3K's are specific for MAP2K's, and MAP2K's are highly specific for MAPK's.

As an isolated module, the MAPK pathway has been shown computationally to be ultrasensitive, where small changes in inputs signals are propagated through the cascade to cause dramatic changes in output signals [93]. Studies in Xenopus later revealed that MAPK ultrasensitivity plays a critical role in bistability at the single-cell level [94]. Other more recent work in MAPK-like modules has suggested that the deactivating MAPK phosphatases were particularly important in controlling response parameters, like response time and signal duration [95]. Experiments with a MAPK-protein kinase C (PKC) signaling network later corroborated the importance of phosphatases in system behavior [96]. For mechanistic modeling, these studies illustrated the power of a well-defined signaling system, which could be studied in detail and abstracted from the rest of the network [97].

The MAPK pathway has also been analyzed within larger networks of pathways, such as growth factor-induced signaling. Drawing from literature-based parameters, multiple studies

have analyzed the theoretical requirements for the MAPK signal adaptation [98, 99]. Internalization-degradation of growth factor receptors, inhibitory phosphorylation of adaptors and upstream kinases, and GTPase recruitment have all been identified as adaptation determinants for the transient-vs.-sustained MAPK signal (see Chapter 1.1.4 and references in [100]). Here, the mechanistic models were a useful tool for quantifying the strength of negative feedback required to explain the transient ERK MAPK signals measured experimentally [19, 101]. Multiple models correctly identified receptor internalization as a critical determinant for EGF-induced transient ERK signaling [19, 100]. Interestingly, a lack of receptor internalization could not be the mechanism for NGF-induced sustained ERK activation, because NGF receptor (TrkA) internalization is required for sustained ERK signaling [102]. Very recently, a combined experimental-computational study proposed EGF- and NGF-specific differences in upstream G proteins [100]. Ras and Rap1 G proteins activate different Raf MAP3K-family members en route to ERK phosphorylation, and Rap1 is specifically activated by NGF. Ras and Rap1 were shown to act as differential and linear transducers by sensing the rate of change and final levels of the input growth factor signal respectively [100].

The latest mechanistic models of growth factor-induced MAPK [19, 100] and other signaling networks [103, 104] have illustrated an important shift toward experiment and prediction. Quantitative analyses that explain disparate biological conclusions remain highly valuable [105, 106]. Yet, there is a growing impetus for models to incorporate new experimental data and validations along with published values [107]. In the next chapter, we discuss some the challenges associated with constructing realistic and useful mechanistic models of signaling.

1.3. Challenges in constructing and validating mechanistic models

Despite experimental advances [108], the biological data for any signaling system is currently insufficient to constrain the unknown (and therefore flexible) parameters of mechanistic models. Moreover, the existing data are fragmented across multiple publications, laboratory notebooks, and cell types [107]. Some have suggested that robustness should be used to evaluate models of large reaction networks [109], but this concept is both vague and

unfulfilling for most biologists. The general uncertainty surrounding mechanistic models has led some to question whether such approaches have really added anything to the qualitative intuition of experimentalists [110]. It is more likely that mechanistic models of signaling networks will be indispensable for quantitatively framing biological information in the future. However, extreme caution is needed when constructing and validating these models to avoid straying from real biological mechanism and into phenomenology.

1.3.1. Computational challenges

Mechanistic models will always be imperfect, but they should strive to reflect the existing biochemical and cell-biological knowledge as faithfully and realistically as possible. In particular, because the system of ODE's describes the aggregate dynamics of many biochemical reactions, the mathematics must be constrained by the biophysical limits of these reactions. For instance, there are important biophysical constraints embedded in any bimolecular protein-protein interaction:

$$A + B \underset{k_r}{\overset{k_f}{\rightleftharpoons}} AB$$

The above reaction scheme is ubiquitous in mechanistic models of signaling networks. An A-B interaction could be a simple binding event or the initial formation of an enzyme-substrate complex before a catalyzed reaction (see below). The challenge for most mechanistic models is that A, B, AB, k_f, and k_r are usually unknown or partially known: A is estimated from a different cell type, k_f is calculated from proteins of another species, etc.

Protein abundances (here, molecules A and B) vary between 10 nM–1 μM for most signaling molecules [103]. There are, of course, exceptions to these ranges, but they can serve as a useful reality check for published estimates, which sometimes neglect to correct for dilution after lysis [111]. The reverse rate constant (k_r) varies widely with the affinity of the interaction, but importantly, the forward rate constant (k_f) is inherently constrained by diffusion. Theoretical and experimental evidence suggests that most k_f values fall between 10^5–10^6 $M^{-1}s^{-1}$ [112], with the fastest, electrostatically assisted associations occurring near the diffusion limit of 10^9–10^{10} M^{-}

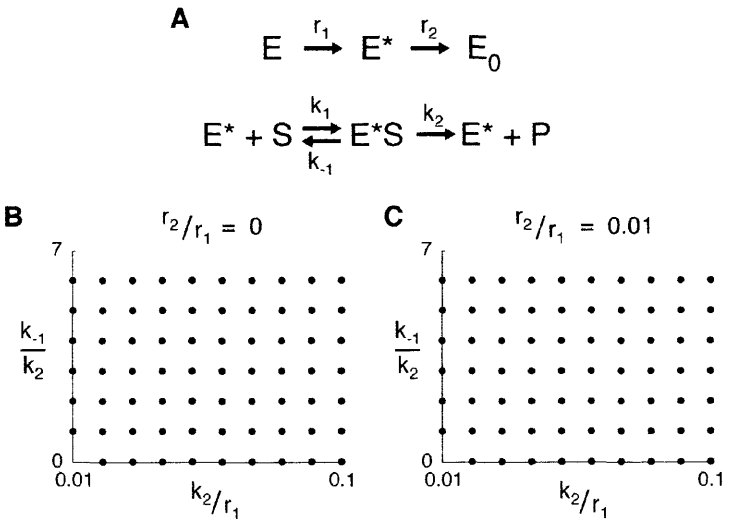

Figure 1-3. Michaelis-Menten kinetics are inappropriate for modeling reaction cascades. (**A**) Reaction scheme of a two-tiered signaling cascade, where the enzyme is subject to activation at a rate r_1 and deactivation at a rate r_2. The active enzyme (E*) then catalyzes the conversion of substrate (S) to product (P). (**B and C**) Comparison between the Michaelis-Menten approximation and the full numerical solution for a range of k_{-1} and k_2 when (B) $r_1/r_2 = 0$ (i.e., no deactivation) or (C) $r_1/r_2 = 0.01$ (i.e., slight deactivation). The Michaelis-Menten equation was compared against the steady-state numerical solution and scored blue (solutions within 10% of each other) or red (greater than 10% discrepancy). The enzyme-to-substrate ratio was kept constant at nine.

$^1s^{-1}$ [113]. This simple diffusional constraint provides an simple means for questioning the validity of models with k_f values as high as 10^{13} $M^{-1}s^{-1}$ [114] or with unlisted reaction parameters [115]. Similar arguments can be made to constrain k_f and k_r further by enforcing their consistency with published binding affinities [116].

In addition to bimolecular interactions, signaling network models are also encoded with many enzyme-catalyzed conversions of substrate (S) to product (P):

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \xrightarrow{k_{cat}} E + P$$

Throughout biochemistry, we are taught how to "simplify" the above reaction scheme by rearranging the rate constants (k_1, k_{-1}, and k_{cat}) into the classic Michaelis-Menten equation:

$$\frac{d[P]}{dt} = \frac{k_{cat}[E]_T[S]_0}{\frac{k_{cat} + k_{-1}}{k_1} + [S]_0} = \frac{V_{\max}[S]_0}{K_M + [S]_0}$$

The two kinetic parameters of this equation (K_M and V_{max}) are widely available for many signaling enzymes and therefore represent a rich source of information for mechanistic models. However, the Michaelis-Menten equation is, itself, a model of enzyme kinetics, with the critical assumptions that the active enzyme concentration is constant ($d([E]+[ES])/dt = 0$) and that enzyme-substrate complex (ES) is at pseudo-steady state ($d[ES]/dt = 0$).

The pseudo-steady state approximation works very well with purified enzymes and substrates during the time scale of most biochemical experiments. However, this same assumption is problematic when considering all ES complexes in a signaling network. Of particular concern is the dynamic behavior of ES complexes in enzymatic cascades, where the enzyme activity is regulated (Figure 1-3A). Using a simple two-reaction cascade with a regulated enzyme, we compared the final product formation estimated by Michaelis-Menten kinetics to that predicted by the full system of ODE's. With constant enzyme activity, the Michaelis-Menten approximation matched the full numerical solution for a broad range of input concentrations (Figure 1-3B). However, when the enzyme was dynamically regulated, there were significant discrepancies between the two models for most starting values of E and S (Figure 1-3C).

Because cells are not bounded by Michaelis-Menten kinetics, it is recommended to use the full set of biochemical ODE's (characterizing E, S, ES, and P dynamics) when modeling signaling networks. Unfortunately, this implies that K_M and V_{max} values from experiments cannot be used directly in the rate equations. Nevertheless, the Michaelis-Menten parameters add important algebraic constraints for optimizing unknown rate constants. For instance, the k_1, k_{-1}, and k_{cat} fitted in a model can be constrained so that $(k_{cat} + k_{-1})/k_1$ must equal the published K_M for the reaction. In this way, K_M and V_{max} aid model construction similarly to the biophysical constraints discussed above.

1.3.2. Experimental challenges

Computational cell biology distinguishes itself from theoretical biology by its firm grounding in experiment and measurement [107]. Biological validation is an essential component of mechanistic models, but not all experiments are equally valid for this purpose. In fact, experiments with deceptive "quantitative" readouts are possibly more damaging to a model

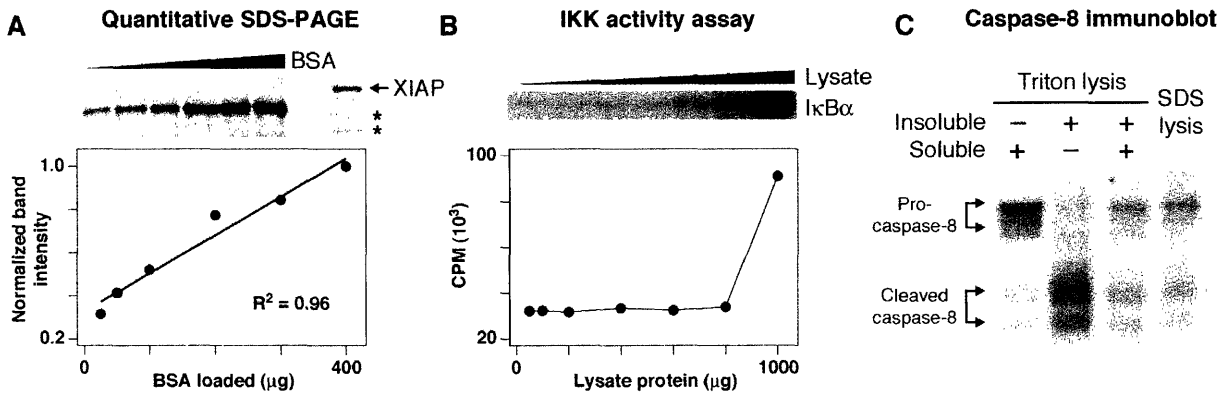

Figure 1-4. Validating mechanistic models requires careful experimental controls. **(A)** Bovine serum albumin (BSA)-based calibration of recombinant X-linked inhibitor-of-apoptosis protein (XIAP) to validate this standard as a calibration for initial condition calculations. Proteins were visualized with SYPRO-Ruby (upper panel) and quantified by fluorescence densitometry (lower panel). Asterisks indicate XIAP cleavage products that should be excluded from the quantification. **(B)** A quantitative, but nonlinear, IκB kinase (IKK) activity assay. Reaction conditions were same as described [5], but with an anti-IKKα antibody from Pharmingen. Phosphorylated IκBα was quantified by autoradiography (upper panel) and liquid scintillation (bottom panel). **(C)** Differential partitioning of procaspase-8 and cleaved caspase-8 into Triton-soluble and insoluble fractions. HT-29 cells were treated with 100 ng/ml TNF for 24 hr and Triton or SDS lysates prepared by standard techniques.

than no data at all. The two types of measurements that are both experimentally tractable and informative for mechanistic models are 1) the starting concentrations, or initial conditions, of signaling molecules and 2) the dynamic response of a signaling molecule to a defined stimulus. Ideally, both approaches should be used to constrain models [19], but unfortunately this is rare.

Quantification of initial conditions is both experimentally challenging and tedious. Because these data are difficult to obtain, we have found that is most productive to prioritize effort on the signaling molecules to which the model is most sensitive [114]. Using quantitative Western blotting, a lysate of known cell density is regressed against a calibration curve from recombinant protein standards of the signaling molecule of interest [19]. Importantly, we have found that commercially available proteins have unreliable stock concentrations, due to cleavage products and contaminants. Therefore, we calibrate each recombinant standard against an albumin standard by quantitative SDS-PAGE and SYPRO-Ruby staining (Figure 1-4A) before proceeding with Western blotting or antibody microarrays [16].

Since most biochemical assays quantify changes relative to a positive or negative control, it is essential for the linearity of any assay to be verified explicitly. Dilution series are a common way to detect nonlinear measurement biases, and we have disqualified several assays because of highly nonlinear response characteristics. This includes assays that measure the final assay

product with quantitative detection formats, such as autoradiography and liquid scintillation (Figure 1-4B). Recently, we have shown that the performance of statistical models can be fragile to nonlinear data [117]. It will be interesting to examine the prediction stability of mechanistic models after similar perturbations to the training data.

A significant logistical challenge for experimentally validating computational models is to retain meaningful biological variation, yet control for spurious experimental variation. These consistency controls provide critical normalization factors to fuse large experimental datasets coherently [12]. Like the assays themselves, the choice of controls is essential, because the results can be disastrous if normalizations are superficial or nonexistent.

One familiar example involves loading controls for Western blots. Structural proteins, such as β-actin and α-tubulin, are frequently used to control for sample-to-sample variations in cell number and protein content. However, in studies where significant cytoskeletal rearrangements are involved, these controls might be inappropriate. This is particularly true for mechanistic models of apoptotic signaling networks [114], because actin is a known target for degradation by apoptotic caspases [118]. In general, protein variation can be controlled by quantifying total protein beforehand using a bicinchoninic acid assay.

Another implicit control for validating mechanistic models is the choice of lysis recipe during sample preparation. Commonly, nonionic detergents (such as Triton X-100, NP-40) are the preferred chemical agent, because the retain protein complexes and activities more effectively. However, in control studies, we have found that certain signaling molecules, such as caspase-8, can partition between the Triton soluble fraction (the clarified extract) and the insoluble fraction (which is usually discarded) (Figure 1-4C). This translocation is widely appreciated in adhesion and migration signaling [119, 120]. But for mechanistic models, it is technical detail with significant implications when fitting biochemical data of this kind. By using clarified extracts alone to analyze partitioned signals, there would appear to be a significant "degradation" mechanism that would then be falsely encoded in the model. This experimental artifact can be resolved with whole cell extracts (soluble + insoluble fractions) for assays like Western blotting and clarified extracts only for functional or coassociation studies.

1.3.3. Summary

The value of large-scale quantitative experiments characterizing signaling is clearly recognized [87]. Combined with new measurement technologies [108], this suggests that high quality datasets will soon be available for modeling the reactions of complex networks [121]. It remains unlikely, however, that these data will be sufficient to constrain all of the parameters embedded in a mechanistic network model (see Chapter 1.3.1). Furthermore, as datasets grow, it becomes increasingly difficult to interpret and reconcile what these data actually mean [122]. It is here that less detailed modeling approaches, which are founded on the measurements themselves rather than on mechanism, can be valuable. In the next chapter, we discuss some recent insights that have been gained from data-driven models of signaling networks.

1.4. Models of data

Data-driven models lie on the opposite end of the spectrum when compared with mechanistic models (Figure 1-2) and are usually rooted in statistics and probability. Data-driven models are powerful because they can handle empirical data and prior knowledge about the underlying phenomena in a flexible fashion. Unlike mechanistic models, data-driven models do not require fine-grained knowledge about the interrelationship between measured components. Model constraints and assumptions are still formalized mathematically, but these methods allow dataset to "speak for themselves" in terms of global variations and covariations. We focus here on two data-rich fields where analysis of biological signals and phenotypes has provided higher-level insight: biological interaction networks and gene expression-based classifications.

Transcriptional and interaction networks are now available for a variety of organisms, including *Caenorhabditis elegans*, *Drosophila melanogaster*, and *Saccharomyces cerevisiae* [123]. Despite concerns about the accuracy of these large datasets [124], many groups have analyzed these connection maps to study network organization and topology [125]. Terms like "scale-free" have been drawn from network theory to describe biological interaction maps [126]. More recently, these data have been mined further by focusing on phenotypically-relevant subsets of the network, such as the DNA-damage subnetwork [127]. These phenotypically-focused subnetworks were shown to be more tightly interconnected than nonessential genes but less tightly interconnected than essential genes [128]. The patterns of the connections,

themselves, have also been shown to be overrepresented in biological networks [129]. Interestingly, these patterns, or "network motifs", are significantly different from the motifs that emerge from engineered networks [130], suggesting distinct organizing principles.

These analyses of biological networks are more accurately data "mining", rather than true data modeling, because they lack predictive ability. That is, one looks for an overrepresented network topology [131, 132] or a network "hub" [133], but predicting the function of this element is not possible from the mining algorithm used to identify it. This inability to predict is endemic to the current interaction data, which is mostly static and without interventions or perturbations [134]. Other data-rich fields, like transcriptional profiling, contain a wealth of these types of measurements, and gene-based predictions have been in development for some time. For various cancer types, expression profiles have been used to predict tumor class [135] as well as patient outcome [136, 137]. An additional benefit to these models is that key predictor variables (here, transcripts) help to suggest biological hypotheses about the link between gene products and the studied tumor [138]. Extensions to this work are currently being pursued in combination with newer technologies, like laser capture microdissection [139] and multicolor flow cytometry [140]. There is tremendous interest for applying similar approaches to protein signaling and expression [141, 142]. However, because of a lack of experimental methods to quantify many proteins in parallel, these studies have been limited to biomarker discovery, rather than predictive modeling. In the next chapter, we detail several techniques developed to provide the kind of quantitative, multiplex protein data required to begin modeling signaling networks from a data-driven perspective.

CHAPTER 2

Techniques and approaches for systematically measuring signaling networks

2.1. Introduction

Complex patterns of signal transduction arise when cells are exposed to combinations of extracellular cues that vary in onset, duration, origin and synchrony. Cells process these cues through an interconnected network of multifunctional, redundant molecules to elicit a set of phenotypic responses that subsequently impact function at the cell, tissue and organ level. In order to develop a molecular understanding of the complex pathophysiology underlying human diseases and utilize this information for prognosis and therapy, a systems-level, network-biology approach should be applied to the signaling networks governing the relevant cell responses [143]. This approach will require frequent temporal sampling of protein activity at critical nodes within parallel signaling pathways inside the cell in a quantitative manner to accurately characterize the flow of information. Such functional measurements are likely to be as valuable, or more valuable, than measurements of simple protein abundance. By quantitatively exploring the functional response of the signaling network to distinct extracellular cues and correlating these molecular events with phenotypic responses, one can construct predictive models of cue-signal and signal-response relationships.

Evolving proteomic approaches to network biology have largely focused on measuring abundances of many proteins at only few time points or under a limited number of experimental conditions [144]. Complementary information on functional protein characteristics, such as enzyme activity, has been lacking in these systematic analyses, in large part because there do not exist quantitatively robust, high-throughput techniques to simultaneously measure multiple protein activities in cells. Initially, this type of data collection on protein functional status should

focus on frequent sampling of a limited number of key molecules that sit at critical nodes in different signaling pathways (Figure 1-1).

2.2. A high-throughput multiplex radioactivity-based kinase activity platform

To characterize cytokine-induced signaling networks more quantitatively, we developed a generalized assay for the multiplex analysis of multiple protein kinase activities in a 96-well format. The procedure utilizes parallel kinase-specific immunopurification steps, followed by rapid quantitative high-throughput activity measurements within the linear-rate regime for each kinase. We applied this technique to measure the activities of five kinases (extracellular-regulated kinase (ERK) [145], Akt, Iκ-B kinase (IKK), c-jun N-terminal kinase 1 (JNK1), and mitogen-activated protein kinase-associated protein kinase 2 (MK2)) in a model system for tumor necrosis factor-alpha (TNF-α)-induced colon epithelial cell death.

The generalized assay format involves four steps, shown in Figure 2-1: i) parallel purification of endogenous kinase from whole cell lysates by immunoprecipitation onto Protein A or G microtiter plates that have been precoated with kinase-specific antibodies, ii) low stringency washes to remove nonpurified lysate components, iii) addition of [γ-^{32}P]ATP and a

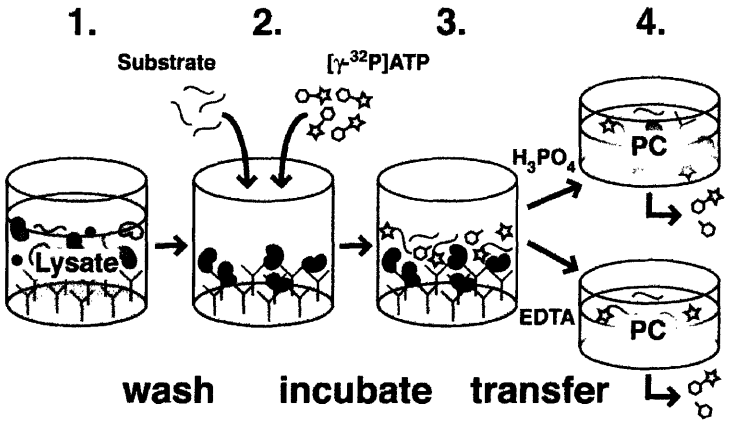

Figure 2-1. General schematic of the high-throughput multiplex kinase activity assay format. Lysates are incubated with Protein A or G microtiter wells precoated with anti-kinase antibody. After several washes, an appropriate substrate and [γ-^{32}P]ATP are added to the plate to initiate an *in vitro* phosphorylation reaction. The reaction is terminated with H_3PO_4 (for Akt and JNK1 assays) or EDTA (for IKK and MK2 assays), and a fraction of the reaction mix is transferred to a phosphocellulose (PC) filter plate and washed to remove free ^{32}P.

kinase-specific substrate to initiate an *in vitro* kinase reaction, and iv) termination of the *in vitro* reaction either by H_3PO_4 or EDTA and liquid transfer to a 96-well phosphocellulose (PC) filter plate to isolate phosphoproteins and remove free [γ-^{32}P]ATP. We developed assays for ERK [145], Akt, JNK1, IKK and MK2 activity in cell extracts, after rigorous optimization and quantitative validation.

To develop the parallel immunopurification step, we experimentally screened multiple commercially available products and identified for each kinase a single, high-affinity antibody that retained enzymatic activity. Coating conditions for each anti-kinase antibody on Protein A/G microtiter plates were individually optimized, revealing that the antibodies maximally bound their intended targets when 50 μl of 10 μg/ml antibody was applied to each well (Figure 2-2A–D, *left panels*), consistent with our estimates of the number of antibody binding sites on the plate surface. Higher coating concentrations of polyclonal antibodies (anti-JNK1, anti-IKKα/β and anti-MK2) reduced the solid-phase avidity for the kinase and caused a decrease in purification efficiency (Figure 2-2B–D, *left panels*). At the optimized coating concentration (Figure 2-2A–D, *left panels, arrows*), the immunopurification was always linear in the amount of kinase purified over a substantial range of lysate concentrations (Figure 2-3A–D). This demonstrates that the antibody capture step linearly reflects kinase abundance in the lysate.

Next, kinase reaction conditions for each assay were optimized by modifying the following *in vitro* parameters: choice of substrate, concentration of substrate, radioactive-to-nonradioactive ATP ratio and reaction duration (see Chapter 6.2.1.). When an individual kinase was determined to effectively phosphorylate multiple substrates (Figure 2-4A,B,D and data not shown), the substrate showing the highest specific activity (CPM/μmol) was selected to maximize sensitivity. Other reaction parameters, such as buffer composition and reaction temperature, were intentionally kept constant to enable assays of different kinases to be performed in parallel on the same microtiter plate in a single step.

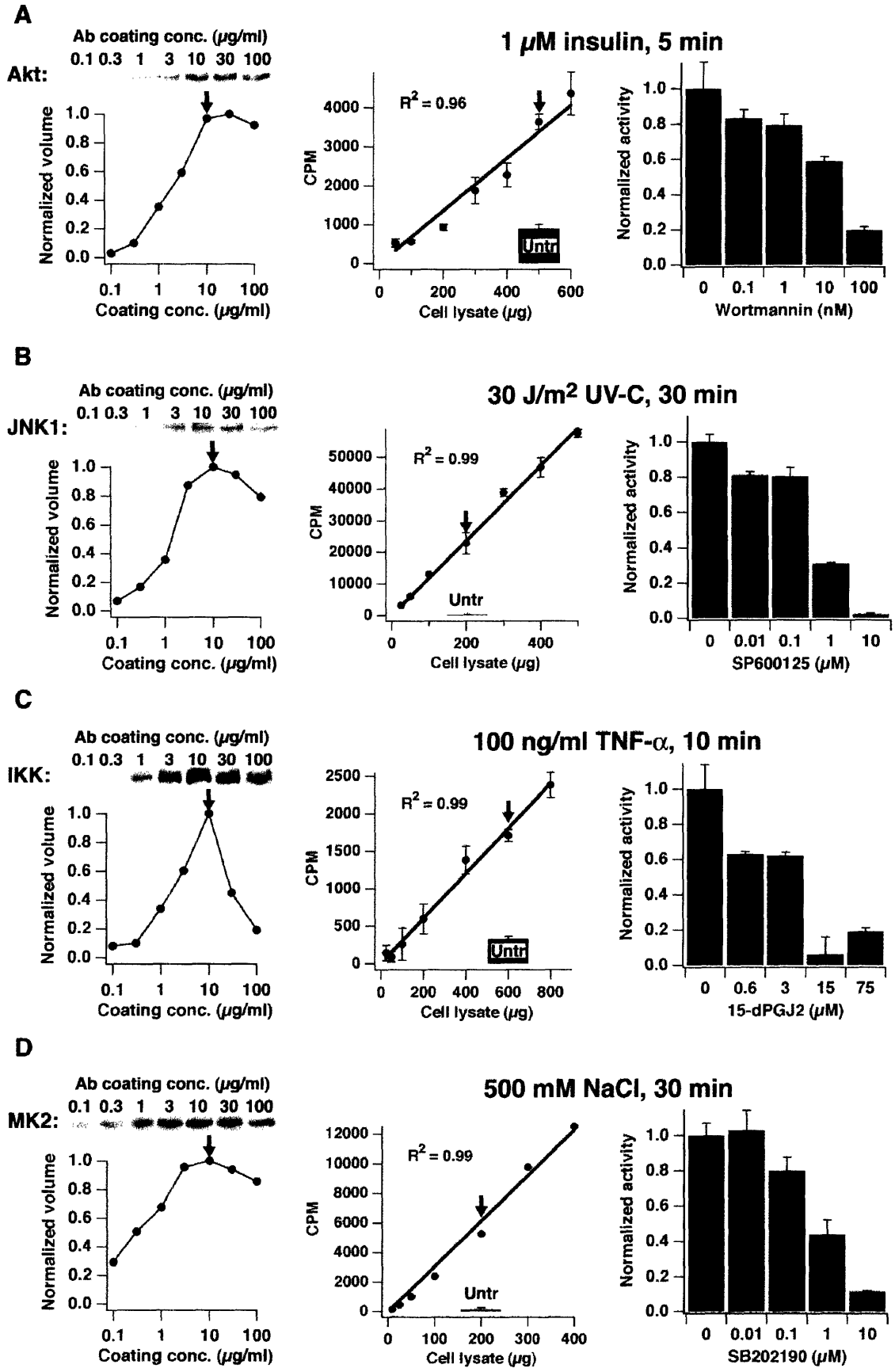

Figure 2-2. (See legend on next page.)

Figure 2-2. (Previous page) High-throughput multiplex kinase activity assays are optimally sensitive, quantitatively linear and specific for the target kinase. **(A)** *Left panel*, Akt was immunopurified from 500 μg of HT-29 lysate on Protein G microtiter plates coated with 50 μl of anti-Akt at various coating concentrations. Plate-bound Akt was analyzed by Western blotting, as described [5], and quantified by densitometry of the band intensity to calculate a normalized volume. *Middle panel*, 50-600 μg of lysate from HT-29 cells treated with 1 μM insulin for 5 min was measured with the high-throughput Akt activity assay using 10 μM Aktide as substrate, as described [5]. The stimulated lysate (*arrow*) was compared with 500 μg untreated control lysate (*solid bar*). *Right panel*, HT-29 cells were pretreated with various concentrations of wortmannin (Sigma) for 1 h then stimulated with 1 μM insulin for 5 min and lysed. Akt activity was quantified as described [5]. **(B)** *Left panel*, JNK1 was immunopurified from 200 μg of HT-29 lysate on Protein A microtiter plates coated with 50 μl of anti-JNK1 at various coating concentrations and analyzed as described in (A). *Middle panel*, 25-500 μg of lysate from HT-29 cells harvested 30 min after treatment with 30 J/m^2 UV-C was measured with the high-throughput JNK1 activity assay using 3 μg ATF-2 as substrate, as described [5]. The stimulated lysate (*arrow*) was compared with 200 μg untreated control lysate (*solid bar*). *Right panel*, plate-bound JNK1 from UV-stimulated HT-29 cells was incubated for 10 min with various concentrations of SP600125 (Calbiochem), and JNK1 activity was quantified as described [5]. **(C)** *Left panel*, IKK was immunopurified from 800 μg of HT-29 lysate on Protein A microtiter plates coated with 50 μl of anti-IKKα/β at various coating concentrations and analyzed as described in (A). *Middle panel*, 50-800 μg of lysate from HeLa cells treated with 100 ng/ml TNF-α for 10 min was measured with the high-throughput IKK activity assay using 10 μg IκBα as substrate, as described [5]. The stimulated lysate (*arrow*) was compared with 600 μg untreated control lysate (*solid bar*). *Right panel*, plate-bound IKK from TNF-stimulated HeLa cells was incubated for 1 h with various concentrations of 15-deoxy-$\Delta^{12,14}$-prostaglandin J_2 (15d-PGJ_2, Calbiochem), and IKK activity was quantified as described [5]. **(D)** *Left panel*, MK2 was immunopurified from 200 μg of HT-29 lysate on Protein G microtiter plates coated with 50 μl of anti-MK2 at various coating concentrations and analyzed as described in (A). *Middle panel*, 10-400 μg of lysate from HT-29 cells treated with 500 mM NaCl for 30 min was measured with the high-throughput MK2 activity assay using 10 μM MK2tide as substrate, as described [5]. The stimulated lysate (*arrow*) was compared with 200 μg untreated control lysate (*solid bar*). *Right panel*, HT-29 cells were pretreated with various concentrations of SB202190 (Calbiochem) for 1 h then stimulated with 500 mM NaCl for 30 min and lysed. MK2 activity was quantified as described [5]. All kinase activities are reported as the mean ± S.E.M. of triplicate samples (error bars for the MK2 assay were smaller than the size of the marker). Western blots were repeated at least twice with similar results; representative images are shown. *Arrows* indicate fixed experimental conditions for the adjacent experiment.

ERK, Akt and JNK1 assays were terminated by adding an equal volume of 75 mm H_3PO_4 to the microtiter reaction well and transferring the well contents to a parallel 96-well PC filter plate for washing and quantitation. MK2 and IKK assays contained unintended phosphorylation products on the plate surface that were released into solution by H_3PO_4 (data not shown). Therefore, these assays were terminated by adding an equal volume of 20 mM EDTA to the microtiter reaction well and transferring the contents to the 96-well PC filter plate preacidified with 75 mM H_3PO_4 in each well. SDS-PAGE analysis of the terminated reaction mixtures transferred to the PC filter verified that phosphorylation was only occurring on the added substrate (Figure 2-4A–E). Therefore individual filters in the wells of the PC plate could be punched out for scintillation counting to accurately quantify kinase activity, eliminating the low-throughput, SDS-PAGE and autoradiography step required in classical immune complex kinase assays.

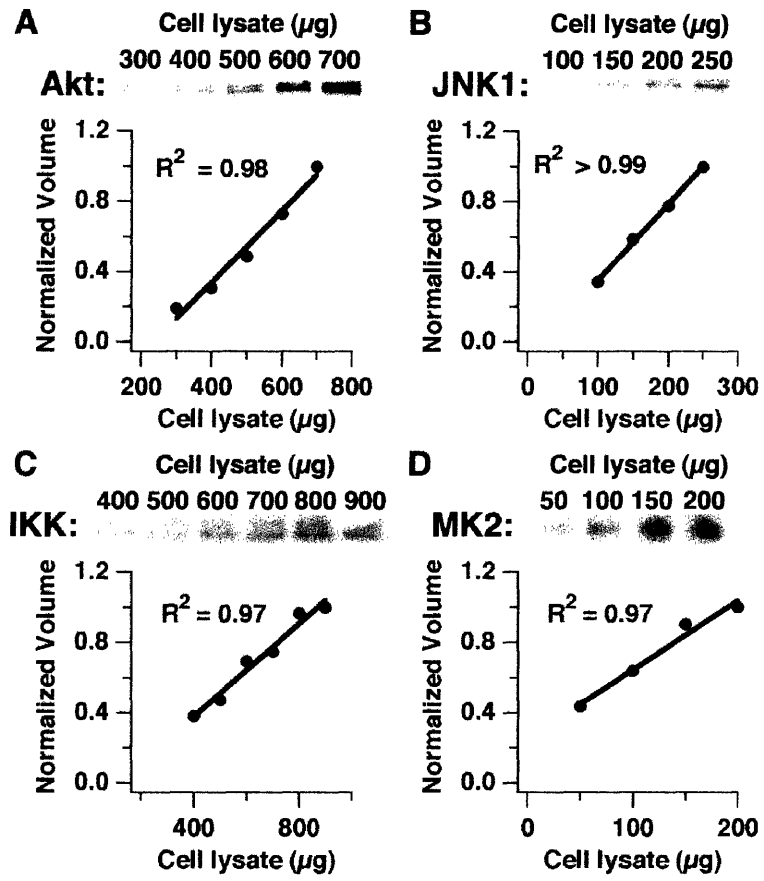

Figure 2-3. Purification of endogenous kinases is linear on antibody-coated microtiter plates. (**A**) 300-700 μg of HT-29 lysate was incubated for 3 h on Protein G microtiter plates coated with 10 μg/ml of anti-Akt. Plate-bound Akt was analyzed by Western blotting, as described [5], and quantified by densitometry of the band intensity to calculate a normalized volume. (**B**) 100-250 μg of HT-29 lysate was incubated for 3 h on Protein A microtiter plates coated with 10 μg/ml of anti-JNK1 and analyzed as described in *A*. (**C**) 400-900 μg of HT-29 lysate was incubated overnight on Protein A microtiter plates coated with 10 μg/ml of anti-IKKα/β and analyzed as described in (A). (**D**) 50-200 μg of HT-29 lysate was incubated for 3 h on Protein G microtiter plates coated with 10 μg/ml of anti-MK2 and analyzed as described (A)

To investigate the sensitivity, dynamic range and linearity of these assays, HT-29 and HeLa cells were treated with known activators of each kinase, and the kinase activities (CPM on the PC filter) measured as a function of different dilutions of the activated lysates (Figure 2-2A–D, *middle panels*). The ^{32}P incorporation was linear over at least an order of magnitude in activity, and the absolute sensitivities of the assays were always below 200 μg, with some assays capable of measuring activity from as little as 10-25 μg of total lysate. This result is comparable to, or better than, existing assays that usually require several hundred μg of lysate for analysis [146-148]. The aggregate sensitivity of the format was more than sufficient to measure ERK, Akt, JNK1, MK2 and IKK activity simultaneously from a single 10 cm plate of HT-29 cells.

The signal-to-noise and reproducibility characteristics for each kinase assay were examined by selecting a single concentration in the middle of the dynamic range of each assay (Figure 2-2A–D, *middle panels, arrows)* and comparing the kinase activities between lysates from stimulated and unstimulated cells. For each agonist, this revealed a relative activation of each endogenous kinase that was comparable in magnitude to that reported in the literature, with intra-assay coefficients of variation always ≤10%. Thus, the assays clearly reflected activation of the endogenous pathways, and the endpoint measures of activity were highly sensitive and reproducible.

In order for these assays to reflect information flow through the signaling network accurately, it was crucial to confirm that our endpoint CPM measurement linearly reflected the kinase activity in the lysate. Therefore, the kinetics of the *in vitro* reaction were examined in detail at a selecting fixed lysate concentration within the linear range (Figure 2-2A–D, *middle panels, arrows*). As shown in Figure 2-5A–D, each phosphorylation reaction displayed linear kinetics up to the time of termination, suggesting that neither the [γ-^{32}P]ATP nor the substrate was significantly depleted over the course of the reaction. Thus, the linearity at every step in the assay procedure (kinase capture, *in vitro* reaction kinetics and endpoint CPM measurement) strongly suggests that these assays are a linear reflection of endogenous kinase activity *in vivo.*

Figure 2-4. (Next page) Terminated *in vitro* reaction products specifically contain the phosphorylated substrate. (**A**) The high-throughput Akt assay, using 10 μM Aktide or 40 μg partially purified histones (Sigma) as substrate, was analyzed by SDS-PAGE and autoradiography after termination of the *in vitro* reaction as described [5]. Note that the phosphopeptide has run off the gel and no other phosphorylated bands are evident. (**B**) The high-throughput JNK1 assay, using 3 μg ATF-2 or 2 μg c-jun (Upstate) as substrate, was analyzed as described in (A). *Single asterisks* indicate ^{32}P incorporation into partially cleaved substrates. (**C**) The high-throughput IKK assay, using 10 μg IκBα as substrate, was analyzed as described in (A). *Double asterisks* indicate nonspecific ^{32}P incorporation (note that these bands are of equal intensity in the blank sample, such that background subtraction will remove their contribution). (**D**) The high-throughput MK2 assay, using 10 μM MK2tide or 5 μg Hsp27 (Upstate) as substrate, was analyzed as described in (A). Note that the phosphopeptide has run off the gel and no other phosphorylated bands are evident. (**E**) The peptide substrates (10 μM Aktide for the Akt assay and 10 μM MK2tide for the MK2 assay) were analyzed on a tricine polyacrylamide gel and autoradiographed after termination of the *in vitro* reaction as described [5].

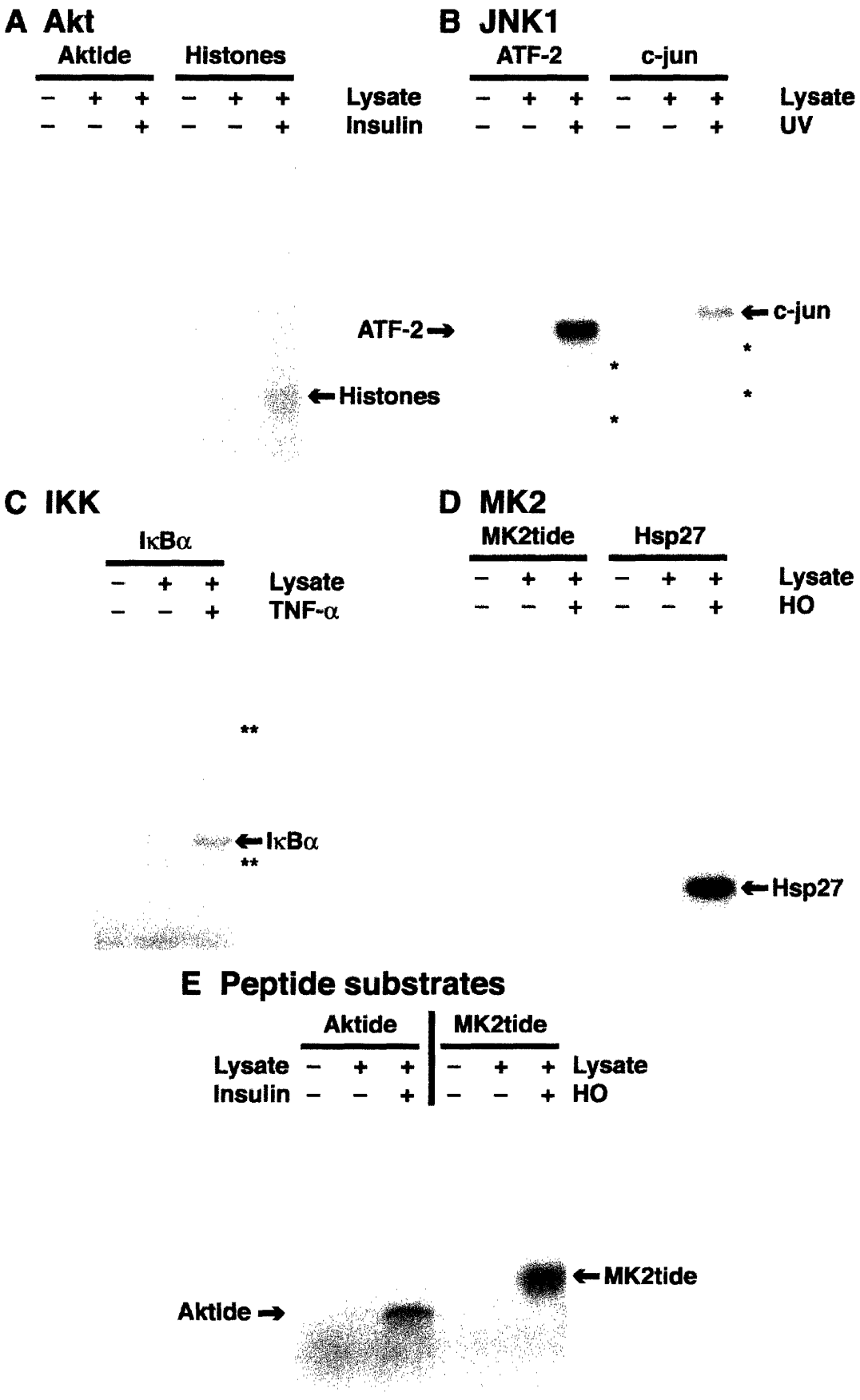

Figure 2-4. (See legend on previous page.)

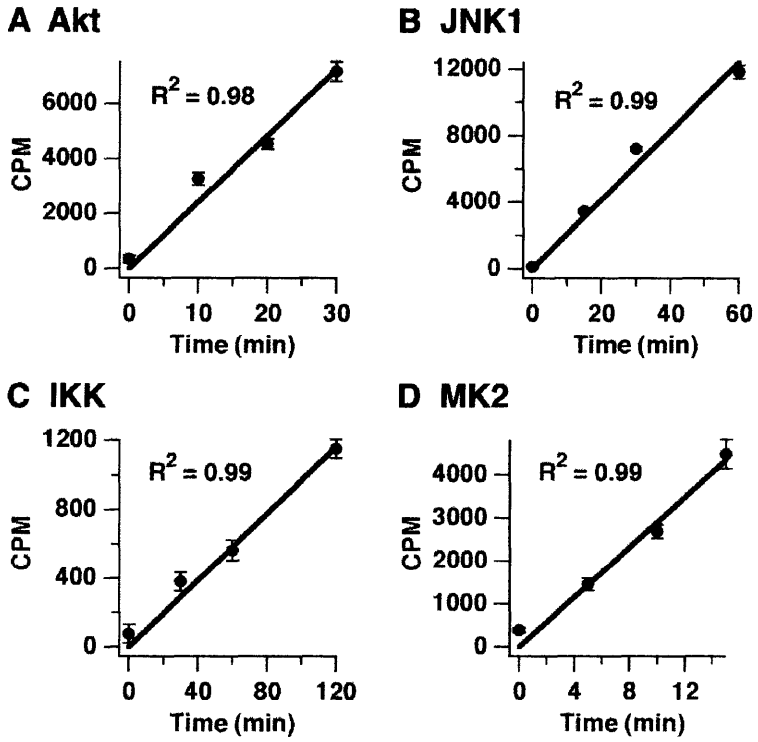

Figure 2-5. Kinetics of the *in vitro* kinase reactions are linear with time. Assays were incubated with fixed amounts of HT-29 or HeLa lysates treated with known activators of the pathway of interest (for details, see Figure 2-2, A–D). *In vitro* reactions were allowed to proceed for the indicated times and analyzed as described [5]. (**A**) Akt assay kinetics with Akt from 500 μg of HT-29 lysate. (**B**) JNK1 assay kinetics with with JNK1 from 200 μg of HT-29 lysate. (**C**) IKK assay kinetics with IKK from 800 μg of HeLa lysate. (**D**) MK2 assay kinetics with MK2 from 200 μg of HT-29 lysate. Western blots were performed at least twice with similar results; representative images are shown.

