Microfluidics-based Single-cell Functional Proteomics Microchip for Portraying Protein Signal Transduction Networks within the Framework of Physicochemical Principles, with Applications in Fundamental and Translational Cancer Research

Thesis by

Wei Wei

In Partial Fulfillment of the Requirements for the degree

of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2014

(Defended February 12, 2014)

To my dearest wife,

Bing Shu

ACKNOWLEDGEMENTS

It would not have been possible to finish this doctoral thesis without the constant help and support from the talented and kind people around me during my study at Caltech, to only some of whom it is possible to give a particular mention here.

Foremost, I would like to express my deepest appreciation to my advisor, Professor Jim Heath. His boundless enthusiasm for science and persistent pursuit of truth exceptionally inspire and enrich me all the time, not to mention his advice, insight and supreme knowledge in the field. He offers everyone in the group sufficient academic freedom, supports and resources, allowing me to pursue my own scientific ideas and shaping me from a knowledge absorber to an independent scientist. Jim's unlimited support, trust and friendship have been invaluable to me on both an academic and a personal level, for which I am extremely grateful. I could not have imagined having a better advisor and mentor for my PhD study.

I have been so fortunate to have had the great opportunity to work with many intelligent collaborators through Jim's organization among Caltech, UCLA, and the Institute for Systems Biology (ISB). I would like to express my sincere gratitude to many senior collaborators, including Professor Paul Mischel, Professor Raphael Levine, Professor Tim Cloughesy, Professor Francoise Remacle, Professor Antoni Ribas, and Professor Leroy Hood for their unsurpassed knowledge and great support in the collaborative projects.

I would like to thank the members of my thesis committee, Professor William Goddard, Professor Julia Greer, Professor William Johnson and Professor Mark Davis, for their time, encouragement and insightful comments.

I would like to thank my friends and colleagues whom I have had the pleasure of working with over the years. Specifically, I am most grateful to Dr. Qihui Shi and Dr. Lidong Qin, who trained me when I joined into Jim's lab as a fresh PhD. Qihui and I worked closely to develop the single-cell functional proteomic microchip and successfully applied it to address fundamental cancer biology problems. We also become very close friends. I thank my great collaborators and good friends, Dr. Young Shik Shin and Dr. Beatrice Gini. We worked closely to apply our microchip technology to the preclinical cancer research. I thank my fellow labmates, Dr. Heather Agnew, Dr. Rosemary Rohde, Dr. Ann Cheung, Dr. Min Xue, Dr. Kiwook Hwang, Dr. Jing Zhou, Dr. Jing Yu, Dr. Chao Ma, Dr. Peigen Cao, Dr. Nataly Balasha, Dr. Jun Wang, Alex Sutherland, and all other current and past members of the Heath group, for all the fun we have had in the last four years. I thank my friends Michael Amori, Ke Sun, and Chenguang Ji for sharing personal lives as well as scientific ideas. I am very grateful to Kevin Kan and Elyse Garlock for their effort in lab management and administration.

My sincere thanks also go to my past and current option representatives, Professor Brent Fultz and Professor Sossina Haile for their mentoring and useful advice in academic affairs. I thank Pam Albertson, Christy Jenstad, Natalie Gilmore, and Laura Flower Kim for their kind help in administrative affairs.

Last, but by no means least, I owe much to my beautiful wife, Bing Shu, for her unlimited love, care and support that enrich my life with joy and happiness throughout. I

thank my parents and parents-in-law for letting me pursue my dream for so long and so

far away from home, and supporting me spiritually all the time.

ABSTRACT

Single-cell functional proteomics assays can connect genomic information to biological function through quantitative and multiplex protein measurements. Tools for single-cell proteomics have developed rapidly over the past 5 years and are providing unique opportunities. This thesis describes an emerging microfluidics-based toolkit for single cell functional proteomics, focusing on the development of the single cell barcode chips (SCBCs) with applications in fundamental and translational cancer research.

The microchip designed to simultaneously quantify a panel of secreted, cytoplasmic and membrane proteins from single cells will be discussed at the beginning, which is the prototype for subsequent proteomic microchips with more sophisticated design in preclinical cancer research or clinical applications. The SCBCs are a highly versatile and information rich tool for single-cell functional proteomics. They are based upon isolating individual cells, or defined number of cells, within microchambers, each of which is equipped with a large antibody microarray (the barcode), with between a few hundred to ten thousand microchambers included within a single microchip. Functional proteomics assays at single-cell resolution yield unique pieces of information that significantly shape the way of thinking on cancer research. An in-depth discussion about analysis and interpretation of the unique information such as functional protein fluctuations and protein-protein correlative interactions will follow.

The SCBC is a powerful tool to resolve the functional heterogeneity of cancer cells. It has the capacity to extract a comprehensive picture of the signal transduction network from single tumor cells and thus provides insight into the effect of targeted therapies on protein

signaling networks. We will demonstrate this point through applying the SCBCs to investigate three isogenic cell lines of glioblastoma multiforme (GBM).

The cancer cell population is highly heterogeneous with high-amplitude fluctuation at the single cell level, which in turn grants the robustness of the entire population. The concept that a stable population existing in the presence of random fluctuations is reminiscent of many physical systems that are successfully understood using statistical physics. Thus, tools derived from that field can probably be applied to using fluctuations to determine the nature of signaling networks. In the second part of the thesis, we will focus on such a case to use thermodynamics-motivated principles to understand cancer cell hypoxia, where single cell proteomics assays coupled with a quantitative version of Le Chatelier's principle derived from statistical mechanics yield detailed and surprising predictions, which were found to be correct in both cell line and primary tumor model.