The current knowledge of intracellular signal transduction is staggeringly complex. To identify network-level properties that affect cell function, it will be necessary to mathematically model the dynamic, multivariate characteristics of signaling proteins within cells [88]. For such models to be realized, quantitative experimental techniques that are both high-throughput and multiplex are needed. The kinase activity assay format presented here represents a first step in this direction. The 96-well, microtiter format is highly versatile, in that it is amenable to scale-up and automated liquid handling, yet tractable for individual scientists and more moderate studies. Importantly, these assays possess linearity, reproducibility, specificity and sensitivity characteristics as good as, or better than, the corresponding low-throughput technique. We anticipate that these functional assays will complement existing proteomic approaches [144] and

find broad applicability towards biological and clinical problems involving signal transduction and human disease.

2.3. A high-throughput multiplex fluorescence-based kinase activity platform

2.3.1. Introduction

As illustrated in the previous chapter, the most common methods for measuring cellular kinase activities involve isolation of the target kinase from cell lysates [149, 150]. Then, *in vitro* kinase assays are performed with protein or peptide substrates and [γ-^{32}P]ATP, and autoradiography or liquid scintillation counting later quantifies the phosphorylated product. Limitations of this format are that the kinase reaction is a "quenched-point" assay, and processing steps needed to isolate the kinase and measure activity are time-consuming and inherently low-throughput. Moreover, sub-physiological ATP concentrations are required (10-50 μM) to avoid overdiluting the [γ-^{32}P]ATP. Such "ATP-starved" conditions diminish the absolute activity of most kinases, which reduces the signal-to-noise ratio of radioactivity-based kinase assays.

To more directly and rapidly analyze crude cell lysates, a continuous, fluorescence-based, homogeneous kinase activity assay is very desirable. Despite numerous reports of kinase sensors in the literature, only a fraction are compatible with unfractionated cell lysates or within cells [151-163]. Moreover, all of these assays either a) exhibit small fluorescence changes, b) involve non-versatile design strategies, or c) require technically challenging microfluidic setups. Thus, despite the recognized importance of protein kinases, there is not a straightforward way to quantify kinase activity from cells.

Previously, a versatile and sensitive fluorescence-based chemosensor strategy was described for monitoring recombinant kinase activity *in vitro* [164]. The chemosensor comprises a small sensing module appended to an optimized peptide substrate for the target kinase. The fluorescence signal is generated when the nonnatural Sox amino acid [4] undergoes chelation-

enhanced fluorescence in the presence of divalent magnesium (Figure 2-6a). The Mg^{2+}-binding affinity of the product phosphopeptide is much greater than the substrate peptide, which results in a large fluorescence increase upon phosphorylation [164].

We found that the sensitivity and selectivity of Sox-based chemosensors are sufficient to measure kinase activities directly from unfractionated cell lysates. Using similar design principles, we engineered new Sox-based fluorescent substrates for three important protein kinases: Akt [165, 166], MK2 [167, 168], and PKA [169, 170]. These chemosensors are excellent reporters of both recombinant enzyme activity *in vitro* and endogenous activity *ex vivo*. We established optimized assay conditions for preferential, quantitative detection of cellular Akt, MK2, and PKA activation. The homogeneous assay format is high-throughput, remarkably straightforward, and reproducible; importantly, it is also compatible with physiological concentrations of ATP (1 mM). This work establishes an important tool for studying the roles of protein kinases in cellular regulation.

2.3.2 Design and validation of fluorescent kinase chemosensors

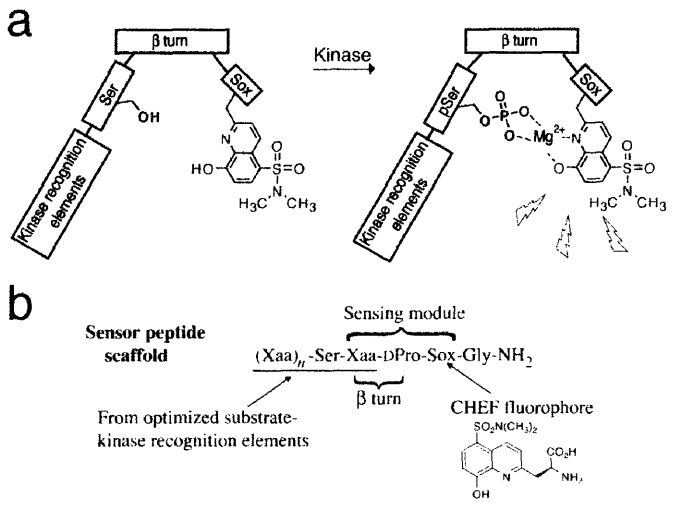

Figure 2-6. Design of fluorescent chemosensors of Akt, MK2 and PKA activity. (**a**) The chemosensor peptide contains three important design modules for fluorescence sensing: critical kinase specificity determinants, the chelation-enhanced fluorophore (CHEF) Sox [4], and a β-turn to preorganize Mg^{2+}-binding between Sox and the incipient phosphate. Phosphorylation increases the affinity of the peptide for Mg^{2+} and the fluorescent signal is generated by chelation of Mg^{2+}. (**b**) The critical kinase specificity determinants in the sensor peptide sequence are derived from an optimized peptide substrate, including one residue in the sensing module as part of the β-turn sequence.

Table 2-1. Chemosensor sequences, sensitivities, and kinetic parameters.

Sensor	Sequence	Fold fluorescence increase	K_M (μM)	V_{max} (μmol/min/mg)
Akt-S1	AcARKRERAYSF-DPro-Sox-Gly-NH_2	9.3	3.8 ± 0.1	0.59 ± 0.08
MK2-S1	AcAHLQRQLSI-DPro-Sox-Gly-NH_2	8.5	21 ± 1	2.3 ± 0.1
PKA-S3	AcLRRASL-DPro-Sox-Gly-NH_2	5.6	2.9 ± 0.3	6.3 ± 0.3

The underlined portion of each chemosensor sequence indicates the series of amino acids from the optimized peptide substrates Aktide [8], MK2tide [11] and Kemptide [13], and the phosphorylatable serine is in bold. Fluorescence enhancements and kinetic parameters (K_M and V_{max}, reported as mean ± s.e.m. for triplicate experiments) were determined as described in [6].

The kinase chemosensor peptides (Akt-S1, MK2-S1 and PKA-S3) were developed from optimized peptide substrates for Akt [8], MK2 [11], and PKA [13] (Figure 2-6b). For each chemosensor, the *N*-terminal portion of the optimized substrate was preserved to retain the consensus motif for substrate recognition. The sensing module was appended *C*-terminal to the target serine. Importantly, the β-turn element in the sensing module tolerates one additional amino acid recognition element for improved specificity and reactivity.

To evaluate the fluorescence increase and kinetic parameters of each substrate, both substrate peptides and phosphoserine (pS)-containing product peptides (Akt-P1, MK2-P1, and PKA-P3) were synthesized. When excited at 360 nm under the final assay conditions, the phosphopeptides were five–ten-fold more fluorescent (λ_{em} = 485 nm) than the corresponding unphosphorylated peptides at identical concentrations (Table 2-1). The absolute fluorescence increases depend on the instrumentation, concentration of Mg^{2+}, presence of other chelators, and ionic strength [6]. Compared with other sensors, this increase is over tenfold larger than most [151-162] and over threefold larger than the most sensitive probe used to date [163]. By fluorescence [164], it was determined that these new chemosensors retained the high activity of the original optimized peptides [8, 11, 13] (Table 2-1).

2.3.3. Development of an Akt-S1 kinase activity assay

The sensitivity and efficacy of these kinase chemosensors prompted investigation of how each would perform in more complex environments. While direct *in vivo* applications are a common next step for most kinase activity sensors [151-162], there are several advantages to

testing sensor properties in cell extracts. For instance: a) sensitivity and specificity can be improved by adding inhibitors of phosphatases and off-target kinases, b) quantification of fluorescence changes is more straightforward, and c) multiple distinct lysate-based assays can be performed from different aliquots of the same sample.

We first developed an Akt-S1 activity assay, because this chemosensor had the largest fluorescence increase after phosphorylation. An initial concern for a homogeneous assay was compatibility between the lysis buffer (containing crude cellular proteins, nonionic detergents, and various inhibitors) and a typical kinase assay buffer [171] (containing kinase and phosphatase inhibitors, ATP, and substrate) (see Chapter 6.2.2). Typically, 40-100 μg of cell lysate in 7.5 vol% of the total reaction was used in the Akt-S1 assay. Radioactive Akt assays [147, 171] typically employ 500 μg lysate and micromolar concentrations of ATP. Here, less lysate is needed because the Akt-S1 assay is compatible with 1 mM ATP. We found that assay sensitivity improves by an order of magnitude when the ATP concentration was increased from 10 μM to 1 mM (Figure 2-7).

Another challenge for a crude lysate assay was the presence of competing cellular processes, such as off-target kinases and phosphatases. To improve specificity for Akt, a cocktail of phosphatase and kinase inhibitors was included in the assay buffer (see Chapter 6.2.2). We confirmed that none of these kinase inhibitors affected recombinant Akt1 activity,

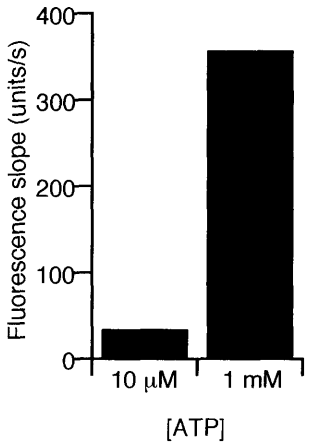

Figure 2-7. Akt-S1 kinase assay sensitivity improves by an order of magnitude when the ATP concentration is increased from 10 μM to 1 mM. Insulin-treated HT-29 lysates were assayed as described [6]. Akt-S1 kinase activity was quantified by calculating the change of fluorescence in the reaction mixture during the initial reaction phase (see Figure 2-9).

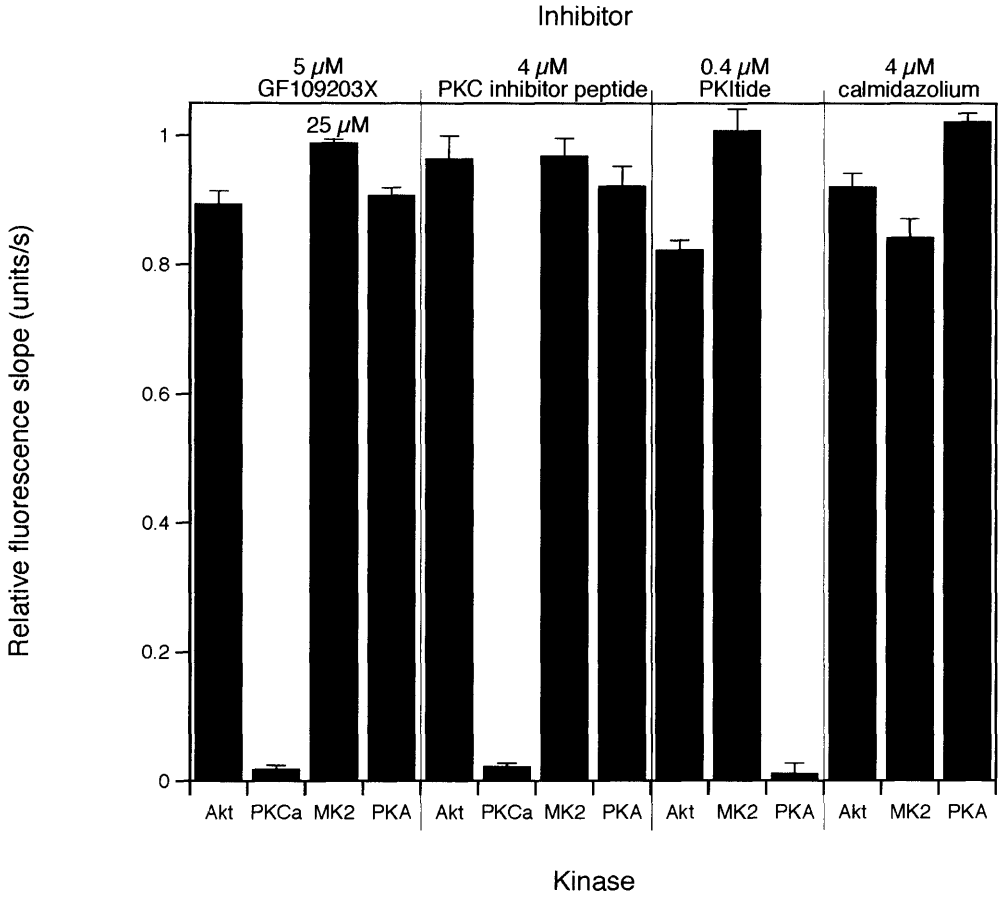

Figure 2-8. Kinase inhibitors used in the assay mixture do not affect recombinant Akt, MK2 and PKA activity, but each inhibitor does completely inhibit its target enzyme. Recombinant Akt1, MK2, PKA and PKC_{α} were assayed with the indicated concentrations of GF109203X, PKC inhibitor, PKItide and calmidazolium. Akt-S1 kinase activity was quantified by calculating the change of fluorescence in the reaction mixture during the initial reaction phase (see Figure 2-9). Fluorescence slopes are normalized a kinase reaction lacking inhibitor. Recombinant Akt1, MK2, PKA assays were conducted as described in [6] with 10 μM chemosensor peptide. PKC_{α} was assayed with 10 μM Ac-Sox-PGSFRRR-NH_2 [14] in Buffer A (see [6]) containing 0.3 mM $CaCl_2$, 5 μg/ml phosphatidylserine (porcine brain, Avanti Polar Lipids) and 1 μg/ml diacylglycerol (dioleoyl, Avanti Polar Lipids). Plotted values indicate mean ± s.e.m for triplicate measurements.

while completely inhibiting their target enzymes (Figure 2-8). For the kinase assay, it was critical to add the small-molecule bisindolylmaleimide PKC inhibitor, GF109203X, which targets all PKC isozymes as well as p70S6 kinase and MAPKAP-K1 [172] (see Chapter 6.2.2).

The Akt-S1 kinase activity assay was optimized in a fluorescence microcuvette and then adapted to a 96-well glass-plate format [6]. To obtain Akt activity measurements, cell lysates

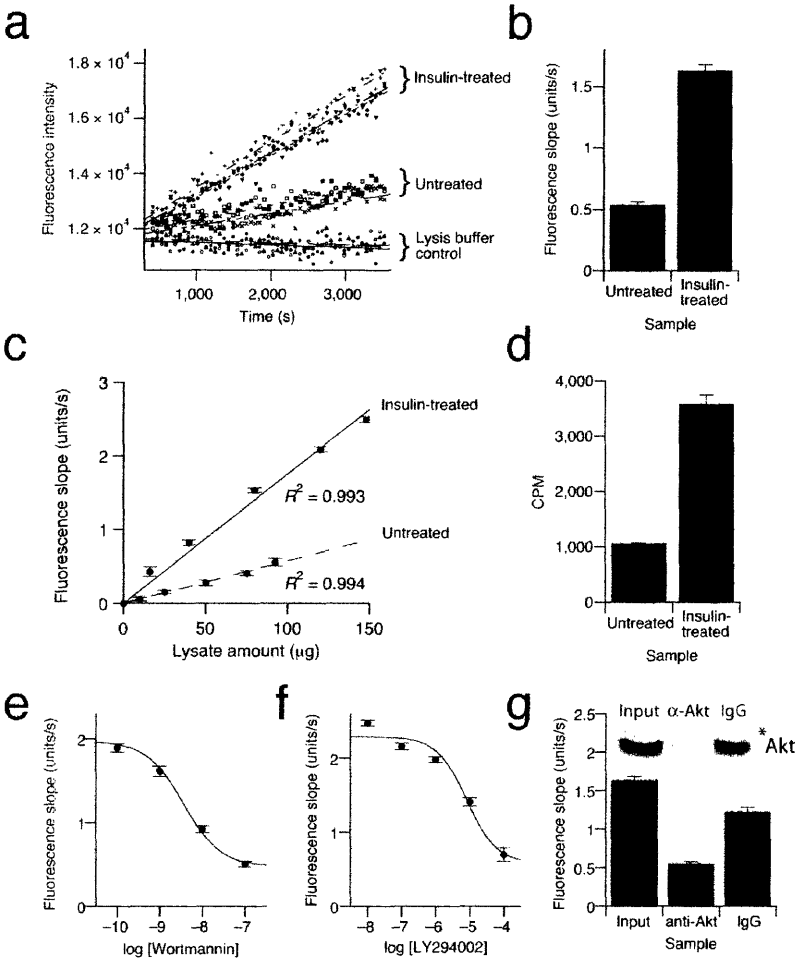

Figure 2-9. Akt-S1 kinase activity is quantitatively linear and preferential for Akt. (**a**) Fluorescence measurements of lysis buffer (●,○,▲), untreated HT-29 lysates (■,□,×) and insulin-stimulated HT-29 lysates (+,▼,◆) over 60 min. R^2 values range from 0.72 to 0.97 (**b**) Summary of fluorescence slopes from data in (a). (**c**) The insulin-treated (—) and untreated (- -) samples were diluted in lysis buffer to vary the total protein content, and the volume of lysis buffer was kept constant throughout all data points. (**d**) Results of an immune complex radioactive assay [5] using the same lysates as in (a) with unmodified Aktide [8]. (**e,f**) HT-29 lysates were pretreated with various concentrations of the PI3K inhibitors (**e**) wortmannin and (**f**) LY294002 for one hour prior to stimulation with 500 ng/ml insulin. The carrier (DMSO) was kept constant at 0.1% for all stimulation conditions. (**g**) Kinase activity was measured from an insulin-treated HT-29 lysate (input), and following immunodepletion of this lysate with anti-Akt1/2 or naïve mouse IgG. Note that the residual activity in the anti-Akt sample is similar to the residual activity in the maximally inhibited samples in (e) and (f). Inset: Western blot for Akt in the measured samples (asterisk indicates a nonspecific band). Plotted values indicate mean ± s.e.m for triplicate measurements. Akt-S1 kinase assays were performed as described in [6].

were added to the complete assay mixture, and fluorescence readings were taken each minute for 60 min (see Chapter 6.2.2). The change in fluorescence is linear during this time (Figure 2-9a), and the slope was used as the measure of activity since it is the best aggregate metric for the increase in fluorescence. All subsequent data were collected in this high-throughput format.

2.3.4. Validation of the Akt-S1 kinase assay

After generically characterizing lysate activity, we determined if the assay format could quantitatively detect changes in activity from cell samples stimulated with known Akt activators. HT-29 cells were treated with insulin for 5 min and lysed by standard techniques. Then, Akt-S1 fluorescence changes were monitored in triplicate with insulin-stimulated or untreated-control lysates (Figure 2-9a). The assay revealed a threefold increase in Akt-S1 kinase activity (Figure 2-9b). Using recombinant active Akt1, the untreated and insulin-treated lysates were found to correspond to about 8 ng and 23 ng active Akt1 equivalents per 93 μg lysate, respectively. The amount of turnover in the insulin-treated samples after 60 min was 7%.

The Akt-S1 fluorescence slope varied linearly with dilutions of both insulin-treated and untreated lysates over at least a fivefold concentration range of cellular protein (Figure 2-9c). In addition, the fold-change between insulin-stimulated and unstimulated activities correlates with

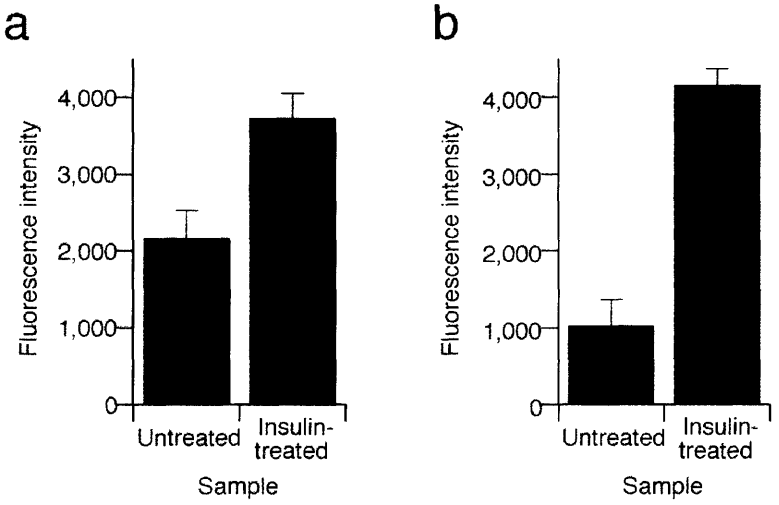

Figure 2-10. Quenched-point fluorescence assays with immunopurified Akt report insulin-stimulated phosphorylation of Akt-S1. (**a**) Fluorescence intensities after 30 min with Akt immunopurified from 1250 μg HT-29 cell lysates. Assays were conducted as described in [6]. (**b**) Fluorescence intensities of Akt-S1 kinase assays after 30 min with 90 μg crude cell lysate. Assays were conducted as described in [6]. Plotted values indicate mean ± s.e.m for triplicate measurements. Quantitatively, the extent of Akt activation by insulin in **a** was smaller and less reproducible than both the Akt-S1 kinase assay in (b and Figure 2-9b) and the radioactive Akt assay (Figure 2-9d). The approach could be optimized with improved immunopurification protocols that are more compatible with fluorescence detection and can overcome diffusional limitations for plate-bound kinase.

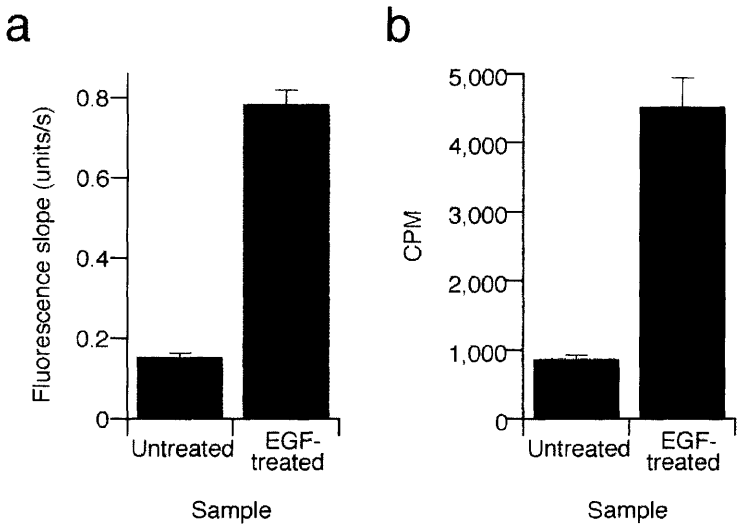

Figure 2-11. Measured Akt-S1 kinase activity in CHO cell lysates is quantitatively comparable to a radioactive assay. (**a**) Akt-S1 fluorescence slopes were monitored over 60 min in triplicate with EGF-stimulated or untreated-control CHO cell lysates. (**b**) Results of a radioactive assay using the same lysates as in (a). Plotted values indicate mean ± s.e.m for triplicate measurements. Lysates were prepared and assays performed as described in [6].

the activation measured by a quantitative radioactive assay [171] (Figure 2-9d). Akt-S1 also reported Akt activity in a quenched-point fluorescence assay with immunopurified Akt (Figure 2-10). These results indicate that the fluorescence-based readings quantitatively report Akt-S1 kinase activity in cell lysates.

Next, we tested whether the Akt-S1 kinase assay was preferential for measuring Akt activity in crude biological samples by using two small-molecule inhibitors (wortmannin and LY294002) of phosphatidylinositol 3-kinase (PI3K). PI3K initiates recruitment of Akt to the plasma membrane; this recruitment is required and sufficient for Akt activation [173]. Cells were preincubated with the indicated concentration of each PI3K inhibitor for one hour, then stimulated with insulin for 5 min. Dose-dependent inhibition of Akt-S1 kinase activity was observed for both wortmannin (Figure 2-9e, IC_{50} = 3.6 nM) and LY294002 (Figure 2-9f, IC_{50} = 8.2 mM). These values agreed with reported IC_{50} values for the upstream PI3K inhibition by wortmannin (5 nM) and LY294002 (1.4 mM), considering differences in ATP concentrations [174]. Since wortmannin and LY294002 inhibit PI3K by different mechanisms [174], sensitivity to both strongly implicates a PI3K pathway in Akt-S1 phosphorylation.

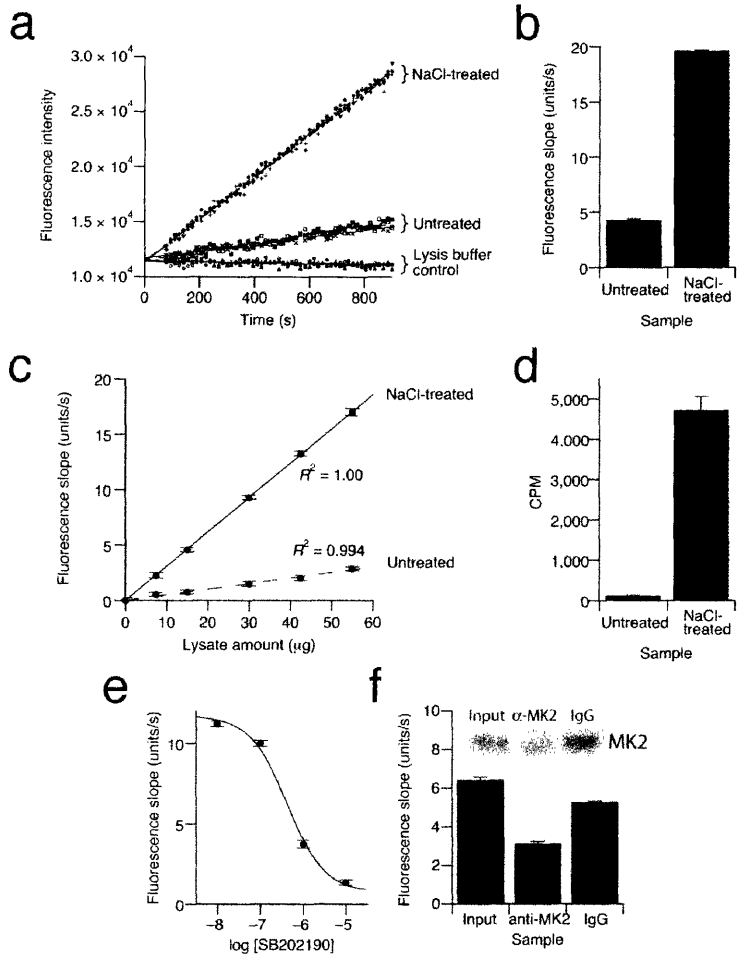

Figure 2-12. MK2-S1 kinase activity is quantitatively linear and preferential for MK2. (**a**) Fluorescence measurements of lysis buffer (●,○,▲), untreated HT-29 lysates (■,□,×) and NaCl-stimulated HT-29 lysates (+,▼,◆) over 15 min. R^2 values range from 0.93 to 1.0 (**b**) Summary of fluorescence slopes from data in (a). (**c**) The NaCl-treated (—) and untreated (- -) samples were diluted in lysis buffer to vary the total protein content, and the volume of lysis buffer was kept constant throughout all data points. (**d**) Results of an immunc complex radioactive assay [5] using the same lysates as in (a) with unmodified MK2tide [11]. (**e**) HT-29 lysates were pretreated with various concentrations of the p38 inhibitor SB202190 for one hour prior to stimulation with 250 mM NaCl. The carrier (DMSO) was kept constant at 0.1% for all stimulation conditions. (**f**) Kinase activity was measured from an NaCl-treated HT-29 lysate (input), and following immunodepletion of this lysate with anti-MK2 or naïve sheep IgG. Inset: Western blot for MK2 in the measured samples. Plotted values indicate mean ± s.e.m for triplicate measurements. MK2-S1 kinase assays were performed as described in [6].

Since other PI3K-dependent kinases might nonspecifically phosphorylate Akt-S1, Akt was removed *in vitro* by immunodepletion. Insulin-stimulated lysates were depleted with an Akt-specific antibody, and Akt-S1 kinase activity was compared to both the input lysate and a naïve mouse IgG immunodepletion. Akt immunodepleted lysates showed threefold less activity than naïve immunodepleted lysates (Figure 2-9g), strongly suggesting that the Akt-S1 signal is predominantly due to Akt-mediated phosphorylation. Immunodepletion was confirmed by

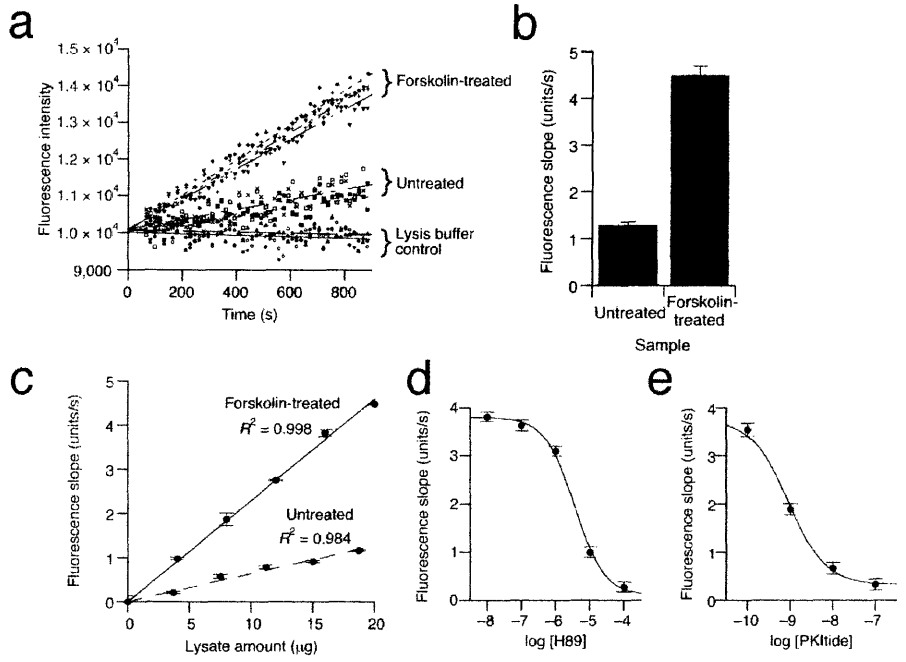

Figure 2-13. PKA-S3 kinase activity is quantitatively linear and preferential for PKA. (**a**) Fluorescence measurements of lysis buffer (●,○,▲), untreated HT-29 lysates (■,□,×) and forskolin-stimulated HT-29 lysates (+,▼,◆) over 15 min. R^2 values range from 0.60 to 0.97 (**b**) Summary of fluorescence slopes from data in **a**. (**c**) The forskolin-treated (—) and untreated (- -) samples were diluted in lysis buffer to vary the total protein content, and the volume of lysis buffer was kept constant throughout all data points. (**d,e**) Forskolin-treated HT-29 lysates were assayed in the presence of various concentrations of (**d**) H89 and (**e**) PKItide. Plotted values indicate mean ± s.e.m for triplicate measurements. PKA-S3 kinase assays were performed as described in [6].

Western blot analysis with the immunodepleting antibody (Figure 2-9g, inset), as well as with an independent anti-Akt antibody (data not shown). Slightly reduced activity in the naïve immunodepletion sample was due to nonspecific loss of Akt (~23% loss, estimated by densitometry). Also, there was a small but detectable amount of residual Akt in the Akt-depleted lysates (~5% of input). Therefore, we consider the immunodepleted samples to be a conservative estimate for the selectivity of the assay.

To test the general applicability of the Akt-S1 kinase assay for other mammalian cell types, we measured Akt-S1 activity in Chinese hamster ovary (CHO) cells. Here, epidermal growth factor (EGF) was used as the Akt agonist (see [6]). Similarly to the insulin-stimulated Akt activity in HT-29 cells (Figure 2-9b,d), we found that EGF-induced Akt activation in CHO cells quantitatively correlated with a radioactive Akt assay [171] (Figure 2-11). Together, these

results indicate that the Akt-S1 assay is a quantitative, selective and generalizable format for studying Akt signaling.

2.3.5. Development of MK2-S1 and PKA-S3 kinase activity assays

After successfully developing an Akt-S1 kinase activity assay, we investigated whether the protocol could be applied more broadly to the other target kinases (Figure 2-6b). To design activity assays for PKA and MK2 in HT-29 cell lysates, the assay mixture from the Akt assay was used as a starting point and then refined for each kinase. We confirmed that the panel of inhibitors for each assay mixture did not affect recombinant MK2 and PKA activities (Figure 2-8).

The MK2-S1 (Figure 2-12) and PKA-S3 (Figure 2-13) activity assays were evaluated in HT-29 lysates similarly to the Akt-S1 activity assay. Both assays were linear for 15 min (Figures 2-13a and 2-14a) and over a fivefold range of lysate concentrations (Figures 2-13c and

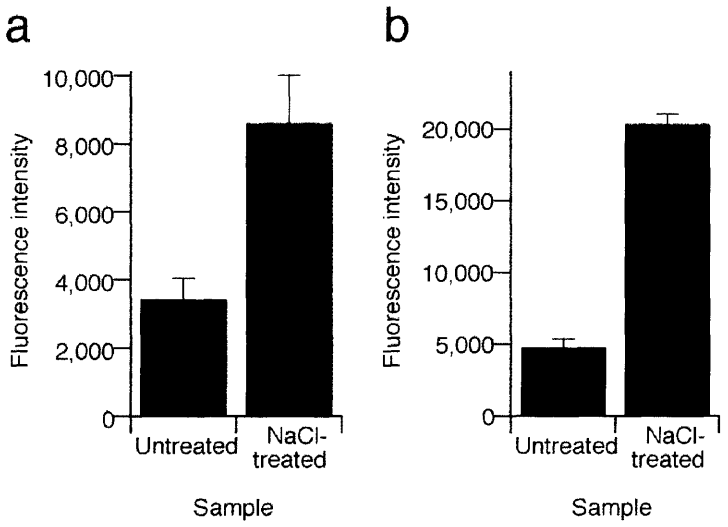

Figure 2-14. Quenched-point fluorescence assays with immunopurified MK2 report osmolarity-stimulated phosphorylation of MK2-S1. (**a**) Fluorescence intensities after 15 min with MK2 immunopurified from 750 μg HT-29 cell lysates. Assays were conducted as described in [6]. (**b**) Fluorescence intensities of MK2-S1 kinase assays after 15 min with 90 μg crude cell lysate. Assays were conducted as described in [6]. Plotted values indicate mean ± s.e.m for triplicate measurements. Quantitatively, the extent of MK2 activation by NaCl in **a** was smaller and less reproducible than both the MK2-S1 kinase assay (b and Figure 2-12b) and the radioactive MK2 assay (Figure 2-12d). The approach could be optimized with improved immunopurification protocols that are more compatible with fluorescence detection and can overcome diffusional limitations for plate-bound kinase.

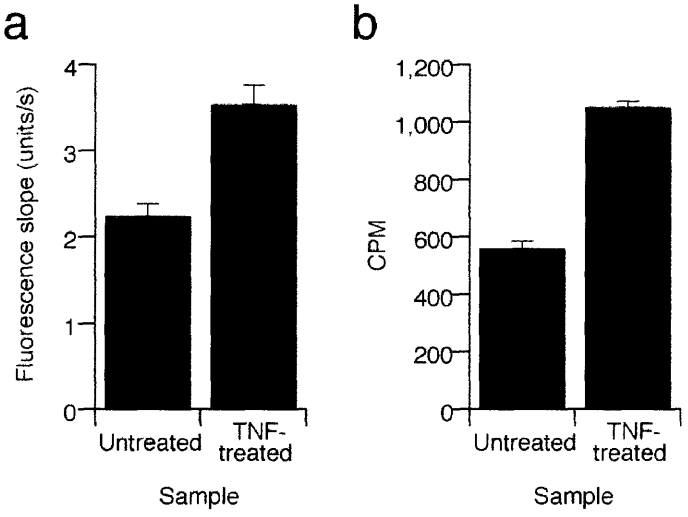

Figure 2-15. Measured MK2-S1 kinase activity in CHO cell lysates is quantitatively comparable to a radioactive assay. (**a**) MK2-S1 fluorescence slopes were monitored over 15 min in triplicate with TNF-stimulated or untreated-control CHO cell lysates. (**b**) Results of a radioactive assay using the same lysates as in (a). Plotted values indicate mean ± s.e.m for triplicate measurements. Lysates were prepared as described in [6]. Assays were conducted as described in [6].

2-14c). MK2 activation by NaCl treatment resulted in a 4.6-fold increase in MK2-S1 phosphorylation (Figure 2-12b). 53 μg of the stimulated lysate contains 51 ng recombinant MK2 equivalents with 18% turnover in 15 min. PKA activation by forskolin treatment resulted in a 3.5-fold increase in PKA-S3 phosphorylation (Figure 2-13b), which corresponds to 12 ng recombinant PKA in 18 mg of HT-29 cell lysate with 9.1% turnover in 15 min.

There was a somewhat weaker correlation between the MK2-S1 assay (Figure 2-12b) and a radioactive MK2 assay (Figure 2-12d), so the chemosensor selectivity was examined in detail. MK2-S1 phosphorylation was abrogated in NaCl-treated lysates when cells were pretreated with SB202190, a small molecule inhibitor of the MK2 kinase, p38 (Figure 2-12e). The apparent IC_{50} value (430 nM) correlates reasonably well with that reported for upstream p38 inhibition (50-100 nM), considering differences in ATP concentrations [172]. When MK2 was removed *in vitro*, the immunodepleted lysate lost 50% activity compared to the input lysate (Figure 2-12f). In addition, Western blot analysis showed that the immunodepleted sample had retained 25% of input MK2, and the naïve immunodepleted sample had nonspecifically lost ~15% of input MK2 (Figure 2-12f, inset). Together, these data suggest that 75% of MK2-S1 activity is due to MK2. The residual 25% nonspecific activity likely explains the higher untreated background activity of

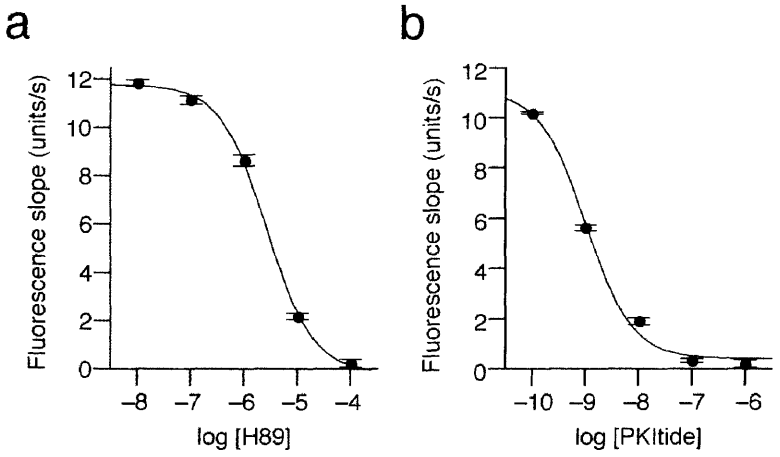

Figure 2-16. Dose-dependent inhibition of recombinant PKA is observed for H89 and PKItide. Recombinant PKA was assayed in the presence of various concentrations of (**a**) H89 (IC_{50} = 2.6 μM) and (**b**) PKItide (IC_{50} = 1.0 μM). Assays were performed as described in [6]. Plotted values indicate mean ± s.e.m for triplicate measurements. These plots match those determined for the PKA-S3 kinase assay (Figure 2-13d,e).

the MK2-S1 assay relative to the radioactive MK2 assay [171] (Figure 2-12d). This could possibly be due to the truncation of additional N-terminal recognition elements in the MK2-S1 peptide relative to the originial MK2tide (Table 2-1 and [11]). MK2-S1 also reported MK2 activity in a quenched-point fluorescence assay with immunopurified MK2 (Figure 2-14). Finally, using the MK2-S1 activity assay, we also detected MK2 activation in CHO cell lysates treated with tumor necrosis factor (TNF) (Figure 2-15). This activation correlated with a radioactive MK2 assay [171] and showed that the assay could be used with diverse stimuli and cell types.

For PKA, an adequate anti-PKA antibody was not commercially available to perform either a radioactive PKA assay or an immunodepletion. Therefore, two mechanistically distinct PKA inhibitors were used to interrogate the selectivity of the PKA-S3 kinase assay: H89, a small molecule ATP-competitive inhibitor, and PKItide, the active fragment of a PKA-specific inhibitor protein [175]. PKA-S3 activity was abolished in a dose-dependent manner with both inhibitors (Figure 2-13d,e). The calculated IC_{50} values (3.6 mM for H89 and 0.91 nM for PKItide) match those for the inhibition of recombinant PKA activity measured under identical conditions (Figure 2-16). This inhibition data provides strong evidence that the PKA-S3 assay is highly preferential.

2.3.6. Summary

The homogeneous, fluorescence-based kinase activity assays reported here have several advantages over existing assays. The Sox-based chemosensors are uniquely quantitative in their ability to estimate enzymatic activity and phosphorylated product generation from cell samples. The amount of lysate activity can easily be related to recombinant protein standards or normalized to untreated control lysates. In addition, the assay is compatible with physiological concentrations of ATP. Together with the solution-phase format, this dramatically increases the sensitivity of the assay. Whereas most immune complex kinase activity assays require 200-500 μg total cell protein (Figure 2-2) [147], the kinase assays presented here can make equivalent measurements at less than 100 μg (Figures 2-9b,d and 2-12b,d). This sensitivity is critical for applications where cell samples are limited, such as high-throughput cell-based screening and clinical diagnostics. In addition, crude cell lysate assays eliminate several hours of manipulations, such as incubation and washing steps, before and after the *in vitro* reaction. During the kinase reaction, the 60 time points collected provide additional activity information and reduce error without extra experimental effort.

An important feature of these lysate-based assays is that selectivity can be improved by pharmacologically inhibiting kinases with overlapping substrate specificity. Because the relevant off-target enzymes will depend upon the peptide chemosensor, kinase-by-kinase optimization will be needed to ensure maximum possible specificity. Although some residual off-target activity is inevitable (here, $< 30\%$ for Akt and 25% for MK2, Figures 2-9g and 2-12g), this singular limitation is outweighed by the many benefits of these fluorescent kinase activity assays, in that they are straightforward, rapid, continuous, non-radioactive, quantitative, and sensitive. The format is conceptually similar to many fluorogenic protease assays [176] that, because of these same benefits, have found widespread use in a number of applications [177]. Finally, our work provides a general protocol for developing assays using other Sox-based chemosensors. Since it is straightforward to multiplex chemosensors to measure several cellular kinase activities in parallel, this assay platform is of immediate and expanding utility in drug discovery and molecular biology.

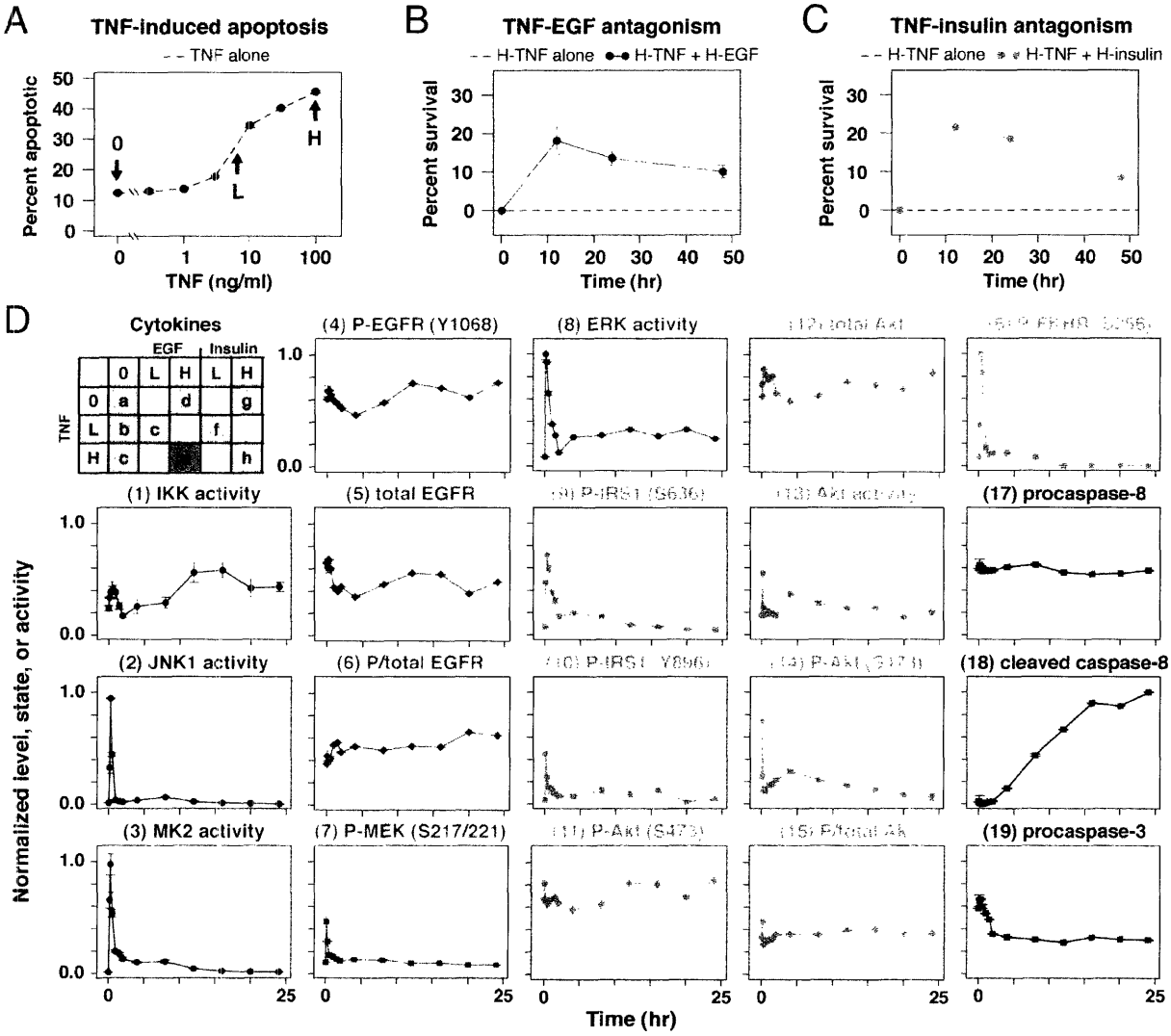

Figure 2-17. A proteomic compendium of antagonistic TNF-EGF-insulin signaling in HT-29 cells. **(A)** Dose-response for HT-29 apoptosis induced by various concentrations of TNF. Apoptosis was measured at 24 hr by annexin V-propidium iodide staining and flow cytometry as described in Chapter 6.2.4. Mock (0), low (L), and high (H) TNF concentrations were defined at 0, 5, and 100 ng/ml (red). **(B and C)** Growth factor antagonism of TNF-induced apoptosis. HT-29 cells were costimulated with 100 ng/ml TNF (H-TNF) and 100 ng/ml EGF (H-EGF, B) or 500 ng/ml insulin (H-insulin, C) and compared against 100 ng/ml TNF alone (red) for apoptosis at 12, 24, and 48 hr as described in (A). **(D)** Signaling network response to 100 ng/ml TNF (H-TNF) and 100 ng/ml EGF (H-EGF). Signals are referenced by number to the targets specified in Figure 1-1. Data are presented as the mean ± S.E.M. of triplicate biological samples as described in Chapter 6.2.4. Nine combinations (a-h) of TNF, EGF, and insulin were similarly measured and are available in [2].

2.4. Proteomic compendium of the TNF-EGF-insulin signaling network

Next, we wanted to use the above kinase activity assay platform (Chapter 2.2) with other quantitative biochemical techniques for studying intracellular signal processing of conflicting cytokine stimuli. To explore the molecular basis of crosstalk between mitogenic and apoptotic signals, human HT-29 colonic adenocarcimona cells were stimulated with TNF in combination with either EGF or insulin. Although we focused this study on HT-29 cells, we confirmed that other cell lines, such as HCT-116 and HeLa, were sensitive to prodeath TNF and prosurvival EGF or insulin treatments[1]. The potency of TNF as a proapoptotic factor in these cell lines is most apparent in combination with immunostimulatory cytokines such as interferon-γ (IFN-γ) [178, 179]. Therefore, cells in this study were pretreated with IFN-γ for 24 hr before cytokine addition.

Nine combinations of TNF, insulin, and EGF, at either subsaturating ("low", "L-") or saturating ("high", "H-") concentrations were added to HT-29 cells without refeeding the medium (see Chapter 6.2.3) (Figure 2-17A–C). In addition, a "mock" treatment ("0"), in which cells were manipulated like any other stimulus but without added cytokine, served as a baseline control (Figure 2-17A and data not shown). At 13 time points spanning 5 min to 24 hr, we prepared triplicate cell extracts and collected 19 protein measurements from each extract (Table 2-2). Based on literature data, the 19 measurements could be roughly characterized as TNF-dependent ("T" signals), EGF-dependent ("E" signals), or insulin-dependent ("I" signals). Since many kinase signals exhibit rapid changes between 0 and 30 min whereas caspases rise slowly in activity over many hours, time points were distributed unevenly, with seven time points concentrated within the first two hours.

Considerable effort was expended to ensure reproducibility and self-consistency of the data [12]. The median coefficient of variation for biological repeats across the full dataset was ~11%, making it possible to reliably compare changes in protein state or activity of ±25%. Inspection of the complete 7,800-measurement dataset revealed both fast-acting signals, such as a transient 50-fold c-jun N-terminal kinase 1 (JNK1) activation that peaked at 30 min, and sustained signals, such as a threefold activation of Akt that lasted for 24 hr. Each cytokine treatment also elicited multiple classes of signals: high TNF activated both T and E signals (data not shown) and high TNF + high EGF elicited T, E and I signals (Figure 2-17D). We therefore conclude that TNF, EGF and insulin signal extensively by crosstalk in HT-29 cells [180].

[1] J.G.A. and P.K.S., unpublished observations.

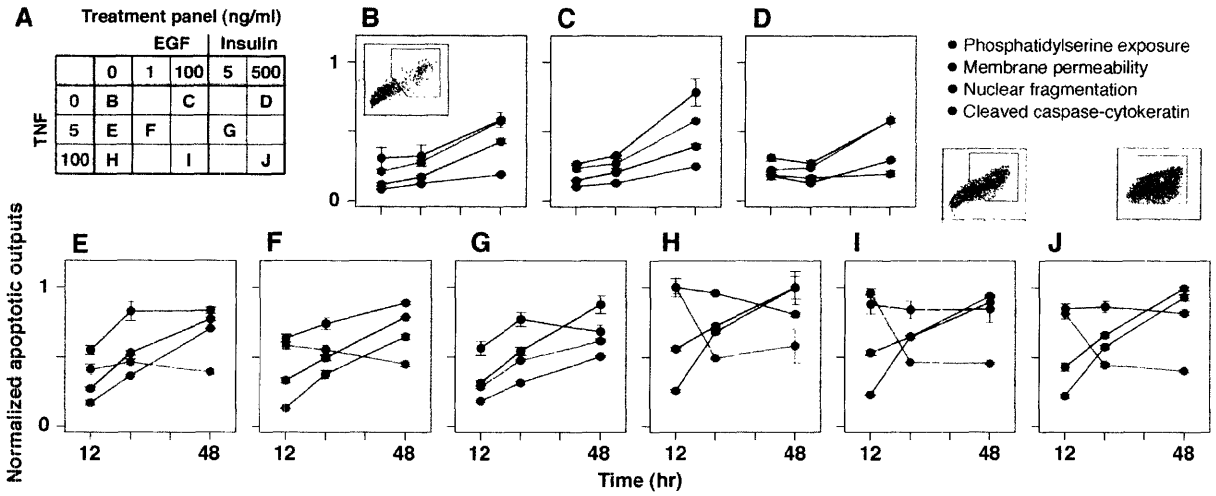

Figure 2-18. A proteomic compendium of TNF-, EGF-, and insulin-induced apoptotic outputs in HT-29 cells. **(A)** Panel of input combinations and concentrations of TNF, EGF, and insulin. Numbers represent the concentration in ng/ml and letters correspond to the subsequent apoptosis subpanel. The carrier (0.02% DMSO) was maintained constant for all treatments. **(B–J)** Apoptotic outputs induced by various combinations and concentrations of TNF, EGF, and insulin. The exact treatment conditions are shown in (A). Data are presented as the mean ± S.E.M. of triplicate biological samples.

In this system, TNF and EGF-insulin are prodeath and prosurvival cytokines respectively [178, 181]. How EGF and insulin antagonize TNF-induced apoptosis is not clear, so we asked whether a complementary set of measurements characterizing cytokine-induced apoptosis could connect the signaling network to the apoptotic phenotype (Figure 2-18A). Apoptosis elicits a variety of cellular changes (outputs) that can be regulated independently [182]. We were concerned that individual parameters used to characterize cell death (e.g., loss of membrane asymmetry as a single output) would only partially reflect the overall cellular response. Therefore, we selected four distinct apoptotic outputs (phosphatidylserine exposure, membrane permeability, nuclear fragmentation, and cytokeratin cleavage) and measured each output response by flow cytometry at 12, 24, and 48 hr after stimulation (Figure 2-18B–J) [1]. Taken together, these output measurements constituted an apoptotic "signature" that characterized early (phosphatidylserine exposure), middle (cytokeratin cleavage and membrane permeability), and late (nuclear fragmentation) outputs of apoptosis.

Table 2-2. Signaling network measurements in the proteomic compendium

Protein name	Function*	Molecular signal†	Assay‡
Inhibitor of nuclear factor-kappaB kinase (IKK)	Ser kinase	Kinase activity	Kinase assay
c-jun N-terminal kinase 1 (JNK1)	Ser-Thr kinase	Kinase activity	Kinase assay
Mitogen-activated protein kinase-associated protein kinase 2 (MK2)	Ser-Thr kinase	Kinase activity	Kinase assay
Epidermal growth factor receptor (EGFR)	Receptor Tyr kinase	Phosphorylation (Tyr1068)	Ab microarray
		Total level	Ab microarray
		Phospho/total ratio	Ab microarray
Mitogen-activated protein kinase and extracellular regulated kinase kinase (MEK)	Dual-specificity kinase	Phosphorylation (Ser217/Ser221)	Western blot
Extracellular regulated kinase (ERK)	Ser-Thr kinase	Kinase activity	Kinase assay
Insulin receptor substrate 1 (IRS1)	Adaptor-scaffold	Phosphorylation (Ser636)	Western blot
		Phosphorylation (Tyr896)	Western blot
Akt	Ser-Thr kinase	Phosphorylation (Ser473)	Ab microarray
		Total level	Ab microarray
		Kinase activity	Kinase assay
		Phosphorylation (Ser 473)	Western blot
		Phospho/total ratio	Ab microarray
Forkhead (FKHR)	Transcription factor	Phosphorylation (Ser256)	Western blot
Caspase-8	Cys protease	Zymogen level	Western blot
		Cleaved level	Western blot
Caspase-3	Cys protease	Zymogen level	Western blot

*Biological functions were assigned based upon the most recognized property of the protein.
†Signals indicate the biochemical property of the protein that was measured.
‡Assay indicates whether the protein was measured by high-throughput kinase activity assay [5], antibody microarray [16], or quantitative Western blotting [12].

The apoptotic signature revealed unique temporal and cytokine dose-dependent features that would have been missed by measuring individual apoptotic outputs alone. Membrane permeability and cytokeratin cleavage measured dead cells cumulatively [183]. Therefore, these outputs increased monotonically with time and TNF dose (Figure 2-18B,E,H) and decreased with EGF or insulin cotreatment at 12 and 24 h (Figure 2-18F,G,I,J). In contrast, phosphatidylserine exposure (before membrane permeability) and nuclear fragmentation (before the formation of apoptotic bodies) were transient cell states [183] that could increase and decrease with time and input dose (Figure 2-18B–J).

Using these distinct output patterns, we observed that high concentrations of EGF and insulin antagonized TNF-induced cell death with similar apoptotic signatures (Figure 2-18I,J). However, low concentrations of TNF with either EGF or insulin elicited different output responses. EGF reduced TNF-induced membrane permeability at 48 h but increased transient phosphatidylserine exposure at 12 h (Figure 2-18E,F; $p < 0.05$). In contrast, insulin reduced TNF-induced membrane permeability at 12 and 48 h, and phosphatidylserine exposure was decreased at 12 h but higher at 48 h (Figure 2-18E,G; $p < 0.05$). These signatures suggested that the different apoptotic outputs could be controlled separately, depending upon the input stimulus. Moreover, it emphasized that connecting the intracellular network to the complete apoptotic signature would require contributions from many molecular signals.

CHAPTER 3

Data-driven approaches for analyzing signaling networks

3.1. Introduction

The new measurement technologies and applications described in the previous chapter allowed us to interrogate signaling networks quantitatively. Yet, large datasets do not directly equate to biological understanding [184]. In this chapter, we describe several approaches used to reduce data complexity and aid hypothesis generation.

3.2. Categorical inspection

3.2.1. Transient vs. sustained signals

The fluorescence-based chemosensors (Chapter 2.3) were used in a multiplex assay format by examining Akt, MK2, and PKA signaling responses to two growth factors: insulin and EGF. HT-29 cells were treated with saturating concentrations of insulin or EGF and lysed in triplicate at various times between zero and two hours. These lysates were then analyzed for Akt, MK2 and PKA activity with their respective kinase assays, and activation levels were calculated by normalizing to the zero minute samples. For each growth factor, 63 independent activity assays were performed in parallel with less than 170 μg cell lysate. By comparison, a similar quantitative radioactive assay [171] for Akt and MK2 would require at least 700 μg cell lysate and would not be possible for PKA.

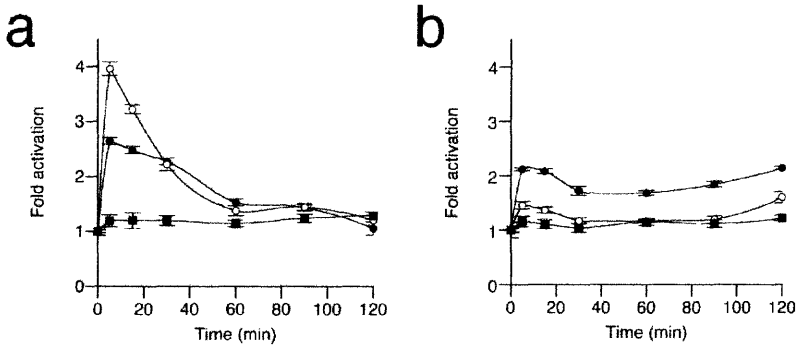

Figure 3-1. The multiplex kinase assay reveals differential responses of Akt, MK2 and PKA in response to stimulation by EGF and insulin. Akt-S1 (●), MK2-S1 (○) or PKA-S3 (■) kinase activities were measured in HT-29 lysates after stimulation by (**a**) EGF or (**b**) insulin for various times as described in [6]. Plotted values indicate mean ± s.e.m for triplicate measurements. Changes in kinase activity were normalized to the baseline (0 min) fluorescence slope.

The multiplexed three-kinase assay revealed qualitatively and quantitatively different responses to EGF (Figure 3-1a) and insulin (Figure 3-1b). EGF activated MK2 more potently and transiently. Akt was also activated transiently in response to EGF, but to a lesser extent, and PKA was not activated at all. In contrast, insulin activated Akt potently, and the response was sustained over the time course. MK2 was slightly activated in response to insulin, and again, PKA was not activated at all. The features of these time courses correlate with those reported for Akt [185, 186] and MK2 [187] (or upstream p38 [188]) activity dynamics in other mammalian cell lines. To the best of our knowledge, there is limited evidence that either EGF or insulin activate PKA.

The transient-vs.-sustained categorization as noted here with Akt is a simple way to classify changes in signaling dynamics. There is an extensive literature on these types of signaling responses in MAPK signaling (see Chapter 1.1.4). These qualitative differences have motivated several computational and experimental studies aimed at understanding their mechanism (see Chapter 1.2). Nevertheless, reducing time-course data into transient or sustained classifications results in data loss, because there do not exist clear quantitative definitions for these types of signaling profiles. How long must a signal persist to be considered "sustained"?

3.2.2. Biphasic signals

Multiphasic signaling patterns are also commonly observed. In a separate study, IFN-sensitized HT-29 cells were treated with 50 ng/ml TNF-α, in the presence or absence of 100 nM insulin cotreatment, and triplicate lysates prepared at 13 time points over 24 h. Since much of the signaling induced by these cues occurs shortly after cytokine addition, time points were more densely sampled in the first four hours (Figure 3-2A,B, *inset*). From these cell extracts, quantitative measurements of ERK, Akt, JNK1, MK2 and IKK kinase activity were performed using the high-throughput multiplex kinase activity assays (Figure 3-2A,B). This set of time

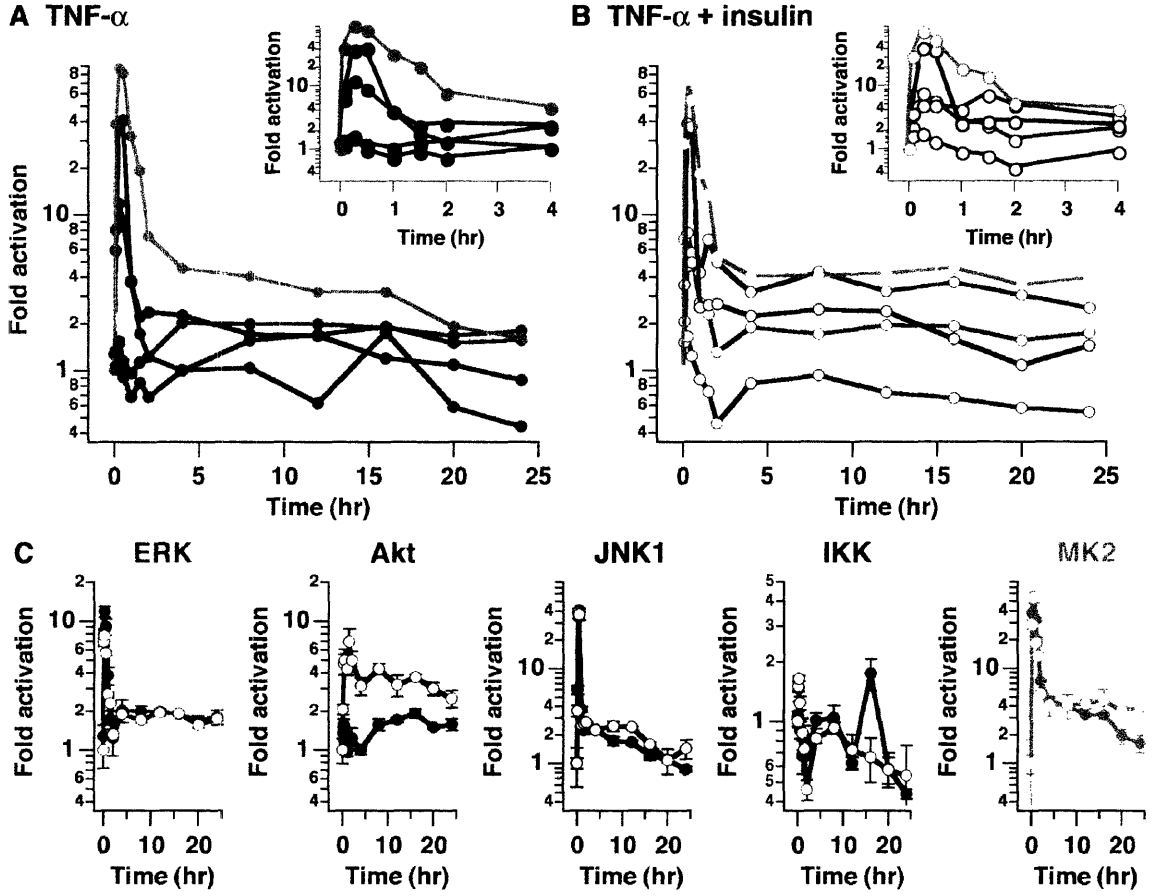

Figure 3-2. TNF-α and TNF-α + insulin treatments elicit quantitatively distinct signaling patterns in HT-29 cells. Endogenous ERK (,), Akt (●, ○), JNK1 (●, ○), IKK (●, ○) and MK2 (,) activities in HT-29 cells in response to: **(A)** 50 ng/ml TNF-α (*solid circles*) and (B) 50 ng/ml TNF-α + 100 nM insulin (*hollow circles*). Lysates were generated at 0, 5, 15, 30, 60, 90 min and 2, 4, 8, 12, 16, 20, 24 h, then measured for kinase activity with the high-throughput multiplex kinase assay. Results are plotted as the mean fold activation of three independent cell extracts. Error bars are omitted for clarity, but are shown in (C). (C) Comparison of ERK, Akt, JNK1, IKK and MK2 activities in response to TNF-α (*solid circles*) and TNF-α + insulin (*hollow circles*) from (A) and (B) plotted as mean fold activation ± S.E.M. of triplicate samples.

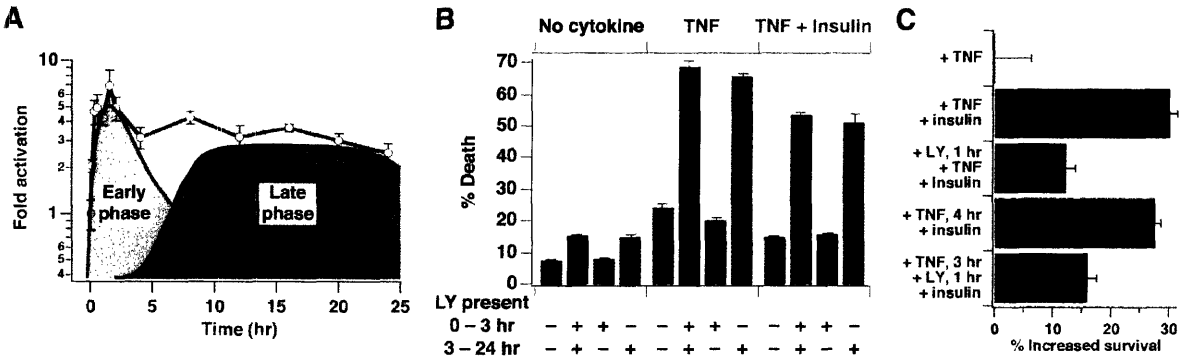

Figure 3-3. Insulin elevates two phases of Akt activity, and the late-phase of Akt activity provides a critical anti-apoptotic signal in HT-29 cells. **(A)** hypothetical two-phase contribution to the sustained Akt activity that was observed in the presence of insulin (Figure 3-2, B and C). **(B)** HT-29 cell death in response to no cytokines (■), 50 ng/ml TNF-α (■), or 50 ng/ml TNF-α + 100 nM insulin ( ) in the presence of 20 μM LY294002 (LY) at the indicated times. Cells were stained for cleaved cytokeratin and cleaved caspase 3 and analyzed by flow cytometry as described in Experimental Procedures. Values represent the percentage of apoptotic cells ± S.E.M. of triplicate samples. **(C)** HT-29 cell survival in response to timed combinations of 50 ng/ml TNF-α, 100 nM insulin and 20 μM LY294002. Values represent the percentage of increased survival ± S.E.M. of triplicate samples, computed as described in Experimental Procedures. The carrier (0.1% DMSO) was kept constant for control experiments.

courses consisted of over 400 independent activity measurements, which would have been extraordinarily laborious and technically impractical by traditional techniques.

To deconstruct the effects of insulin therapy on the signaling network in the context of TNF-α signaling, a pairwise comparison was performed for each kinase under TNF-α with and without insulin costimulation. As shown in Figure 3-2C, the dynamics of activation of some pathways, such as the ERK pathway (Figure 3-2C, green), were essentially superimposable. Thus insulin does not appear to influence activation of the ERK pathway by TNF-α under these conditions. Other TNF-induced kinase activities were affected by insulin in a time-dependent manner. JNK1 activity, for example, was larger at intermediate times (8-16 h), while IKK activity was smaller at 16 h, and MK2 activity was larger at late times (16-24 h). Although potentially important, these transient signaling differences did not clearly reflect the insulin-induced survival phenotype observed at the cellular level [5].

In contrast to the subtle influences of insulin on ERK, JNK1, IKK and MK2 signaling, it was readily apparent that insulin dramatically augmented Akt activity, inducing a rapid increase within five minutes and sustaining activation for 24 h, whereas with TNF-α alone, the Akt response was much smaller in magnitude (Figure 3-2C, *red*). In addition to these quantitative differences, we noted that the dynamic TNF-induced Akt response was also qualitatively

different without insulin, in that the course of activation was clearly biphasic, with a brief initial peak, followed by a small (*i.e.*, twofold), sustained increase in activity after four hours. If these data had not been rigorously quantitative, or had the signaling network not been sampled frequently in time, this increase in activity would likely have been missed. By virtue of the high-throughput activity assays, however, it was clear that Akt activity after four hours was significantly upregulated in comparison to baseline activity ($p < 0.05$ for all times after four hours, Student's *t*-test).

These findings suggested that a significant component of the insulin-induced anti-apoptotic effect was mediated by the high sustained Akt activity, and that this activity profile was derived from the superposition of an early, elevated phase of activity that was transient with a late, elevated phase of activity that was sustained (Figure 3-3A). Conversely, in cells treated with TNF-α alone, these two phases were reduced in intensity and separated in time. Since Akt is thought to provide strong pro-survival signals [189], we hypothesized that one or both of these temporal components of Akt activation were involved in controlling the phenotypic response of HT-29 cells to TNF-α and insulin.

To investigate this, we used a reversible, selective inhibitor (LY294002 [190]) of PI 3-kinase (PI 3-K), an upstream activator of Akt. LY294002 was added or removed at critical times to eliminate the early phase of Akt activity, the late phase of Akt activity or both, and apoptosis was quantified by flow cytometry with an anti-cleaved caspase 3 antibody and the M30 antibody [191] against caspase-cleaved cytokeratin. We have found this double stain to be the most sensitive, quantitatively reproducible measure of apoptosis in HT-29 cells[2].

When LY294002 was added one hour before cytokine addition and tonically maintained to block Akt activity over the entire time course, a dramatic increase in TNF-induced cell death was observed (Figure 3-3B, *red*). Surprisingly, if LY294002 was present only for the first three hours of TNF-α addition, thus permitting only the late phase of Akt activity, there was no increased cell death relative to treatments without inhibitor. In contrast, when LY294002 was first added at three hours after cytokine addition to selectively abolish late-phase Akt activity, we observed an increase in cell death equivalent to that observed when the inhibitor was present for the entire 24 h. An identical pattern was observed in response to TNF-α + insulin (Figure 3-3B, *green*), whereas LY294002 had only a small effect on basal apoptosis (Figure 3-3B, *blue*).

[2] J.G. Albeck, unpublished observations.

These data suggest that late-phase Akt activation is necessary to restrain the percentage of cells undergoing programmed cell death in response to TNF-α in the presence or absence of insulin.

The importance of late Akt activity in controlling TNF-induced cell death, in conjunction with the strong, sustained late Akt activation in the presence of insulin, led to the hypothesis that the insulin-induced increase in cell survival was due, in part, to this increase in late-phase Akt signaling. If true, then addition of insulin up to four hours after the pro-death stimulus should elicit an equivalent reduction in the extent of TNF-induced apoptosis. Indeed, this was observed; the delayed addition of insulin four hours following TNF-α addition resulted in the same increase in HT-29 cell survival as observed when insulin was added simultaneously with TNF-α (Figure 3-3C). Moreover, pretreatment with LY294002 one hour before insulin addition significantly reduced the survival response, regardless whether insulin was added simultaneously or four hours after TNF-α treatment (Figure 3-3C). These experiments strongly implicate late-phase signaling, along a PI 3-K dependent/Akt pathway, as a critical pro-survival mechanism of insulin in colon epithelia.

Later studies revealed that the secondary activation of Akt was dependent upon refeeding during stimulation[3]. When TNF was spiked into the existing medium rather than added with a fresh medium change, the late phase of TNF-induced Akt activity disappeared and cell death increased significantly[4]. Together, these corroborate that late-phase Akt signaling is a potent prosurvival signal in HT-29 cells. Importantly, the identification of a TNF-induced biphasic Akt activation led more readily to a testable hypothesis than the transient-vs.-sustained Akt signaling induced by EGF and insulin (Chapter 3.2.1). This study provided the initial evidence that future studies should reduce signaling data in a way that retained time-dependent information.

3.3. Clustering through principal components analysis

3.3.1. Directed data acquisition and biological significance

[3] K.A. Janes and S. Gaudet, unpublished observations.
[4] K.A. Janes and J.G. Albeck, unpublished observations.

One way to analyze time-dependent information is to treat each time point as a separate observation. Data-driven clustering of time-point observations is most effectively achieved with a data matrix that contains informative, quantitative measurements. Our work focuses on the regulation of programmed cell death in response to the prodeath factor TNF-α and the survival factor insulin. We directed our experimental measurements to a subset of the downstream kinases, caspases, and other regulators that are known to influence cell death or survival (Figure 3-4). In particular, we focused on those proteins whose signaling varies in response to TNF-α, insulin, or both, and whose perturbation, by genetic or other means, affects cell survival. By imposing both these criteria, we increase the likelihood that the data we collect will make a

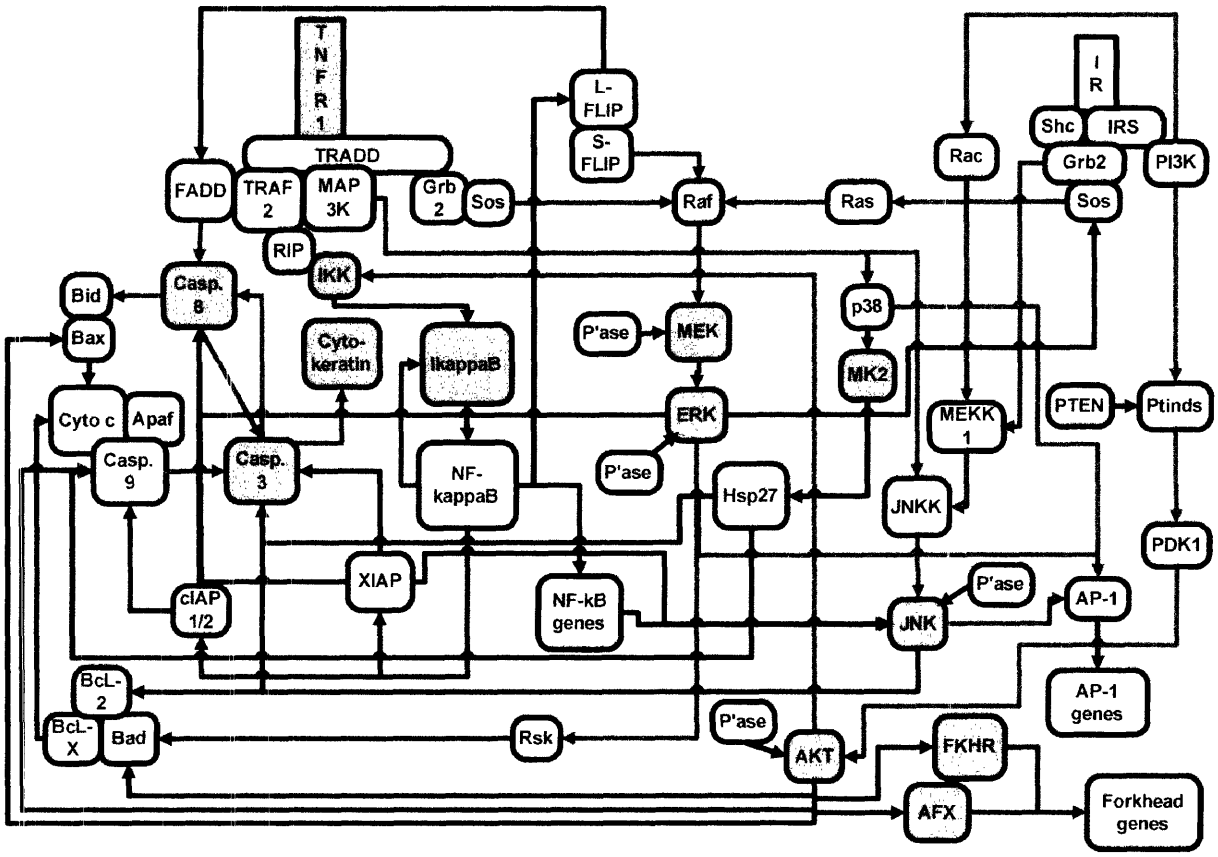

Figure 3-4. Directed data acquisition of the biomolecular signaling network downstream of prodeath cues (TNF-α) and prosurvival cues (insulin) [5]. Network state is determined by many component characteristics, including protein expression levels, protein-protein interactions, protein enzymatic activities, and protein locations. Approximately 30 of these characteristics are measured in quantitative and dynamic terms for multivariate systems analysis. Green arrows indicate activating interactions, red arrows indicate inhibitory interactions, and blue arrows indicate transcriptional interactions. Shaded nodes highlight proteins whose characteristics were measured experimentally. Diagram is not implied to be comprehensive (*e.g.*, location-dependent interactions have been abstracted).

useful contribution to models of apoptotic signaling.

In our experimental analysis of apoptosis, we have examined four types of protein-based signals: phosphorylation, enzymatic activity, abundance and cleavage state. Protein phosphorylation is an important mechanism of signal transduction in cells [69] and measurements of the phosphorylation states of signaling proteins, and the enzymatic activity of the protein kinases that catalyze these phosphorylation events, represent one source of information on death and survival signals. Similarly, measurement of protein abundance captures important information on both intracellular signaling proteins and cytokine receptors. Finally, measurement of the cleavage states of caspases, the biochemical executors of apoptosis [42], makes it possible to gauge how far cells have progressed towards death.

The specific data points collected in this paper come from three types of measurements, which each characterize one or more of these four types of signals. In one set of data, we explored signaling dynamics in depth by quantifying nearly 30 protein states, levels of abundance, and activity from HT-29 cells treated with 50 ng/ml TNF-α alone or in combination with 100 nM insulin at various time points subsequent to cytokine addition (Figure 3-5). The kinase activities of five protein kinases (ERK, Akt, JNK, IKK, and MK2) were measured with a kinase activity assay (Figure 3-5c,d) [5], in which kinases are immunoprecipitated from cells and mixed with protein or peptide substrates. Three of these kinases (ERK, MK2, and JNK) are components of mitogen-activated protein kinase cascades, Akt is a key mediator of survival signals, and IKK regulates the nuclear factor-κB transcription factor (Figure 3-4). Regulatory phosphorylation events were monitored for MEK, Akt, and JNK by Western blotting (Figure 3-5a,b). Western blotting was also used to examine the phosphorylation of several Akt and IKK substrates, as well as the cleavage state of caspase 8 and the executioner caspase, caspase 3. One important feature of the measurements is that they yield heterogeneous data with regard to experimental technique and biological significance. In spite of this heterogeneity, the data all capture signals that meet the criteria for multivariate analysis: signals that vary in response to cues and signals that, when perturbed, affect cell response.

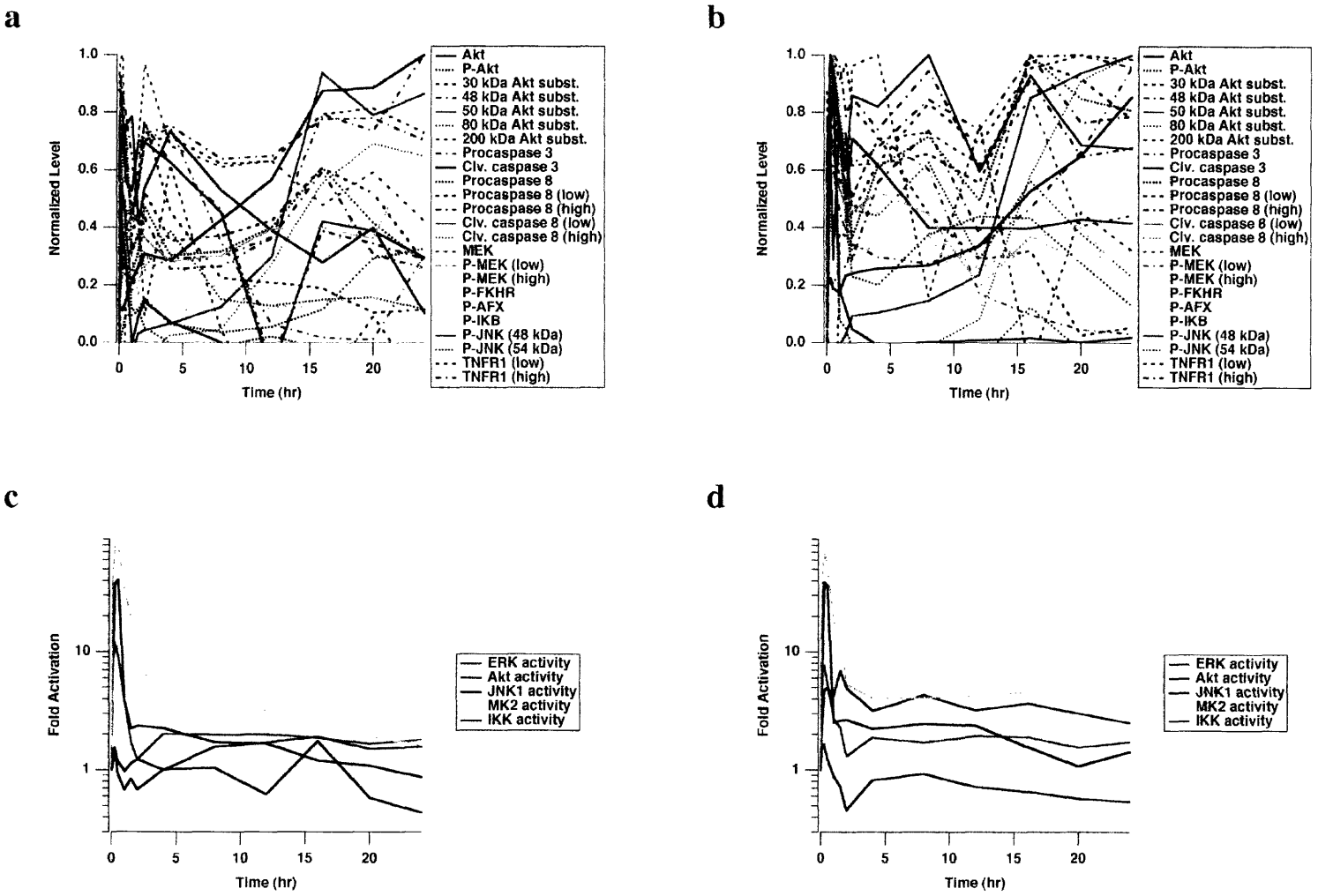

Figure 3-5. TNF-α alone and TNF-α + insulin cotreatment elicit distinct signaling patterns in HT-29 cells. Protein state, level ((**a**) and (**b**)), and activity ((**c**) and (**d**)) measurements in HT-29 cells over 24 hours treated with 50 ng/ml TNF-α ((**a**) and (**c**)) or both 50 ng/ml TNF-α and 100 nM insulin ((**b**) and (**d**)). For (**a**) and (**b**), "P-" signifies phosphorylation state. Note the difference in scale between (a–b) and (c–d). Activity measurements are reprinted from [5] with permission.

3.3.2. Compact representations of signaling by principal component analysis

As a first approach to reducing the complexity of our data, we used principal component analysis (PCA), a nondirected multivariate analysis technique. We cast our data as a matrix of M rows, representing m_i experimental samples, and N columns, representing n_i measurements on the experimental samples. For multivariate analysis to work most effectively, this matrix should have as few empty elements as possible, implying that the same measurements be performed on all samples.

Here, all 29 measurements were performed on all 26 samples (13 time points per treatment × 2 treatments), leaving no empty elements in the data matrix. The different time points were collected from separate tissue culture plates, so the individual time points from both treatments were blindly considered as independent samples for the purposes of the multivariate analysis. Since all of the measurements were made on samples prepared under identical conditions, we can safely include these in the same sample row. However, not all measurements

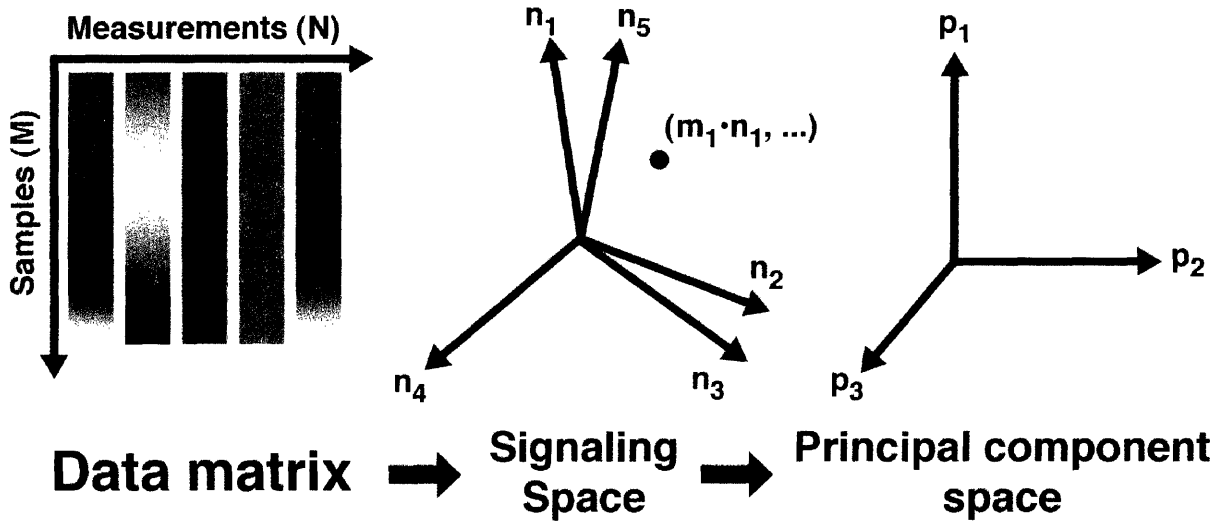

Figure 3-6. Schematic of data reduction via principal components analysis. A data matrix of N signaling measurements can be represented as an N-dimensional signaling space and reduced via principal components analysis to a P-dimensional principal component space. Column shade represents the magnitude of the measurement in the different samples. Note the colinearity of axes in signaling space and the orthogonality of axes in principal component space.

used the same cell extract material (since the different assays require different extract procedures; see [3]), so it is possible that some of the variation in the measurements might be due to variability in the lysate preparations.

An M × N matrix is frequently viewed as a flat, two-dimensional spreadsheet of numbers; alternatively, the data can be thought to represent a multidimensional "signaling" space, where the different protein measurements form a set of N axes (Figure 3-6). The M samples are then vectors describing coordinates in the N-dimensional space, such that the projection of a sample vector m_i on to axis n_i represents one measurement. In this N-dimensional space, the axes are not orthogonal, meaning that the measurements defining these axes are partially redundant. Moreover, the m_i sample vectors project on some axes more than others, because some measurements have a wider range of variation than others.

PCA reduces the dimensionality of the data space, and the size of the M × N matrix, by factoring it into the product of an M × P scores matrix, containing the same number of scores as samples in the data, and a P × N loadings matrix, containing the same number of loadings as original measurements. This decomposition represents the data in P dimensions, with p_i principal components, or latent variables, of the system. These principal components form a new, orthogonal coordinate system in which the scores matrix defines how far the samples project along the coordinate system of the p_i principal components, and the loadings matrix defines how the principal component axes "point" relative to the original measurement space; these two matrices enable one to flip from principal component space to signaling space, and vice versa, by matrix multiplication and division, respectively. The multiplication of a row in the scores matrix and a column in the loadings matrix is called an outer product.