The third part of the thesis demonstrates the application of this technology in the preclinical cancer research to study the GBM cancer cell resistance to molecular targeted therapy. Physical approaches to anticipate therapy resistance and to identify effective therapy combinations will be discussed in detail. Our approach is based upon elucidating the signaling coordination within the phosphoprotein signaling pathways that are hyperactivated in human GBMs, and interrogating how that coordination responds to the perturbation of targeted inhibitor. Strongly coupled protein-protein interactions constitute most signaling cascades. A physical analogy of such a system is the strongly coupled atom-atom interactions in a crystal lattice. Similar to decomposing the atomic interactions into a series of independent normal vibrational modes, a simplified picture of signaling network coordination can also be achieved by diagonalizing protein-protein correlation or

covariance matrices to decompose the pairwise correlative interactions into a set of distinct linear combinations of signaling proteins (i.e. independent signaling modes). By doing so, two independent signaling modes – one associated with mTOR signaling and a second associated with ERK/Src signaling have been resolved, which in turn allow us to anticipate resistance, and to design combination therapies that are effective, as well as identify those therapies and therapy combinations that will be ineffective. We validated our predictions in mouse tumor models and all predictions were borne out.

In the last part, some preliminary results about the clinical translation of single-cell proteomic chips will be presented. The successful demonstration of our work on human-derived xenografts provides the rationale to extend our current work into the clinic. It will enable us to interrogate GBM tumor samples in a way that could potentially yield a straightforward, rapid interpretation so that we can give therapeutic guidance to the attending physicians within a clinical relevant time scale. The technical challenges of the clinical translation will be presented and our solutions to address the challenges will be discussed as well. A clinical case study will then follow, where some preliminary data collected from a pediatric GBM patient bearing an EGFR amplified tumor will be presented to demonstrate the general protocol and the workflow of the proposed clinical studies.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Chapter 3

Chapter 4

C h a p t e r   1

Introduction

1.1 NANOTECHNOLOGY INNOVATIVE TOOLS FOR CANCER RESEARCH

Cancer is the leading cause of death worldwide. The early detection and treatment of cancer has been a long-term bottleneck in the clinic. Over the past decade, however, nanotechnology offers a wealth of innovative tools that lead to advances in early detection[1], molecular imaging[2, 3], assessment of therapeutic efficacy[4, 5], targeted and multifunctional therapeutics[6, 7], and prevention and control of this complex disease[8, 9].

Nanotechnology is being applied to cancer in two broad areas: the development of nanovectors, such as nanoparticles, which can be loaded with drugs or imaging agents and then targeted to tumors, and high-throughput biomolecular profiling devices that can detect a large number of different molecular species at the same time[10].

Thousands of nanovectors are currently under study, which include but not limited to liposomes[11, 12], nanoparticles[13-15], polymeric micelles[16, 17], dendrimers[18] and quantum dots[2, 3]. Fundamental advantage of nanovectors is the multifunctionality that combines the avoidance of biobarriers, tumor targeting and controlled release. On the other hand, various micro- or nanofabrication technologies, such as photolithography, soft lithography, etc. enable high precision patterning of biological molecules on substrates, which in turn become the technological foundation of high-throughput and multiplexed platform for biomolecule detection[19-27]. Microarrays[19, 20, 23], as a prime example, are used for molecular diagnostics, genotyping and biomarker-guided therapeutic targeting. Microfluidic-based

microchips further integrate highly versatile and extremely miniaturized microfluidic platform with biomolecule microarray and therefore enable quantitative measurements of large numbers of biological signatures from minuscule amount of blood[1] or even single cells taken from disease tissues[4, 24-27].

Combined, such technologies opens up new avenues for studying, diagnosing and treating cancer at a systems-level to facilitate predictive, preventive and personalized medicine in which early and accurate tumor detection and subsequent identification of effective therapy strategies lead to rapid initiation of smart treatment tailored to each patient's tumor molecular profile[9, 28].

1.2 CELLULAR HETEROGENEITY AND SINGLE-CELL TECHNOLOGY

Cancer is a highly heterogeneous disease[29]. Phenotypic and functional heterogeneity arise among cancer cells within the same tumor as a consequence of genetic change, environmental differences, and reversible changes in cell properties[30]. Much of our knowledge of biology is based upon ensemble measurements under the assumption that ensemble averages reflect the dominant biological mechanism operating within individual cells in a population[31]. After decades of research on single cells, however, it has been widely recognized that cell-to-cell variations are always present in any population of "seemingly identical" cells[32], and the traditional ensemble average measurements may not correctly present the behaviors of any individual cell when the measured properties are of binary nature[33, 34], cell-cycle dependent[35, 36] or dominated by a small subset of cells[37-40]. Although this heterogeneity is often ascribed to some process (such as stochastic gene expression), it is also intrinsic to the finite nature of a single cell[41]. Some cell-to-cell variations are due to biochemical noise and may not have functional significance[42, 43]. This heterogeneity, however, is not always without consequences; for example, it can contribute to the diversity of an immune response[44] or to the emergence of therapeutic resistance in cancers[39, 45].

Determining whether observed heterogeneity has functional significance requires a framework for quantifying heterogeneity and assessing its information content[31]. Measuring parameters of interest at single cell resolution is therefore desired, especially for cancer cells which are genetically unstable and usually comprised of multiple subpopulations of cells with different functional activities[40, 46].