Most computational algorithms determine the principal components iteratively, with the first component capturing as much of the measurement information (*i.e.*, the changes in measurement values over all samples) as possible. Then, the outer product of the first score-loading vector pair is subtracted from the original data matrix, and the second principal component is calculated to capture as much of the residual information as possible. To avoid biasing the decomposition toward measurements with disproportionately large absolute magnitudes or dynamic ranges, we subtract the mean, calculated across all samples for each measurement separately, and divide by the measurement variance. These preprocessing techniques (called mean centering and variance scaling, respectively) eliminate difficulties

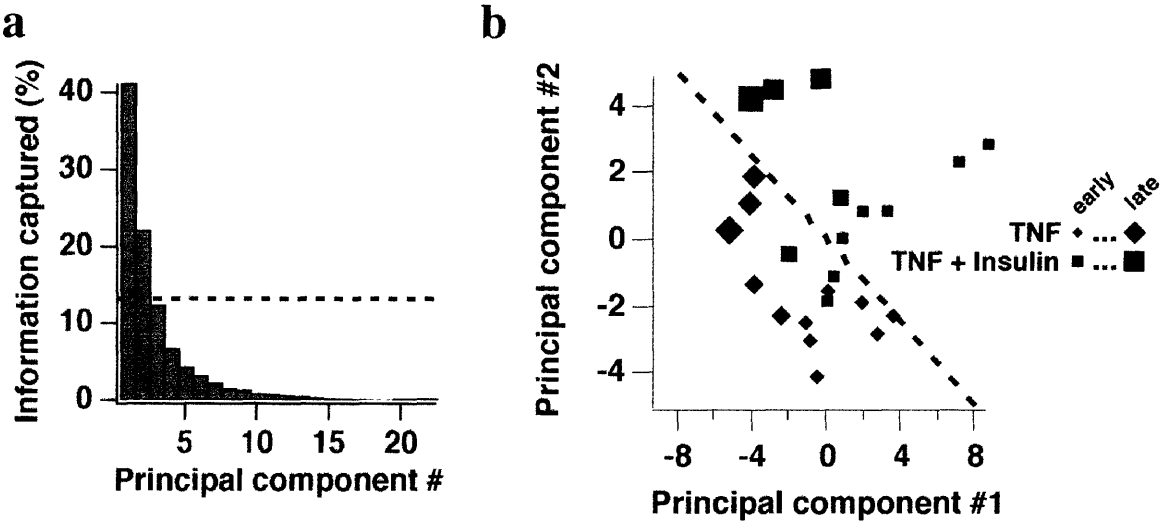

Figure 3-7. Principal component analysis reduces a complex data set of heterogeneous signaling measurements to a data-driven model that retains the differences in the original measurement set. (**a**) Information captured as a function of increasing principal component number. The hatched line represents the arbitrary information cutoff for the two-component PCA model. Higher components were very fluctuant in their scores and loadings and did not indicate consistent data-wide variations in the measurement set (data not shown), suggesting that these components were fitting measurement error and noise. (**b**) Scores plot for the two-component PCA model, showing the projection of each sample along principal components #1 and #2 identified by the PCA model. The hatched line highlights the data-driven separation of the TNF-α (◆) and TNF-α + insulin (■) samples. The size of the markers indicates increasing time.

associated with measurement heterogeneity by putting all data evenly on a unitless measurement axis with a mean of zero and variance of one, thereby enabling scale-free, inter-measurement comparisons in the decomposition procedure.

If the data matrix is completely decomposed by iteration, then P will equal the lesser of M and N, and the product of the scores and loadings matrices will exactly equal the original data set. However, because the principal components were iteratively calculated to maximize information capture, good approximations of the original data set can be obtained with only the first few principal components. In this way, PCA generates a data-driven model, comprising the most significant principal components, and extracts the most salient features of a data set while removing spurious fluctuations (usually from measurement error and noise). The critical question in PCA is whether a significant fraction of the measurement information ($\geq 60\%$) can be captured effectively when P is much less than N.

To begin to determine the most critical factors in our data set for influencing the death-vs.-survival decision, we performed PCA on a complete set of preprocessed protein data and examined the amount of information captured as a function of increasing principal component

number. In Figure 3-7a, we see that the first two principal components together capture over 62% of the information in the original data matrix, with progressively smaller contributions from higher dimensions. Often with PCA, the residual information from higher components (in this model, 38% of the total) is postulated to mostly contain random experimental fluctuations. However, with only 29 measurements, our sample set was too small to permit a rigorous treatment of signal to noise (see Figure 3-7a legend).

Therefore, we further examined the quality of the two-component PCA model by plotting how the different samples mapped along the two principal components via their scores (Figure 3-7b). The size of the data markers represents increasing time; we see that that earlier time points fall largely south and southeast on the two principal components in Figure 3-7b, whereas later time points are located more north and northwest. Importantly, when we applied an objective test to the TNF-α (diamonds) and TNF-α + insulin (squares) time courses, we found that the two sets of measurements were separated in a highly significant fashion, with $p < 0.005$ that the data represents two distinct populations by a two-dimensional Kolmogorov-Smirnov test [192], illustrated schematically by the hatched line in Figure 3-7b. This result contrasts the separation achieved by a PCA model with only one principal component (when the null hypothesis was not rejected by this same test, with $p = 0.19$), and it therefore supports the two-component model as a compact, but not oversimplified, representation of the sample set.

The information contained in the reduced model (see first two columns in Figure 3-7a) has general implications for the study of protein networks. Starting with heterogeneous measurements from a complex signaling network (Figure 3-4), we created a model that efficiently reduced the dimensional complexity from 29 to 2 dimensions (Figure3-7a). This reduced model was informative, in that it clearly determined from the differences in protein state (but without any prior knowledge) that the data derived from two sets of cells that had been exposed to different death and survival mediators.

As mentioned in Chapter 3.3.1, the original measurements themselves contain real biological meaning, it is therefore valuable to explore the contribution of individual signals to the data-wide model, as well as the interplay between signals; the relationship between the multivariate models and the original signaling measurements is contained in the loadings matrix. Upon inspection of the loadings, we first found that none of the reduced axes (the principal components) corresponded to an individual measurement (data not shown). Rather, the axes

were composed of combinations of measurements: the first principal component points largely toward early, prosurvival signals, such as phosphorylated Akt substrates and phospho-IκB, whereas the second component points more toward late, prodeath signals, such as cleaved and procaspase levels (data not shown). Second, we found that not all measurements were equally informative. As mentioned earlier, high-dimensional data usually contains measurements that are collinear, and we identified several groups of measurements that with equally weighted loadings (data not shown), indicating that these signals were redundant in their information content. Importantly, this grouping was consistent with the known biology. For example, the regulatory phosphorylation of Akt and the enzymatic activity of Akt were highly collinear, and the phosphorylation states of many Akt substrates were clustered. We can therefore conclude that PCA has achieved an effective data fusion of heterogeneous sources of data.

Overall, our results with PCA suggest that a multivariate, but by no means comprehensive, set of signaling measurements is sufficient to classify phenotypically distinct samples of cells, whereas no single measurement has this power of discrimination. PCA also appears to be effective at fusing diverse data sources into a single consistent model. More generally, the analysis suggests that there exists a set of highly informative measurements that, in combination, compactly describes the relevant variations in signaling network state and might in the future obviate the need to characterize all of the signals in the network.

3.4. Classification through partial least squares regression

3.4.1. Identification of insulin-induced survival signals

Our next step in data analysis focused on attempts to incorporate prior knowledge about cell treatments rather than to derive them *post facto*. In general, data analysis is much more powerful if prior knowledge can be captured effectively; the specific prior knowledge in our data is that some cells were treated only with TNF-α whereas others were cotreated with insulin. HT-29 cells die in response to TNF-α [178, 193] and are rescued from TNF-induced death when cotreated with insulin [181]. To focus on this death-vs.-survival decision, we used PLS, since it

is designed to generate data-driven models that relate a matrix, or block, of independent measurements to a block of dependent measurements or classifications.

Cue, signal, and response measurements can act as independent or dependent blocks, but it is conceptually helpful to organize the data so that cause-effect relationships are appropriately directed in relation to the original biological setting (*i.e.*, so that an independent block of signals affects a dependent block of responses, rather than the reverse). In our case, the independent block contains the levels and activities of proteins, and the dependent block contains the binary information on cytokine treatments that constitutes our prior knowledge about the outcome:

- Samples that are associated with cell death (*i.e.*, with TNF-α treatment) ≡ 0
- Samples that are associated with cell survival (*i.e.*, with TNF-α + insulin treatment) ≡ 1

For the classification, we assigned time points from each treatment to one of the two classes, ignoring the time component. This class architecture highlights proteins whose state, level, or activity is consistently different between the two responses at all time points. (In future work, we expect to have sufficient data to look at each time point individually.) Since our model is meant to bring out the differences between the two outcome classes, it belongs to a subcategory of PLS called discriminant partial least squares regression (DPLSR).

To perform DPLSR, the original M samples × N measurements data matrix is supplemented with a separate, dependent block of M samples × one classification. Principal components are extracted iteratively from the independent block, analogously to PCA, and the dependent block might also be decomposed, if there were multiple classifications. The most important modification is that, whereas PCA extracts latent variables to maximize the information captured from the independent block, DPLSR decomposes the data matrix to maximize the correlation between the principal components and the (possibly decomposed) classification of the dependent block. The principal components are related to the dependent block by additional coefficients, called weights, that quantify the contribution of each component to the class model through an "inner product".

Whereas PCA generates a reduced model of the variations in the independent block alone, the decomposition criterion for DPLSR focuses on the relationship between the independent and dependent blocks, highlighting the measurements that strongly covary with the known outcome and deemphasizing those that do not. In other words, it is insufficient for a measurement to simply change from sample to sample—to be regarded as significant in the

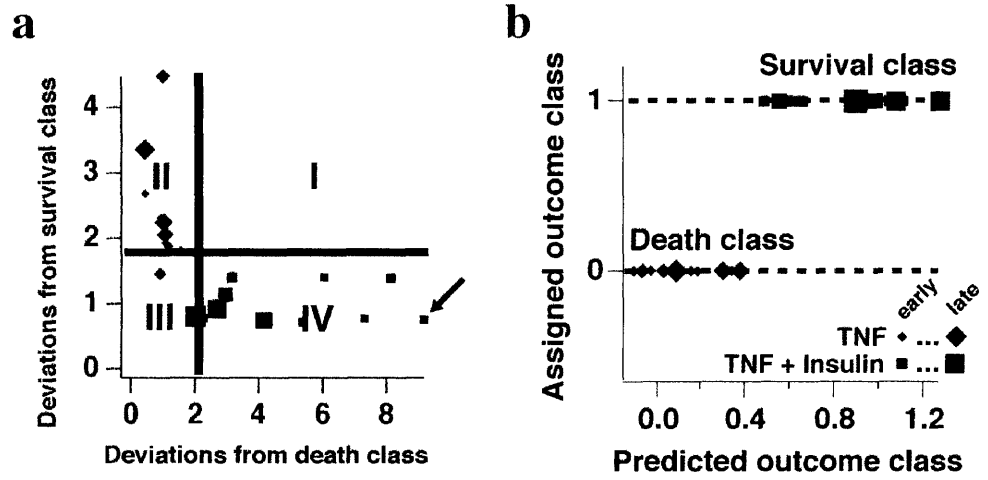

Figure 3-8. Discriminant partial least squares regression explicitly distinguishes distinct phenotypic outcomes from a data set of heterogeneous signaling measurements. **(a)** Cooman's plot for the DPLSR model. The four quadrants represent: (I) neither class, (II) the death class, (III) both classes, and (IV) the survival class for the TNF-α (◆) and TNF-α + insulin (■) samples. The solid lines represent 95% confidence intervals for excluding a sample from the model. The size of the markers indicates increasing time. The arrow highlights the 30 min TNF-α + insulin time point as the most clearly classified survival sample. **(b)** Observed-vs.-predicted plot for the DPLSR model. The hatched lines represent the original class assignments for the TNF-α (◆) and TNF-α + insulin (■) samples. The size of the markers indicates increasing time. Perfect recapitulation of the data set would collapse the TNF-α samples on (0, 0) and the TNF-α + insulin samples on (1, 1). Note that the two sample sets do not overlap horizontally.

DPLSR model, these changes must covary with the changes in apoptotic response. This is a powerful quality of PLS, since it is frequently difficult to identify the determinative changes in a particular response, when the characteristics of many other proteins are simultaneously varying.

We decomposed our signaling data set with DPLSR and the aforementioned binary classification system, after preprocessing the independent and dependent block as described earlier. The decomposition criterion makes the data in the original dependent block more valuable than the independent block for assessing information capture, so we concentrated on model capture of the class structure rather than capture of the signaling variation, which was the focus of PCA. As with the PCA model, however, we observed that the first two components were highly informative (78% capture of the class structure) and that information content fell rapidly with increasing component number (data not shown).

To cross-validate the two-component DPLSR model, we first used a Cooman's plot (Figure 3-8a) to determine the ability of the model to discriminate between two samples classes. The abscissa and ordinate mark the deviations of the samples from the two classes. The four quadrants demarcate the samples that, according to the model, would fall into: (I) neither class,

(II) the death class, (III) both classes, or (IV) the survival class. We see in Figure 3-8a that the TNF-α samples (diamonds) fall largely in (II) and TNF-α + insulin samples (squares) fall in (IV). Thus, the DPLSR model allows good discrimination of the two outcomes, with only a few samples falling into (III) and no samples in (I). Interestingly, the sample with the largest deviation from the death class (TNF-α + insulin at 30 min, marked by an arrowhead in Figure 3-8a) corresponds to the time point at which the largest number of measurements differ in a pairwise comparison of the TNF-α and TNF-α + insulin samples (data not shown).

Although both PLS and PCA reduce data complexity, PLS is generally more powerful because of the predictive ability of the resulting models. After a PLS model has been generated linking independent and dependent blocks, one can attempt to predict which dependent variable is linked to a particular set of measurements. PLS model predictions are achieved in three steps: 1) an experiment is projected onto the reduced principal component space of the PLS model; 2) in this principal component space, the model then uses the inner product, consisting of scores, loadings, and weights, to calculate a predicted dependent component; and 3) if the dependent block has been reduced in the model, the predicted dependent component is recomposed to form a predicted dependent block.

We used this predictive power as an additional form of cross-validation to evaluate how well the PLS model recapitulated the original outcome classes from the independent block of signaling measurements. In Figure 3-8b, for each sample, we plotted the observed (assigned) class against the class predicted by the DPLSR model using only the independent block of the sample. These predictions were not *de novo* predictions, since all the data are part of the training set. Nevertheless, these predicted outcome classes do provide an indication of model quality: if the model were to perfectly fit the training set, then all of the TNF-α treated samples would collapse upon (0,0) and the TNF-α + insulin treated samples upon (1,1) in the observed-vs.-predicted plot. Instead, the data spread laterally from these coordinates, because we retained only two principal components in the PLS model, but we considered the two-component model valid, since it achieved the desired separation of the samples, with no overlap between the two sample classes.

As mentioned before, our original motivation for developing the DPLSR model was to quantify the contributions of various signaling measurements to the outcome classification, and we therefore examined the measurement space of the PLS model. In the PLS model, the

contributions of a measurement take the form of a regression coefficient that includes the multiplicative and additive influences of the loadings and scores of the measurement in the two principal components that are retained, along with the weights relating these principal components to the class outcome. The columns in Figure 3-9 shows the regression coefficients for all state, level, and activity measurements included in the model. Positive regression coefficients suggest correlations with survival, by contributing to a "1" classification by addition, and negative regression coefficients suggest correlations with death, by contributing to a "0" classification by subtraction.

Next, we assessed the significance of these regression coefficients in the context of the model, the class structure, and the original data set. Techniques for estimating significance in PLS models vary, but we had success with a Monte Carlo-type strategy, where the rows of the dependent block (here, the classifications) are shuffled randomly, and a DPLSR model is created

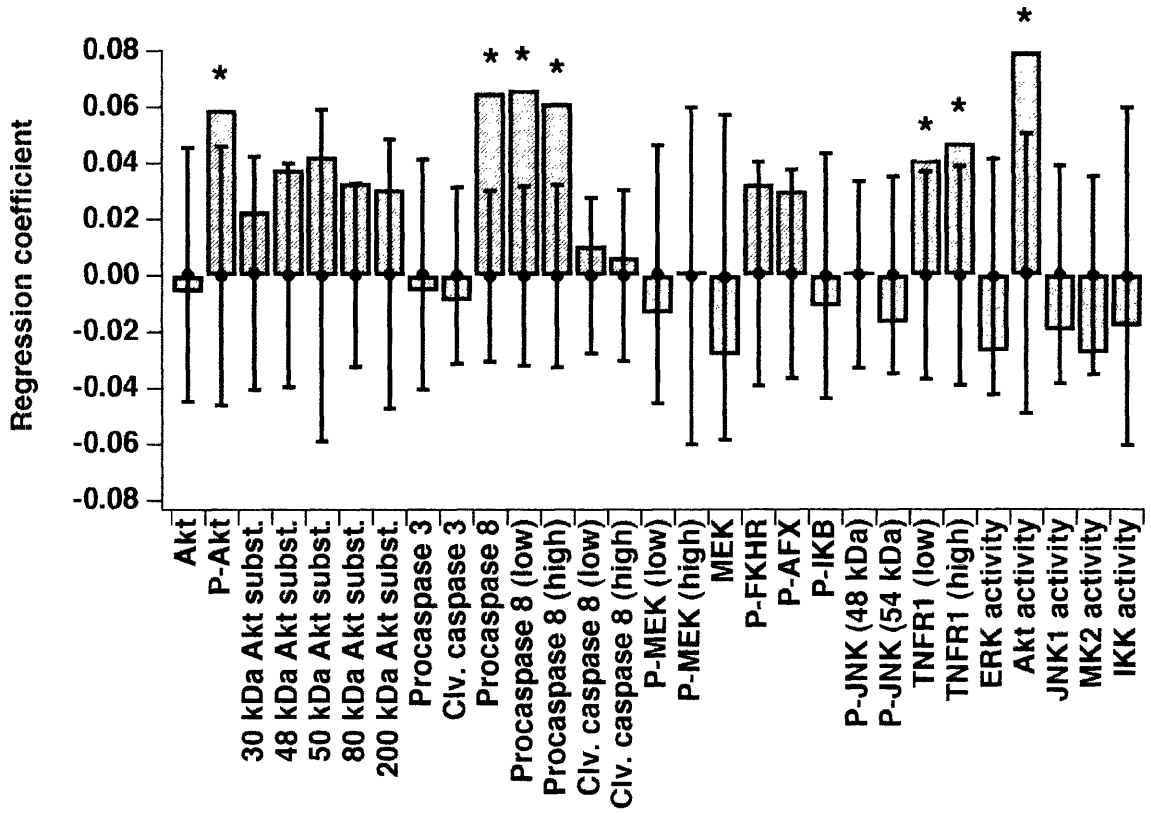

Figure 3-9. Discriminant partial least squares regression identifies Akt phosphorylation state, procaspase 8 levels, TNFR1 levels, and Akt activity as important signals for cell survival. The columns represent the model regression coefficients, whereas the points represent the mean naïve coefficient (± S.D.) from 500 block shuffling iterations. The asterisks indicate measurements with regression coefficients outside of those expected by a naïve model.

with a new set of regression coefficients [194]. These regression coefficients are naïve, since the prior knowledge about the outcome of these samples is lost by the randomization procedure. After several hundred shuffling iterations, one can gain an indication of each mean naïve regression coefficient (usually near zero) in a particular data structure and, importantly, the variation of the regression coefficient about its mean. The results of these shuffling iterations are included in as markers in Figure 3-9, with asterisks indicating the seven "outcome classifiers." These measurements have regression coefficients substantially different from the naïve coefficients that would be expected by chance classification.

Most significantly, we found that, despite the simple classifications in the DPLSR model, the multicomponent properties of the apoptosis network were clear (Figure 3-10). When we generated DPLSR models with partial data sets consisting of certain isolated classifiers, such as the levels of TNF receptor 1 (TNFR1, leftmost column in Figure 3-10), we recapitulated the class structure poorly (38% capture). In contrast, with the complete set of outcome classifiers, we generated a model of nearly the same quality as when the entire data set was used (74% vs. 78% capture, compare third and fourth column in Figure 3-10). By this analysis, we identified those signals that the data support as a "minimal model" for predicting the dependent block. Therefore, both the PCA and DPLSR models support the hypothesis that measurement of individual molecules is not sufficient to capture network properties, but an appropriate set of information-rich measurements can, in combination, create useful models of the signaling network.

A number of the survival classifiers found by DPLSR were intuitively reasonable from a biological perspective: Akt phosphorylation state, procaspase 8 level (both forms), and Akt activity. Akt is phosphorylated in response to insulin treatment [189], and that phosphorylation leads to an increase in catalytic activity [195, 196] to constitute a potent prosurvival signal [197]. Caspase 8 is an initiator caspase along one of the death pathways [42]. Early in the TNF-α time course, the inactive pro- form was processed to the active, cleaved form (Figure 3-5a). In the TNF-α + insulin time course, however, this cleavage was blocked, retaining high levels of the procaspase 8 zymogen (Figure 3-5b).

In contrast, it was intriguing and counterintuitive that the model identified the level of expression of TNFR1 as a survival classifier. Comparison of the original time courses showed that the DPLSR result was supported by the original sample set (Figure 3-5a,b), perhaps

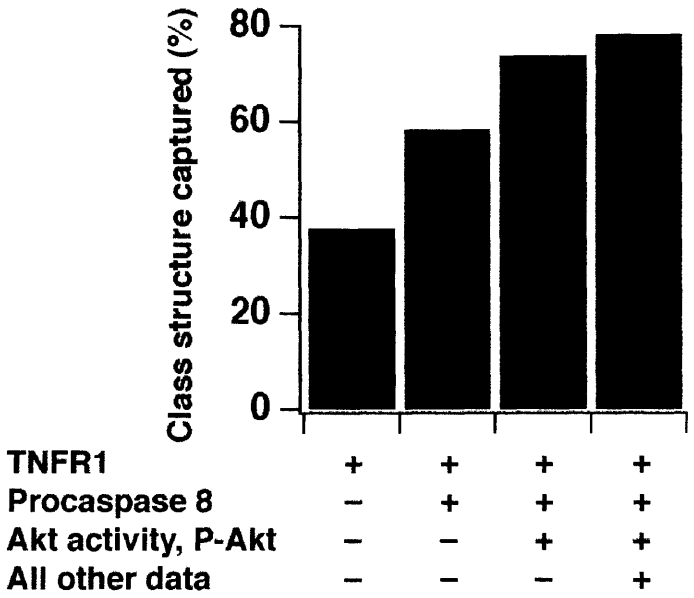

Figure 3-10. A combined subset of critically important survival signals quantitatively captures outcome classes to the same extent as an entire data set. Smaller subsets of the data were modeled individually by DPLSR and characterized by their ability to capture the death-vs.-survival class structure. Note that the data set consisting of measurements of TNFR1, procaspase 8, and Akt, in combination, captures nearly as much of the assigned class structure as all 29 measurements.

suggesting that insulin competes with, or inhibits, the internalization and downregulation of these receptors [198]. Alternatively, there may be other potentiating effects of insulin on TNFR1 expression [199]. The identification of TNFR1 levels as a predictor for outcome emphasizes that DPLSR models can incorporate prior knowledge without biasing against novel or counterintuitive signal-response relationships.

In addition to extracting outcome classifiers, the DPLSR model also identified measurements with zero, or near zero, contribution to any class discrimination (*e.g.*, 48 kDa JNK phosphorylation state, total Akt level). Thus, in the same way that PCA filters noise and identifies redundant measurements within a data set, PLS models deemphasize signals that do not productively contribute to a correlation with the dependent block. This is useful from a practical standpoint, since future experiments can be made more efficient and less costly by eliminating these measurements.

Interestingly, the DPLSR model exclusively highlighted survival classifiers from this set of signaling measurements, with no death classifiers significantly outside of random fluctuations in the data. We believe that failure to identify death classifiers is simply a reflection of the two

treatment conditions that generated the samples: TNF-α treatment alone, placed in the death class, and TNF-α + insulin cotreatment, placed in the survival class. Insulin was expected to activate a number of pathways that DPLSR would identify as survival classifiers, but since TNF-α was present as a cue in both the death and survival outcomes, insulin would need to rapidly attenuate a TNF-induced pathway at all time points for the pathway to be highlighted as a death classifier. We expect that our models will extract additional classifiers, as we work toward supplementing the data set with more diverse cue combinations (*e.g.*, insulin alone, no cytokines).

3.4.2. Construction of a cytokine-signal map

To generate a compact data-driven view of signals induced by TNF, EGF, and insulin from the compendium described in Chapter 2.4, we constructed a cytokine-signal map by again using DPLSR [3]. Briefly, each time course of the 19 measured signals was integrated from

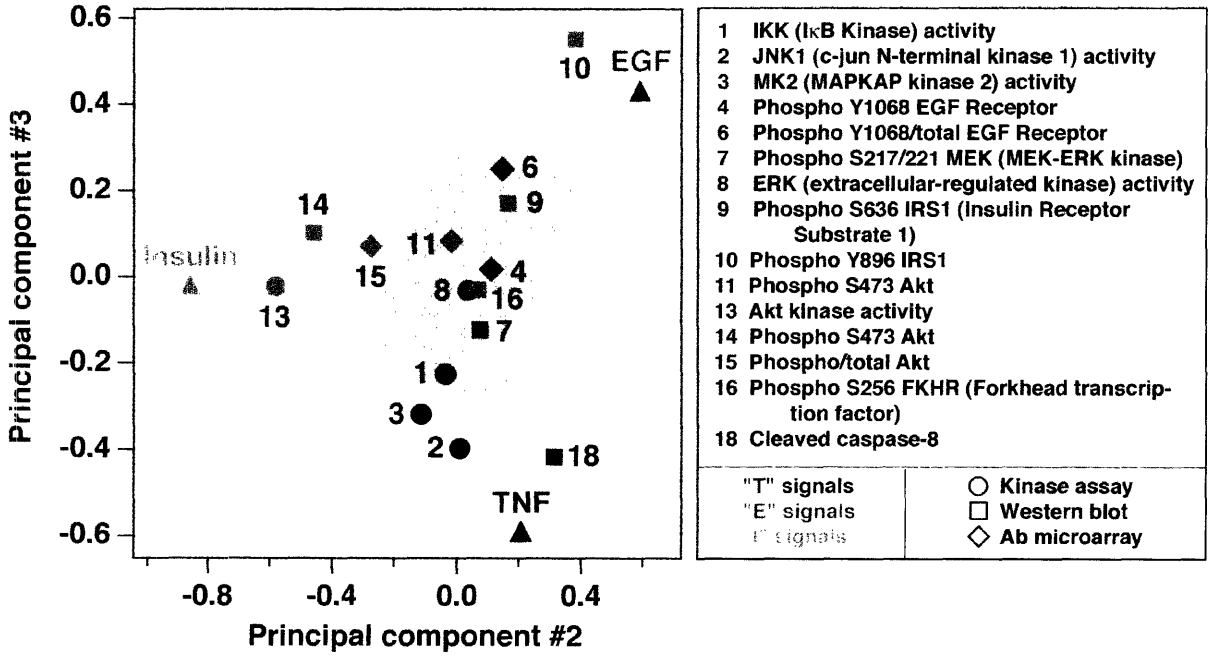

Figure 3-11. DPLSR mapping of intracellular crosstalk in the network shared by TNF, EGF, and insulin. Data were mapped as described in the text [3]. The numbers, colors, and markers are identical to those in Figure 1-1. The gray box indicates the crosstalk region shared by TNF, EGF, and insulin.

0–24 hr to capture the magnitude and duration of the signal [200]. For each combination of TNF, EGF and insulin, the 19 integrated signals for that cytokine stimulus together formed a "treatment vector" indicating how strongly the signaling network was activated. We then constructed a second vector of "cytokine classes" in which the three vector elements classified the input levels of TNF, EGF, or insulin for that treatment (mock ≡ 0, low ≡ 1, high ≡ 2). Finally, a DPLSR map was created in which the treatment and cytokine-class vectors were projected onto a common set of vectors, called principal components, that maximize the covariation between the treatment vectors and the cytokine-class vectors (see [3] for details).

The first principal component in the DPLSR mapping identified a generic baseline signal for all cytokines, whereas the second and third principal components discriminated TNF, EGF, and insulin most effectively (data not shown). By plotting the cytokine classes and integrated signals along these two principal components, we could relate cytokine treatments to the activities of various kinases, caspases, and other measured signals in a simple fashion (Figure 3-11). This mapping proved to be particularly important for the subsequent discovery of the TNF-induced autocrine circuits described in Chapter 4.2.

3.5. Predictive modeling through partial least squares regression

Our ultimate goal for data-driven modeling was to predict cell-death decisions from intracellular signaling patterns activated by TNF, EGF, and insulin. Studying multiple input stimuli in this way requires information about the network as a whole; otherwise, intracellular changes in signal transduction molecules ("molecular signals" hereafter) can appear paradoxical. For example, JNK is a protein that has been reported to be proapoptotic [78], antiapoptotic [77], or uninvolved in apoptosis [201] for different cell systems. To investigate this JNK-apoptosis link further, we added multiple combinations of TNF and EGF to HT-29 cells and then measured JNK phosphorylation (P-JNK) and apoptosis (Figure 3-12A) [1]. The TNF-EGF input stimuli established a four-dimensional signal-response "surface" connecting the P-JNK molecular signal to an apoptotic output (Figure 3-12B). "Slices" through this surface mimicking single TNF or EGF inputs could recapitulate any of the previously reported correlations between P-JNK and apoptosis (Figure 3-12C–E) [77, 78, 201]. This indicated that individual molecular signals, like

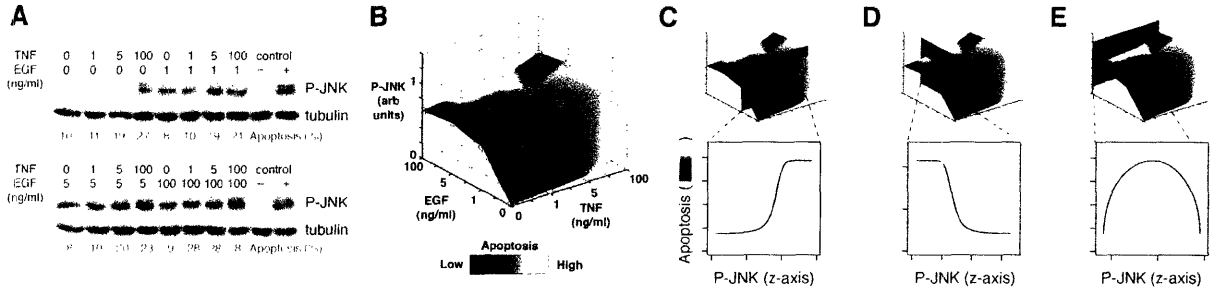

Figure 3-12. Individual molecular signals are poorly predictive of apoptotic outputs. (**A**) TNF- and EGF-induced JNK phosphorylation (P-JNK) and apoptosis. P-JNK and apoptosis were analyzed by quantitative Western blotting and cleaved caspase-cytokeratin at 15 min and 24 hr after cytokine stimulation, respectively. For Western blotting, tubulin was included as a loading control. (**B**) Response surface for P-JNK (z-axis) and apoptosis (color bar) defined by the data in (A). (**C–E**) P-JNK was classified as (C) proapoptotic, (D) antiapoptotic, or (E) uninvolved in apoptosis depending on whether P-JNK increased, decreased, or remained unchanged with increasing apoptosis.

P-JNK, cannot uniquely determine complex output responses, like apoptosis. Quantitative experiments [5] that dynamically sample many critical signals were needed.

Since individual molecular signals correlate with particular apoptotic outputs for only a limited set of experimental conditions (Figure 3-12C–E). We therefore sought to determine whether molecular signals in combination could quantitatively predict apoptosis more globally and capture the entire apoptotic signature as a response. To pursue this, we designed a mathematical formalism that could identify the "information variables" within each molecular signal that mapped most closely with the output responses. The resulting multisignal-response map would allow us to project the input stimuli along only the most relevant network variables for apoptosis and then use these variables to predict apoptotic responses to new stimuli outside the training set.

It was unclear what aspect of a dynamically sampled molecular signal would fill the role of an information variable for apoptosis. For a time course of kinase activity, it was not known if the maximum activity, the rate of rise of activity, or the time when peak activity occurred would contain useful information. Therefore, we defined a panel of time-dependent signaling "metrics", which were candidate information variables that could be empirically derived from any dynamic signal (Figure 3-13A). Each of the metrics (e.g., ERK activity at 5 min, ERK activity at 15 min, peak ERK activity, and P-MEK at 5 min) can be represented by an "axis" along which particular multi-input stimuli project (Figure 3-13C).

In total, we extracted 30–40 metrics (Table 3-1) from a time course to form a composite "metric set" (Figure 3-13A). Metric sets were defined for all 19 molecular signals and then

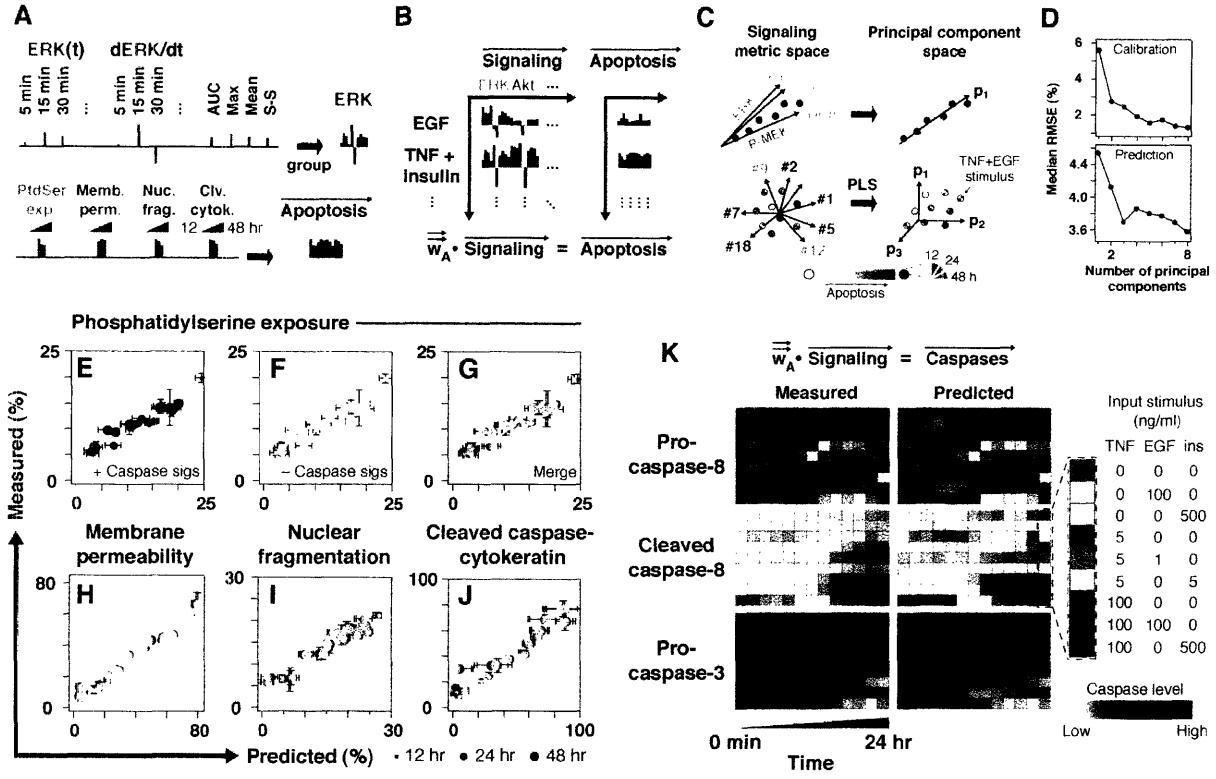

Figure 3-13. The PLS signal-response model correctly predicts apoptosis and caspase activation induced by TNF, EGF, and insulin. **(A)** Conversion of signaling time courses into signaling metrics and concatenation of apoptotic outputs into an apoptotic vector. From each dynamic network measurement [2], the following metrics were extracted: individual time points (t), instantaneous derivatives (d/dt), area under the curve (AUC), and the maximum, mean, and steady-state (S-S) values. In addition, the activation slope and decay rates for signaling peaks were incorporated [1]. **(B)** Construction of the PLS apoptosis model. Signaling metric vectors were concatenated into a single vector for each treatment condition and regressed against the apoptotic response vectors as described [3]. w_A is the coefficient matrix of the apoptosis regression model. **(C)** Reduction of a signaling metric space into a principal component space and organization of the apoptosis vector through the principal component space. Collinear axes like peak ERK activity, ERK activity at 5 min, and P-MEK levels at 5 min were reduced into a single principal component axis (p_1). In the full model, the four apoptotic outputs at 12, 24, and 48 hr were oriented along three principal component axes (p_1, p_2, and p_3). Each point corresponds to one multi-input stimulus, such as TNF + EGF. **(D)** RMSE of calibration and prediction as a function of increasing number of principal components. An optimum model of three components was selected. **(E–J)** Correlation plots between measured apoptotic outputs and cross-validated predictions of the PLS model with caspase signals (red) and without caspase signals (yellow). The merged overlay is shown in green, and marker size corresponds to the response time point. Data are presented as the mean ± S.E.M., and model uncertainties were estimated by jack-knifing [3]. **(K)** Construction of the PLS caspase model and heat-map comparisons between measured caspase levels and cross-validated model predictions. Signaling metric vectors from non-caspase molecular signals were regressed against the 0–24 hr time-point measurements of procaspase-8, cleaved caspase-8, and procaspase-3. w_C is the coefficient matrix of the caspase regression model. The nine rows in the measured and predicted heat map correspond to the nine treatment combinations shown in Figure 2-18A.

consolidated as individual axes that together define a 660-dimensional signaling metric space. The projection of a multi-input stimulus along these axes forms a "signaling vector" of 660 column elements corresponding to the 30–40 metrics from each of the 19 molecular signals

(Figure 3-13B). The apoptotic outputs were similarly concatenated to form an apoptosis vector, where the 12 column elements are the four apoptotic outputs at three time points each (Figure 3-13A). Each input stimulus has its own particular signaling and apoptosis vectors, so these were calculated separately for the nine treatment combinations (Figure 2-18A).

Our goal was then to identify the best mapping of the 660-dimensional signaling metric space onto the 12-dimensional output response space. A linear mapping would require the 7920-element coefficient matrix (12 rows × 660 columns) that best transformed the signaling space into the response space (Figure 3-13B), a calculation that would be impossible given only nine

Table 3-1. Signaling metrics extracted from dynamic network measurements

Metric class*	Metrics extracted
Time point	0 min
	5 min
	15 min
	30 min
	60 min
	90 min
	2 hr
	4 hr
	8 hr
	12 hr
	16 hr
	20 hr
	24 hr
Instantaneous derivative	0 min
	5 min
	15 min
	30 min
	60 min
	90 min
	2 hr
	4 hr
	8 hr
	12 hr
	16 hr
	20 hr
	24 hr
Summary metrics	Area under the curve
	Maximum signal
	Mean signal
	Steady-state signal
Peak metrics	Area under the curve
	Activation slope
	Decay rate

*See [1] for a complete description and definition of the signaling metrics.

multi-input conditions (Figure 2-18A).

To simplify the mapping, we first assumed that some of the signaling metric dimensions were either redundant or irrelevant. For instance, the peak activity of ERK would contain the same information as the ERK activity at 5 min, when the ERK peak occurred (Figure 3-13C). Likewise, other axes (e.g., metrics at 0 min, prior to stimulus) scrambled the projections because they pointed in directions that were unrelated to apoptosis (Figure 3-13C). These uninformative axes could be deemphasized or eliminated without any loss of information about the apoptotic outputs. Second, we assumed that informative axes could be compressed into linear combinations that retained the apoptotic information of the original axes but were simpler. For instance, proteins like MEK and ERK are part of the same signaling pathway and were thus activated similarly. Although not identical, the information in many MEK and ERK metrics pointed in collinear directions that could be combined into a MEK-ERK "super axis" that retained the projection with fewer dimensions (Figure 3-13C).

To reduce unnecessary axes and condense important axes mathematically in this way, we used partial least squares (PLS) regression, which simplifies dimensions based upon their covariance with a specified dependent variable [3, 202]. The apoptosis vector specified the 12 dependent variables (Figure 3-13A, B) to distill from the original network space those metrics that together best oriented the measured apoptotic outputs (Figure 3-13C). PLS modeling, like singular value decomposition [203], specifies an orthogonal set of "principal components", which are the super axes that contain linear combinations of the original 660 metric dimensions weighted by their contribution to the apoptotic outputs (Figure 3-13D). Principal components are calculated iteratively so that each PLS dimension is regressed against the residual information about the apoptotic outputs that was not captured by the preceding component. After several iterations, more dimensions become undesirable, because the residual information is so small that new principal components are essentially capturing spurious fluctuations in the outputs, like measurement error and noise [3].

To optimize the number of model dimensions, we examined the root mean-squared error (RMSE) between the measured apoptosis vector and the values from models with increasing numbers of principal components. First, all of the input treatments were included during the model training to assess the RMSE of data fitting, or calibration. Next, each treatment was individually withheld from the training set to construct a cross-validation model, in which an

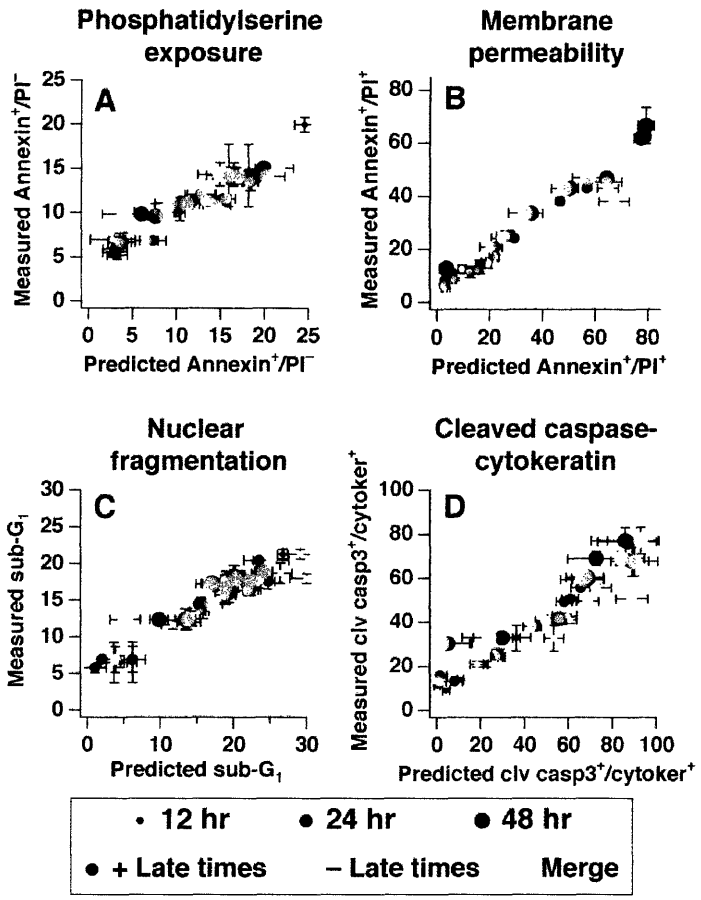

Figure 3-14. The partial least squares model does not require late time points to make predictions of apoptosis. Correlation plots between measured apoptotic indices and cross-validated predictions of (**A**) phosphatidylserine exposure, (**B**) membrane permeability, (**C**) nuclear fragmentation, and (**D**) cleaved caspase-cytokeratin. The full PLS model (red) was compared with a truncated PLS model (yellow) that omitted signaling measurements after four hours. The merged overlay is shown in green, and marker size corresponds to the response time point. Data are presented as the mean ± S.E.M., and model uncertainties were estimated by jack-knifing [3].

RMSE of prediction could be assessed by bootstrapping. The calibrated RMSE decreased monotonically, but the predicted RMSE minimized after three principal components (Figure 3-13D), identifying the optimal model complexity [202].

Using the three-component apoptosis model, we examined the correlation between the measured apoptotic outputs and the cross-validated predictions for each treatment [3]. We found an extremely high correlation for all four apoptotic outputs and all time points, with an overall predictive ability of 94% (Figure 3-13E–J). This illustrated that the signaling measurements contained enough network information to predict outputs for TNF, EGF, and insulin stimuli that had not been used during the model training. Although the 12 individual apoptotic outputs were quantitatively different (Figure 2-18B–J), their stimulus-dependent changes were entirely

captured by a single PLS model of the intracellular network (Figure 3-13E–J). Furthermore, this network-level information was absolutely essential—an equivalent PLS model given the multi-input concentrations (Figure 2-18A) instead of signaling metrics predicted only 43% of the apoptotic outputs [9].

The original signaling network measurements [2] contained several direct effectors of apoptosis, such as caspases (Table 2-2) [42]. The model would obviously be less valuable if these late-effector signals were providing all of the predictive power to the model. To test whether caspase metrics were required for accurate predictions, we removed all of the caspase signals from the model and rederived the principal component axes. No significant change in the predictions was observed (Figure 3-13E–J). Thus, non-caspase signals contained more than enough information to predict the apoptotic outputs quantitatively. Furthermore, the molecular signals activated before the initial onset of apoptosis were adequate for predicting the apoptotic signature. We found that a separate PLS apoptosis model derived exclusively from signaling measurements at times 0–4 hr after cytokine addition was accurate within 81% (Figure 3-14).

Next, to examine the relationship between other molecular signals and late effector caspases directly, we removed the apoptotic outputs altogether and defined the procaspase-8, cleaved caspase-8, and procaspase-3 time-point measurements as a new set of outputs (Figure 3-13K). Using the remaining network measurements, the PLS model predicted the caspase response dynamics within 80% accuracy (Figure 3-13K). Together, this suggested that both the caspase effector signals and the final cellular output responses were encoded by the upstream signaling network.