1.3 ONCOGENIC SIGNALING NETWORK IN CANCER CELLS.

Cellular activities and functions are governed by a complex system of communication mechanism called cell signaling which ensures cells to perceive and correctly respond to their microenvironment and serves as the basis of development, tissue repair and immunity as well as normal tissue homeostasis. Alterations and errors in cell signaling processes are responsible for disease such as cancer, autoimmunity and diabetes[47]. Traditional biology has focused on studying individual parts of cell signaling pathways. The advancements in systems biology and "omics" (proteomics, genomics, transcripomics, metabolomics, etc.) technologies have yielded large inventories of genes, transcripts, proteins and metabolites and have shaped our views on signal transduction[48]. The signaling pathways are not simple linear paths that pipe the signal from cell surface to nucleus, but are organized as networks. Portraying the underlying structure of the signaling networks, understanding how changes in these networks may affect the transmission and flow of information and exploring how these networks will respond to external perturbations are crucial to transform how disease is understood, attached and possibly prevented.

For many cancers, genomic surveys are revealing a landscape of altered signal transduction cascades that often cluster along a set of druggable core pathways. In fact, these pathways contain molecular targets for newer generations of cancer therapies[49]. However, the translation of genomic data into effective clinical treatments has been confounded because non-genetic cell-to-cell variability is profound in drug responses and resistance development. A recent editorial[50] has pointed out that capturing the functional protein signaling network may prove valuable for this purpose, because those signaling

proteins, "not the genes *per se,* are responsible for the phenotypes of tumors and for the
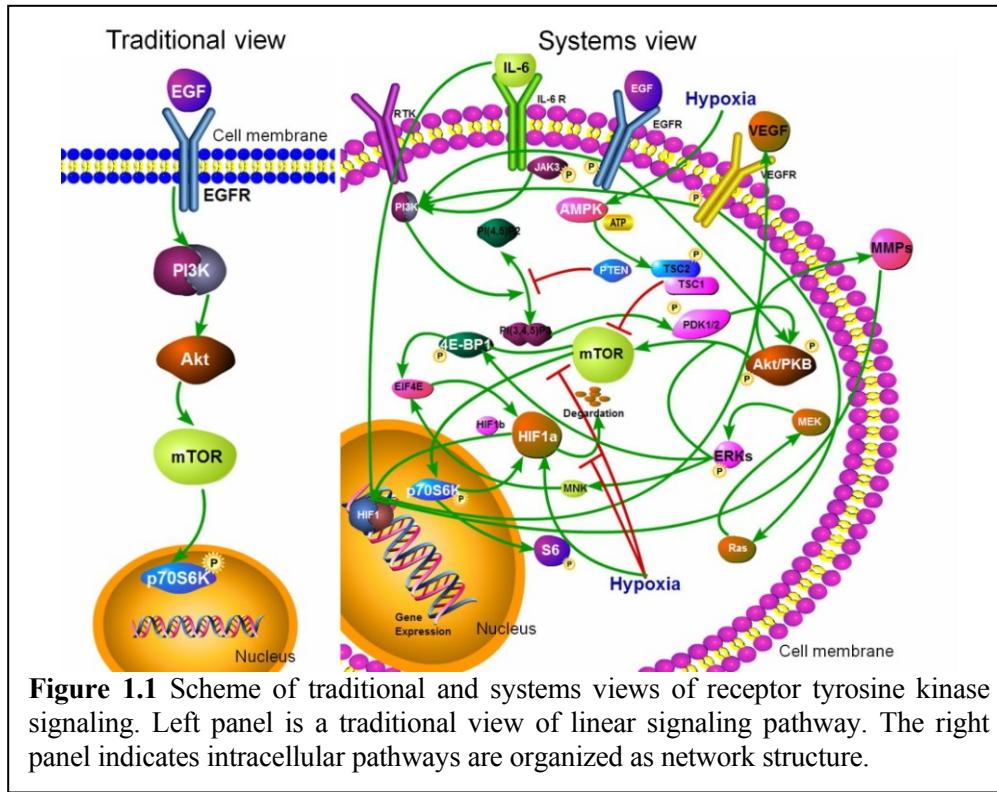
Figure 1.1 Scheme of traditional and systems views of receptor tyrosine kinase signaling. Left panel is a traditional view of linear signaling pathway. The right panel indicates intracellular pathways are organized as network structure.

emergence of therapeutic resistance." Single cell proteomics therefore provides the most

direct approach for elucidating protein signaling network structure and coordination, and

builds the natural bridge connecting signaling events to biological functions[27].

1.4 MICROFLUIDIC-BASED MICROCHIP PLATFORMS FOR SINGLE-CELL PROTEOMICS

Single cell proteomics has evolved over more than 50 years, dating back to the invention of the Coulter counter[51], which is the precursor of the first cytometer[52]. The increasing availability of lasers, photon detectors, high speed electronics, bioconjugation chemistries, and dye molecules fed into the development of fluorescence flow cytometry (FFC)[53], fluorescence activated cell sorting (FACS)[54] and Enzyme-linked immunosorbent spot (ELISpot)[55]. Most recently, mass spectrometry have been harnessed for the development of mass cytometry[56]. Excepting ELISpot, the dominant applications of these tools have been sorting or enumerating cellular phenotypes based upon labeling and measurements of surface marker (membrane) proteins. Intracellular staining (ICS) techniques[57] have opened FFC and mass cytometry up to the analysis of at least a few cytoplasmic functional proteins per cells. Each of those techniques has strongly influenced the development and/or specific application of the more recent microchip tools.