In the HT-29 system, two regulated autocrine stimuli—transforming growth factor-α (TGF-α) and interleukin-1α (IL-1α)—cooperate with TNF to activate the signaling network (Figure 3-15A) [2]. Whether these autocrine circuits contributed significantly to the TNF-induced apoptotic signature was unknown. We therefore challenged the PLS model to predict what the apoptotic signature would be when these autocrine TGF-α and IL-1α circuits were disrupted pharmaceutically with either an antagonistic EGFR antibody (C225) or IL-1 receptor antagonist (IL-1ra) and the cells then treated with TNF. All 19 molecular signals were measured from 0–24 hr [2] and provided as input data to the PLS model. We then measured the actual TNF + C225- and TNF + IL-1ra-induced apoptotic signatures experimentally [1] and compared these with the model predictions. This experiment was a particularly stringent test of the model,

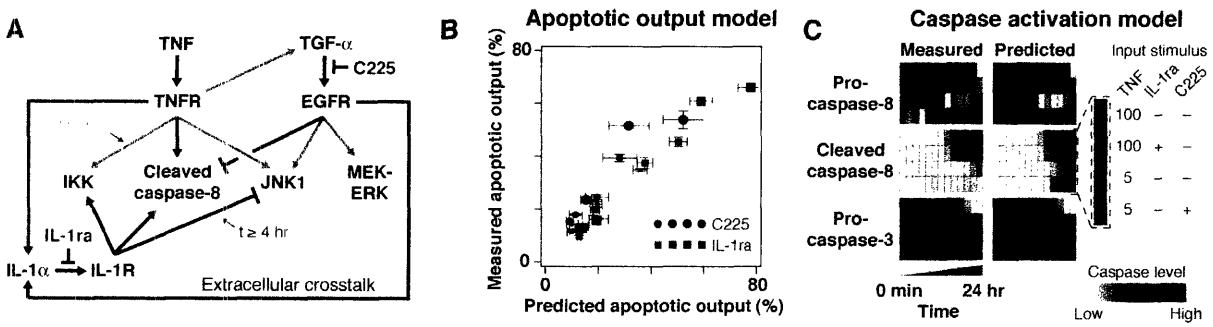

Figure 3-15. The PLS signal-response model correctly predicts apoptotic outputs induced by autocrine stimuli. (**A**) Diagram of TNF-induced autocrine circuits from the studies presented in [2]. Orange arrows indicate fast pathways (before four hours) and purple arrows indicate slow pathways (after four hours). C225 and IL-1ra were used as pharmacological inhibitors of autocrine TGF-α and IL-1α, respectively. (**B**) Correlation plot between measured apoptotic outputs and predictions of the autocrine circuit perturbations. Apoptotic outputs were presented as Figure 3-13E–J for the C225 (circle) and IL-1ra (square) perturbations. Marker color corresponds to the apoptotic index (Figure 2-18B–J) and size corresponds to the time point (Figure 3-13E–J). (**C**) Heat-map comparisons between measured caspase levels and model predictions for the autocrine circuit perturbations. The unperturbed caspase dynamics were included for comparison.

because neither TGF-α nor IL-1α stimuli had been explicitly specified in the original training set (Figure 2-18A).

Strikingly, we observed an 84% correlation between the measured apoptotic outputs and the model predictions when the two autocrine loops were disrupted individually (Figure 3-15B), indicating that the model could predict the contributions of cytokines other than TNF, EGF, and insulin. Furthermore, we found that upstream molecular signals alone were also highly predictive of downstream effector caspase activation after TNF + C225 and TNF + IL-1ra stimuli (Figure 3-15C). Therefore, the contributions of autocrine TGF-α and IL-1α to TNF-induced apoptosis had been implicitly and correctly incorporated throughout both the full apoptosis model and the caspase activation model.

CHAPTER 4

Biological discovery through data-driven modeling

4.1. Introduction

Data-driven approaches to map, classify, and predict measurements are all useful for interpreting large biological datasets. However, the critical test for these methods is in their value as tools for discovery. In this chapter, we describe model-driven discovery of extracellular, intracellular, and network mechanisms for how cells make decisions to die or survive when stimulated with conflicting input stimuli.

4.2. TNF-induced crosstalk through regulated autocrine cytokines

By inspecting the positions of signals relative to cytokines on the DPLSR map (Chapter 3.4.2), we could determine the extent to which signals covaried with particular treatments. We found that some covariations were as expected: the cleavage of caspase-8 (Figure 3-11, signal 18), a proapoptotic protease activated directly by the TNF receptor [38], mapped closely to TNF. Similarly, three overlapping measures of Akt (signals 13-15) mapped closely to insulin, which is a powerful inducer of Akt [25]. Sometimes, the map position suggested unexpected biological regulation; for example, the phosphorylation of IRS1 on Y896 (P-IRS1 (Y896), signal 10) was closely associated with EGF and not insulin [12]. Of greatest interest was a cluster of signals that lay midway between TNF, EGF, and insulin, which implied covariance of these signals with two or more cytokines (Figure 3-11, gray box). When the Euclidean distance was calculated between cytokine cues and E signals in this central cluster—EGFR (Figure 3-11, signal 4), ERK (signal 8), and MAPK-ERK kinase (MEK, signal 7)—we found that TNF and EGF were nearly equidistant, but insulin was farther away (see [2]). Inspection of the individual cytokine time

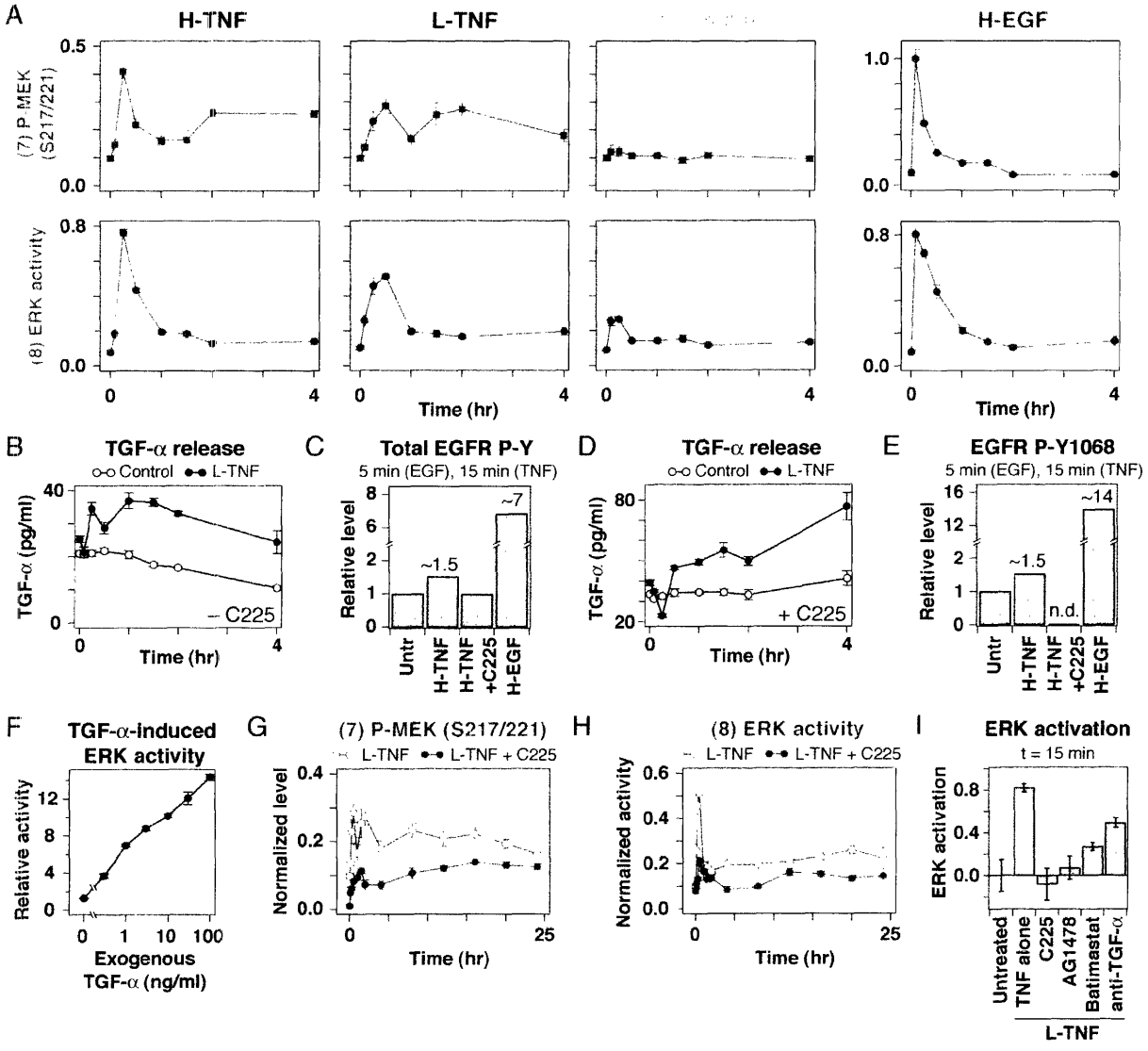

Figure 4-1. TNF activates an early-phase TGF-α autocrine circuit to crosstalk through a EGFR-MEK-ERK signaling pathway. (**A**) Comparison of P-MEK (upper) and ERK (lower) signaling dynamics induced by 100 ng/ml TNF (H-TNF, left), 5 ng/ml TNF (L-TNF, center left), and 500 ng/ml insulin (H-insulin, center right) compared to 100 ng/ml EGF (H-EGF, right). EGFR phosphorylation was similarly increased in response to TNF (data not shown). (**B**) TGF-α release in response to 5 ng/ml TNF. The control treatment was a mock stimulation with carrier only. (**C**) Densitometric analysis of Western blots for total phosphotyrosine (P-Y) of immunoprecipitated EGFR from HT-29 cells treated with 100 ng/ml TNF (H-TNF) for 15 min with or without pretreatment with C225. 100 ng/ml EGF (H-EGF) treatment for 5 min was included as a positive control. P-Y band densities were normalized to the untreated (Untr) sample for comparison. (**D**) TGF-α release in response to 5 ng/ml TNF in the presence of 10 μg/ml C225 pretreatment for 1 hr before stimulation. The control treatment was a mock stimulation with carrier only. (**E**) Densitometric analysis of Western blots for phospho-EGFR on Tyr1068 (P-Y1068). Treatments were identical to those described in (C). P-Y1068 band densities were normalized to the untreated (Untr) sample for comparison. P-Y1068 was not detectable (n.d.) for the H-TNF + C225 treatment. (**F**) ERK activity induced by recombinant TGF-α. HT-29 cells were stimulated with various concentrations of exogenous TGF-α for 5 min and analyzed for ERK activity. (Legend continued on next page.)

Figure 4-1 (cont). (G and H) TNF-induced E signaling through EGFR. ERK (G) and pMEK (H) responses to 5 ng/ml TNF with or without 10 μg/ml C225 pretreatment. (I) Perturbation of TNF-induced ERK activation by pretreatment with 10 μg/ml C225, 1 μM AG1478, 10 μM batimastat, or 1 μg/ml anti-TGF-α for 1 hr before stimulation with 5 ng/ml TNF (L-TNF) for 15 min. 0.1% DMSO was added as a control pretreatment (TNF alone). ERK activity from untreated cells was included as a baseline control, and ERK activation was defined as the increase in ERK activity compared to untreated cells. For (A), (B), (D), and (F–I), data are presented as the mean ± S.E.M. of triplicate biological samples as described in [2].

courses confirmed that TNF activated EGFR, ERK, and MEK similarly to EGF, whereas insulin did not (Figure 4-1A and data not shown). Others have also noted crosstalk among TNF and EGF in more focused studies with different cell types [204, 205]. Together, this supports the paradoxical conclusion that EGFR and its downstream targets MEK and ERK are activated with similar strengths and kinetics by TNF and EGF.

4.2.1. Rapid activation of a TGF-α autocrine circuit by TNF

A direct intracellular link from TNF to ERK has not been definitively established, but TNF can induce the shedding of EGF-family ligands in mammary epithelial cells when stimulated for 24 hr [206]. HT-29 cells shed ligands of the EGF family [207, 208], so we asked whether TNF could potentiate this shedding and thereby activate MEK and ERK via EGFR. The EGF-family ligands known to act as autocrine factors include transforming growth factor-α (TGF-α), amphiregulin (AR), and heparin-binding epidermal growth factor (HB-EGF). We measured the levels of these three EGF-family ligands before and after TNF addition by using quantitative ELISA's. We found all three ligands at biologically significant levels in conditioned medium from HT-29 cells, but only TGF-α was significantly upregulated by TNF (Figure 4-1B and Figure 4-2). Furthermore, this TNF-induced TGF-α release was significantly faster than previous reports [206], peaking at 1 hr after TNF addition (Figure 4-1B). We conclude that TNF treatment of HT-29 cells is associated with the rapid release of TGF-α into the medium.

EGFR activation itself causes shedding of TGF-α into the medium [209], so it was unclear whether TGF-α release following TNF treatment was a cause or consequence of EGFR activation. To distinguish between these possibilities, cells were treated with C225 antibody, which blocks the interaction of ligands with EGFR, and then stimulated with TNF. As expected, the 1.5-fold increase in EGFR tyrosine phosphorylation associated with TNF stimulation of HT-

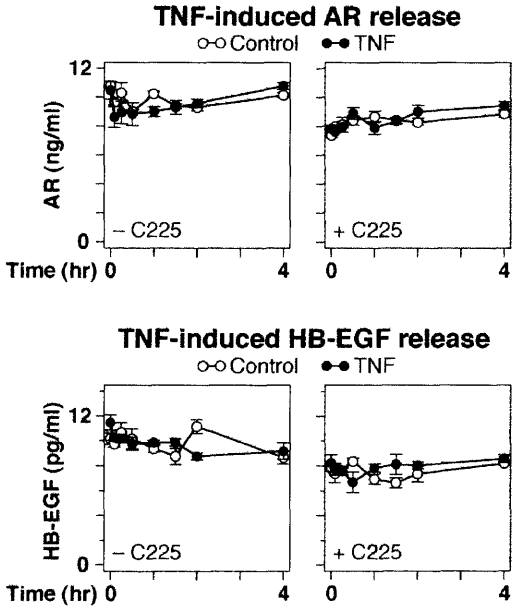

Figure 4-2. TNF-induced AR and HB-EGF release in HT-29 cells. Cells were stimulated with 5 ng/ml TNF with or without pretreatment with 10 μg/ml of a function-blocking anti-EGFR antibody (C225), and AR and HB-EGF levels were measured as described in [2]. The control treatment was a mock stimulation with carrier only. Data are presented as the mean ± S.E.M. of three independent biological samples.

29 cells was blocked by C225 treatment (Figure 4-1C). C225 treatment also elevated the levels of TGF-α in conditioned medium from 20 to 30 pg/ml by blocking the basal uptake of TGF-α into cells via EGFR (Figure 4-1D). The inhibition of this constitutive autocrine circuit was also evident in the dramatic decrease of baseline EGFR phosphorylation on Y1068 (Figure 4-1E). Nevertheless, the ability of exogenous TNF to increase TGF-α levels twofold (from 30 to 60 pg/ml over 2 hr) was not affected by C225 treatment (Figure 4-1D). Although rather complex, these data are completely consistent with current understanding of both constitutive and inducible EGFR autocrine circuits (see Figure 4-3 and [90]). Taken together, this indicates that a basal TGF-α autocrine circuit in HT-29 cells is significantly upregulated by TNF addition, consistent with previous findings [206] but on a much faster time scale than has been reported previously.

After establishing that autocrine TGF-α signaling in HT-29 cells could be blocked with C225 antibody, we asked whether TNF-induced MEK and ERK activation, like EGFR phosphorylation (Figures 4-1C,E) was inhibited by C225 and therefore EGFR-dependent. We observed dramatic inhibition of both acute (fourfold reduction at 30 min, $p < 0.005$) and sustained MEK-ERK activation (twofold reduction from 4-24 hr, $p < 10^{-6}$) in response to TNF

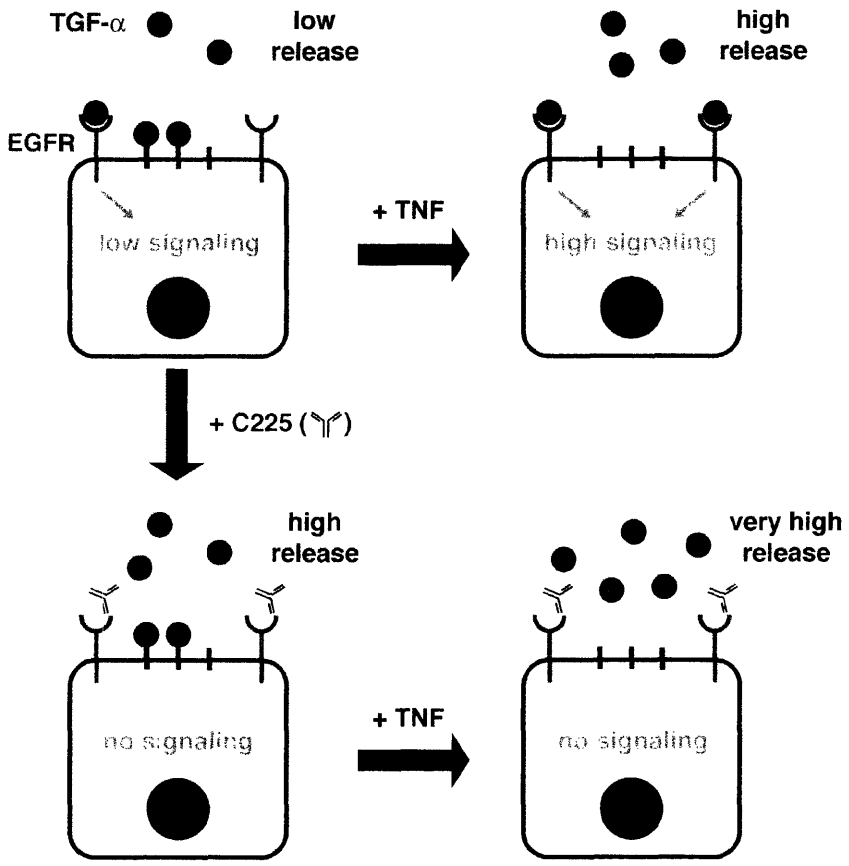

Figure 4-3. TNF rapidly upregulates a basal TGF-α autocrine circuit. In resting cells, low levels of TGF-α are released into the medium, which leads to partial EGFR occupation and "low" basal signaling (upper left). Upon TNF stimulation, TGF-α release is upregulated, and free TGF-α binds unoccupied EGFR's to increase signaling (upper right). When cells are treated with a function-blocking anti-EGFR antibody (C225), both the basal and TNF-induced autocrine circuits are disrupted. This reduces EGFR signaling to below basal levels and increases the levels of TGF-α in the medium (lower left). Nevertheless, EGFR blockade by C225 does not affect TGF-α release, so TNF treatment still increases TGF-α release in C225-treated cells (lower right). The remaining signaling not blocked by C225 could be due to an alternative pathway of TNF-induced ERK activation.

treatment when cells were pretreated with C225 antibody (Figure 4-1G,H). By using a standard curve to relate ERK activity to exogenous TGF-α levels, we estimated the effective levels of TGF-α in TNF-treated cells (Figures 4-1F and 4-4). The fivefold ERK activation observed with 100 ng/ml TNF was equivalent to 400 pg/ml of exogenous TGF-α, even though only 40 pg/ml of TGF-α was detected in conditioned medium (Figure 4-1B). This implied a significantly greater TGF-α concentration in the vicinity of cells compared to the bulk medium, consistent with local TGF-α release and uptake [210].

Some TNF-dependent activation of MEK and ERK was still observed in C225-treated cells (~25% of the unperturbed activation). This perhaps reflected the inability of C225 to fully

block EGFR ligand binding [211] or possibly the existence of an additional pathway between TNF and MEK-ERK [212]. To determine whether the bulk of the low TNF-stimulated activation of ERK was indeed mediated by EGFR, we also showed that ERK activation was blocked by a small molecule inhibitor (AG1478) of the EGFR kinase, an inhibitor (Batimastat) of the matrix-metalloproteases involved in shedding TGF-α from the plasma membrane [213], and, to a lesser extent, a neutralizing anti-TGF-α antibody. From these data we conclude that an

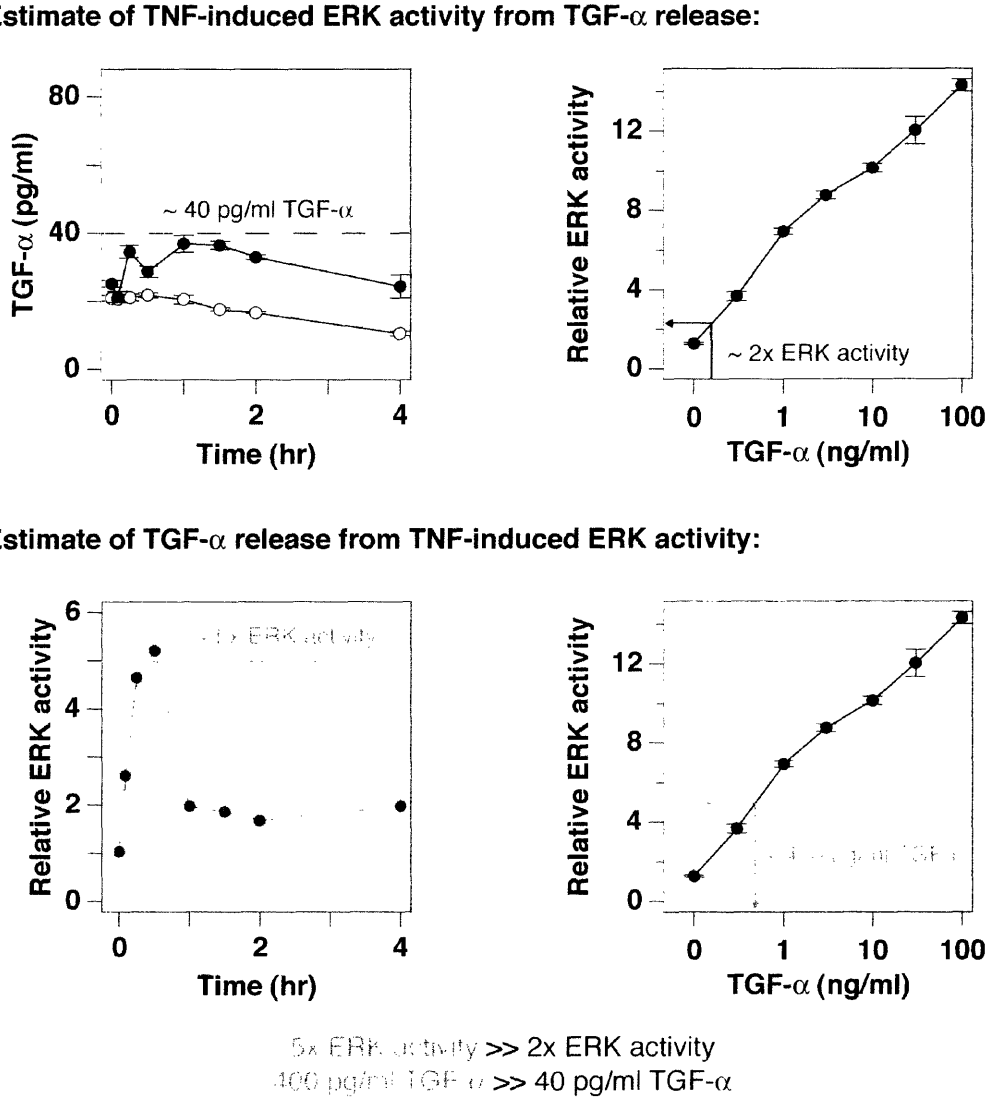

Figure 4-4. The TNF-induced TGF-α autocrine circuit is sensed locally. A peak TNF-induced TGF-α release of ~40 pg/ml (upper left, same as Figure 4-1B) was compared against the recombinant TGF-α-induced ERK activity (upper right, same as Figure 4-1F) to estimate ~2x ERK activation after TNF treatment. This estimate was much smaller than the peak TNF-induced ERK activity, which was measured ~5x (lower left, same as Figure 4-1A). Using the ~5x ERK activity, it was estimated that cells were locally sensing ~400 pg/ml TGF-α, assuming equivalent bioactivity between endogenous TGF-α and recombinant TGF-α.

autocrine circuit involving TGF-α and EGFR is the primary mechanism by which TNF activates MEK-ERK signaling at both short (before 1 hr) and long (after 4 hr) time scales in HT-29 cells. The 10-min interval between the peak of direct ERK activation by EGF at 5 min and indirect activation by TNF at 15 min is an estimate of the minimum time required to establish the TGF-α autocrine circuit (Figure 4-1A). The speed of autocrine TGF-α in this system is comparable to the near-instantaneous autocrine circuit triggered by G-protein coupled receptors through HB-EGF [214].

4.2.2. TNF activates a late-phase IL-1α autocrine circuit

Activation of IKK (Figure 3-11, signal 1) was a second signal that occupied an unexpectedly central position on the DPLSR map. IKK is activated by binding to the DISC complex, which assembles on the intracellular domain of TNF receptor (Figure 1-1). In the "canonical" NF-κB pathway, IKK phosphorylates and inactivates IκB, an NF-κB inhibitor, which allows NF-κB to translocate into the nucleus and induce gene expression [43]. Inspection of signaling time courses revealed that IKK was activated 5–30 minutes after TNF addition, along with other T signals induced by the DISC, such as the stress kinases JNK1 and MK2 (Figure 4-5A–C). However, IKK also exhibited a sustained period of activation between 4 and 24 hr when JNK1 and MK2 activities had returned to baseline. IKK activation after 4 hr was associated with sustained translocation the p65 subunit of NF-κB into the nucleus (Figure 4-5D–F), implying that late-phase IKK activity was relevant for signaling.

The disappearance of JNK1 and MK2 signaling by 2 hr suggested that late IKK activation might not be mediated by DISC or TNFR (Figures 4-5A–C). A similar sustained phase of NF-κB signaling occurs in keratinocytes following ultraviolet radiation and involves an interleukin-1α (IL-1α)-dependent autocrine circuit [215]. Since IL-1 is a potent IKK agonist [216], we determined whether TNF treatment of HT-29 cells was associated with IL-1 release by analyzing conditioned medium for IL-1 cytokines. An eightfold increase in the soluble IL-1α, but not IL-1β, was observed 4–25 hr after TNF treatment (Figure 4-5G). HT-29 cells express the IL-1 receptor (IL-1R) [217], and we found that exogenous IL-1α activated IKK at concentrations as low as 30 pg/ml (Figure 4-5H), consistent with the known sensitivity of IL-1R signaling

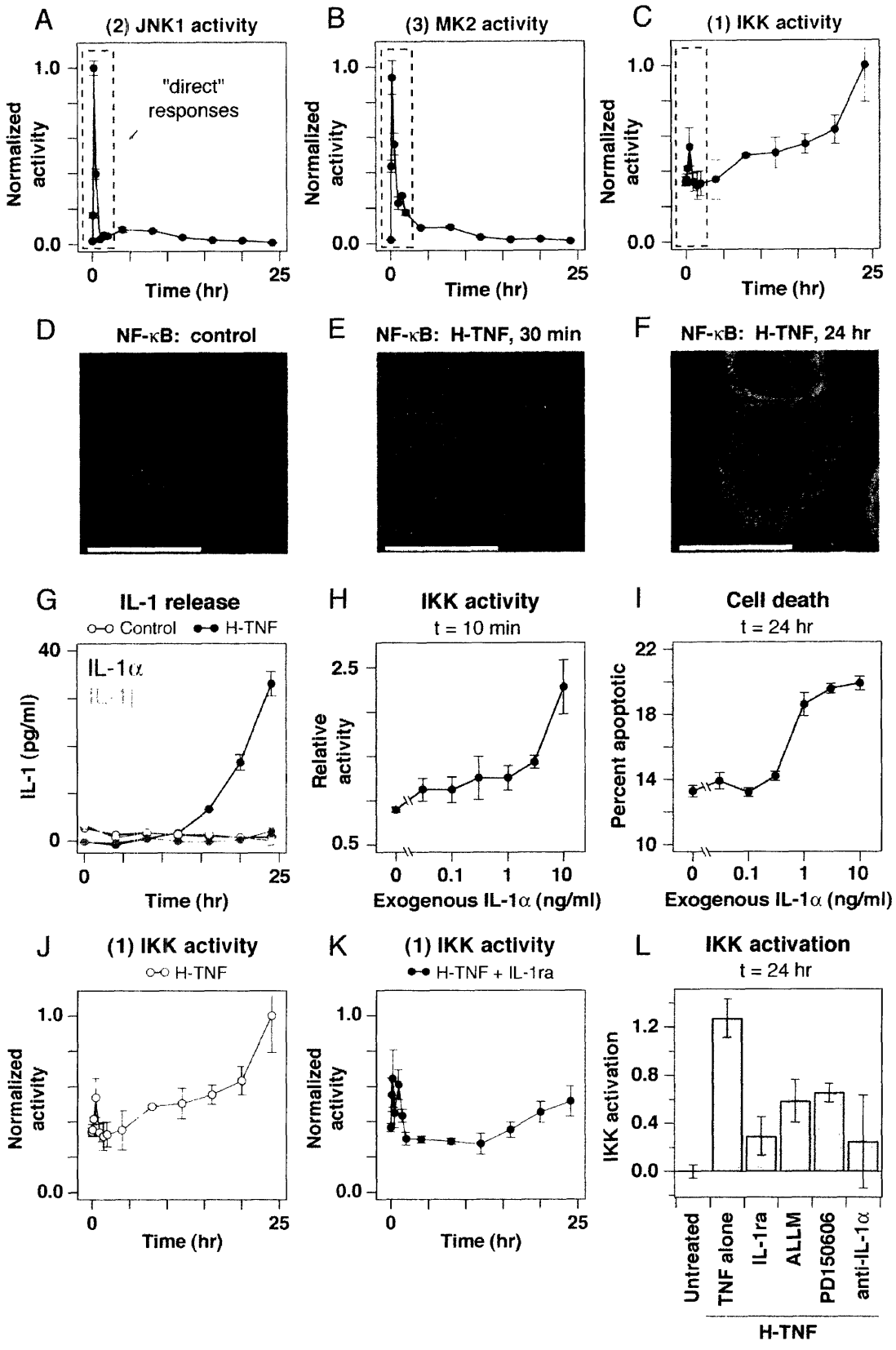

Figure 4-5. (See legend on next page.)

Figure 4-5. (Previous page) TNF activates a late-phase IL-1α autocrine circuit to crosstalk through the IKK-NF-κB signaling pathway. **(A–C)** TNF-induced T signaling after stimulation with 100 ng/ml TNF. JNK1 activity (A), MK2 activity (B), and IKK activity (C) dynamics were compared against a 2 hr gate corresponding to "direct" responses (dotted). **(D–F)** Indirect immunofluorescence images of the p65 subunit of NF-κB in untreated cells (D) and in cells stimulated with 100 ng/ml TNF (H-TNF) for 30 min (E) and 24 hr (F). Scale bar = 15 μm. **(G)** IL-1α (purple) and IL-1β (orange) release in response to 100 ng/ml TNF (H-TNF). The control treatment was a mock stimulation with carrier only. **(H)** IKK activity induced by recombinant IL-1α. HT-29 cells were stimulated with various concentrations of exogenous IL-1α for 10 min and analyzed for IKK activity. **(I)** Dose-response for HT-29 apoptosis induced by various concentrations of IL-1α. Apoptosis was measured as described in Figure 1-1. **(J and K)** IKK response to 100 ng/ml TNF (H-TNF) with (K) or without (J) 10 μg/ml IL-1ra cotreatment. **(L)** Perturbation of TNF-induced IKK activation by cotreatment with 10 μg/ml IL-1ra, 25 μM ALLM, 25 μM PD150606, 1 μg/ml anti-IL-1α, or 1 μg/ml anti-IL-1β during stimulation with 100 ng/ml TNF (H-TNF) for 24 h. IKK activity from untreated cells was included as a baseline control, and IKK activation was defined as the increase in IKK activity compared to untreated cells. For (A–C) and (G–L), data are presented as the mean ± S.E.M. of triplicate biological samples as described in [2].

[216]. Finally, exogenous IL-1α also induced apoptosis in HT-29 cells, demonstrating that IL-1α is a prodeath factor in this experimental system (Figure 4-5I). To exclude the possibility that TNF-induced IL-1 release was affected by rebinding to the HT-29 IL-1R's (Figure 4-6), we blocked binding with IL-1 receptor antagonist (IL-1ra) and remeasured TNF-stimulated IL-1 release. No significant change in IL-1α release was observed in the presence of IL-1ra (Figure 4-7). We therefore conclude that HT-29 cells secrete IL-1α in response to TNF, that HT-29 cells contain functional IL-1R's, and that IL-1R activation is proapoptotic.

To determine if late IKK induction by TNF was mediated by secreted IL-1α, IKK activity was measured from cells stimulated by high TNF with or without saturating levels of IL-1ra (Figure 4-5J,K). The rapid twofold increase in IKK activity 15–30 min after TNF stimulation was unaltered by IL-1ra addition, but the sustained activation after 4 hr was reduced twofold ($p < 0.001$). The residual IL-1α-independent IKK activity observed at very late times (Figure 4-5K) could be due to incomplete IL-1R blockade by IL-1ra [216] or other mechanisms, such as IKK activation by MEKK2 [218]. Significant reduction in TNF-induced IKK activity at 24 hr was also observed with ALLM and PD150606—two structurally distinct calpain inhibitors that block processing and release of IL-1α from cells [219]—as well as IL-1α neutralizing antibodies (Figure 4-5L). From these data, we conclude that early activation of IKK by TNF occurs in an IL-1α–independent manner, consistent with the canonical TNF-induced NF-κB pathway [43]. In contrast, sustained IKK activation after 4 hr (which is quantitatively more significant than early activation in these cells) is mediated largely by the binding of autocrine

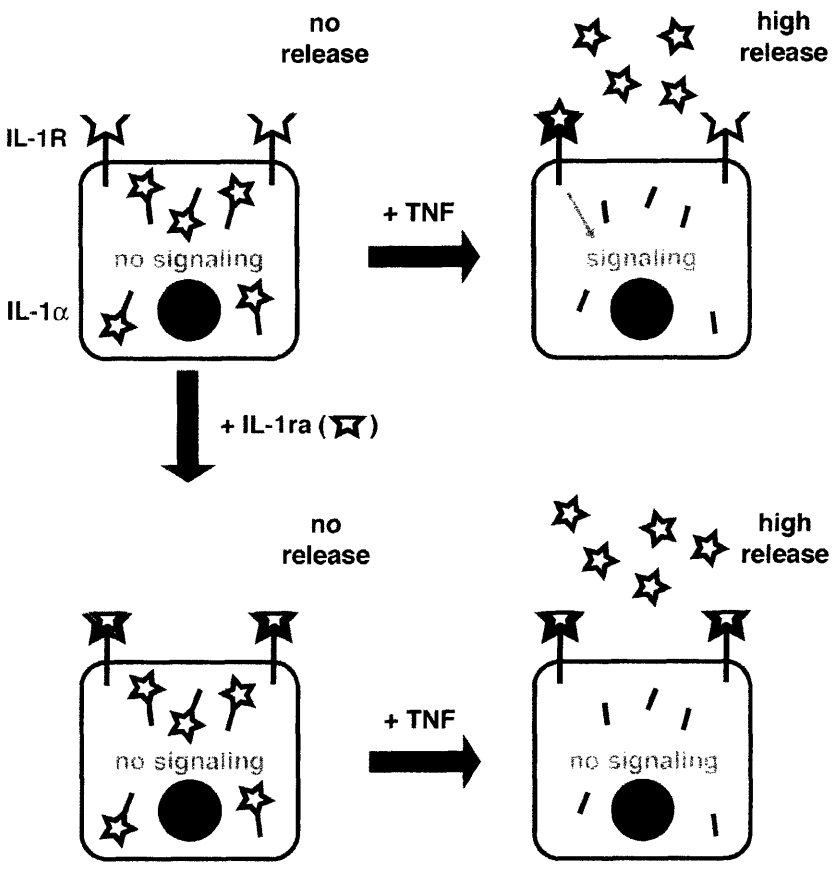

Figure 4-6. Tumor necrosis factor (TNF) activates an interleukin-1α (IL-1α) circuit. In resting cells, IL-1α is retained in the cytoplasm as an inactive zymogen (upper left). Upon TNF stimulation, processed IL-1α is released into the medium, and a small fraction binds the IL-1R to cause signaling (upper right). When cells are treated with IL-1ra, the IL-1R's are blocked, but there is no change in baseline signaling, because the IL-1α autocrine circuit has not been activated (lower left). When TNF activates the IL-1α circuit in the presence of IL-1ra, signaling through the IL-1R does not occur (lower right). Since only a small fraction of autocrine IL-1α was captured without IL-1ra (upper right), there is not a measurable difference in TNF-induced IL-1α release in the presence or absence of IL-1ra (Figure 4-7).

IL-1α to IL-1R. Furthermore, the requirement for calpain in this response suggests that this autocrine circuit is regulated by the release of cytoplasmic proIL-1α [219]. By calibrating the measured late IKK activation with exogenous IL-1α (Figure 4-5H), we estimated that HT-29 cells could be locally exposed to high (i.e., ng/ml) concentrations of IL-1α at 24 hr (Figure 4-8). Whether residual IKK activation in the presence of IL-1ra therefore reflects incomplete receptor blockade or the existence of additional regulatory mechanisms remains to be established. Nonetheless, we can conclude that IKK is activated in TNF-treated HT-29 cells by a direct pathway before 1 hr and by an autocrine IL-1α pathway after 4 hr.

4.2.3. TNF-induced autocrine circuits quantitatively affect direct TNF signals

Autocrine TGF-α and IL-1α were critical for ERK-MEK and late IKK signaling mediated by TNF (Figures 4-1F,G, and 4-5K), so we next asked if other signals might be similarly autocrine-dependent. In particular, we were interested in the T signals: JNK1, MK2, and the caspases. We therefore treated cells with TNF and C225 antibody or IL-1ra and measured the entire set of protein signals from 0–24 hr (see [12] for details).

Many T signals were unexpectedly dependent upon the TNF-induced autocrine circuits. For instance, although JNK1 was located near TNF on the DPLSR map (Figure 3-11, signal 2), JNK1 was strongly influenced by the TGF-α autocrine cascade. C225 treatment significantly reduced both early and late JNK1 activity by twofold ($p < 10^{-11}$, Figure 4-9A). MK2 was also reduced significantly at late times ($p < 10^{-6}$) and to a much lesser extent at early times (Figure 4-9B). Reduced signaling through these stress kinases is probably due to reduced activation of MEKK1, a MAP3K that is regulated by both TNFR–complex I (Figure 1-1) and EGFR–Ras [220]. In contrast, MK2 signaling was completely unaffected by IL-1ra treatment, and JNK1

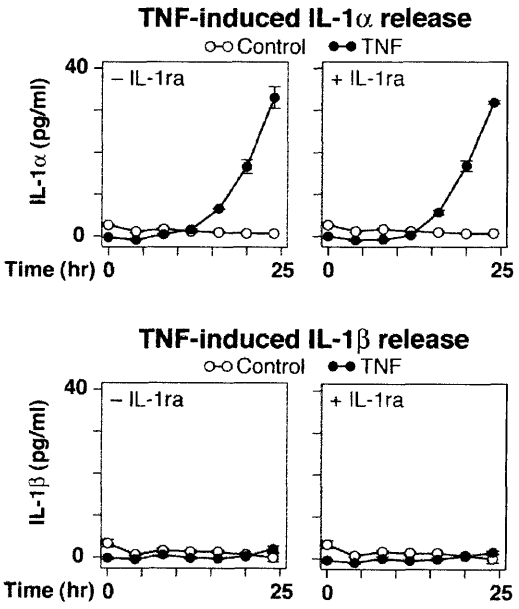

Figure 4-7. TNF-induced IL-1 release in HT-29 cells. Cells were stimulated with 100 ng/ml TNF with or without cotreatment with 10 μg/ml IL-1ra, and IL-1α and IL-1β levels were measured as described in [2]. The control treatment was a mock stimulation with carrier only. Data are presented as the mean ± S.E.M. of three independent biological samples.

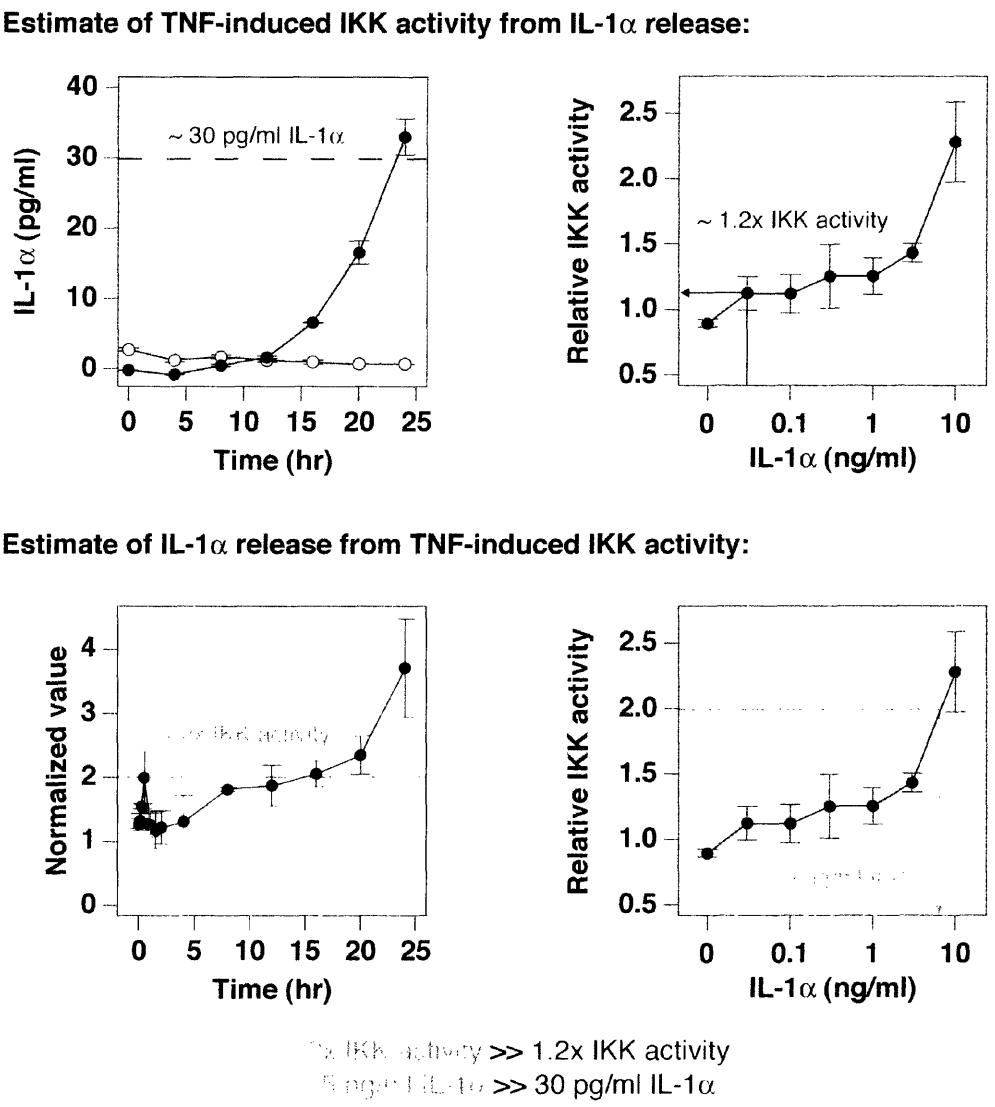

Figure 4-8. The TNF-induced IL-1α autocrine circuit is sensed locally. A peak TNF-induced IL-1α release of ~30 pg/ml (upper left, same as Figure 4-5G) was compared against the recombinant IL-1α-induced IKK activity (upper right, same as Figure 4-5H) to estimate ~1.2x IKK activation after TNF treatment. This estimate was much smaller that the steady-state TNF-induced IKK activity, which was measured ~2x (lower left, same as Figure 4-5C). Using the ~2x IKK activity, it was estimated that cells were locally sensing ~5 ng/ml IL-1α, assuming equivalent bioactivity between endogenous IL-1α and recombinant IL-1α.

signaling was significantly increased after 4 hr ($p < 10^{-8}$, Figure 4-9C,D). Elevated JNK1 activity is likely a consequence of reduced late IKK signaling in the presence of IL-1ra (Figure 4-5K), because sustained IKK-NF-κB signaling has been shown to negatively regulate JNK1 [65]. This illustrates that JNK1 and MK2 stress-activated pathways are quantitatively and distinctly dependent upon TNF-induced autocrine cascades.