Over the past 25 years, microfluidic techniques[58] have emerged for the manipulation, sorting and analysis of small biological samples, ranging from microliters of blood[1] to single cells[59, 60]. On-chip assays range from cell counting to molecular measurements, with applications that span broadly across the field of biology and biomedicine[27, 61, 62]. Common advantages of microfluidic tools are that they can often be cheaply manufactured in large quantities, they can handle very small amount numbers of cells and require only tiny quantities of expensive reagents, and they are highly versatile for meeting different purposes, and so on. Many of these features are briefly characterized and compared against other non-microchip single cell proteomics tools in Table 1.1[27].

One of the newer technologies to emerge, and to be adapted to microchip platforms, is that of single cell functional proteomics. Functional proteins include the secreted cytokines, chemokines, proteases, and granulocytes that are commonly associated with immune cell function, but also include catalytically active (phosphorylated) kinases and associated effector proteins that participate in intracellular signaling cascades. Examples of such cascades involve the hyper-activated phosphoprotein signaling pathways that are commonly associated with tumorigenic activity in cancer cells, and are consequently targeted by anticancer therapies. In their active states, these signaling pathways have functional consequences that can be associated with the various hallmarks of cancer.

For single cell functional proteomics, microfluidic-based platforms fall into two groups: those in which the cells are stained to identify specific proteins, and those for which proteins are released from the cells and measured using surface immunoassays.

Table 1.1 Single-cell functional proteomics tools[27]

Technique	Number and types of protein assayed	Throughput	Detection limit	Statistical accuracy and signal quantification	Notes and features
Flow cytometry methods					
Fluorescence flow cytometry[53]	Around 15 proteins (mostly membrane proteins, a few cytoplasmic proteins)	10^4 cells/s	500 copies per cell	90% phenotyping accuracy; relative protein abundance	Standard for sorting and enumeration of cellular phenotypes. Secretion blocked and cell fixed for cytoplasmic proteins
Mass flow cytometry[56]	Around 35 membrane and intracellular proteins, likely expandable	10^3 cells/s	$>10^3$ copies per cell	Good cell counting statistics; relative protein abundance	Cells handled in bulk prior to analysis. Secretion blocked and cells fixed for cytoplasmic proteins
Surface methods					
ELISpot[55]	1-3 secreted proteins	Not available	6 spots per 10^5 cells	Quantitative for percentage active cells	Cells secrete proteins onto antibody coated surface; secretion activity correlated with cell location
Microfluidics technologies					
Image cytometry[63, 64]	3-4 membrane or intracellular proteins and cell size	10^3-10^4 cells per chip	10^5 fluorophores per μm2	Good cell counting statistics; relative protein abundance	Cell are fixed and stained (in bulk) with fluorescent antibodies; protein assay and cell

					location spatially correlated
Cell array[65-68]	1 intracellular proteins	$<10^3$ cells per chip	Not available	Good cell counting; relative protein abundance	Single cells separated and imaged on chip; continuous monitoring of cell physiology
Micro-droplet[69-71]	1 membrane or intracellular protein	10^2 μdrops/s	Not defined	Good cell sampling statistics	Cells entrained in microdroplets; microdroplet composition control permits screening cells
Micro-engraving[72-75]	3 secreted plus 3 membrane proteins	10^4-10^5 cells per chip	Not available	Very good cell number statistics; relative protein abundance	Cells isolated in microwells; surface immunoassays; proteins colorimetrically detected; secretome kinetics from single cells; proteomic and functional assays from same cell.
Single cell barcode chips[1, 4, 24, 26, 27, 44, 76-78]	About 20 secreted, membrane or cytoplasmic proteins, expandable	10^3-10^5 cells per chip	10^2 copies	Good cell counting statistics; absolute quantification; 10% measurement error per protein per cell	Cells insolated in microchambers; miniature antibody arrays yield spatial separation of specific protein assays; proteomic and functional assays from the same cell; single cells or defined small cell populations accessed

The first group includes a image cytometry, cell-array, and micro-droplet techniques. Early variations of such tools detected proteins from single cells by imaging stained cells, or by following the labeled cells or cell-encapsulation droplets through a microfluidic channel designed to allow fluorescence detection. These were basically microchip versions of FFC or FACS[79]. More recent approaches have significantly diverged to take advantage of some of the unique aspects of microfluidics. For example, cells can be spatially segregated into large arrays[65-68], or they can be entrained within arrays of drops[69-71]. Such manipulations are followed by immunostaining of membrane proteins, followed by automated imaging to quantify single cell fluorescence signals. These approaches can offer control over the cell environment before analysis, which make them attractive screening

tools[63, 64]. One disadvantage of these and other cell-staining approaches is that they have limited multiplexing capacity, which might be overcome through integrating super resolution imaging[80] with a conceptual extension of optical barcodes[81].

The most advanced microfluidic single cell proteomics tools use surface-immobilized antibodies for separating protein detection from cell manipulation. This approach is conceptually similar to ELISpot, but has capabilities that can in many ways surpass those cytometry tools. Separating protein assays from the cells implies that individual proteins can be spatially, rather than colorimetrically, identified, and that sandwich ELISA-type assays can be used. Of course, cell staining of proteins can be still simultaneously carried out. The result is a significantly higher level of multiplexing and, for some proteins, absolute quantification. Moreover, intracellular, membrane and secreted proteins may be assayed from the same cell. The chambers in which the cells are isolated can potentially accommodate multiple cells and/or cell types, thus permitting measurements of cellular interactions. Finally, these platforms allow the integration of functional assays (e.g. cell mobility) with protein assays.