Apoptotic signaling through caspases was also affected by disrupting the TGF-α and IL-1α autocrine circuits. Cleavage of the initiator caspase-8 was significantly increased by C225

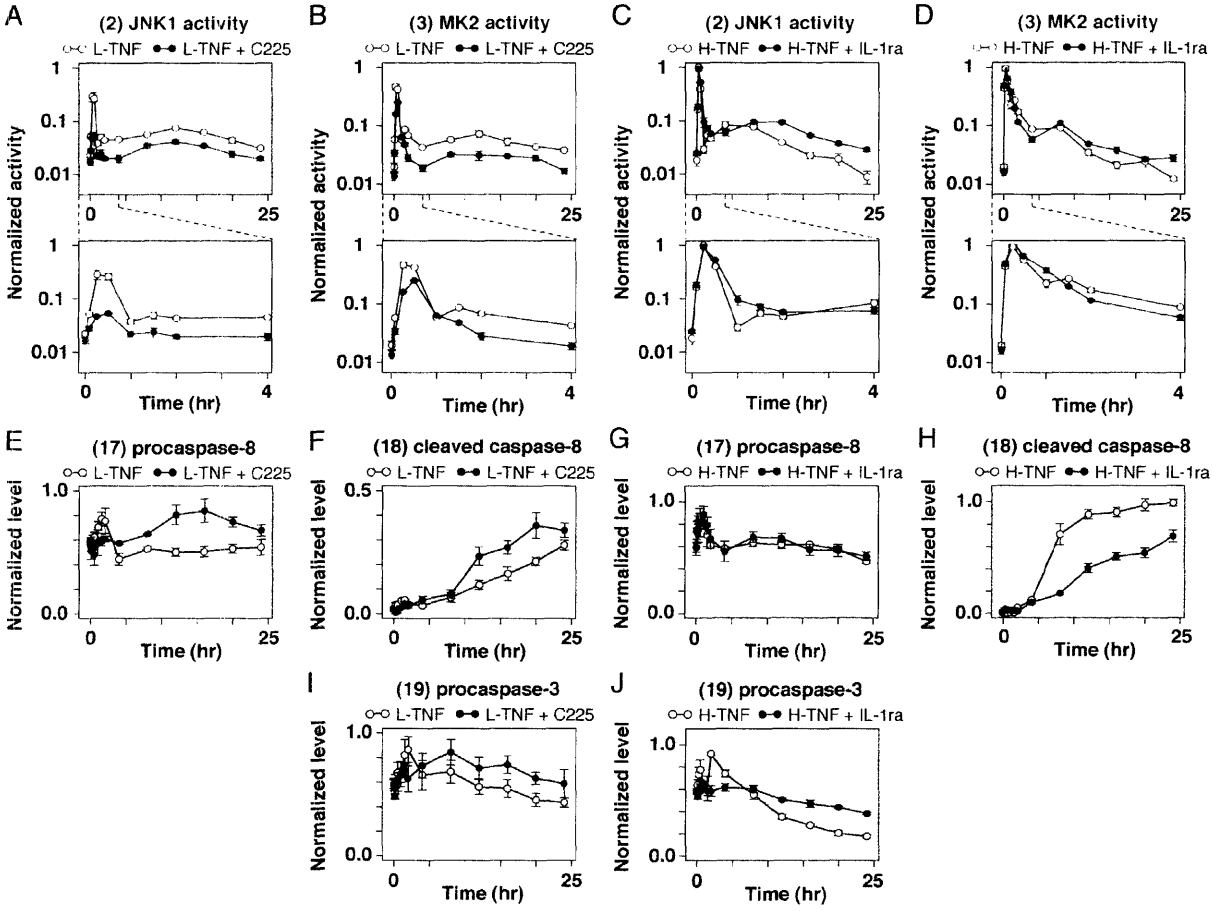

Figure 4-9. The TGF-α and IL-1α circuits quantitatively affect TNF-induced JNK1, MK2, and caspase signaling. **(A–D)** Quantitative changes in TNF-induced stress kinase signaling after perturbation of the TGF-α and IL-1α autocrine circuits. JNK1 (A and C) and MK2 (B and D) responses to 5 ng/ml TNF (L-TNF) with or without 10 μg/ml C225 pretreatment (A and B) or 100 ng/ml TNF (H-TNF) with or without 10 μg/ml IL-1ra cotreatment (C and D). **(E–J)** Quantitative changes in TNF-induced caspase signaling after perturbation of the TGF-α and IL-1α autocrine circuits. Procaspase-8 (E and G), cleaved caspase-8 (F and H), and procaspase-3 (I and J) responses to 5 ng/ml TNF (L-TNF) with or without 10 μg/ml C225 pretreatment (E, F, and I) or 100 ng/ml TNF (H-TNF) with or without 10 μg/ml IL-1ra cotreatment (G, H, and J).

antibody ($p < 0.001$, Figure 4-9F), suggesting stronger prodeath signaling in the absence of autocrine TGF-α. However, C225-mediated increases in caspase-8 cleavage were accompanied by increased levels of the precursor zymogen, procaspase-8 ($p < 10^{-5}$, Figure 4-9E). Recent microarray studies of TNF-treated HT-29 cells have shown that the procaspase-8 mRNA is upregulated threefold after TNF addition, suggesting that transcription is important for controlling procaspase-8 levels[5]. In contrast to C225, IL-1ra did not affect procaspase-8 but significantly decreased prodeath signaling by reducing cleaved caspase-8 levels ($p < 10^{-10}$, Figure

[5] K.A.J. and P.K.S., unpublished observations.

4-9G,H). Thus, TGF-α and IL-1α exert opposing control on caspase-8 processing in HT-29 cells.

Interestingly, although both C225 and IL-1ra perturbed initiator caspase-8 signaling (Figure 4-9F,H), the executioner caspase precursor, procaspase-3, was only affected by IL-1ra (Figure 4-9I,J). Autocrine IL-1α blockade significantly reduced the processing and disappearance of procaspase-3 after 4 hr ($p < 10^{-9}$, Figure 4-9J) because of reduced active caspase-8 (Figure 4-9H). In contrast, procaspase-3 was not substantially affected by the increases in caspase-8 cleavage induced by C225 (Figure 4-9I), raising the possibility that caspase-8 signaling was blocked by a modulator, like an inhibitor of apoptosis (IAP) protein [221]. Taken together, these results show that directly induced stress and apoptotic signals from TNF are quantitatively dependent upon contributions from the early TGF-α and late IL-1α autocrine circuits.

4.2.4. Coupling of TGF-α and IL-1α Autocrine Circuits

Treatment of HT-29 cells with TNF induces two autocrine circuits, one in which TGF-α binds to EGFR within 30 min of TNF addition and one in which IL-1α binds to IL-1R after 4 hr (Figures 4-1 and 4-5). To determine whether the two circuits were linked, we blocked EGFR with C225 and then measured soluble IL-1α levels and IKK activity in response to TNF. C225 completely prevented IL-1α release triggered by low concentrations of TNF (Figure 4-10A) and significantly inhibited IL-1α release induced by high TNF (Figure 4-11). Expectedly, then, we found that low TNF-induced late IKK activity was significantly reduced by twofold at late times, when the IL-1α circuit is normally operative ($p < 0.001$, Figure 4-10B). More surprising was that early TNF-induced IKK was also inhibited ($p < 10^{-10}$), implying that TGF-α cooperated with TNF to activate IKK–NF-κB. EGF-family members activate NF-κB in certain breast cancer cells [222], and more recently, IKKα has been shown to be required for EGF-induced histone H3 phosphorylation [223]. Our IKK assays measure both IKKα and IKKβ [5] and might therefore quantify a mixture of NF-κB activation and histone H3 phosphorylation in the early phase [224, 225]. Regardless, the IL-1α–IL-1R–IKK autocrine circuit after 4 hr (Figure 4-5G) was quantitatively dependent on the TGF-α–EGFR circuit established at earlier times (Figure 4-1B).

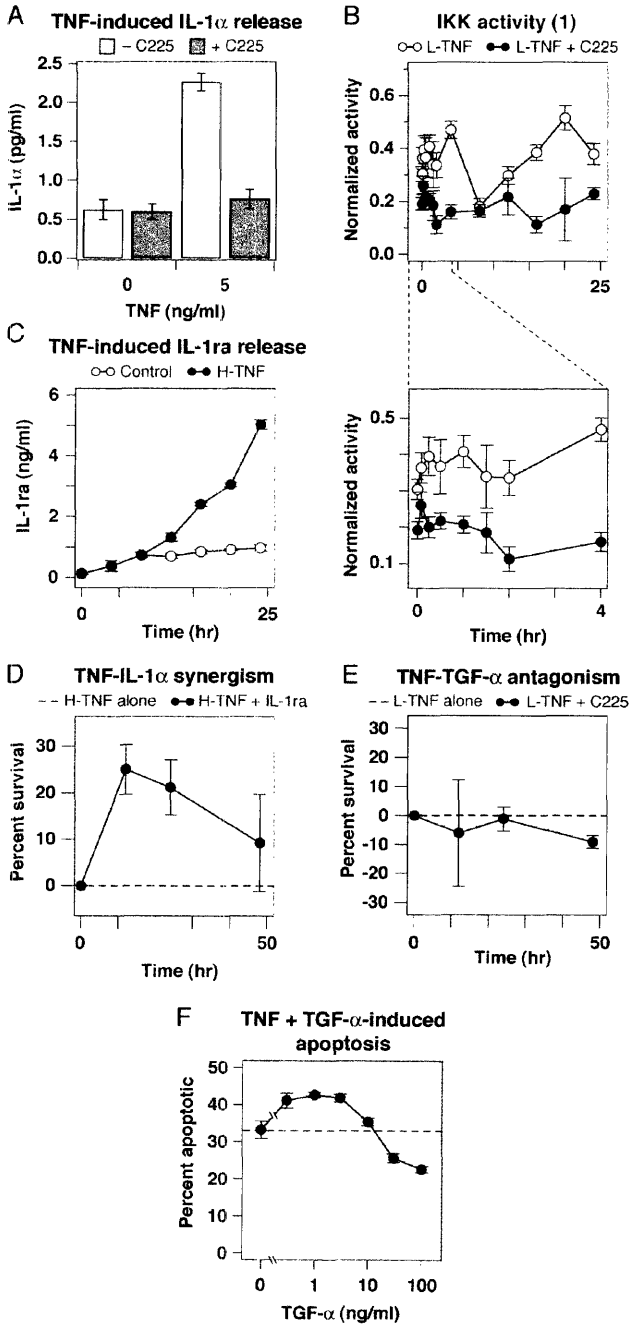

Figure 4-10. TGF-α and IL-1α autocrine circuits are coupled with autocrine IL-1ra to control TNF-induced apoptosis. **(A)** IL-1α release induced by zero or 5 ng/ml TNF with or without 10 μg/ml C225 pretreatment. **(B)** IKK response to 5 ng/ml TNF (L-TNF) with or without 10 μg/ml C225 pretreatment. **(C)** IL-1ra release in response to 100 ng/ml TNF. The control treatment was a mock stimulation with carrier only. **(D)** IL-1α-mediated synergism with TNF-induced apoptosis. HT-29 cells were cotreated with 100 ng/ml TNF (H-TNF) and 10 μg/ml IL-1ra and compared against H-TNF alone (red) for apoptosis at 12, 24, and 48 hr as described in Figure 2-17B. **(E)** TGF-α-mediated antagonism of TNF-induced apoptosis. HT-29 cells were pretreated with 10 μg/ml C225 then stimulated with 5 ng/ml TNF (L-TNF) and compared against L-TNF alone (red) for apoptosis at 12, 24, and 48 hr as described in Figure 2-17B. **(F)** TGF-α-mediated antagonism for TNF-induced apoptosis. HT-29 cells were treated with various concentrations of TGF-α and 5 ng/ml TNF (L-TNF) and compared against L-TNF alone (red) for apoptosis at 24 hr as described in Figure 2-17B. Data are presented as the mean ± S.E.M. of triplicate biological samples as described in [2].

Interestingly, this crosstalk was not reciprocal, because adding exogenous IL-1α alone did not increase the levels of secreted TGF-α (data not shown). Moreover, exogenous TGF-α by itself was not sufficient to provoke IL-1α release, implying that the IL-1α autocrine circuit requires signaling from both TNF and TGF-α (data not shown). Together, this shows that the TGF-α and IL-1α autocrine circuits induced by TNF treatment of HT-29 cells are coupled to each other unidirectionally.

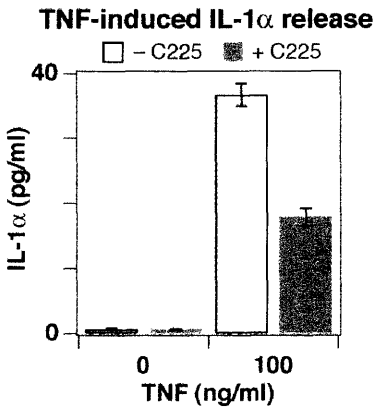

Figure 4-11. The TGF-α autocrine circuit is not absolutely required for TNF-induced IL-1α release. Cells were stimulated with zero or 100 ng/ml TNF with or without 10 μg/ml C225 pretreatment, and IL-1α release into the medium was measured at 24 h as described in [2]. Data are presented as the mean ± S.E.M. of three independent biological samples.

We considered it likely that this TNF-induced autocrine cascade contained additional components. Related transcriptional profiling experiments of TNF-treated HT-29 cells revealed strong upregulation of the IL-1ra gene[6], and IL-1ra ELISA's showed that IL-1ra began to accumulate significantly in the medium of TNF-treated cells after 12 hr (Figure 4-10C). As described above, IL-1ra blocks signaling through IL-1R [226] and its release into the medium constitutes a third, presumably prosurvival, step in the TNF-induced autocrine cascade. These data demonstrate that crosstalk exists between multiple time- and dosage-dependent autocrine circuits activated by TNF.

TGF-α is a growth factor and we therefore expected TNF to be more potent as an inducer of cell death when the TGF-α–EGFR autocrine interaction was blocked. However, only modest changes in TNF-induced apoptosis were observed in the presence of C225 (Figure 4-10D). From the signaling studies, we observed an increase in TNF-induced caspase-8 cleavage upon C225 treatment, as well as reduced ERK activity (Figures 4-1H and 4-9F), both of which are expected to increase cell killing [42, 227]. However, C225 also a reduced JNK1 and MK2 signaling at various times, which should increase cell survival [75].

One way to explain these conflicting findings is by noting that C225 treatment blocks both signaling from EGFR (Figures 4-1G,H) as well as secondary signaling from the coupled IL-1α autocrine circuit (Figure 4-10A). To estimate what fraction of TNF-induced apoptosis could

[6] K.A.J. and P.K.S., unpublished observations.

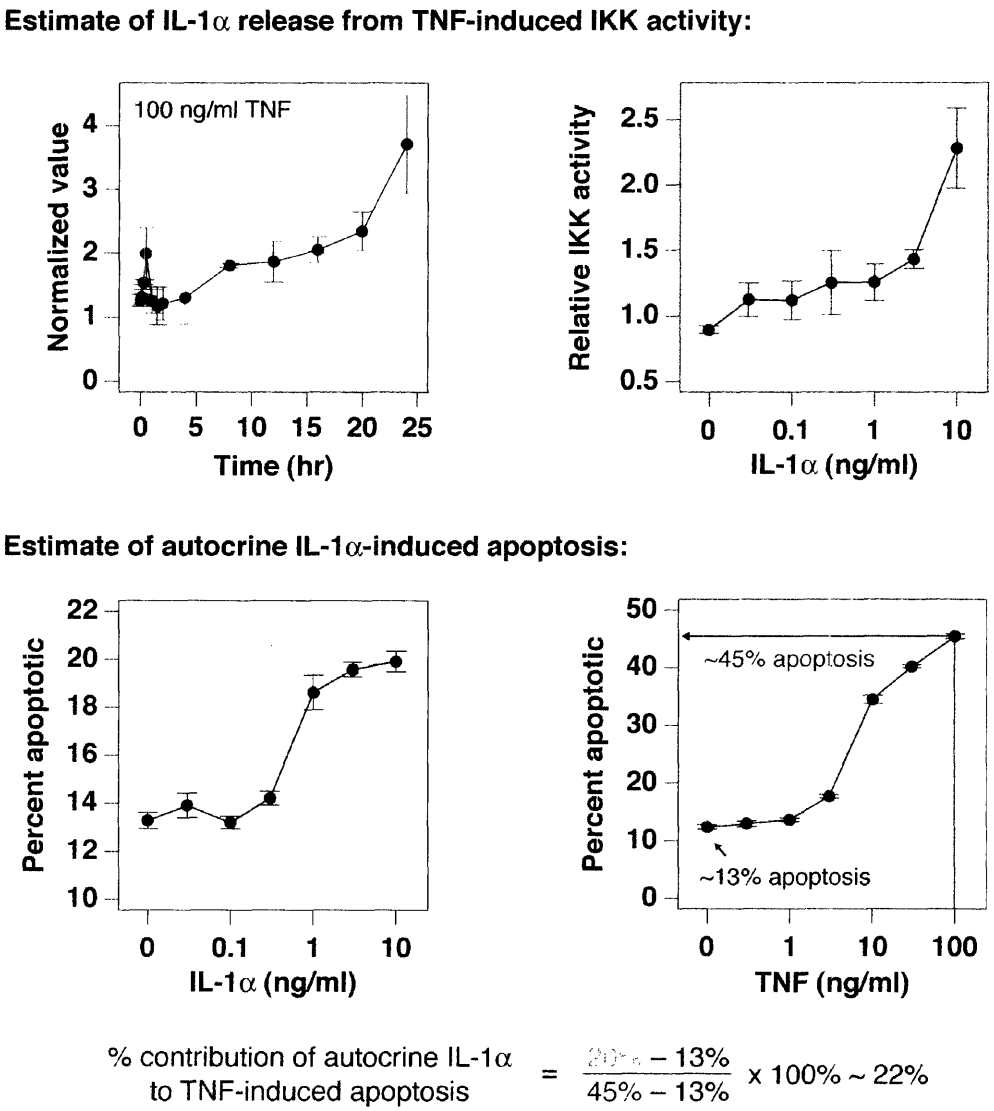

Figure 4-12. Calibrated estimate for the contribution of autocrine IL-1α to TNF-induced apoptosis. The local estimate of autocrine IL-1α (top, same as Figure 4-8) was compared against recombinant IL-1α-induced apoptosis (lower left, same as Figure 4-5I) to estimate ~20% apoptosis. This local concentration of IL-1α was induced by 100 ng/ml TNF (see Figure 4-5), which caused ~45% apoptosis (lower right, same as Figure 2-17A). After correcting for ~13% basal apoptosis caused by mock treatment, we estimated that ~22% of TNF-induced apoptosis could be due to autocrine IL-1α.

be a result of autocrine IL-1α, we indexed the levels of TNF-induced IL-1α to a calibration curve relating exogenous IL-1α concentrations to fraction of cells killed (see Figure 4-12 for details). We estimated that the IL-1R autocrine circuit was responsible for ~25% of the total apoptosis induced by high TNF, which was consistent with our measured IL-1ra-mediated increases in survival at 12 hr (Figure 4-10E). This degree of protection from cell killing is similar to that provided by high levels of exogenous EGF or insulin (Figures 2-17B,C).

Therefore, blocking the TGF-α autocrine circuit does not have a net effect on TNF-induced cell death because a downstream proapoptotic circuit involving IL-1α is also inhibited.

From these data, we might expect exogenous TGF-α itself to have a complex effect on cells. To examine this directly, cells were treated with low TNF and in the presence of varying concentrations of exogenous TGF-α and measured the extent of apoptosis. At very low concentrations (0.01 to 1 ng/ml) TGF-α was observed to potentiate the apoptotic effects of TNF, but at higher concentrations apoptosis was blocked (Figure 4-10F). Based on data presented above, we interpret dosage-dependent changes in TGF-α biology to reflect tight coupling between EGFR and the proapoptotic IL-1α autocrine circuit. Considered together, these data strongly suggest that autocrine cascades have an important function in cell physiology and represent an extracellular extension to trans-antagonistic intracellular signals induced by TNF [38]. Moreover, our findings show how receptor blockade, in this case by a clinically important therapeutic antibody, can substantially perturb an intracellular network, but not manifest itself as a significant change in cell response.

4.3. Identification of a late MK2 prosurvival signaling mechanism

We next compared the full TNF-induced apoptotic signature when the TGF-α and IL-1α autocrine circuits were blocked by C225 antibody and IL-1ra. Interestingly, IL-1α disruption decreased most of the apoptotic outputs relative to TNF alone (Figure 4-13A), suggesting that autocrine IL-1α was an extracellular positive-feedback circuit that augmented the apoptotic response to TNF. In contrast, TGF-α disruption by C225 led to large changes in TNF-induced activation of the network but did not lead to any clear overall changes in the apoptotic outputs (Figure 4-13A). Some outputs, like phosphatidylserine exposure at 12–48 hr, were increased by C225 ($p < 10^{-5}$), whereas others, like cleaved cytokeratin at 12 hr, were decreased ($p < 0.05$), and 6 of the 12 apoptotic outputs did not change significantly [1, 9].

In these cells, autocrine TGF-α causes the early activation of prosurvival signals, like MEK-ERK, from 15–30 min and the late activation of prodeath signals through IL-1α from 12–24 hr (Figure 3-15A) [2]. The net lack of an effect of autocrine TGF-α on apoptosis suggested that late prodeath signals were offset by some unknown late prosurvival signal. Using

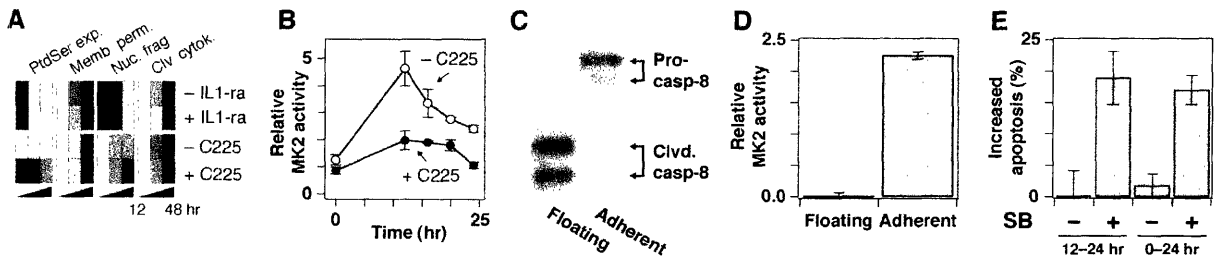

Figure 4-13. The PLS signal-response model suggests late MK2 activity as a TGF-α-induced prosurvival signal. (**A**) Heat-map comparison of TNF-induced apoptotic outputs with and without autocrine IL-1α and autocrine TGF-α. Mean outputs are color-scaled from white (smallest) to black (largest). (**B**) Comparison of late MK2 activities induced by 5 ng/ml TNF with or without C225 to block autocrine TGF-α. Time courses were extracted from the dataset presented in [2]. (**C**) Caspase-8 immunoblot of floating and adherent cells after treatment with 100 ng/ml TNF for 24 hr to confirm separation of apoptotic and viable cells. It was also verified by DAPI staining that isolation of floating and adherent cells had purified the apoptotic and viable subpopulations by over 96% [9]. (**D**) MK2 activity assay of floating and adherent cells after treatment with 100 ng/ml TNF for 20 hr. Data were normalized to basal MK2 activity as described [12]. (**E**) Change in cleaved cytokeratin after treatment with 100 ng/ml TNF for 24 hr with or without SB202190 pretreatment for 12–24 hr or 0–24 hr. Inhibitor experiments were performed as described [5]. For (E), (G), and (H), data are presented as the mean ± S.E.M. of triplicate biological experiments [2].

the PLS model, we searched for candidate molecular signals and identified MK2. In the model, early MK2 kinase activity was positively correlated with apoptosis and therefore prodeath [9]. However, late MK2 signaling after 12 hr was anticorrelated with apoptosis and was inhibited by C225 (Figure 4-13B) [2, 9].

If late MK2 activity was actually a prosurvival signal, then MK2 activity at these times should be present only in viable cells. To investigate this, we compared MK2 signaling in live- vs.-apoptotic cells separately by analyzing the floating and adherent subpopulations of TNF-treated cells (Figure 4-13C). All of the MK2 activity was observed in the viable cells (Figure 4-13D). To verify the importance of late MK2 activity as a prosurvival signal, we rapidly inhibited the MK2-activating kinase, p38, with a specific small-molecule [228, 229]. Inhibition of p38 starting at 12 hr after TNF treatment caused a significant increase in cell death at 24 hr, which was identical to that observed when p38 was inhibited throughout the entire time course of TNF treatment (Figure 4-13E). Thus, the PLS model correctly identified late-phase MK2 activity (or another p38 substrate) as a prosurvival signal in this system.

Our model-driven observation that late MK2 activity was TGF-α-dependent (Figure 4-13B) was intriguing, because the p38 activator, MKK3, is transcriptionally upregulated by growth factors in a Grb2- (and, by extension, MEK-ERK)- dependent fashion [28]. We have found in the HT-29 system that MKK3 is upregulated severalfold in response to TNF [9], raising

the possibility that MEK-ERK signaling through autocrine TGF-α (Figure 3-15A) upregulates late MK2 activity (Figure 4-13B) by transcription. Mechanistic biological hypotheses like these would not have been easily recognized without the data-driven mathematical formalism provided by the PLS model.

4.4. Network mechanism of growth factor-mediated inhibition of TNF-induced apoptosis

In the full model, 660 metrics derived from 7182 signaling measurements were used to predict apoptosis (Figures 3-13E–J and 3-14B). In retrospect, it was not clear that this breadth of signaling information was actually necessary for prediction. If not, then it implied that the apoptotic outputs could be predicted equally well by a reduced number of metrics derived from smaller, more tractable experiments.

To investigate the minimal experimental requirements that were sufficient for network coverage, we constructed models given only a fraction of the available signaling metrics. First, we examined the metrics in the model with the largest relative contributions to the most influential principal components [1]. We found that a model containing only the top 20 most informative metrics (Table 4-1) was nearly as predictive of apoptosis as a full model that used all 660 metrics (Figure 4-14A). A model given fewer than these top 20 metrics was significantly less predictive [9]. The need for at least 20 metrics explicitly supported our hypothesis that individual signaling measurements would not broadly predict cell response (Figure 3-12C–E).

The most noteworthy feature of the top 20 metrics was that they were completely nonobvious. The list of metrics included activation slopes of Akt and IRS1 phosphorylation and integrated peaks of JNK1 and IKK signaling but did not include any caspase metrics (Table 4-1). Since 20 metrics were as predictive as entire list of 660 metrics (Figure 4-14A), it suggested that the apoptotic information was redundantly encoded in the original model. To investigate this, we sequentially removed the top 20, 50, 100, etc. metrics and then recalculated the prediction of apoptosis. Surprisingly, up to the top 350 metrics could be eliminated before the apoptosis model significantly lost predictive ability (Figure 4-14A). This implied that the full biological

system was redundantly encoded with the stimulus-specific information required to mediate all of the apoptotic outputs [230, 231].

All of the above PLS models (Figure 4-14A) included metrics derived from all 19 molecular signals. From an experimental standpoint, it was also important to determine how many individual molecular signals needed to be sampled to broadly predict apoptosis. Intriguingly, when we ranked molecular signals based on the average information contained in their derived metrics, we found that the top three molecular signals belonged to the MAPK cascades: JNK1, MK2, and ERK [9]. A model given only metrics from these three signals performed nearly as well as the full PLS model in predicting the autocrine perturbations (Figure 4-14B). At the same time, a model without JNK1, MK2 and ERK and containing metrics derived from only the seven least informative molecular signals was also equally predictive (Figure 4-14B). Together, these findings suggest that 3–7 relevant molecular signals in a defined intracellular network are sufficient to predict network-dependent output responses. In agreement with these estimates, we found that the maximum per-signal prediction efficiency occurred at

Table 4-1. Top 20 most informative signaling metrics of the PLS model

VIP*	Protein†	Signal†	Metric‡
1.560	IRS-1	P-Ser636	Decay rate, peak #2
1.443	Akt	P-Ser473	Activation slope, peak #3
1.368	IRS-1	P-Ser636	Activation slope, peak #2
1.351	JNK1	Activity	Area under the curve, peak #2
1.343	IKK	Activity	Area under the curve, peak #2
1.297	MK2	Activity	4 hr time point
1.295	MK2	Activity	8 hr time point
1.291	JNK1	Activity	8 hr time point
1.287	IKK	Activity	Steady-state
1.283	JNK1	Activity	Decay rate, peak #2
1.281	MK2	Activity	Area under the curve
1.280	MK2	Activity	Mean
1.280	Akt	P-Ser473	Area under the curve, peak #3
1.276	MK2	Activity	30 min time point
1.274	IKK	Activity	Area under the curve
1.274	IKK	Activity	12 hr time point
1.274	IKK	Activity	20 hr time point
1.273	JNK1	Activity	Mean
1.271	IKK	Activity	16 hr time point
1.271	MK2	Activity	Maximum§

*Variable importance in the projection [1].
†See Table 2-2 and [2] for a complete description of the network measurements.
‡See Table 3-1 and [1] for a complete description of the signaling metrics.
§The top caspase metric, mean cleaved caspase-8, was only 34th overall [9].

4–5 molecular signals in the network (Figure 4-14B, inset).

What biological insight could be gained from the full PLS model? To answer this, we examined the relationship between all of the signaling network measurements and the input treatments through the model's first two principal components [3], which predicted 92% of the apoptotic outputs [9]. These components formed a pair of axes defining the two-dimensional slice through the signaling dataset that most accurately predicted the apoptosis dataset [3]. We found that certain signals and treatments were clearly overrepresented in these dimensions. The first principal component, Axis #1, was heavily oriented toward stress and apoptotic pathways: early JNK1 activity, early MK2 activity, and late cleaved caspase-8 metrics (Table 4-2). In contrast, the second principal component, Axis #2, appeared to constitute a global survival signal

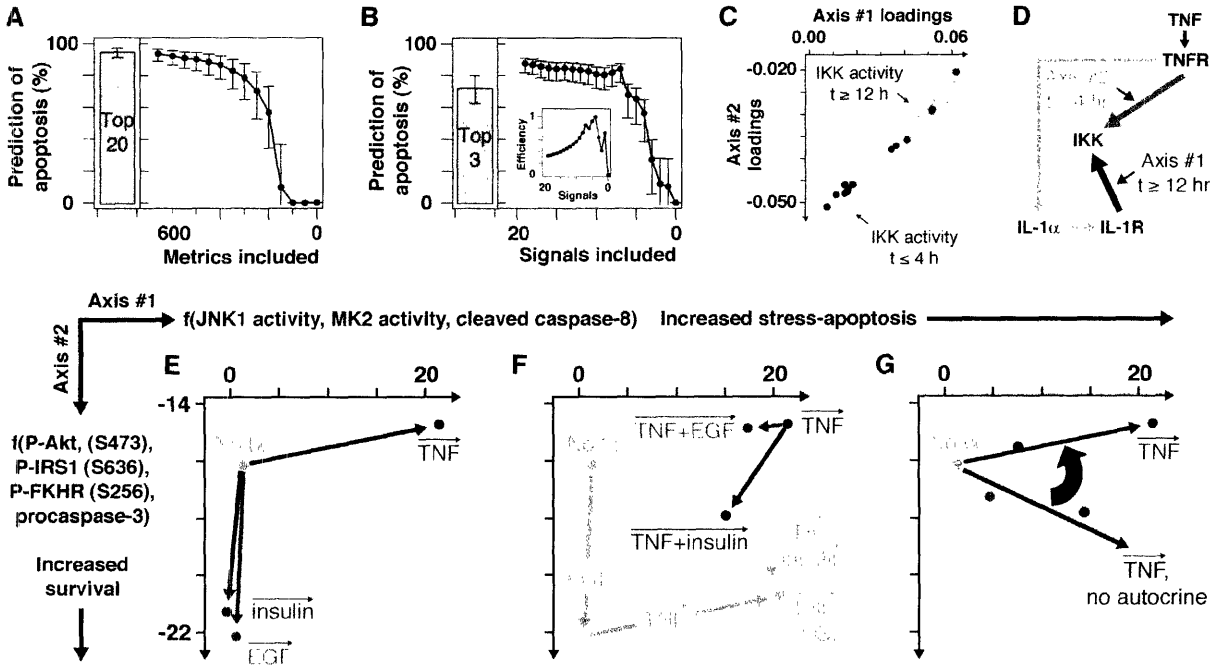

Figure 4-14. Combinations of critical intracellular signals identify a stress-apoptosis axis and a survival axis for cytokine-mediated apoptosis. **(A)** Decrease in TNF-EGF-insulin apoptotic prediction with decreasing number of signaling metrics included. Metrics were eliminated based upon the variable importance of the projection (VIP) [1]. **(B)** Decrease in autocrine apoptotic prediction with decreasing number of molecular signals included. Metrics were eliminated based upon the mean VIP of each signal [1]. (Inset) Predictive efficiency as a function of molecular signals included. To calculate efficiency, the TNF-EGF-insulin apoptotic predictions in (B) were divided by the number of molecular signals in the submodel. **(C)** Model loadings of IKK activity time-point metrics along Axis #1 and #2. Early and late IKK time points are shaded in orange and purple, respectively. The opposite time-dependent loading pattern was observed for MK2 [9]. **(D)** Correspondence between Axis #1-Axis #2 and indirect-direct activation of IKK from TNF. Pathways are excerpted from Fig. 3A. **(E–G)** Projection of selected treatment combinations onto the first two components (Axis #1 and #2) of the PLS model. Samples were projected using the model scores as described [3]. For (A) and (B), data are presented as the central prediction ± 90% Fisher Z-transformed confidence intervals.

Table 4-2. Top 20 loadings in the first principal component of the PLS model

w_1c_1*	Protein†	Signal†	Metric‡
0.0967	Caspase-8	Cleavage	2 hr derivative
0.0951	EGFR	P-Tyr1068/total	16 hr derivative
0.0942	MK2	Activity	5 min derivative
0.0932	JNK1	Activity	5 min derivative
0.0932	Akt	P-Ser473	Activation slope, peak #3
0.0920	JNK1	Activity	Activation slope, peak #1
0.0919	MK2	Activity	15 min time point
0.0910	MK2	Activity	Area under the curve, peak #1
0.0909	JNK1	Activity	15 min time point
0.0902	JNK1	Activity	Maximum
0.0900	Caspase-8	Cleavage	4 hr derivative
0.0889	MK2	Activity	90 min time point
0.0888	JNK1	Activity	30 min time point
0.0880	MK2	Activity	60 min time point
0.0877	Caspase-8	Cleavage	12 hr time point
0.0870	Caspase-8	Cleavage	16 hr time point
0.0868	Caspase-8	Cleavage	8 hr derivative
0.0867	MK2	Activity	30 min time point
0.0861	Akt	Total	20 hr derivative
0.0854	Caspase-8	Cleavage	Steady-state

*Loading weights of the first principal component [1].
†See Table 2-2 and [2] for a complete description of the network measurements.
‡See Table 3-1 and [1] for a complete description of the signaling metrics.

by including P-Akt, P-IRS1, P-FKHR, and procaspase-3 metrics (Table 4-3). Furthermore, the contributions of many signals to the two axes depended upon the time point when the molecular signal was activated. Sometimes, these time-dependent changes corresponded to distinct mechanisms of activation. Early IKK activity directly mediated by TNF was predominantly prosurvival along Axis #2 (Figure 4-14C). In contrast, IKK activity after 12 hr, which occurs indirectly via autocrine IL-1α (Figure 3-15A) [2], was identified by the model as prostress-proapoptotic by its contribution along Axis #1 (Figure 4-14C). Thus, the PLS model axes had separated the time-varying contributions of IKK according to its upstream activators (Figure 4-14D).

Using these axes, we could now reanalyze the signaling contributions to apoptosis when cells processed single and combination cytokines. Expectedly, we found that TNF treatment alone projected strongly along prodeath Axis #1, whereas isolated EGF and insulin treatments mapped exclusively on prosurvival Axis #2 (Figure 4-14E). This reinforced our original

Table 4-3. Top 20 loadings in the second principal component of the PLS model

w_2c_2*	Protein†	Signal†	Metric‡
-0.0695	IKK	Activity	Activation slope, peak #2
-0.0691	EGFR	Total	60 min derivative
-0.0622	Akt	P-Ser473	24 hr time point
-0.0615	MK2	Activity	Decay rate, peak #2
-0.0613	Akt	P-Ser473	8 hr derivative
-0.0608	FKHR	P-Ser256	16 hr time point
-0.0605	Akt	P-Ser473	16 hr time point
-0.0605	EGFR	Total	Area under the curve, peak #3
-0.0603	Akt	P-Ser473	Steady-state
-0.0600	Procaspase-3	Zymogen	12 hr time point
-0.0599	Procaspase-3	Zymogen	Steady-state
-0.0599	Procaspase-3	Zymogen	20 hr time point
-0.0596	Procaspase-3	Zymogen	24 hr time point
-0.0595	Procaspase-3	Zymogen	16 hr time point
-0.0593	Caspase-8	Cleavage	90 min time point
-0.0590	FKHR	P-Ser256	12 hr time point
-0.0587	Akt	P-Ser473	20 hr time point
-0.0586	FKHR	P-Ser256	Steady-state
-0.0583	Akt	P-Ser473	12 hr time point
-0.0581	IRS1	P-Tyr896	24 hr time point§

*Loading weights of the second principal component [1].
†See Table 2-2 and [2] for a complete description of the network measurements.
‡See Table 3-1 and [1] for a complete description of the signaling metrics.
§Five P-IRS1 (Ser636) metrics were identified within the top 30 loadings of the second principal component [9].

intuition that TNF and EGF-insulin stimuli act upon orthogonal and antagonistic signaling axes for apoptosis.

In contrast, the multi-input projections were remarkably different from what would be predicted by a summation of the single-input treatments (Figure 4-14F, gray). TNF, EGF, and insulin each lost a fraction of their original projections along the two axes, indicating that the input stimuli were antagonized when added in combination (Figure 4-14F). Importantly, the TNF + EGF and TNF + insulin projections were now distinctly separated from one another (Figure 4-14F), implying different network mechanisms. EGF appeared to antagonize TNF-induced apoptosis by specifically reducing the projection along the stress-apoptosis Axis #1 without any change along Axis #2 (Figure 4-14F). In contrast, insulin actively promoted prosurvival signaling along Axis #2 while also inhibiting stress-apoptosis signaling along Axis #1. This separation was unexpected, because TNF + EGF and TNF + insulin stimuli elicited nearly identical apoptotic signatures (Figure 2-18I,J). Therefore, analyzing the multi-input

stimuli through these canonical model-derived "network axes" (Figure 4-14E) helped to reveal the different strategies used by EGF and insulin to antagonize TNF-induced apoptosis.

Finally, to determine the contribution of the TNF-induced autocrine circuits in the model, we mapped the TNF + C225 and TNF + IL-1ra treatments (Figure 4-14G). We found that the projection of TNF along the stress-apoptosis axis (Figure 4-14G) was enforced by the autocrine circuits, which increased the contribution along Axis #1 and decreased the contribution along Axis #2. This is consistent with the notion that regulated autocrine circuits provide microenvironment-dependent feedback to cells during phenotypic decision processes, such as death-survival [2, 90]. Furthermore, it directly illustrated that complex environmental stimuli were entirely contained within the two biologically meaningful axes that the PLS model had distilled from the original 660-dimensional signaling metric space (Figure 4-14E).

CHAPTER 5

Conclusions and future directions

5.1. Quantitative high-throughput methods for studying signaling

The current knowledge of intracellular signal transduction is staggeringly complex. To model how network-level properties affect cell function, quantitative experimental techniques that are both high-throughput and multiplex are needed. The kinase activity assay formats presented here represent a first step in this direction. A 96-well, microtiter format is highly versatile, in that it is amenable to scale-up and automated liquid handling, yet tractable for individual scientists and more moderate studies. Importantly, these assays possess linearity, reproducibility, specificity and sensitivity characteristics as good as, or better than, the corresponding low-throughput technique. We anticipate that these functional assays will complement existing proteomic approaches [144] and find broad applicability towards biological and clinical problems involving signal transduction and human disease.

There remains tremendous potential for extending a paradigm of applying engineering design to quantitative biology. There are, of course, many other phosphorylation cascades and other types of potentially informative signals [232, 233]. It seems immediately feasible to design quantitative activity assays for other important classes of signaling enzymes, such as phosphatases and exchange factors. Other types of signals, like protein localization and protein-protein interactions, lack even a low-throughput technique that is quantitative. It would be interesting to consider ways to analyze these signals and compare their predictive value to protein state, level, and activity measurements.

An additional direction to pursue is toward more sensitive signaling assays that are still quantitative. The homogeneous, fluorescence-based kinase activity assays reported here are one example of how sensitivity can be improved. The Sox-based chemosensors are uniquely quantitative in their ability to estimate enzymatic activity and phosphorylated product generation

from cell samples. The amount of lysate activity can easily be related to recombinant protein standards or normalized to untreated control lysates (Figure 3-1). In addition, the assay is compatible with physiological concentrations of ATP. Together with the solution-phase format, this dramatically increases the sensitivity of the assay. Whereas most immune complex kinase activity assays require 200-500 μg total cell protein[147], the kinase assays presented here can make equivalent measurements at less than 100 μg (Figures 2-9b,d and 2-12b,d). This sensitivity is critical for applications where cell samples are limited, such as high-throughput cell-based screening and clinical diagnostics. In addition, crude cell lysate assays eliminate several hours of manipulations, such as incubation and washing steps, before and after the *in vitro* reaction. During the kinase reaction, the 60 time points collected provide additional activity information and reduce error without extra experimental effort.

An important feature of these lysate-based assays is that selectivity can be improved by pharmacologically inhibiting kinases with overlapping substrate specificity. Because the relevant off-target enzymes will depend upon the peptide chemosensor, kinase-by-kinase optimization will be needed to ensure maximum possible specificity. Although some residual off-target activity is inevitable (here, < 30% for Akt and 25% for MK2, Figures 2-9g and 2-12g), this singular limitation is outweighed by the many benefits of these fluorescent kinase activity assays, in that they are straightforward, rapid, continuous, non-radioactive, quantitative, and sensitive. The format is conceptually similar to many fluorogenic protease assays [176] that, because of these same benefits, have found widespread use in a number of applications [177]. Together with the radioactivity-based techniques, these assay platforms are of immediate and expanding utility in drug discovery and molecular biology.

5.2. Autocrine crosstalk in the response of human epithelial cells to apoptotic and mitogenic stimuli

Here we analyzed a proteomic compendium of time-varying changes in cellular signaling induced by pro- and antiapoptotic cytokines in HT-29 colonic epithelial cells. Our overall aim was to determine how the decision between cell proliferation and apoptosis is controlled by opposing death and survival signals. Previous work on this topic focused on crosstalk among

intracellular signaling proteins with pro- and antiapoptotic functions [58]. For example, Akt has been shown to phosphorylate Bad and caspase-9 in response to prosurvival cytokines, which downregulate critical prodeath responses [59, 60]. Conversely, it has been shown that prosurvival EGF-family receptors are targeted for degradation by caspases [62].

Our work highlights the equivalent importance of extracellular crosstalk in cell fate decisions. Specifically, we identify three regulated and overlapping autocrine circuits that play critical roles in controlling TNF-induced apoptosis in HT-29 cells. We find this autocrine cascade is rapidly induced and can activate intracellular signals that are as quantitatively significant as signals immediately downstream of activated TNFR. We propose that crosstalk among autocrine cytokines constitutes a generally important form of biological regulation linking cellular physiology to the extracellular environment.

5.2.1. A tripartite TNF-induced autocrine cascade

TNF triggers a sequential three-part autocrine cascade that plays out over at least 24 hr (Figure 5-1A). Together with TNF, this cascade stacks layers of pro- and antiapoptotic signals that control the death-survival decision. First, proapoptotic signals immediately downstream of TNFR binding are induced within 15 min. A few minutes later, an autocrine TGF-α circuit is established, leading to prosurvival signaling through EGFR. The combination of TNF and TGF-α causes the release of IL-1α starting at 4 hr, which activates prodeath signaling through IL-1R. Finally, upregulation of IL-1ra 4-8 hr later negatively regulates IL-1R-mediated signaling, and presumably constitutes a final antiapoptotic stimulus. The time-dependent interaction of these cascades establishes a robust network of positive and negative signals to set the level of apoptosis in a self-limiting fashion.

The first intracellular step in TNF signaling is the recruitment of TRAF2, FADD, and other adaptor proteins to the DISC on activated TNFR's [38, 39]. Shortly after DISC assembly and 15 min after TNF addition, intracellular kinases such as JNK1, MK2, and IKK are activated by complex I (Figure 5-1A,B). In HT-29 cells these kinases are presumably induced by well-established pathways involving TRAF2-MAP3K's and TRAF2/5-RIP respectively [38]. MAP3K's, such as MEKK1 and ASK1, have been implicated as TNF-dependent activators of

the MKK7 and MKK3 kinases that phosphorylate JNK1 and p38/MK2 [40]. Currently, it is unclear how RIP activates the IKK complex [39].