One of base technologies in this class of platforms is the microengraving approach developed by Love's group[72-74] uses small volume microwells in an array format to isolate and culture single cells. A "microengraved" (antibody-coated) substrate is used to cap the microwell array and to capture secreted proteins. Proteins are detected using sandwich-type ELISA immunoassays. Different fluorophores colorimetrically distinguish between different detection antibodies to allow the simultaneous detection of about three secreted proteins. The microengraved substrate can be replaced multiple times *in situ*, thus enabling kinetic studies at the single cell level. The multiplexing capacity of the microengraving

method can be increased using fluorophore-labeled antibody staining of membrane proteins; fluorescence imaging of the captured cells yields information on membrane protein levels (to identify cellular phenotypes), and the microengraved substrate assays for secreted proteins (to assess cellular function).

A related approach is the single cell barcode chips (SCBCs) which are a highly versatile and information rich tool for single cell functional proteomic analysis. SCBCs are based upon isolating individual cells, or defined numbers of cells, within microchambers, each of which is equipped with a large antibody microarray (the barcode), with between a few hundred to 10^4 individual microchambers included within a single microchip. Depending upon the application, SCBC microchamber volume are designed to be between 0.1–2 nanoliters[25, 26, 44], and microchamber design and operation protocols can permit sandwich-type ELISA immunoassay of cytoplasmic, secreted, or membrane proteins with a measurement error of ~10% for a given protein level[24, 82]. The development and applications of SCBCs is the central topic of this thesis and will be discussed in detail in the subsequent chapters.

1.5 PHYSICAL APPROACHES TOWARD UNDERSTANDING CANCER

The advent of the "omics" age has not only triggered a revolution in technology, but also the way of thinking[48]. The ability to routinely study thousands of genes and proteins enables physical scientists to consider complex biological events as systems of interacting units which can be understood through statistical mechanics, thermodynamics and chemical kinetics. In this perspective, human body can be viewed as a highly coordinated system of interacting molecular networks, among which protein signaling networks provide the most direct access to understand diseases such as cancer since proteins are actual performers of a vast array of functions within living organisms.

Recently developed single cell functional proteomic microchips further grant people to quantify functional proteins, which are often transient and low-abundance targets, in a high-throughput and multiplex fashion at single cell resolution. Functional protein are typically generated, released or activated following stimulation, and their production is closely relevant to how cells process information and respond to perturbation and is often the end results of a series of stochastic events. In other words, they are the opposite of housekeeping proteins such as actin that are always present in abundant and reasonably stable concentrations. As a result, the abundances, kinetics and their statistical distributions of functional proteins contains ample information of molecular interactions within cells and can reflect changes in cellular activity, such as immune-cell activation or the activation or inhibition of protein signaling. Many physical approaches, such as fluctuation dissipation theorem, linear noise approximation, potential landscape, etc. have been developed to

understand functional protein fluctuations. A mean-field model[26] will also be introduced in Chapter 3 to better demonstrate the biological significance of functional protein fluctuation.

Single cell functional proteomics allow simultaneously probing a panel of key functional proteins relevant to the problem of study. By analogy with the concept of phase space in statistical mechanics, the collections of such extracted features (protein levels) allow an individual cell to be represented as a point in (often high-dimensional) feature space, with each axis representing a different measurement. Therefore, populations of cells are transformed into distributions of points in feature space[31], which allows to identify the steady-state of the cell population (or more rigorously, the steady-state of the signaling coordination) through establishing a single cell ensemble and seeking the probability distribution function that is of maximum entropy[82]. The functional protein interactions can be quantified by protein covariance matrices, which may further couple with linear perturbation theory to predict how the levels of protein would respond to a weak perturbation[26]. These will be discussed in detail in Chapter 3.

On the other hand, studying cancer cells from different patients or from the same patient but, under different states during the course of therapy (such as treatment naive, drug responsive, drug resistant) is thus reduced to the problem of identifying patterns of distinct cellular behaviors in feature space. A couple of analytical and computational approaches of decomposing heterogeneous distributions and identifying dominative patterns can be used to extract a set of collective behaviors of functional proteins in the cancer cells, which is similar to decompose the strongly coupled atomic interactions in a

crystal into a series of normal vibrational modes. This has important implications on how to design effective targeted therapy and will be included in Chapter 4.

Cancer cells are highly heterogeneous. Population heterogeneity can arise from factors such as the stochastic nature of intracellular events controlled by low-copy-number transcription factors[83] or through cell-cell interactions[37, 84]. From a traditional biology perspective, this heterogeneity causes cancer to be viewed as a complex (or "hard to understand") disease. Biologists seek to classify the population into different phenotypes and study them separately. However, both phenotypical and functional heterogeneity may be highly dynamic. Even if one starts with a homogeneous phenotype of a cancer cell population, cell-cell interactions can rapidly render the population heterogeneous[85, 86] just like collisions render the velocity distribution of gas molecules Maxwellian.

A heterogeneous tumor, however, as viewed by a physical scientist, might appear as a stable 'organ', with a stability that emerges exactly because of the heterogeneity of the cellular components. Thus, fluctuation measurements can capture cellular heterogeneity, while simultaneously providing a measure of the stability of the organelle, tumor, etc., that is comprised of those cells, and providing a bridge to statistical physics models with predictive capacity. This picture, in turn, can provide insight into how to disrupt that robust state by targeting the signaling networks essential for tumor maintenance, as well as anticipating mechanisms of resistance. These approaches, which are unique to single cell measurements and are one of the central topics of this thesis, contrasts with traditional biology thinking that discards the heterogeneity of the system in favor of a more streamlined (but ultimately non-predictive) description.