In parallel with these well-established pathways, TNF also drives the release of membrane-bound pro-TGF-α in a metalloprotease-dependent fashion. While the precise mechanism of metalloprotease activation is unclear, ERK has been implicated in the growth factor-stimulated release of autocrine TGF-α in other settings [209], possibly via metalloprotease phosphorylation [234]. The p38 pathway has also been implicated in the constitutive release of TGF-α and cytokine-stimulated shedding of HB-EGF [235]. However, TNF does not induce HB-EGF shedding in HT-29 cells even though it strongly stimulates p38 activity. Thus, p38 and ERK must differentially control the release of EGF-family members by as-yet unknown

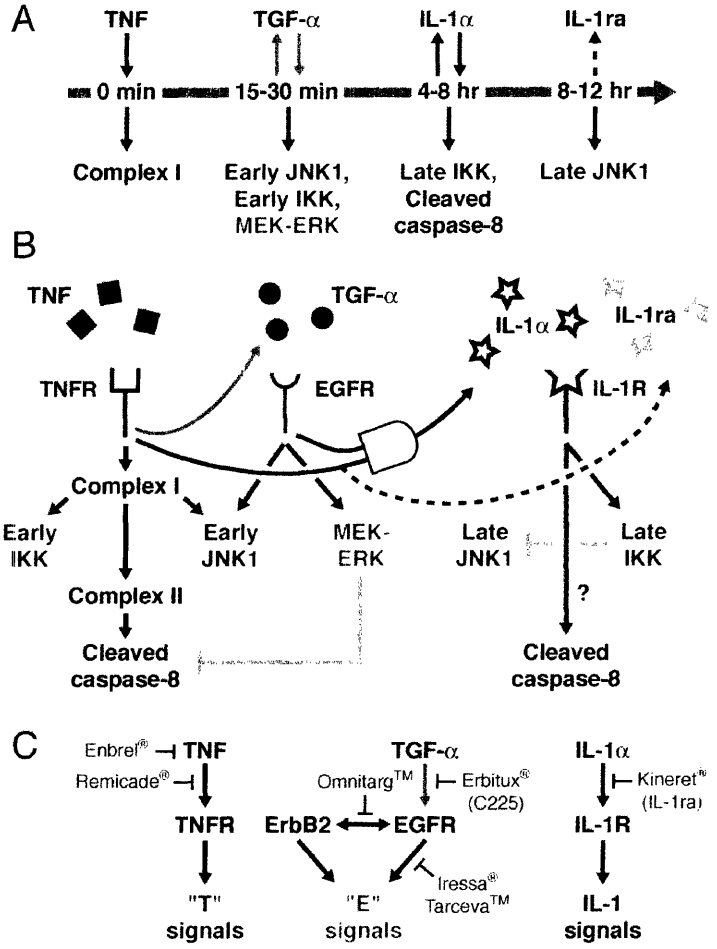

Figure 5-1. Model of TNF-induced extracellular crosstalk. **(A)** Sequence of the time-dependent TNF-induced autocrine cascade established by TGF-α, IL-1α, and IL-1ra. **(B)** Direct and autocrine-indirect intracellular signals activated by TNF. **(C)** Approved pharmaceuticals that target different points of the TNF-induced autocrine cascade.

mechanisms of metalloprotease regulation.

The release of TGF-α into the extracellular medium activates signaling through EGFR and subsequently induces the MEK-ERK pathway (Figure 5-1A). We find that at least 75% of TNF-induced MEK-ERK activation in HT-29 cells is caused by autocrine TGF-α. This is likely to be an underestimate of the extent of autocrine dependence because EGFR blockade by antibodies is incomplete [211] and residual signals from activated receptors are highly amplified [19]. ERK is known to inhibit the activation of caspase-8 [61] consistent with our observation that cleaved caspase-8 levels rise when TGF-α autocrine signaling (and thus, ERK activation) is blocked. Early JNK1 signaling triggered by TNF also appears to be mediated partly by TGF-α, since EGFR blockade reduces JNK1 activity ~ 50%. Thus, the TGF-α–EGFR autocrine circuit appears to directly stimulate multiple intracellular signals that are both pro- and antiapoptotic.

Eight hr after TNF addition and seven hr after maximal TGF-α release, IL-1α begins to accumulate in the medium as a consequence of calpain-mediated proteolysis [219]. IL-1α binds the IL-1R, activating IKK and the NF-κB pathway. JNK1 signaling is reduced ~ 50% by IL-1α, probably due to Gadd45β upregulation by NF-κB [65]. Finally, autocrine IL-1α promotes the cleavage of initiator and effector caspases (Figure 5-1A,B) and leads to increased apoptosis via an as-yet unknown mechanism [236]. Finally, 4–8 hr after IL-1α shedding begins, IL-1ra is released as the third cytokine in the cascade (Figure 4-10F). IL-1ra blocks IL-1R and thereby attenuates the IL-1α autocrine cascade (Figure 5-1B). The overall effect of the IL-1α autocrine circuit is to increase proapoptotic signals (e.g., cleaved caspase-8) and decrease others (e.g., JNK1). The induction of IL-1ra makes the circuit self-limiting and shows how specific receptor antagonists can cause unexpected changes throughout a signaling network.

5.2.2. Induction of intracellular signals by direct and autocrine-indirect processes

Our data support the idea that response of cells to TNF involves a combination of "direct" signals, which lie immediately downstream of the TNFR–DISC complex, and "autocrine-indirect" signals, which are mediated by TNF-released cytokines. TGF-α shedding and DISC formation occur simultaneously, so certain TNF-direct and TGF-α-indirect signals

overlap extensively. Moreover, autocrine signals are transduced within 15–30 min, nearly as rapidly as direct signals, and are as significant quantitatively. Overall, we observe four types of regulation. TNF-induced signals such as MK2 are primarily regulated in a direct fashion by the DISC whereas others, such as MEK and ERK, are regulated almost entirely by autocrine mechanisms[7]. A third set of signaling proteins, such as JNK1, are activated both by TNF-direct and autocrine-indirect mechanisms. Finally, signals such as IKK have the very interesting property of exhibiting direct regulation at early times and autocrine-indirect regulation at later times. Because the target specificity of IKK-induced NF-κB has been shown to change over time as a result of dimer exchange [237], it is possible that early and late phases of IKK activation might result in the activation of distinct sets of genes.

5.2.3. Extracellular crosstalk in the TNF-induced network

Crosstalk in intracellular networks is well recognized [180], but extracellular crosstalk involving autocrine signals has not been extensively documented. A straightforward example of such crosstalk in the current study is the sequential TNF-dependent release of TNF-α, IL-1α, and IL-1ra (Figure 5-1B). Activation of the IL-1α circuit requires prior TGF-α release, and this TGF-α–IL-1α crosstalk is TNF-dependent, because exogenous TGF-α alone does not cause IL-1α release. The TGF-α–IL-1α crosstalk is also directional: exogenous IL-1α alone is insufficient for TGF-α release[8]. The link between IL-1α and IL-1ra appears purely TNF-dependent, since exogenous IL-1α caused a very modest release of IL-1ra. Likewise, TNF-induced IL-1ra release was only slightly affected by TGF-α blockade[9]. In general, we observe a close link between the presence of a cytokine and activation of its cognate receptor. Thus, the directionality and interdependency among autocrine signals is likely to be enforced through regulated proteolysis of precursors, rather than by the modification of receptors ("transactivation"). Further study of the regulatory metallo- and calpain proteases seems warranted.

---

[7] K.A. Janes and P.K. Sorger, unpublished observations.
[8] Ibid.
[9] Ibid.

5.2.4. Biological and clinical significance

Why does the response of cells to TNF involve complex self-limiting extracellular signals? Overlapping positive and negative feedback has been proposed as a means for optimizing the sensitivity and stability of biological systems over a wide ranges of inputs [66]. A role for autocrine circuits in sensing spatial ranges and local environmental cues has also been documented [90]. Thus, multi-step, contingent activation of autocrine signals by TNF might serve to ensure the orderly processing of intracellular and extracellular cues and allow microenvironment-sensitive control over cell death (Figure 5-1B). It seems highly likely that the extent of crosstalk between autocrine and intracellular circuits, and the quantitative significance of various cues, will vary will cell type. In the current study, we use transformed cells grown *in vitro*. Although a fairly artificial experimental setting, immortalized cells were clearly advantageous for initial studies, since the construction of a compendium requires methodological developments and large numbers of cells. It will now be important, however, to examine variations in autocrine-indirect and TNF-direct signaling among primary cells, such as hepatocytes and adipocytes, in which TNF has important physiological and pathological functions [238, 239].

Since its discovery three decades ago as a endotoxin-induced serum factor with tumoricidal activity, TNF has remained an important therapeutic target in a variety of human diseases [240, 241]. Neutralizing anti-TNF antibodies (e.g., Remicade®) and decoy receptors (e.g., Enbrel®) are now used to treat inflammatory bowel disease and rheumatoid arthritis (Figure 5-1C) [45, 242]. However, clinical trials with sepsis and cancer, once promising targets for TNF-directed therapies, were disappointing and highlighted the puzzling inefficacy and side-effects of cytokine-directed therapy [243, 244].

Our experiments provide some insight into this problem by suggesting that secondary autocrine cascades must be considered. One simple example in our current data is with the anti-EGFR antibody C225 (known commercially as Erbitux®). C225 treatment of TNF-stimulated HT-29 cells does not change the level of TNF-induced apoptosis, but extensive changes nonetheless occur in both intracellular and extracellular signaling. Thus, it is not that Erbitux is inactive against HT-29 cells or TNF-induced apoptosis, but rather that the proapoptotic effects of blocking TGF-α autocrine signaling are offset by reduced IL-1α mediated signaling.

Elucidating such interactions among autocrine signals over a variety of tissues appears daunting, but it has the potential to uncover differences between normal and diseased tissues that can be exploited by combination treatments with larger therapeutic indexes than single-drug treatments. Indeed, combination therapies involving anti-TNF biologics and IL-1ra are already being explored for the treatment of rheumatoid arthritis [245].

5.3. Canonical network axes for cytokine-induced apoptosis

By using a systems approach that combines quantitative experiments with data-driven modeling, we have identified two canonical network axes—a stress-apoptosis axis and a survival axis—that together define a reduced signaling space for apoptosis. These axes capture the time-dependent, intracellular signal processing of individual and combination stimuli, as well as autocrine-feedback stimuli (Figure 4-14E–G). Our work illustrates how a complex signaling network can be reduced computationally to a much simpler empirical model that is directly tied to biological mechanism (Figures 4-13B–E and 4-14C,D).

Whether these network axes apply to signaling in general is unknown but clearly warrants further study. Extensions of this work that could examine the generality of the model predictions include:

- Other TNF-family ligands, such as TNF-related apoptosis-inducing ligand (TRAIL) [246]
- Other cell types, such as HeLa, in which TNF-induced cell death is antagonized by growth factors[10]
- Combinations with adhesion signals, like fibronectin and laminin [247]
- Sensitizing agents other than IFN-γ, such as adenoviral infection [248]
- Different apoptosis-inducing stimuli, like DNA-damaging agents [249]
- Therapeutically relevant pathway inhibitors, such as those targeting the EGFR [250]
- Coupling of other output responses with apoptosis, such as proliferation [251]

[10] J.G. Albeck, unpublished observations.

These careful-yet-ambitious extensions would reveal whether the stress-apoptosis and survival axes (Chapter 4.4) reappear as important dimension in other complex biological systems. One can envision a scenario where perhaps 5–10 of these canonical axes are sufficient to predict most of the recognized cellular phenotypes.

5.4. Other data-driven modeling approaches

This thesis has focused largely on multivariate modeling of signaling datasets through PLS regression. It is exciting to speculate that new insights will be gained by applying other data-driven techniques to the same proteomic compendium. There exist a number of other matrix decomposition approaches that reduce dimensionality with different optimization constraints. These include non-negative matrix factorization [252], where basis vectors are constrained to be additive and not subtractive; independent component analysis [253], where axes must correspond to dimensions in the original measurement set; and network component analysis [254], where reduction is constrained by a known connectivity matrix. Compared with the orthogonality constraint of PCA and PLS, these techniques will project the dataset differently in a way that may be more informative for certain applications.

Bayesian network analysis is another data-driven approach that could be readily applied to the TNF-EGF-insulin proteomic compendium [17, 255]. This probabilistic method could be used in an unbiased way to identify from the data the most likely connectivity graph that maps the measured signaling proteins. Such calculations are computationally intensive and require large datasets, but it has recently been shown that resampling from replicated data can improve the power to converge upon a consensus graph confidently [256]. Because the signaling measurements were focused upon a well-understood network (Figure 1-1), it is also possible to include this prior information to score competing models that vary in their edge connectivities [17]. Using Bayesian networks in this fashion could be useful for clarifying regions of the network that are not as well characterized, such as the posttranslation regulation of IRS1 [257].

5.5. The role of transcription in apoptosis-survival cell decisions

Unlike many studies of TNF, which inhibit gene expression by various means [80], our studies permitted cytokine-induced transcripts and proteins to affect the signaling network measurements. Because TNF-induced HT-29 apoptosis occurs on the same time scale as transcription[11], we expect that the apoptosis-survival cell decision contains both transcription-dependent and transcription-independent signals [258, 259]. Consequently, it is likely that the predictive PLS model (Chapter 3.5) took a "shortcut" through the transcriptional branch of the decision process. It would be possible to omit transcription if contributions from important transcripts had been incorporated into the measured signaling proteins. For instance, roles of key

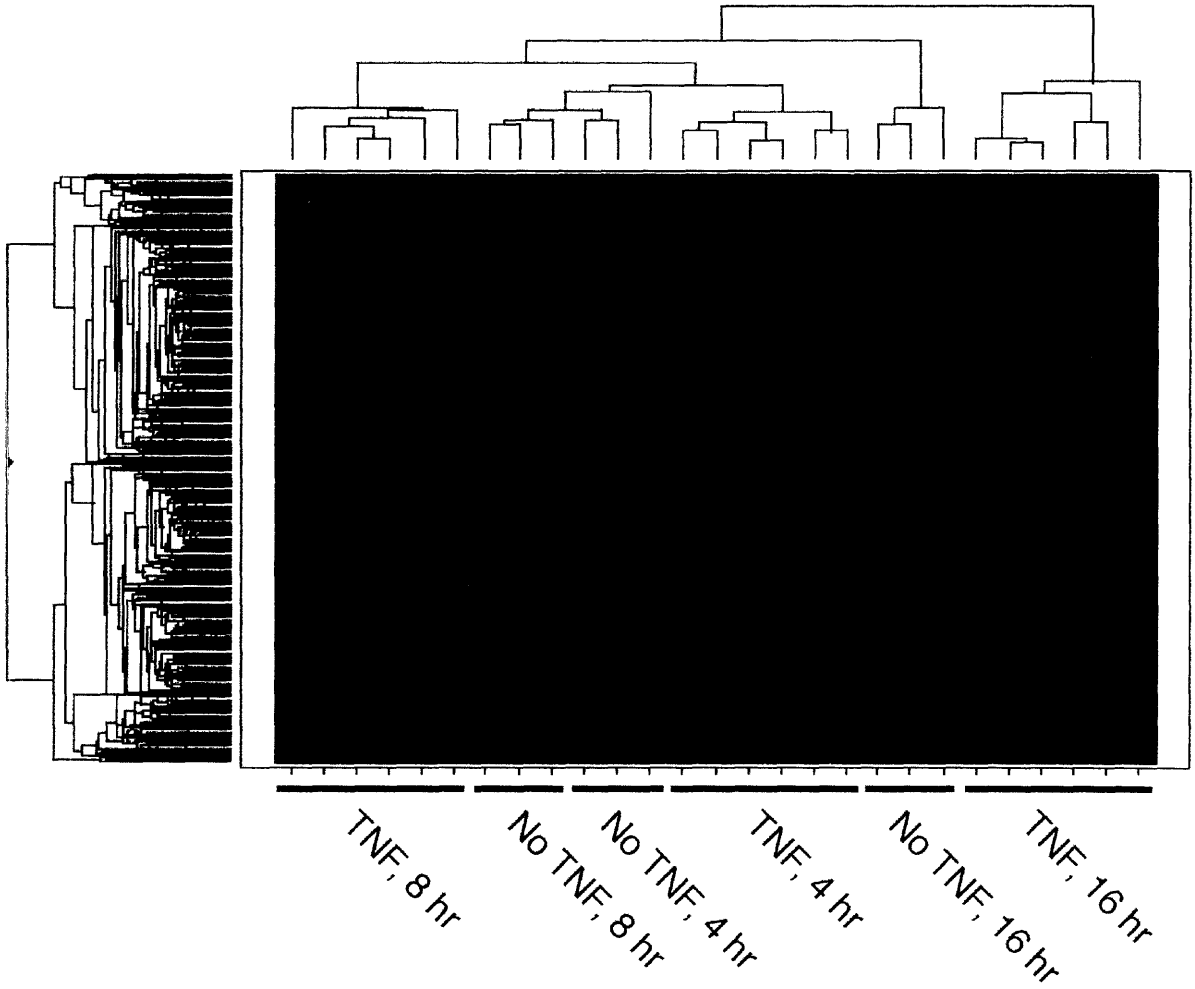

Figure 5-2. Preliminary clustering of transcriptional profiles of HT-29 cells. Note that the TNF-treated time points form distinct clusters relative to the EGF- and insulin-only time points.

IAP proteins [221] could be embedded in the measured caspase cleavage states.

Nonetheless, it will be important to interrogate the role of gene expression in the HT-29 system at a level of detail similar to that achieved for protein signals. In very recent work, we have transcriptionally profiled the responses of HT-29 cells to the nine multi-input stimuli specified in the signaling and apoptosis datasets[12]. Preliminary analyses of these measurements revealed a much stronger clustering of TNF-treated samples compared with EGF- and insulin-treated samples (Figure 5-2). Interestingly, this contrasts the DPLSR mapping of the protein data (Figure 3-11), in which TNF and EGF were more similar compared to insulin. Future studies with these profiles will investigate the interrelationship between protein signals and gene expression responses, as well as examining the relative predictive power of transcriptional data compared to protein data.

5.6. Physiological model systems

The philosophy of "systems biology" [85, 260] has received much fanfare in the scientific community. However, it has been correctly noted that the long-term acceptance of systems biology is heavily dependent upon the emergence of a "success story" [261]. Namely, these systems-level approaches must provide useful biological information to help design real therapies that improve human health. For this to occur, the quantitative engineering approaches like those developed in this thesis need to be applied to real model systems of human physiology and disease.

At present, such model systems must be chosen with care, because it is unclear how our approaches "scale" in systems with less reproducible data, smaller biological samples, and greater uncertainty about the underlying network. However, recent successes in integrating microarray data from diverse tumor biopsies to predict patient survival [262] suggest that these types of experiments are possible. Achieving quantitative signal-response models of human disease will suggest new strategies for therapy and move biology from an anecdotal to a predictive era of discovery.

---

[11] K.A. Janes, unpublished observations.
[12] K.A. Janes and R. Fry, unpublished data.

CHAPTER 6

Appendices

6.1. References

1. Materials and methods are available as supporting material on Science online.
2. Janes, K. A., Gaudet, S., Albeck, J. G., Nielsen, U. B., Lauffenburger, D. A., and Sorger, P. K. (2005). Autocrine crosstalk in the response of human cells to apoptotic and mitogenic stimuli. In preparation.
3. Janes, K. A., Kelly, J. R., Gaudet, S., Albeck, J. G., Sorger, P. K., and Lauffenburger, D. A. (2004). Cue-signal-response analysis of TNF-induced apoptosis by partial least squares regression of dynamic multivariate data. *J Comput Biol* **11,** 544-61.
4. Shults, M. D., Pearce, D. A., and Imperiali, B. (2003). Modular and tunable chemosensor scaffold for divalent zinc. *J. Am. Chem. Soc.* **125,** 10591-10597.
5. Janes, K. A., Albeck, J. G., Peng, L. X., Sorger, P. K., Lauffenburger, D. A., and Yaffe, M. B. (2003). A high-throughput quantitative multiplex kinase assay for monitoring information flow in signaling networks: application to sepsis-apoptosis. *Mol Cell Proteomics* **2,** 463-73.
6. Shults, M. D., Janes, K. A., Lauffenburger, D. A., and Imperiali, B. (2005). A multiplexed homogeneous fluorescence-based assay for protein kinase activity in cell lysates. *Nat Methods* **2,** 277-284.
7. Rives, A. W., and Galitski, T. (2003). Modular organization of cellular networks. *Proc Natl Acad Sci U S A* **100,** 1128-33.
8. Obata, T., Yaffe, M. B., Leparc, G. G., Piro, E. T., Maegawa, H., Kashiwagi, A., Kikkawa, R., and Cantley, L. C. (2000). Peptide and protein library screening defines optimal substrate motifs for Akt/PKB. *J. Biol. Chem.* **275,** 36108-36115.
9. Janes, K. A., Albeck, J. G., and Gaudet, S., unpublished observations.
10. Janes, K. A., Albeck, J. G., Gaudet, S., Nielsen, U. B., Sorger, P. K., and Lauffenburger, D. A. (2004). Multivariate approaches for revealing biological signal-response relationships: analysis of the death-vs.-survival response to TNF-alpha, EGF, and insulin. In "Keystone Symposium on Biological Discovery Using Diverse High-Throughput Data" Steamboat Springs, CO.
11. Manke, I. A., Nguyen, A., Lim, D., Stewart, M. Q., Elia, A. E. H., and Yaffe, M. B. (2005). MAPKAP Kinase-2 is a cell cycle checkpoint kinase that regulates the G_2/M transition and S-phase progression in response to UV irradiation. *Mol. Cell* **17,** 37-48.

12. Gaudet, S., Janes, K. A., Lewis, C. L., Albeck, J. G., Lauffenburger, D. A., and Sorger, P. K. (2005). A quantitative compendium of signaling and apoptotic responses to prodeath and prosurvival cytokines. in preparation.
13. Kemp, B. E., Graves, D. J., Benjamini, E., and Krebs, E. G. (1977). Role of multiple basic residues in determining the substrate specificity of cyclic AMP-dependent protein kinase. *J. Biol. Chem.* **252,** 4888-4894.
14. Shults, M. D., and Imperiali, B. (2003). Versatile fluorescence probes of protein kinase activity. *J Am Chem Soc* **125,** 14248-9.
15. Gomez, S. M., and Rzhetsky, A. (2002). Towards the prediction of complete protein--protein interaction networks. *Pac Symp Biocomput,* 413-24.
16. Nielsen, U. B., Cardone, M. H., Sinskey, A. J., MacBeath, G., and Sorger, P. K. (2003). Profiling receptor tyrosine kinase activation by using Ab microarrays. *Proc Natl Acad Sci U S A* **100,** 9330-5.
17. Sachs, K., Gifford, D., Jaakkola, T., Sorger, P., and Lauffenburger, D. A. (2002). Bayesian network approach to cell signaling pathway modeling. *Sci STKE* **2002,** PE38.
18. Huang, S., and Ingber, D. E. (2000). Shape-dependent control of cell growth, differentiation, and apoptosis: switching between attractors in cell regulatory networks. *Exp Cell Res* **261,** 91-103.
19. Schoeberl, B., Eichler-Jonsson, C., Gilles, E. D., and Muller, G. (2002). Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nat Biotechnol* **20,** 370-5.
20. Janeway, C. A., Travers, P., Walport, M., and Shlomchik, M. J. (2001). "Immunobiology," 5th ed., Garland, New York.
21. Downward, J. (2001). The ins and outs of signalling. *Nature* **411,** 759-62.
22. Sporn, M. B., and Roberts, A. B. (1988). Peptide growth factors are multifunctional. *Nature* **332,** 217-9.
23. Fischer, O. M., Streit, S., Hart, S., and Ullrich, A. (2003). Beyond Herceptin and Gleevec. *Curr Opin Chem Biol* **7,** 490-5.
24. Madhani, H. D. (2001). Accounting for specificity in receptor tyrosine kinase signaling. *Cell* **106,** 9-11.
25. Avruch, J. (1998). Insulin signal transduction through protein kinase cascades. *Mol Cell Biochem* **182,** 31-48.
26. Yarden, Y., and Sliwkowski, M. X. (2001). Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol* **2,** 127-37.
27. Halfon, M. S., Carmena, A., Gisselbrecht, S., Sackerson, C. M., Jimenez, F., Baylies, M. K., and Michelson, A. M. (2000). Ras pathway specificity is determined by the integration of multiple signal-activated and tissue-restricted transcription factors. *Cell* **103,** 63-74.
28. Fambrough, D., McClure, K., Kazlauskas, A., and Lander, E. S. (1999). Diverse signaling pathways activated by growth factor receptors induce broadly overlapping, rather than independent, sets of genes. *Cell* **97,** 727-41.
29. Klinghoffer, R. A., Mueting-Nelsen, P. F., Faerman, A., Shani, M., and Soriano, P. (2001). The two PDGF receptors maintain conserved signaling in vivo despite divergent embryological functions. *Mol Cell* **7,** 343-54.

30. Maina, F., Pante, G., Helmbacher, F., Andres, R., Porthin, A., Davies, A. M., Ponzetto, C., and Klein, R. (2001). Coupling Met to specific pathways results in distinct developmental outcomes. *Mol Cell* **7,** 1293-306.
31. Pawson, T. (2004). Specificity in signal transduction: from phosphotyrosine-SH2 domain interactions to complex cellular systems. *Cell* **116,** 191-203.
32. Whitmarsh, A. J., and Davis, R. J. (2000). Regulation of transcription factor function by phosphorylation. *Cell Mol Life Sci* **57,** 1172-83.
33. Yaffe, M. B., and Elia, A. E. (2001). Phosphoserine/threonine-binding domains. *Curr Opin Cell Biol* **13,** 131-8.
34. Songyang, Z., Blechner, S., Hoagland, N., Hoekstra, M. F., Piwnica-Worms, H., and Cantley, L. C. (1994). Use of an oriented peptide library to determine the optimal substrates of protein kinases. *Curr Biol* **4,** 973-82.
35. Biondi, R. M., and Nebreda, A. R. (2003). Signalling specificity of Ser/Thr protein kinases through docking-site-mediated interactions. *Biochem J* **372,** 1-13.
36. Songyang, Z., Shoelson, S. E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R. J., and et al. (1993). SH2 domains recognize specific phosphopeptide sequences. *Cell* **72,** 767-78.
37. Elia, A. E., Rellos, P., Haire, L. F., Chao, J. W., Ivins, F. J., Hoepker, K., Mohammad, D., Cantley, L. C., Smerdon, S. J., and Yaffe, M. B. (2003). The molecular basis for phosphodependent substrate targeting and regulation of Plks by the Polo-box domain. *Cell* **115,** 83-95.
38. Baud, V., and Karin, M. (2001). Signal transduction by tumor necrosis factor and its relatives. *Trends Cell Biol* **11,** 372-7.
39. Chen, G., and Goeddel, D. V. (2002). TNF-R1 signaling: a beautiful pathway. *Science* **296,** 1634-5.
40. Wajant, H., Pfizenmaier, K., and Scheurich, P. (2003). Tumor necrosis factor signaling. *Cell Death Differ* **10,** 45-65.
41. Micheau, O., and Tschopp, J. (2003). Induction of TNF receptor I-mediated apoptosis via two sequential signaling complexes. *Cell* **114,** 181-90.
42. Nicholson, D. W., and Thornberry, N. A. (1997). Caspases: killer proteases. *Trends Biochem Sci* **22,** 299-306.
43. Karin, M., and Ben-Neriah, Y. (2000). Phosphorylation meets ubiquitination: the control of NF-[kappa]B activity. *Annu Rev Immunol* **18,** 621-63.
44. Micheau, O., Lens, S., Gaide, O., Alevizopoulos, K., and Tschopp, J. (2001). NF-kappaB signals induce the expression of c-FLIP. *Mol Cell Biol* **21,** 5299-305.
45. Rutgeerts, P., Van Assche, G., and Vermeire, S. (2004). Optimizing anti-TNF treatment in inflammatory bowel disease. *Gastroenterology* **126,** 1593-610.
46. Chailler, P., and Menard, D. (1999). Ontogeny of EGF receptors in the human gut. *Front Biosci* **4,** D87-101.
47. Singh, P., and Rubin, N. (1993). Insulinlike growth factors and binding proteins in colon cancer. *Gastroenterology* **105,** 1218-37.
48. Gibson, S., Tu, S., Oyer, R., Anderson, S. M., and Johnson, G. L. (1999). Epidermal growth factor protects epithelial cells against Fas-induced apoptosis. Requirement for Akt activation. *J Biol Chem* **274,** 17612-8.

49. Garcia-Lloret, M. I., Yui, J., Winkler-Lowen, B., and Guilbert, L. J. (1996). Epidermal growth factor inhibits cytokine-induced apoptosis of primary human trophoblasts. *J Cell Physiol* **167,** 324-32.
50. Akca, H., Akan, S. Y., Yanikoglu, A., and Ozes, O. N. (2003). Suppression of TNF-alpha mediated apoptosis by EGF in TNF-alpha sensitive human cervical carcinoma cell line. *Growth Factors* **21,** 31-9.
51. Bedard, S., Marcotte, B., and Marette, A. (1998). Insulin inhibits inducible nitric oxide synthase in skeletal muscle cells. *Diabetologia* **41,** 1523-7.
52. Wu, Y., Tewari, M., Cui, S., and Rubin, R. (1996). Activation of the insulin-like growth factor-I receptor inhibits tumor necrosis factor-induced cell death. *J Cell Physiol* **168,** 499-509.
53. Goetze, S., Blaschke, F., Stawowy, P., Bruemmer, D., Spencer, C., Graf, K., Grafe, M., Law, R. E., and Fleck, E. (2001). TNFalpha inhibits insulin's antiapoptotic signaling in vascular smooth muscle cells. *Biochem Biophys Res Commun* **287,** 662-70.
54. Weiner, F. R., Smith, P. J., Wertheimer, S., and Rubin, C. S. (1991). Regulation of gene expression by insulin and tumor necrosis factor alpha in 3T3-L1 cells. Modulation of the transcription of genes encoding acyl-CoA synthetase and stearoyl-CoA desaturase-1. *J Biol Chem* **266,** 23525-8.
55. Qian, H., Hausman, D. B., Compton, M. M., Martin, R. J., Della-Fera, M. A., Hartzell, D. L., and Baile, C. A. (2001). TNFalpha induces and insulin inhibits caspase 3-dependent adipocyte apoptosis. *Biochem Biophys Res Commun* **284,** 1176-83.
56. Kelly, K., Cochran, B. H., Stiles, C. D., and Leder, P. (1983). Cell-specific regulation of the c-myc gene by lymphocyte mitogens and platelet-derived growth factor. *Cell* **35,** 603-10.
57. Owen, N. E., and Villereal, M. L. (1983). Lys-bradykinin stimulates Na+ influx and DNA synthesis in cultured human fibroblasts. *Cell* **32,** 979-85.
58. Kyriakis, J. M. (2001). Life-or-death decisions. *Nature* **414,** 265-6.
59. Datta, S. R., Dudek, H., Tao, X., Masters, S., Fu, H., Gotoh, Y., and Greenberg, M. E. (1997). Akt phosphorylation of BAD couples survival signals to the cell-intrinsic death machinery. *Cell* **91,** 231-41.
60. Cardone, M. H., Roy, N., Stennicke, H. R., Salvesen, G. S., Franke, T. F., Stanbridge, E., Frisch, S., and Reed, J. C. (1998). Regulation of cell death protease caspase-9 by phosphorylation. *Science* **282,** 1318-21.
61. Holmstrom, T. H., Schmitz, I., Soderstrom, T. S., Poukkula, M., Johnson, V. L., Chow, S. C., Krammer, P. H., and Eriksson, J. E. (2000). MAPK/ERK signaling in activated T cells inhibits CD95/Fas-mediated apoptosis downstream of DISC assembly. *Embo J* **19,** 5418-28.
62. Benoit, V., Chariot, A., Delacroix, L., Deregowski, V., Jacobs, N., Merville, M. P., and Bours, V. (2004). Caspase-8-dependent HER-2 cleavage in response to tumor necrosis factor alpha stimulation is counteracted by nuclear factor kappaB through c-FLIP-L expression. *Cancer Res* **64,** 2684-91.
63. Lotem, J., and Sachs, L. (1999). Cytokines as suppressors of apoptosis. *Apoptosis* **4,** 187-96.
64. Tang, G., Yang, J., Minemoto, Y., and Lin, A. (2001). Blocking caspase-3-mediated proteolysis of IKKbeta suppresses TNF-alpha-induced apoptosis. *Mol Cell* **8,** 1005-16.

65. De Smaele, E., Zazzeroni, F., Papa, S., Nguyen, D. U., Jin, R., Jones, J., Cong, R., and Franzoso, G. (2001). Induction of gadd45beta by NF-kappaB downregulates pro-apoptotic JNK signalling. *Nature* **414,** 308-13.
66. Stelling, J., Sauer, U., Szallasi, Z., Doyle, F. J., 3rd, and Doyle, J. (2004). Robustness of cellular functions. *Cell* **118,** 675-85.
67. Zaslaver, A., Mayo, A. E., Rosenberg, R., Bashkin, P., Sberro, H., Tsalyuk, M., Surette, M. G., and Alon, U. (2004). Just-in-time transcription program in metabolic pathways. *Nat Genet* **36,** 486-91.
68. McAdams, H. H., and Shapiro, L. (2003). A bacterial cell-cycle regulatory network operating in time and space. *Science* **301,** 1874-7.
69. Hunter, T. (2000). Signaling--2000 and beyond. *Cell* **100,** 113-27.
70. Werlen, G., Hausmann, B., Naeher, D., and Palmer, E. (2003). Signaling life and death in the thymus: timing is everything. *Science* **299,** 1859-63.
71. Marshall, C. J. (1995). Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation. *Cell* **80,** 179-85.
72. Murphy, L. O., Smith, S., Chen, R. H., Fingar, D. C., and Blenis, J. (2002). Molecular interpretation of ERK signal duration by immediate early gene products. *Nat Cell Biol* **4,** 556-64.
73. 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,** 144-53.
74. Davis, R. J. (2000). Signal transduction by the JNK group of MAP kinases. *Cell* **103,** 239-52.
75. Varfolomeev, E. E., and Ashkenazi, A. (2004). Tumor necrosis factor: an apoptosis JuNKie? *Cell* **116,** 491-7.
76. Guo, Y. L., Baysal, K., Kang, B., Yang, L. J., and Williamson, J. R. (1998). Correlation between sustained c-Jun N-terminal protein kinase activation and apoptosis induced by tumor necrosis factor-alpha in rat mesangial cells. *J Biol Chem* **273,** 4027-34.
77. Lamb, J. A., Ventura, J. J., Hess, P., Flavell, R. A., and Davis, R. J. (2003). JunD mediates survival signaling by the JNK signal transduction pathway. *Mol Cell* **11,** 1479-89.
78. Lei, K., and Davis, R. J. (2003). JNK phosphorylation of Bim-related members of the Bcl2 family induces Bax-dependent apoptosis. *Proc Natl Acad Sci U S A* **100,** 2432-7.
79. Kamata, H., Honda, S., Maeda, S., Chang, L., Hirata, H., and Karin, M. (2005). Reactive oxygen species promote TNFalpha-induced death and sustained JNK activation by inhibiting MAP kinase phosphatases. *Cell* **120,** 649-61.
80. Ventura, J. J., Cogswell, P., Flavell, R. A., Baldwin, A. S., Jr., and Davis, R. J. (2004). JNK potentiates TNF-stimulated necrosis by increasing the production of cytotoxic reactive oxygen species. *Genes Dev* **18,** 2905-15.
81. Deng, Y., Ren, X., Yang, L., Lin, Y., and Wu, X. (2003). A JNK-dependent pathway is required for TNFalpha-induced apoptosis. *Cell* **115,** 61-70.
82. Lipp, P., Thomas, D., Berridge, M. J., and Bootman, M. D. (1997). Nuclear calcium signalling by individual cytoplasmic calcium puffs. *Embo J* **16,** 7166-73.
83. Berridge, M. J. (1990). Calcium oscillations. *J Biol Chem* **265,** 9583-6.
84. Brent, R. (2000). Genomic biology. *Cell* **100,** 169-83.
85. Kitano, H. (2002). Systems biology: a brief overview. *Science* **295,** 1662-4.

86. Ideker, T., and Lauffenburger, D. (2003). Building with a scaffold: emerging strategies for high- to low-level cellular modeling. *Trends Biotechnol* **21,** 255-62.
87. Endy, D., and Brent, R. (2001). Modelling cellular behaviour. *Nature* **409,** 391-5.
88. Asthagiri, A. R., and Lauffenburger, D. A. (2000). Bioengineering models of cell signaling. *Annu Rev Biomed Eng* **2,** 31-53.
89. Papin, J. A., Hunter, T., Palsson, B. O., and Subramaniam, S. (2005). Reconstruction of cellular signalling networks and analysis of their properties. *Nat Rev Mol Cell Biol* **6,** 99-111.
90. Wiley, H. S., Shvartsman, S. Y., and Lauffenburger, D. A. (2003). Computational modeling of the EGF-receptor system: a paradigm for systems biology. *Trends Cell Biol* **13,** 43-50.
91. Chen, Z., Gibson, T. B., Robinson, F., Silvestro, L., Pearson, G., Xu Be, B. E., Wright, A., Vanderbilt, C., and Cobb, M. H. (2001). MAP Kinases. *Chem Rev* **101,** 2449-76.
92. Kyriakis, J. M., and Avruch, J. (2001). Mammalian mitogen-activated protein kinase signal transduction pathways activated by stress and inflammation. *Physiol Rev* **81,** 807-69.
93. Huang, C. Y., and Ferrell, J. E., Jr. (1996). Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proc Natl Acad Sci U S A* **93,** 10078-83.
94. 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,** 895-8.
95. Heinrich, R., Neel, B. G., and Rapoport, T. A. (2002). Mathematical models of protein kinase signal transduction. *Mol Cell* **9,** 957-70.
96. 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,** 1018-23.
97. Brent, R. (2004). A partnership between biology and engineering. *Nat Biotechnol* **22,** 1211-4.
98. Asthagiri, A. R., and Lauffenburger, D. A. (2001). A computational study of feedback effects on signal dynamics in a mitogen-activated protein kinase (MAPK) pathway model. *Biotechnol Prog* **17,** 227-39.
99. 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,** 169-74.
100. 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,** 365-73.
101. Asthagiri, A. R., Nelson, C. M., Horwitz, A. F., and Lauffenburger, D. A. (1999). Quantitative relationship among integrin-ligand binding, adhesion, and signaling via focal adhesion kinase and extracellular signal-regulated kinase 2. *J Biol Chem* **274,** 27119-27.
102. York, R. D., Molliver, D. C., Grewal, S. S., Stenberg, P. E., McCleskey, E. W., and Stork, P. J. (2000). Role of phosphoinositide 3-kinase and endocytosis in nerve growth factor-induced extracellular signal-regulated kinase activation via Ras and Rap1. *Mol Cell Biol* **20,** 8069-83.
103. Lee, E., Salic, A., Kruger, R., Heinrich, R., and Kirschner, M. W. (2003). The roles of APC and Axin derived from experimental and theoretical analysis of the Wnt pathway. *PLoS Biol* **1,** E10.

104. Hoffmann, A., Levchenko, A., Scott, M. L., and Baltimore, D. (2002). The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation. *Science* **298,** 1241-5.
105. DiMilla, P. A., Barbee, K., and Lauffenburger, D. A. (1991). Mathematical model for the effects of adhesion and mechanics on cell migration speed. *Biophys J* **60,** 15-37.
106. Lander, A. D., Nie, Q., and Wan, F. Y. (2002). Do morphogen gradients arise by diffusion? *Dev Cell* **2,** 785-96.
107. Sorger, P. K. (2005). A reductionist's systems biology. *Curr Opin Cell Biol* **17,** 9-11.
108. Johnson, S. A., and Hunter, T. (2005). Kinomics: methods for deciphering the kinome. *Nat Methods* **2,** 17-25.
109. Morohashi, M., Winn, A. E., Borisuk, M. T., Bolouri, H., Doyle, J., and Kitano, H. (2002). Robustness as a measure of plausibility in models of biochemical networks. *J Theor Biol* **216,** 19-30.
110. Ingolia, N. T., and Murray, A. W. (2004). The ups and downs of modeling the cell cycle. *Curr Biol* **14,** R771-7.
111. Kornberg, A. (2003). Ten commandments of enzymology, amended. *Trends Biochem Sci* **28,** 515-7.
112. Northrup, S. H., and Erickson, H. P. (1992). Kinetics of protein-protein association explained by Brownian dynamics computer simulation. *Proc Natl Acad Sci U S A* **89,** 3338-42.
113. Schreiber, G., and Fersht, A. R. (1996). Rapid, electrostatically assisted association of proteins. *Nat Struct Biol* **3,** 427-31.
114. Bentele, M., Lavrik, I., Ulrich, M., Stosser, S., Heermann, D. W., Kalthoff, H., Krammer, P. H., and Eils, R. (2004). Mathematical modeling reveals threshold mechanism in CD95-induced apoptosis. *J Cell Biol* **166,** 839-51.
115. Fussenegger, M., Bailey, J. E., and Varner, J. (2000). A mathematical model of caspase function in apoptosis. *Nat Biotechnol* **18,** 768-74.
116. Schuck, P., and Minton, A. P. (1996). Kinetic analysis of biosensor data: elementary tests for self-consistency. *Trends Biochem Sci* **21,** 458-60.
117. Janes, K. A., Albeck, J. G., Sorger, P. K., Lauffenburger, D. A., and Yaffe, M. B. (2005). A systems model of intracellular signaling defines two predictive network axes for cytokine-induced apoptosis. In preparation.
118. Mashima, T., Naito, M., and Tsuruo, T. (1999). Caspase-mediated cleavage of cytoskeletal actin plays a positive role in the process of morphological apoptosis. *Oncogene* **18,** 2423-30.
119. Wang, J. S., Coburn, J. P., Tauber, A. I., and Zaner, K. S. (1997). Role of gelsolin in actin depolymerization of adherent human neutrophils. *Mol Biol Cell* **8,** 121-8.
120. Hinck, L., Nathke, I. S., Papkoff, J., and Nelson, W. J. (1994). Dynamics of cadherin/catenin complex formation: novel protein interactions and pathways of complex assembly. *J Cell Biol* **125,** 1327-40.
121. Butler, D. (2001). Data, data, everywhere. *Nature* **414,** 840-1.
122. Ball, P. (2002). Picture this. *Nature* **418,** 11-3.
123. Bork, P., Jensen, L. J., von Mering, C., Ramani, A. K., Lee, I., and Marcotte, E. M. (2004). Protein interaction networks from yeast to human. *Curr Opin Struct Biol* **14,** 292-9.

124. von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S. G., Fields, S., and Bork, P. (2002). Comparative assessment of large-scale data sets of protein-protein interactions. *Nature* **417,** 399-403.
125. Vidal, M. (2005). Interactome modeling. *FEBS Lett* **579,** 1834-8.
126. Babu, M. M., Luscombe, N. M., Aravind, L., Gerstein, M., and Teichmann, S. A. (2004). Structure and evolution of transcriptional regulatory networks. *Curr Opin Struct Biol* **14,** 283-91.
127. Begley, T. J., Rosenbach, A. S., Ideker, T., and Samson, L. D. (2004). Hot spots for modulating toxicity identified by genomic phenotyping and localization mapping. *Mol Cell* **16,** 117-25.
128. Said, M. R., Begley, T. J., Oppenheim, A. V., Lauffenburger, D. A., and Samson, L. D. (2004). Global network analysis of phenotypic effects: protein networks and toxicity modulation in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **101,** 18006-11.
129. Shen-Orr, S. S., Milo, R., Mangan, S., and Alon, U. (2002). Network motifs in the transcriptional regulation network of Escherichia coli. *Nat Genet* **31,** 64-8.
130. Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Network motifs: simple building blocks of complex networks. *Science* **298,** 824-7.
131. Kelley, B. P., Sharan, R., Karp, R. M., Sittler, T., Root, D. E., Stockwell, B. R., and Ideker, T. (2003). Conserved pathways within bacteria and yeast as revealed by global protein network alignment. *Proc Natl Acad Sci U S A* **100,** 11394-9.
132. Sharan, R., Suthram, S., Kelley, R. M., Kuhn, T., McCuine, S., Uetz, P., Sittler, T., Karp, R. M., and Ideker, T. (2005). Conserved patterns of protein interaction in multiple species. *Proc Natl Acad Sci U S A* **102,** 1974-9.
133. Han, J. D., Bertin, N., Hao, T., Goldberg, D. S., Berriz, G. F., Zhang, L. V., Dupuy, D., Walhout, A. J., Cusick, M. E., Roth, F. P., and Vidal, M. (2004). Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* **430,** 88-93.
134. Barrios-Rodiles, M., Brown, K. R., Ozdamar, B., Bose, R., Liu, Z., Donovan, R. S., Shinjo, F., Liu, Y., Dembowy, J., Taylor, I. W., Luga, V., Przulj, N., Robinson, M., Suzuki, H., Hayashizaki, Y., Jurisica, I., and Wrana, J. L. (2005). High-throughput mapping of a dynamic signaling network in mammalian cells. *Science* **307,** 1621-5.
135. Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Bloomfield, C. D., and Lander, E. S. (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286,** 531-7.
136. Shipp, M. A., Ross, K. N., Tamayo, P., Weng, A. P., Kutok, J. L., Aguiar, R. C., Gaasenbeek, M., Angelo, M., Reich, M., Pinkus, G. S., Ray, T. S., Koval, M. A., Last, K. W., Norton, A., Lister, T. A., Mesirov, J., Neuberg, D. S., Lander, E. S., Aster, J. C., and Golub, T. R. (2002). Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nat Med* **8,** 68-74.
137. Pomeroy, S. L., Tamayo, P., Gaasenbeek, M., Sturla, L. M., Angelo, M., McLaughlin, M. E., Kim, J. Y., Goumnerova, L. C., Black, P. M., Lau, C., Allen, J. C., Zagzag, D., Olson, J. M., Curran, T., Wetmore, C., Biegel, J. A., Poggio, T., Mukherjee, S., Rifkin, R., Califano, A., Stolovitzky, G., Louis, D. N., Mesirov, J. P., Lander, E. S., and Golub, T. R. (2002). Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature* **415,** 436-42.