1.6 THESIS OVERVIEW

Single-cell functional proteomics assay connect genomic information to biological function through quantitative and multiplex protein measurements. This thesis will focus on the development of microfluidic-based single-cell functional proteomics microchip and its applications in fundamental cancer biology and preclinical cancer research. Preliminary attempts of translating this microfluidic-based microchip into a valuable clinical toolkit with diagnostic and prognostic capacities will also be included.

Parallel proteomic assays across many different single cells yield unique pieces of information that are not readily disclosed by traditional biology methods. They also significantly shape the way of thinking on cancer itself. The availability of large proteomic dataset allows us to understand cancer using fundamental physicochemical principles such as statistical mechanics, thermodynamics and chemical kinetics. This in turn fosters the development of physical and computational approaches to make predictions on, for example, targeted therapy resistance and effective drug combinations.

Starting from Chapter 2, I will describe a microchip designed to quantify the levels of a dozen secreted, cytoplasmic and membrane proteins from single cells. We use the platform to assess protein-protein interactions associated with the EGF-receptor-mediated PI3K signaling pathway. Single-cell sensitivity is achieved by isolating a defined number of cells in 2nL volume chamber, each of which is patterned with a miniature antibody array. The cells are lysed on-chip, and the levels of released proteins are assayed using the antibody microarrays. We investigate three isogenic cell lines representing the cancer glioblastoma multiforme (GBM), at the basal level, under EGF stimulation, and under

erlotinib inhibition plus EGF stimulation. The measured protein abundances are consistent with previous work, and single-cell analysis uniquely reveals single-cell heterogeneity, and different types of strength of protein-protein interactions. This platform helps provide a comprehensive picture of altered signal transduction networks in tumor cells and provides insight into the effect of targeted therapies on protein signaling networks. (Chapter 2 has been taken in part from *Proc Natl Acad Sci USA* **109**, 419-424 (2012)).

Chapter 3 further applies the single-cell microchip to the study the transition of tumor hypoxia. Hypoxia is a near-universal feature of cancer, promoting glycolysis, cellular proliferation, and angiogenesis. The molecular mechanisms of hypoxic signaling have been intensively studied, but the impact of changes in oxygen partial pressure (pO_2) on the state of signaling network is less clear. In GBM cancer cell model, we examined the response of signaling networks to targeted pathway inhibition between 21% and 1% pO_2. We used a microchip technology that facilitates quantification of a panel of functional proteins from statistical number of single cells. We find that near 1.5% pO_2, the signaling network associated with mammalian target of rapamycin (mTOR) complex 1 (mTORC1)—a critical component of hypoxic signaling and a compelling cancer drug target—is deregulated in a manner such that it will be unresponsive to mTOR kinase inhibitors near 1.5% pO_2, but will respond at higher or lower pO_2 values. These predictions were validated through experiments on bulk GBM cell line cultures and on neurosphere cultures of a human-origin GBM xenograft tumor. We attempt to understand this behavior through the use of a quantitative version of Le Chatelier's principle derived from statistical mechanics, as well as through a steady-state kinetic model of protein interactions, both of which indicate that

hypoxia can influence mTORC1 signaling as a switch. The Le Chatelier approach also indicates that this switch may be thought of as a type of phase transition. Our analysis indicates that certain biologically complex cell behaviors may be understood using fundamental, thermodynamics-motivated principles. (Chapter 3 has been taken in part from *Proc Natl Acad Sci USA* **110**, E1352-1360 (2013)).

Chapter 4 demonstrates the application of this technology in the preclinical cancer research to study the cancer cell resistance to molecular targeted therapy and corresponding physical approaches to anticipate therapy resistance and identify effective therapy combinations. GBM is an aggressive tumor for which there are no effective surgical or pharmacologic treatments. GBM also serves as a prototype of advanced stage cancer. While GBM tumors contain druggable targets, resistance to single-agent targeted therapy is rapid and almost universal. Combination therapies that can anticipate resistance may provide a solution, but identifying effective combinations is largely an unmet challenge. We empirically derived signaling network inferences from quantitative functional proteomic analysis of statistical numbers of single cell separated from the glioblastoma-derived mouse model of mTOR kinase inhibitor resistance. Our approach is based upon elucidating the detailed signaling coordination within the phosphoprotein signaling pathways that are hyperactivated in human GBMs, and interrogating how that coordination responds to the perturbation of targeted inhibitor. We assayed for key elements of the phosphoprotein signaling pathways associated with GBM tumor growth and maintenance. Analysis of how the signaling coordination responses to the targeted inhibitor reveals a rapid adaptation to the presence of the drug, with compensation that occurs via the activation of alternative signaling pathways. The analysis allows us to anticipate resistance,

and to design combination therapies that are effective, as well as identify those therapies and therapy combinations that will be ineffective. The analysis also unveils a general and very fast-acting resistance mechanism.