138. Ebert, B. L., and Golub, T. R. (2004). Genomic approaches to hematologic malignancies. *Blood* **104,** 923-32.
139. Ma, X. J., Wang, Z., Ryan, P. D., Isakoff, S. J., Barmettler, A., Fuller, A., Muir, B., Mohapatra, G., Salunga, R., Tuggle, J. T., Tran, Y., Tran, D., Tassin, A., Amon, P., Wang, W., Enright, E., Stecker, K., Estepa-Sabal, E., Smith, B., Younger, J., Balis, U., Michaelson, J., Bhan, A., Habin, K., Baer, T. M., Brugge, J., Haber, D. A., Erlander, M. G., and Sgroi, D. C. (2004). A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. *Cancer Cell* **5,** 607-16.
140. Irish, J. M., Hovland, R., Krutzik, P. O., Perez, O. D., Bruserud, O., Gjertsen, B. T., and Nolan, G. P. (2004). Single cell profiling of potentiated phospho-protein networks in cancer cells. *Cell* **118,** 217-28.
141. Xu, X. Q., Leow, C. K., Lu, X., Zhang, X., Liu, J. S., Wong, W. H., Asperger, A., Deininger, S., and Eastwood Leung, H. C. (2004). Molecular classification of liver cirrhosis in a rat model by proteomics and bioinformatics. *Proteomics* **4,** 3235-45.
142. Yanagisawa, K., Shyr, Y., Xu, B. J., Massion, P. P., Larsen, P. H., White, B. C., Roberts, J. R., Edgerton, M., Gonzalez, A., Nadaf, S., Moore, J. H., Caprioli, R. M., and Carbone, D. P. (2003). Proteomic patterns of tumour subsets in non-small-cell lung cancer. *Lancet* **362,** 433-9.
143. Nguyen, A., and Yaffe, M. B. (2003). Proteomics and systems biology approaches to signal transduction in sepsis. *Crit Care Med* **31,** S1-6.
144. Tyers, M., and Mann, M. (2003). From genomics to proteomics. *Nature* **422,** 193-7.
145. Asthagiri, A. R., Horwitz, A. F., and Lauffenburger, D. A. (1999). A rapid and sensitive quantitative kinase activity assay using a convenient 96-well format. *Anal Biochem* **269,** 342-7.
146. Whitmarsh, A. J., and Davis, R. J. (2001). Analyzing JNK and p38 mitogen-activated protein kinase activity. *Methods Enzymol* **332,** 319-36.
147. Hill, M. M., and Hemmings, B. A. (2002). Analysis of protein kinase B/Akt. *Methods Enzymol* **345,** 448-63.
148. Kupfer, R., and Scheinman, R. I. (2002). Measurement of IKK activity in primary rat T cells: rapid activation and inactivation. *J Immunol Methods* **266,** 155-64.
149. Hall, J. P., and Davis, R. J. (2002). Analysis of c-Jun N-terminal kinase regulation and function. *Methods Enzymol.* **345,** 413-425.
150. Whitmarsh, A. J., and Davis, R. J. (2001). Analyzing JNK and p38 mitogen-activated protein kinase activity. *Methods Enzymol.* **332,** 319-336.
151. Li, H., Sims, C. E., Kaluzova, M., Stanbridge, E. J., and Allbritton, N. L. (2004). A quantitative single-cell assay for protein kinase B reveals important insights into the biochemical behavior of an intracellular substrate peptide. *Biochemistry* **43,** 1599-1608.
152. Kawai, Y., Sato, M., and Umezawa, Y. (2004). Single color fluorescent indicators of protein phosphorylation for multicolor imaging of intracellular signal flow dynamics. *Anal. Chem.* **76,** 6144-6149.
153. Schleifenbaum, A., Stier, G., Gasch, A., Sattler, M., and Schultz, C. (2004). Genetically encoded FRET probe for PKC activity based on pleckstrin. *J. Am. Chem. Soc.* **126,** 11786-11787.
154. Sasaki, K., Sato, M., and Umezawa, Y. (2003). Fluorescence indicators for Akt/protein kinase B and dynamics of Akt activity visualized in living cells. *J. Biol. Chem.* **278,** 30945-30951.

155. Violin, J. D., Zhang, J., Tsien, R. Y., and Newton, A. C. (2003). A genetically encoded fluorescent reporter reveals oscillatory phosphorylation by protein kinase C. *J. Cell Biol.* **161,** 899-909.
156. Kunkel, M. T., Ni, Q., Tsien, R. Y., Zhang, J., and Newton, A. C. (2004). Spatio-temporal dynamics of protein kinase B/Akt signaling revealed by a genetically-encoded fluorescent reporter. *J. Biol. Chem.*, E-publication December 6.
157. Sato, M., Ozawa, T., Inukai, K., Asano, T., and Umezawa, Y. (2002). Fluorescent indicators for imaging protein phosphorylation in single living cells. *Nat. Biotechnol.* **20,** 287-294.
158. Ting, A. Y., Kain, K. H., Klemke, R. L., and Tsien, R. Y. (2001). Genetically encoded fluorescent reporters of protein tyrosine kinase activities in living cells. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 15003-15008.
159. Zhang, J., Ma, Y., Taylor, S. S., and Tsien, R. Y. (2001). Genetically encoded reporters of protein kinase A activity reveal impact of substrate tethering. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 14997-15002.
160. Kurokawa, K., Mochizuki, N., Ohba, Y., Mizuno, H., Miyawaki, A., and Matsuda, M. (2001). A pair of fluorescent resonance energy transfer-based probes for tyrosine phosphorylation of the CrkII adaptor protein *in vivo*. *J. Biol. Chem.* **276,** 31305-31310.
161. Nagai, Y., Miyazaki, M., Aoki, R., Zama, T., Inouye, S., Hirose, K., Iino, M., and Hagiwara, M. (2000). A fluorescent indicator for visualizing cAMP-induced phosphorylation *in vivo*. *Nat. Biotechnol.* **18,** 313-316.
162. Higashi, H., Sato, K., Ohtake, A., Omori, A., Yoshida, S., and Kudo, Y. (1997). Imagining of cAMP-dependent protein kinase activity in living neural cells using a novel fluorescent substrate. *FEBS Lett.* **414,** 55-60.
163. Yeh, R.-H., Yan, X., Cammer, M., Bresnick, A. R., and Lawrence, D. S. (2002). Real time visualization of protein kinase activity in living cells. *J. Biol. Chem.* **277,** 11527-11532.
164. Shults, M. D., and Imperiali, B. (2003). Versatile fluorescence probes of protein kinase activity. *J. Am. Chem. Soc.* **125,** 14248-14249.
165. Brazil, D. P., Yang, Z.-Z., and Hemmings, B. A. (2004). Advances in protein kinase B signalling: AKTion on multiple fronts. *Trends Biochem. Sci.* **29,** 233-242.
166. Downward, J. (2004). PI 3-kinase, Akt and cell survival. *Semin. Cell Dev. Biol.* **15,** 177-182.
167. Roux, P. P., and Blenis, J. (2004). ERK and p38 MAPK-activated protein kinases: a family of protein kinases with diverse biological functions. *Microbiol. Mol. Biol. Rev.* **68,** 320-344.
168. Kotlyarov, A., Yannoni, Y., Fritz, S., Laab, K., Telliez, Jean-Baptiste, Pitman, D., Lin, L.-L., and Gaestel, M. (2002). Distinct cellular functions of MK2. *Mol. Cell. Biol.* **22,** 4827-4835.
169. Shabb, J. B. (2001). Physiological substrates of cAMP-dependent protein kinase. *Chem. Rev.* **101,** 2381-2411.
170. Taylor, S. S., Yang, J., Wu, J., Haste, N. M., Radzio-Andzelm, E., and Anand, G. (2004). PKA: a portrait of protein kinase dynamics. *Biochim. Biophys. Acta* **1697,** 259-269.
171. Janes, K. A., Albeck, J. G., Peng, L. X., Sorger, P. K., Lauffenburger, D. A., and Yaffe, M. B. (2003). A high-throughput quantitative multiplex kinase assay for monitoring information flow in signaling networks. *Mol. Cell. Proteomics* **2,** 463-473.

172. Davies, S. P., Reddy, H., Caivano, M., and Cohen, P. (2000). Specificity and mechanism of action of some commonly used protein kinase inhibitors. *Biochem. J.* **351,** 95-105.
173. Andjelkovic, M., Alessi, D. R., Meier, R., Fernandez, A., Lamb, N. J., Frech, M., Cron, P., Cohen, P., Lucocq, J. M., and Hemmings, B. A. (1997). Role of translocation in the activation and function of protein kinase B. *J. Biol. Chem.* **272,** 31515-31524.
174. Coffer, P. J. (2000). Phosphatidylinositol 3-kinase signalling: a tale of two kinase activities. *In* "Protein Kinase Functions" (J. R. Woodgett, Ed.), Oxford University Press, Oxford.
175. Glass, D. B., Cheng, H.-C., Mende-Mueller, L., Reed, J., and Walsh, D. A. (1989). Primary structural determinants essential for potent inhibition of cAMP-dependent protein kinase by inhibitory peptides corresponding to the active portion of the heat-stable inhibitor protein. *J. Biol. Chem.* **264,** 8802-8810.
176. Smith, R. E., Bissell, E. R., Mitchell, A. R., and Pearson, K. W. (1980). Direct photometric or fluorometric assay of proteinases using substrates containing 7-amino-4-trifluoromethylcoumarin. *Thromb. Res.* **17,** 393-402.
177. Gurtu, V., Kain, S. R., and Zhang, G. (1997). Fluorometric and colorimetric detection of caspase activity associated with apoptosis. *Anal. Biochem.* **251,** 98-102.
178. Abreu-Martin, M. T., Vidrich, A., Lynch, D. H., and Targan, S. R. (1995). Divergent induction of apoptosis and IL-8 secretion in HT-29 cells in response to TNF-alpha and ligation of Fas antigen. *J Immunol* **155,** 4147-54.
179. Fransen, L., Van der Heyden, J., Ruysschaert, R., and Fiers, W. (1986). Recombinant tumor necrosis factor: its effect and its synergism with interferon-gamma on a variety of normal and transformed human cell lines. *Eur J Cancer Clin Oncol* **22,** 419-26.
180. Dumont, J. E., Dremier, S., Pirson, I., and Maenhaut, C. (2002). Cross signaling, cell specificity, and physiology. *Am J Physiol Cell Physiol* **283,** C2-28.
181. Remacle-Bonnet, M. M., Garrouste, F. L., Heller, S., Andre, F., Marvaldi, J. L., and Pommier, G. J. (2000). Insulin-like growth factor-I protects colon cancer cells from death factor-induced apoptosis by potentiating tumor necrosis factor alpha-induced mitogen-activated protein kinase and nuclear factor kappaB signaling pathways. *Cancer Res* **60,** 2007-17.
182. Del Bino, G., Darzynkiewicz, Z., Degraef, C., Mosselmans, R., Fokan, D., and Galand, P. (1999). Comparison of methods based on annexin-V binding, DNA content or TUNEL for evaluating cell death in HL-60 and adherent MCF-7 cells. *Cell Prolif* **32,** 25-37.
183. Darzynkiewicz, Z., Juan, G., Li, X., Gorczyca, W., Murakami, T., and Traganos, F. (1997). Cytometry in cell necrobiology: analysis of apoptosis and accidental cell death (necrosis). *Cytometry* **27,** 1-20.
184. Pennisi, E. (2003). Systems biology. Tracing life's circuitry. *Science* **302,** 1646-9.
185. Cross, D. A. E., Alessi, D. R., Cohen, P., Andjelkovic, M., and Hemmings, B. A. (1995). Inhibition of glycogen synthase kinase-3 by insulin mediated by protein kinase B. *Nature* **378,** 785-789.
186. Cross, D. A. E., Watt, P. W., Shaw, M., van der Kaay, J., Downes, C. P., Holder, J. C., and Cohen, P. (1997). Insulin activates protein kinase B, inhibits glycogen synthase kinase-3 and activates glycogen synthase by rapamycin-insensitive pathways in skeletal muscle and adipose tissue. *FEBS Lett.* **406,** 211-215.

187. Hazzalin, C. A., Cuenda, A., Cano, E., Cohen, P., and Mahadevan, L. C. (1997). Effects of the inhibition of p38/RK MAP kinase on induction of five *fos* and *jun* genes by diverse stimuli. *Oncogene* **15**.
188. Neve, R. M., Holbro, T., and Hynes, N. E. (2002). Distinct roles for phosphoinositide 3-dinase, mitogen-activated protein kinase and p38 MAPK in mediating cell cycle progression of breast cancer cells. *Oncogene* **21,** 4567-4576.
189. Lawlor, M. A., and Alessi, D. R. (2001). PKB/Akt: a key mediator of cell proliferation, survival and insulin responses? *J Cell Sci* **114,** 2903-10.
190. Vlahos, C. J., Matter, W. F., Hui, K. Y., and Brown, R. F. (1994). A specific inhibitor of phosphatidylinositol 3-kinase, 2-(4-morpholinyl)-8-phenyl-4H-1-benzopyran-4-one (LY294002). *J Biol Chem* **269,** 5241-8.
191. Leers, M. P., Kolgen, W., Bjorklund, V., Bergman, T., Tribbick, G., Persson, B., Bjorklund, P., Ramaekers, F. C., Bjorklund, B., Nap, M., Jornvall, H., and Schutte, B. (1999). Immunocytochemical detection and mapping of a cytokeratin 18 neo-epitope exposed during early apoptosis. *J Pathol* **187,** 567-72.
192. Fasano, G., and Franceschini, A. (1987). A Multidimensional Version of the Kolmogorov-Smirnov Test. *Monthly Notices of the Royal Astronomical Society* **225,** 155-170.
193. Wilson, C. A., and Browning, J. L. (2002). Death of HT29 adenocarcinoma cells induced by TNF family receptor activation is caspase-independent and displays features of both apoptosis and necrosis. *Cell Death Differ* **9,** 1321-33.
194. Rohlf, F. J., and Corti, M. (2000). Use of two-block partial least-squares to study covariation in shape. *Syst Biol* **49,** 740-53.
195. Kohn, A. D., Takeuchi, F., and Roth, R. A. (1996). Akt, a pleckstrin homology domain containing kinase, is activated primarily by phosphorylation. *J Biol Chem* **271,** 21920-6.
196. Alessi, D. R., Andjelkovic, M., Caudwell, B., Cron, P., Morrice, N., Cohen, P., and Hemmings, B. A. (1996). Mechanism of activation of protein kinase B by insulin and IGF-1. *Embo J* **15,** 6541-51.
197. Dudek, H., Datta, S. R., Franke, T. F., Birnbaum, M. J., Yao, R., Cooper, G. M., Segal, R. A., Kaplan, D. R., and Greenberg, M. E. (1997). Regulation of neuronal survival by the serine-threonine protein kinase Akt. *Science* **275,** 661-5.
198. Kull, F. C., Jr. (1988). The TNF receptor in TNF-mediated cytotoxicity. *Nat Immun Cell Growth Regul* **7,** 254-65.
199. Dunger, A., Schroder, D., Augstein, P., Witstruck, T., Wachlin, G., Vogt, L., Ziegler, B., and Schmidt, S. (1995). Impact of metabolic activity of beta cells on cytokine-induced damage and recovery of rat pancreatic islets. *Acta Diabetol* **32,** 217-24.
200. Asthagiri, A. R., Reinhart, C. A., Horwitz, A. F., and Lauffenburger, D. A. (2000). The role of transient ERK2 signals in fibronectin- and insulin-mediated DNA synthesis. *J Cell Sci* **113,** 4499-510.
201. Abreu-Martin, M. T., Palladino, A. A., Faris, M., Carramanzana, N. M., Nel, A. E., and Targan, S. R. (1999). Fas activates the JNK pathway in human colonic epithelial cells: lack of a direct role in apoptosis. *Am J Physiol* **276,** G599-605.
202. Geladi, P., and Kowalski, B. R. (1986). Partial Least-Squares Regression - a Tutorial. *Analytica Chimica Acta* **185,** 1-17.

203. Alter, O., Brown, P. O., and Botstein, D. (2000). Singular value decomposition for genome-wide expression data processing and modeling. *Proc Natl Acad Sci U S A* **97,** 10101-6.
204. Izumi, H., Ono, M., Ushiro, S., Kohno, K., Kung, H. F., and Kuwano, M. (1994). Cross talk of tumor necrosis factor-alpha and epidermal growth factor in human microvascular endothelial cells. *Exp Cell Res* **214,** 654-62.
205. Hirota, K., Murata, M., Itoh, T., Yodoi, J., and Fukuda, K. (2001). Redox-sensitive transactivation of epidermal growth factor receptor by tumor necrosis factor confers the NF-kappa B activation. *J Biol Chem* **276,** 25953-8.
206. Chen, W. N., Woodbury, R. L., Kathmann, L. E., Opresko, L. K., Zangar, R. C., Wiley, H. S., and Thrall, B. D. (2004). Induced autocrine signaling through the epidermal growth factor receptor contributes to the response of mammary epithelial cells to tumor necrosis factor alpha. *J Biol Chem* **279,** 18488-96.
207. Anzano, M. A., Rieman, D., Prichett, W., Bowen-Pope, D. F., and Greig, R. (1989). Growth factor production by human colon carcinoma cell lines. *Cancer Res* **49,** 2898-904.
208. Culouscou, J. M., Garrouste, F., Remacle-Bonnet, M., Bettetini, D., Marvaldi, J., and Pommier, G. (1988). Autocrine secretion of a colorectum-derived growth factor by HT-29 human colon carcinoma cell line. *Int J Cancer* **42,** 895-901.
209. Fan, H., and Derynck, R. (1999). Ectodomain shedding of TGF-alpha and other transmembrane proteins is induced by receptor tyrosine kinase activation and MAP kinase signaling cascades. *Embo J* **18,** 6962-72.
210. Batsilas, L., Berezhkovskii, A. M., and Shvartsman, S. Y. (2003). Stochastic model of autocrine and paracrine signals in cell culture assays. *Biophys J* **85,** 3659-65.
211. Lauffenburger, D. A., Oehrtman, G. T., Walker, L., and Wiley, H. S. (1998). Real-time quantitative measurement of autocrine ligand binding indicates that autocrine loops are spatially localized. *Proc Natl Acad Sci U S A* **95,** 15368-73.
212. Schievella, A. R., Chen, J. H., Graham, J. R., and Lin, L. L. (1997). MADD, a novel death domain protein that interacts with the type 1 tumor necrosis factor receptor and activates mitogen-activated protein kinase. *J Biol Chem* **272,** 12069-75.
213. Peschon, J. J., Slack, J. L., Reddy, P., Stocking, K. L., Sunnarborg, S. W., Lee, D. C., Russell, W. E., Castner, B. J., Johnson, R. S., Fitzner, J. N., Boyce, R. W., Nelson, N., Kozlosky, C. J., Wolfson, M. F., Rauch, C. T., Cerretti, D. P., Paxton, R. J., March, C. J., and Black, R. A. (1998). An essential role for ectodomain shedding in mammalian development. *Science* **282,** 1281-4.
214. Prenzel, N., Zwick, E., Daub, H., Leserer, M., Abraham, R., Wallasch, C., and Ullrich, A. (1999). EGF receptor transactivation by G-protein-coupled receptors requires metalloproteinase cleavage of proHB-EGF. *Nature* **402,** 884-8.
215. Bender, K., Gottlicher, M., Whiteside, S., Rahmsdorf, H. J., and Herrlich, P. (1998). Sequential DNA damage-independent and -dependent activation of NF-kappaB by UV. *Embo J* **17,** 5170-81.
216. Dinarello, C. A. (1997). Interleukin-1. *Cytokine Growth Factor Rev* **8,** 253-65.
217. Panja, A., Goldberg, S., Eckmann, L., Krishen, P., and Mayer, L. (1998). The regulation and functional consequence of proinflammatory cytokine binding on human intestinal epithelial cells. *J Immunol* **161,** 3675-84.

218. Schmidt, C., Peng, B., Li, Z., Sclabas, G. M., Fujioka, S., Niu, J., Schmidt-Supprian, M., Evans, D. B., Abbruzzese, J. L., and Chiao, P. J. (2003). Mechanisms of proinflammatory cytokine-induced biphasic NF-kappaB activation. *Mol Cell* **12,** 1287-300.
219. Kobayashi, Y., Yamamoto, K., Saido, T., Kawasaki, H., Oppenheim, J. J., and Matsushima, K. (1990). Identification of calcium-activated neutral protease as a processing enzyme of human interleukin 1 alpha. *Proc Natl Acad Sci U S A* **87,** 5548-52.
220. Minden, A., Lin, A., McMahon, M., Lange-Carter, C., Derijard, B., Davis, R. J., Johnson, G. L., and Karin, M. (1994). Differential activation of ERK and JNK mitogen-activated protein kinases by Raf-1 and MEKK. *Science* **266,** 1719-23.
221. Salvesen, G. S., and Duckett, C. S. (2002). IAP proteins: blocking the road to death's door. *Nat Rev Mol Cell Biol* **3,** 401-10.
222. Biswas, D. K., Cruz, A. P., Gansberger, E., and Pardee, A. B. (2000). Epidermal growth factor-induced nuclear factor kappa B activation: A major pathway of cell-cycle progression in estrogen-receptor negative breast cancer cells. *Proc Natl Acad Sci U S A* **97,** 8542-7.
223. Anest, V., Cogswell, P. C., and Baldwin, A. S., Jr. (2004). IkappaB kinase alpha and p65/RelA contribute to optimal epidermal growth factor-induced c-fos gene expression independent of IkappaBalpha degradation. *J Biol Chem* **279,** 31183-9.
224. Anest, V., Hanson, J. L., Cogswell, P. C., Steinbrecher, K. A., Strahl, B. D., and Baldwin, A. S. (2003). A nucleosomal function for IkappaB kinase-alpha in NF-kappaB-dependent gene expression. *Nature* **423,** 659-63.
225. Yamamoto, Y., Verma, U. N., Prajapati, S., Kwak, Y. T., and Gaynor, R. B. (2003). Histone H3 phosphorylation by IKK-alpha is critical for cytokine-induced gene expression. *Nature* **423,** 655-9.
226. Dinarello, C. A. (2000). The role of the interleukin-1-receptor antagonist in blocking inflammation mediated by interleukin-1. *N Engl J Med* **343,** 732-4.
227. Ballif, B. A., and Blenis, J. (2001). Molecular mechanisms mediating mammalian mitogen-activated protein kinase (MAPK) kinase (MEK)-MAPK cell survival signals. *Cell Growth Differ* **12,** 397-408.
228. Rouse, J., Cohen, P., Trigon, S., Morange, M., Alonso-Llamazares, A., Zamanillo, D., Hunt, T., and Nebreda, A. R. (1994). A novel kinase cascade triggered by stress and heat shock that stimulates MAPKAP kinase-2 and phosphorylation of the small heat shock proteins. *Cell* **78,** 1027-37.
229. Lee, J. C., Laydon, J. T., McDonnell, P. C., Gallagher, T. F., Kumar, S., Green, D., McNulty, D., Blumenthal, M. J., Heys, J. R., Landvatter, S. W., Strickler, J. E., McLaughlin, M. M., Siemens, I. R., Fisher, S. M., Livi, G. P., White, J. R., Adams, J. L., and P.R., Y. (1994). A protein kinase involved in the regulation of inflammatory cytokine biosynthesis. *Nature* **372,** 739-46.
230. Yadav, D., and Sarvetnick, N. (2003). Cytokines and autoimmunity: redundancy defines their complex nature. *Curr Opin Immunol* **15,** 697-703.
231. Ozaki, K., and Leonard, W. J. (2002). Cytokine and cytokine receptor pleiotropy and redundancy. *J Biol Chem* **277,** 29355-8.
232. Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002). The protein kinase complement of the human genome. *Science* **298,** 1912-34.
233. Hanahan, D., and Weinberg, R. A. (2000). The hallmarks of cancer. *Cell* **100,** 57-70.

234. Diaz-Rodriguez, E., Montero, J. C., Esparis-Ogando, A., Yuste, L., and Pandiella, A. (2002). Extracellular signal-regulated kinase phosphorylates tumor necrosis factor alpha-converting enzyme at threonine 735: a potential role in regulated shedding. *Mol Biol Cell* **13,** 2031-44.
235. Takenobu, H., Yamazaki, A., Hirata, M., Umata, T., and Mekada, E. (2003). The stress- and inflammatory cytokine-induced ectodomain shedding of heparin-binding epidermal growth factor-like growth factor is mediated by p38 MAPK, distinct from the 12-O-tetradecanoylphorbol-13-acetate- and lysophosphatidic acid-induced signaling cascades. *J Biol Chem* **278,** 17255-62.
236. Bohler, T., Waiser, J., Hepburn, H., Gaedeke, J., Lehmann, C., Hambach, P., Budde, K., and Neumayer, H. H. (2000). TNF-alpha and IL-1alpha induce apoptosis in subconfluent rat mesangial cells. Evidence for the involvement of hydrogen peroxide and lipid peroxidation as second messengers. *Cytokine* **12,** 986-91.
237. Saccani, S., Pantano, S., and Natoli, G. (2003). Modulation of NF-kappaB activity by exchange of dimers. *Mol Cell* **11,** 1563-74.
238. Ruan, H., and Lodish, H. F. (2003). Insulin resistance in adipose tissue: direct and indirect effects of tumor necrosis factor-alpha. *Cytokine Growth Factor Rev* **14,** 447-55.
239. Streetz, K., Leifeld, L., Grundmann, D., Ramakers, J., Eckert, K., Spengler, U., Brenner, D., Manns, M., and Trautwein, C. (2000). Tumor necrosis factor alpha in the pathogenesis of human and murine fulminant hepatic failure. *Gastroenterology* **119,** 446-60.
240. Palladino, M. A., Bahjat, F. R., Theodorakis, E. A., and Moldawer, L. L. (2003). Anti-TNF-alpha therapies: the next generation. *Nat Rev Drug Discov* **2,** 736-46.
241. Carswell, E. A., Old, L. J., Kassel, R. L., Green, S., Fiore, N., and Williamson, B. (1975). An endotoxin-induced serum factor that causes necrosis of tumors. *Proc Natl Acad Sci U S A* **72,** 3666-70.
242. Sfikakis, P. P., and Kollias, G. (2003). Tumor necrosis factor biology in experimental and clinical arthritis. *Curr Opin Rheumatol* **15,** 380-6.
243. Reinhart, K., and Karzai, W. (2001). Anti-tumor necrosis factor therapy in sepsis: update on clinical trials and lessons learned. *Crit Care Med* **29,** S121-5.
244. Anderson, G. M., Nakada, M. T., and DeWitte, M. (2004). Tumor necrosis factor-alpha in the pathogenesis and treatment of cancer. *Curr Opin Pharmacol* **4,** 314-20.
245. Zwerina, J., Hayer, S., Tohidast-Akrad, M., Bergmeister, H., Redlich, K., Feige, U., Dunstan, C., Kollias, G., Steiner, G., Smolen, J., and Schett, G. (2004). Single and combined inhibition of tumor necrosis factor, interleukin-1, and RANKL pathways in tumor necrosis factor-induced arthritis: effects on synovial inflammation, bone erosion, and cartilage destruction. *Arthritis Rheum* **50,** 277-90.
246. Aggarwal, B. B. (2003). Signalling pathways of the TNF superfamily: a double-edged sword. *Nat Rev Immunol* **3,** 745-56.
247. Comoglio, P. M., Boccaccio, C., and Trusolino, L. (2003). Interactions between growth factor receptors and adhesion molecules: breaking the rules. *Curr Opin Cell Biol* **15,** 565-71.
248. Miller, K. E., Wong, Y. L., Van Parijs, L., and Lauffenburger, D. A. (2004). E1/E3-deleted adenoviral vector sensitization of human epithelial cells to TNFalpha-mediated apoptosis results from saturation of Akt signaling protection effect. *J Virol,* in review.

249. Hickman, M. J., and Samson, L. D. (2004). Apoptotic signaling in response to a single type of DNA lesion, O(6)-methylguanine. *Mol Cell* **14,** 105-16.
250. Mendelsohn, J., and Baselga, J. (2000). The EGF receptor family as targets for cancer therapy. *Oncogene* **19,** 6550-65.
251. Hipfner, D. R., and Cohen, S. M. (2004). Connecting proliferation and apoptosis in development and disease. *Nat Rev Mol Cell Biol* **5,** 805-15.
252. Lee, D. D., and Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature* **401,** 788-91.
253. Liebermeister, W. (2002). Linear modes of gene expression determined by independent component analysis. *Bioinformatics* **18,** 51-60.
254. Liao, J. C., Boscolo, R., Yang, Y. L., Tran, L. M., Sabatti, C., and Roychowdhury, V. P. (2003). Network component analysis: reconstruction of regulatory signals in biological systems. *Proc Natl Acad Sci U S A* **100,** 15522-7.
255. Friedman, N. (2004). Inferring cellular networks using probabilistic graphical models. *Science* **303,** 799-805.
256. Woolf, P. J., Prudhomme, W., Daheron, L., Daley, G. Q., and Lauffenburger, D. A. (2005). Bayesian analysis of signaling networks governing embryonic stem cell fate decisions. *Bioinformatics* **21,** 741-53.
257. White, M. F. (2002). IRS proteins and the common path to diabetes. *Am J Physiol Endocrinol Metab* **283,** E413-22.
258. Tran, S. E., Meinander, A., and Eriksson, J. E. (2004). Instant decisions: transcription-independent control of death-receptor-mediated apoptosis. *Trends Biochem Sci* **29,** 601-8.
259. Pahl, H. L. (1999). Activators and target genes of Rel/NF-kappaB transcription factors. *Oncogene* **18,** 6853-66.
260. Ideker, T., Galitski, T., and Hood, L. (2001). A new approach to decoding life: systems biology. *Annu Rev Genomics Hum Genet* **2,** 343-72.
261. Ideker, T. (2004). Systems biology 101--what you need to know. *Nat Biotechnol* **22,** 473-5.
262. Beer, D. G., Kardia, S. L., Huang, C. C., Giordano, T. J., Levin, A. M., Misek, D. E., Lin, L., Chen, G., Gharib, T. G., Thomas, D. G., Lizyness, M. L., Kuick, R., Hayasaka, S., Taylor, J. M., Iannettoni, M. D., Orringer, M. B., and Hanash, S. (2002). Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med* **8,** 816-24.
263. Obata, T., Yaffe, M. B., Leparc, G. G., Piro, E. T., Maegawa, H., Kashiwagi, A., Kikkawa, R., and Cantley, L. C. (2000). Peptide and protein library screening defines optimal substrate motifs for AKT/PKB. *J Biol Chem* **275,** 36108-15.
264. Gupta, S., Campbell, D., Derijard, B., and Davis, R. J. (1995). Transcription factor ATF2 regulation by the JNK signal transduction pathway. *Science* **267,** 389-93.
265. Manke, I. A., Nguyen, A., Lim, D., Stewart, M. Q., Elia, A. E., and Yaffe, M. B. (2005). MAPKAP kinase-2 is a cell cycle checkpoint kinase that regulates the G2/M transition and S phase progression in response to UV irradiation. *Mol Cell* **17,** 37-48.
266. Geleziunas, R., Ferrell, S., Lin, X., Mu, Y., Cunningham, E. T., Jr., Grant, M., Connelly, M. A., Hambor, J. E., Marcu, K. B., and Greene, W. C. (1998). Human T-cell leukemia virus type 1 Tax induction of NF-kappaB involves activation of the IkappaB kinase alpha (IKKalpha) and IKKbeta cellular kinases. *Mol Cell Biol* **18,** 5157-65.
267. Efron, B., and Tibshirani, R. J. (1993). "An Introduction to the Bootstrap," Chapman and Hall, London.

6.2. Experimental protocols

6.2.1. High-throughput radioactivity-based kinase activity assays

The microtiter-based kinase activity assays were performed with the following antibodies: anti-ERK1/2 CT (Upstate), anti-Akt PH domain (Upstate), anti-JNK1 (Santa Cruz), anti-IKKα/β (Santa Cruz) or anti-MAPKAP kinase 2 (Upstate). Protein A or G microtiter strips (Pierce) were coated overnight with 10 μg/ml anti-kinase antibody and washed three times with blocking buffer (1% bovine serum albumin (Sigma) in 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.05% Triton X-100). Cell lysates (50 μg for ERK, 500 μg for Akt, 200 μg for JNK1, 200 μg for MK2, 600 μg for IKK) were added for 3 h (ERK, Akt, JNK1 and MK2) or overnight (IKK), then washed two times with wash buffer (50 mM Tris-HCl (pH 7.5), 150 mM NaCl) and two times with kinase wash buffer (20 mM Tris-HCl (pH 7.5), 15 mM $MgCl_2$, 5 mM β-glycerophosphate, 1 mM EGTA, 0.2 mM Na_3VO_4, 0.2 mM DTT). The wells were resuspended in 20 μl kinase wash buffer and warmed to 37°C. 20 μl kinase assay buffer (kinase wash buffer plus 0.4 μM PKA inhibitor, 4 μM PKC inhibitor, 4 μM calmidazolium, 0-25 μM cold ATP, 1-5 μCi [γ-^{32}P]ATP) was added to the wells, followed by 20 μl of substrate (40 μg myelin basic protein for ERK, 10 μM Aktide [263] for Akt, 3 μg GST-ATF2(1-109) [264] for JNK1, 10 μM MK2tide [265] for MK2, 10 μg GST-IκBα(1-62) [266] for IKK) to initiate the reaction. The kinase reactions were allowed to proceed for 15-120 min at 37°C, then terminated by 60 μl 75 mM H_3PO_4 or 20 mM EDTA. Exact conditions for each kinase assay are detailed in Table I. For EDTA-terminated reactions, 40 μl of the terminated reaction was transferred to a phosphocellulose filter plate (Millipore) containing 100 μl 75 mM H_3PO_4 and mixed, whereas H_3PO_4-terminated reactions

Table 6-1. Experimental conditions for the individual in vitro kinase assays[a]

Kinase	Antibody	Substrate	ATP (μM)	[γ-^{32}P]ATP (μCi)	Reaction time (min)	Termination
ERK	Anti-ERK1/2 CT	Myelin basic protein	25	1	60	H_3PO_4
Akt	Anti-Akt PH dom.	Aktide	10	5	30	H_3PO_4
JNK1	Anti-JNK1	GST-ATF2	10	2	60	H_3PO_4
IKK	Anti-IKKα/β	GST-IκBα	0	5	120	EDTA
MK2	Anti-MAPKAP K2	MK2tide	25	2	15	EDTA

[a] Other experimental parameters were maintained as described in Experimental Procedures.

were added directly to the filter plates. The terminated reaction contents were vacuum-filtered and washed five times with 75 mM H_3PO_4 and three times with 70% EtOH. The filters were punched into vials and the radioactivity incorporated was measured by liquid scintillation (Cytoscint, ICN). The results from blank wells, containing only lysis buffer during the immunopurification step, were subtracted to remove nonspecific contributions, with the exception of Akt, where this was not necessary.

6.2.2. High-throughput fluorescence-based kinase activity assays

Activity assays were performed in a glass 96-well plate (Zinsser #3600500 with 250 μl or Zinsser #3600501 with 150 μl per well) at 30°C in a fluorescence plate reader in Buffer D with 4 μM PKC inhibitor, 4 μM calmidazolium (Sigma), 10 μM substrate peptide, 1 mM ATP (Calbiochem, Low Metals Grade) and for Akt, 0.4 μM PKItide, 5 μM GF 109203X (Calbiochem); for MK2, of 0.4 μM PKItide, 25 μM GF 109203X; and for PKA, 5 μM GF 109203X. This assay buffer was prepared in bulk at 30°C and aliquoted to each well to ensure equal concentrations of the chemosensor. To begin each reaction, 7.5 vol% lysis buffer (Buffer C) or lysate (18.5 μl in 250 μl for Akt, and 11 μl in 150 μl for MK2 and PKA) was added and the contents of each well mixed gently. 60 data points were collected over each reaction. Akt-S1 activity was monitored over 60 min typically with 93 μg lysate. MK2-S1 and PKA-S3 activities were monitored for 15 min typically with 50 μg and 20 μg lysate, respectively. To quantify product formation, different turnover amounts (5%, 10%, 15%) were simulated in triplicate and lysis buffer was added as for a blank sample. For recombinant enzyme comparison, different amounts (≥ 3) of recombinant enzyme were added with lysis buffer to begin the reaction. For PKA inhibition studies, the indicated concentration of H89 or PKItide was included in the assay buffer. Chemosensor sensitivities were determined by comparing the absolute fluorescence of 10 μM substrate and 10 μM corresponding product phosphopeptides in triplicate under the optimized assay conditions.

6.2.3. Signaling network measurements

Briefly, 19 quantitative signaling network measurements—IKK activity, JNK1 activity, MK2 activity, P-EGFR (Y1068), total EGFR, P/total EGFR, P-MEK (S217/221), ERK activity, P-IRS1 (S636), P-IRS1 (Y896), P-Akt (S473) by Western blot, total Akt, Akt activity, P-Akt (S473) by antibody microarray, total Akt, P/total Akt, P-FKHR (S256), procaspase-8, cleaved caspase-8, and procaspase-3—were compiled from triplicate biological samples treated with 0, 5, 100 ng/ml TNF and 0, 1, 100 ng/ml EGF or 0, 5, 500 ng/ml insulin for 0, 5, 15, 30, 60, 90 min or 2, 4, 8, 12, 16, 20, 24 hr. Exact treatment combinations are shown in Figures 2-17 and 2-18. For autocrine perturbations, cells were pretreated with 10 μg/ml C225 for one hr or cotreated with 10 μg/ml IL-1ra, then stimulated with 5 or 100 ng/ml TNF [2].

6.2.4. Apoptosis measurements

HT-29 cells were plated in 24-well plates at a density of 50,000 cells/cm^2, and were sensitized and treated identically to the cells used in the corresponding signaling measurements on the same day [2]. After 12, 24, or 48 hours of cytokine treatment, the cells were trypsinized until all cells were detached from the plate. The growth medium was combined with the trypsinized cells to ensure capture of both floating and adherent cells in each well. The combined pool of cells from each well was split into thirds; one third was analyzed for phosphatidylserine exposure and membrane permeability, one third was fixed with methanol and analyzed for cleaved caspase-cytokeratin, and one third was fixed with methanol (MeOH) and analyzed for nuclear fragmentation. For the phosphatidylserine-membrane permeability analysis, the cells were washed once with Annexin Binding Buffer (ABB, 10 mM HEPES, 140 mM NaCl, 2.5 mM $CaCl_2$) and stained with Alexa Fluor 488-conjugated Annexin-V (Molecular Probes) and 1 ug/ml propidium iodide (PI) for 10 minutes at room temperature. Excess ABB was added and the cells were analyzed by a Becton Dickinson FACScan or FACSCalibur flow cytometer. For the cleaved caspase-cytokeratin assay, MeOH-fixed cells were stored at -20 C for up to 1 week. After centrifugation to remove the MeOH, cells were washed in PBS + 0.1% Tween-20 (PBS-T), and were then stained with anti-cleaved caspase-3 (Cell Signaling) and anti-cleaved cytokeratin (Roche) antibodies in PBS-T + 1% bovine serum albumin (PBS-TB) for 1 hour at room temperature. The cells were washed and stained with secondary antibodies, Alexa

Fluor 488-conjugated donkey anti-mouse IgG and Alexa Fluor 647-conjugated donkey anti-rabbit IgG (Molecular Probes) in PBS-TB for 1 hour at room temperature. The cells were then washed and analyzed on a Becton Dickinson FACSCalibur flow cytometer. For the nuclear fragmentation assay, MeOH cells were stored at -20 C for up to one week. After centrifugation to remove the MeOH, cells were washed once with PBS-T, and then incubated with 1 mg/ml RNAse A in PBS-T for 2.5 hours. PI was then added at a final concentration of 50 ug/ml, and the cells were analyzed on a FACSCalibur flow cytometer to quantify nuclear fragmentation. For all flow cytometry assays, data was analyzed using FlowJo software (TreeStar).

To fuse quantitative measurements on different days, 100 ng/ml TNF and mock treatments were always included as fixed positive and negative controls. Experimental samples were then scaled by the TNF (membrane permeability and cleaved caspase-cytokeratin) or mock (phosphatidylserine exposure and nuclear fragmentation) index for that day. By cross-validation we verified that the magnitude and dynamics of these normalized apoptotic indices were highly reproducible, with $R^2 = 0.79–0.98$ [9]. The entire dataset of 1440 flow cytometry runs is available in the Supporting Online Material.

6.2.5. Metric extraction

For each signaling time course, the following signaling metrics were derived:

a. Thirteen time-point metrics, defined as the mean signal at each point in the time course.
b. Thirteen instantaneous-derivative metrics, defined as the forward slope between the current time point and the subsequent time point. For the final time point (24 hr), the instantaneous derivative was set to zero.
c. Four summary metrics, which included the area under the curve for the entire time course, the maximum signal, the mean signal, and the steady-state signal (defined as the mean of the final four time points).
d. A variable number of peak metrics. For each signal a tolerance value (defined as a percentage of the dynamic range of the signal) was set for the algorithm to classify an upward fluctuation as a peak. Tolerances were set at: 50% for Akt activity, IKK activity, and all antibody microarray measurements; 5% for JNK1 activity and MK2 activity; and 20% for all other signaling network measurements. For each peak identified, three

metrics were extracted: area under the curve, activation slope (defined as the best linear fit between the preceding local minimum and the peak maximum), and decay rate (defined as the best exponential fit between the peak maximum and the proceeding local minimum).

These signaling metrics are summarized in table S2. Metric extraction algorithms were coded in MATLAB and are available upon request.

6.2.6. PLS model construction and validation

Prior to all analyses, the signaling and apoptosis matrices were variance scaled to nondimensionalize the different measurements. The PLS model was constructed in the SIMCA-P 10.0 (Umetrics) software suite according to the following iterative formulas:

$$E_1 = X - t_1 p_1^T;\ E_2 = E_1 - t_2 p_2^T,\ t_2 = E_1 w_1;\ E_i = E_{i-1} - t_i p_i^T,\ t_i = E_{i-1} w_i$$
$$F_1 = Y - b_1 t_1 q_1^T;\ F_2 = F_1 - b_2 t_2 q_2^T;\ F_i = F_{i-1} - b_i t_i q_i^T$$

where E_i represents the residual of the i^{th} principal component, with score vector t_i, weight vector w_i, and loading vector p_i, and ^{T} represents transpose. F_i represents the residuals of the i^{th} dependent principal component, with score vector t_i and loading vector q_i, and b_i represents the coefficient characterizing the inner relation between the independent and dependent principal components. Model predictions were made by leave-one-out crossvalidation for the TNF-EGF-insulin treatments and by unbiased prediction for the autocrine perturbations [3]. Model uncertainties were calculated by jack-knifing [267]. Signaling axes and treatment mappings (Figure 4-14E–G) were plotted using w_1, w_2 and t_1, t_2 respectively after a 60° subspace rotation [203]. Centered and scaled coefficients were used as the regression weights.