The human-derived GBM model recapitulates the heterogeneity, invasive growth, and a drug response profile reflective of clinical behavior. We sought to elucidate the general mechanism of resistance by considering two resistance mechanisms. The first, Darwinian-like selection, occurs when a drug targeted at the dominant tumor cell population generates an environment suitable for a sub-population of cancer cells to flourish. The second mechanism is one in which the same tumor cells that initially respond to the drug adapt by altering their protein signaling networks. We analyzed the tumor model at 3 stages: control, responding to an mTOR kinase inhibitor, and resistant to that inhibitor. Analysis of the effect of the mTOR inhibitor resolves two independent signaling modes – one associated with mTOR signaling and a second associated with ERK/Src signaling. This suggested that drugging one target from each mode would provide an effective treatment. We tested 3 therapy combinations expected to be effective, and 4 expected to be ineffective, in mouse tumor models. All predictions were borne out: the effective therapy combinations completely halted tumor growth until the point of drug release, with no apparent side effects. We also identified that cellular adaption, rather than Darwinian evolution, led to resistance. This finding increases the clinical relevance of this work; this resistance mechanism is not readily identified via deep sequencing, but it can be detected via a few-day *in vitro* analysis using single cell functional proteomics. A retrospective analysis of tumor tissues from all treatment combinations further revealed that the mTOR signaling mode was driving tumor growth, while ERK/Src signaling was the dominant resistance

mechanism. We also show that this type of analysis can be done on a clinically relevant time-scale (Chapter 4 has been taken in part from a manuscript that is currently under review in *Nature Medicine*).

In Chapter 5, some preliminary results about the clinical translation of single cell proteomic chips will be presented. The hypothesis is that there exists a sufficient pharmacy to treat many GBM patients, and appropriately designed assays can inform, at the individual patient level, how those drugs can be combined for effective therapy. A key challenge is that those diagnostic assays must resolve the functional heterogeneity within a given patient's tumor. Single cell functional proteomics on statistical numbers of single cells therefore becomes a perfect candidate. Compared to model cell lines, clinical samples always have a low purity and weak functional protein expression. To meet the clinical challenges, the surface chemistries of the SCBCs has been intensively optimized to improve the assay sensitivity. This includes the use of on-chip poly-L-lysine (PLL) treatment and a covalent binding method to immobilize the DNAs to the PLL surface. A protocol on the single cell proteomic analysis of patient biopsy samples has also been developed, tested and standardized to ensure the assay reproducibility and robustness. A case study of a pediatric GBM patient sample will be discussed in detail for demonstrating the process of anticipating potential resistance and identifying the effective therapy combination within a clinical relevant time-scale.

1.7 REFERENCES

1. Fan, R. et al. Integrated barcode chips for rapid, multiplexed analysis of proteins in microliter quantities of blood. *Nat Biotechnol* **26**, 1373-8 (2008).

2. Alivisatos, P. The use of nanocrystals in biological detection. *Nat Biotechnol* **22**, 47-52 (2004).

3. Michalet, X. et al. Quantum dots for live cells, in vivo imaging, and diagnostics. *Science* **307**, 538-44 (2005).

4. Ma, C. et al. Multifunctional T-cell analyses to study response and progression in adoptive cell transfer immunotherapy. *Cancer Discov* **3**, 418-29 (2013).

5. Zhou, M. et al. Assessment of therapeutic efficacy of liposomal nanoparticles mediated gene delivery by molecular imaging for cancer therapy. *J Biomed Nanotechnol* **8**, 742-50 (2012).

6. Bender, E. Nanoparticle theories slowly turn into practice. *Cancer Discov* **1**, 280 (2011).

7. Schluep, T. et al. Preclinical efficacy of the camptothecin-polymer conjugate IT-101 in multiple cancer models. *Clin Cancer Res* **12**, 1606-14 (2006).

8. Ferrari, M. Cancer nanotechnology: opportunities and challenges. *Nat Rev Cancer* **5**, 161-71 (2005).

9. Misra, R., Acharya, S. & Sahoo, S.K. Cancer nanotechnology: application of nanotechnology in cancer therapy. *Drug Discov Today* **15**, 842-50 (2010).

10. Cancer nanotechnology: small, but heading for the big time. *Nat Rev Drug Discov* **6**, 174-5 (2007).

11. Gabizon, A., Goren, D., Cohen, R. & Barenholz, Y. Development of liposomal anthracyclines: from basics to clinical applications. *J Control Release* **53**, 275-9 (1998).

12. Torchilin, V. Antibody-modified liposomes for cancer chemotherapy. *Expert Opin Drug Deliv* **5**, 1003-25 (2008).

13. Perez, J.M., Simeone, F.J., Saeki, Y., Josephson, L. & Weissleder, R. Viral-induced self-assembly of magnetic nanoparticles allows the detection of viral particles in biological media. *J Am Chem Soc* **125**, 10192-3 (2003).

14. Portney, N.G. & Ozkan, M. Nano-oncology: drug delivery, imaging, and sensing. *Anal Bioanal Chem* **384**, 620-30 (2006).

15. Mohanty, C. & Sahoo, S.K. The in vitro stability and in vivo pharmacokinetics of curcumin prepared as an aqueous nanoparticulate formulation. *Biomaterials* **31**, 6597-611 (2010).

16. Torchilin, V.P., Lukyanov, A.N., Gao, Z. & Papahadjopoulos-Sternberg, B. Immunomicelles: targeted pharmaceutical carriers for poorly soluble drugs. *Proc Natl Acad Sci U S A* **100**, 6039-44 (2003).

17. Rawat, M., Singh, D., Saraf, S. & Saraf, S. Nanocarriers: promising vehicle for bioactive drugs. *Biol Pharm Bull* **29**, 1790-8 (2006).

18. Menjoge, A.R., Kannan, R.M. & Tomalia, D.A. Dendrimer-based drug and imaging conjugates: design considerations for nanomedical applications. *Drug Discov Today* **15**, 171-85 (2010).

19. Fodor, S.P. et al. Light-directed, spatially addressable parallel chemical synthesis. *Science* **251**, 767-73 (1991).

20.   Lee, K.B., Park, S.J., Mirkin, C.A., Smith, J.C. & Mrksich, M. Protein nanoarrays generated by dip-pen nanolithography. *Science* **295**, 1702-5 (2002).

21.   Rosenblatt, K.P. et al. Serum proteomics in cancer diagnosis and management. *Annu Rev Med* **55**, 97-112 (2004).

22.   Wu, G. et al. Bioassay of prostate-specific antigen (PSA) using microcantilevers. *Nat Biotechnol* **19**, 856-60 (2001).

23.   Bailey, R.C., Kwong, G.A., Radu, C.G., Witte, O.N. & Heath, J.R. DNA-encoded antibody libraries: a unified platform for multiplexed cell sorting and detection of genes and proteins. *J Am Chem Soc* **129**, 1959-67 (2007).

24.   Shi, Q. et al. Single-cell proteomic chip for profiling intracellular signaling pathways in single tumor cells. *Proc Natl Acad Sci U S A* **109**, 419-24 (2012).

25.   Wang, J. et al. Quantitating cell-cell interaction functions with applications to glioblastoma multiforme cancer cells. *Nano Lett* **12**, 6101-6 (2012).

26.   Wei, W. et al. Hypoxia induces a phase transition within a kinase signaling network in cancer cells. *Proc Natl Acad Sci U S A* **110**, E1352-60 (2013).

27.   Wei, W. et al. Microchip platforms for multiplex single-cell functional proteomics with applications to immunology and cancer research. *Genome Med* **5**, 75 (2013).

28.   Heath, J.R., Davis, M.E. & Hood, L. Nanomedicine targets cancer. *Sci Am* **300**, 44-51 (2009).

29.   Rubin, H. The significance of biological heterogeneity. *Cancer Metastasis Rev* **9**, 1-20 (1990).

30.   Meacham, C.E. & Morrison, S.J. Tumour heterogeneity and cancer cell plasticity. *Nature* **501**, 328-37 (2013).

31.   Altschuler, S.J. & Wu, L.F. Cellular heterogeneity: do differences make a difference? *Cell* **141**, 559-63 (2010).

32.   Beck, M. et al. The quantitative proteome of a human cell line. *Mol Syst Biol* **7**, 549 (2011).

33.   Spencer, S.L., Gaudet, S., Albeck, J.G., Burke, J.M. & Sorger, P.K. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature* **459**, 428-32 (2009).

34.   Ferrell, J.E., Jr. & Machleder, E.M. The biochemical basis of an all-or-none cell fate switch in Xenopus oocytes. *Science* **280**, 895-8 (1998).

35.   Son, S. et al. Direct observation of mammalian cell growth and size regulation. *Nat Methods* **9**, 910-2 (2012).

36.   Strickfaden, S.C. et al. A mechanism for cell-cycle regulation of MAP kinase signaling in a yeast differentiation pathway. *Cell* **128**, 519-31 (2007).

37.   Yuan, T.L., Wulf, G., Burga, L. & Cantley, L.C. Cell-to-cell variability in PI3K protein level regulates PI3K-AKT pathway activity in cell populations. *Curr Biol* **21**, 173-83 (2011).

38.   Campbell, L.L. & Polyak, K. Breast tumor heterogeneity: cancer stem cells or clonal evolution? *Cell Cycle* **6**, 2332-8 (2007).

39.   Cohen, A.A. et al. Dynamic proteomics of individual cancer cells in response to a drug. *Science* **322**, 1511-6 (2008).

40.   Marusyk, A., Almendro, V. & Polyak, K. Intra-tumour heterogeneity: a looking glass for cancer? *Nat Rev Cancer* **12**, 323-34 (2012).

41.   Hill, T.L. Thermodynamics of small systems (Dover Publications, New York, 1994).

42.   Newman, J.R. et al. Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature* **441**, 840-6 (2006).

43.   Raser, J.M. & O'Shea, E.K. Control of stochasticity in eukaryotic gene expression. *Science* **304**, 1811-4 (2004).

44.   Ma, C. et al. A clinical microchip for evaluation of single immune cells reveals high functional heterogeneity in phenotypically similar T cells. *Nat Med* **17**, 738-43 (2011).

45.   Sharma, S.V. et al. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell* **141**, 69-80 (2010).

46.   Irish, J.M. et al. Single cell profiling of potentiated phospho-protein networks in cancer cells. *Cell* **118**, 217-28 (2004).

47.   Weinberg, R.A. The biology of cancer (Garland Science, New York, 2013).

48.   Kholodenko, B., Yaffe, M.B. & Kolch, W. Computational approaches for analyzing information flow in biological networks. *Sci Signal* **5**, re1 (2012).

49.   Logue, J.S. & Morrison, D.K. Complexity in the signaling network: insights from the use of targeted inhibitors in cancer therapy. *Genes Dev* **26**, 641-50 (2012).

50.   Yaffe, M.B. The scientific drunk and the lamppost: massive sequencing efforts in cancer discovery and treatment. *Sci Signal* **6**, pe13 (2013).

51.   Coulter, W.H. (Means for counting particles suspended in a fluid, US Patent 2656508, 1953).

52. Fulwyler, M.J. Electronic separation of biological cells by volume. *Science* **150**, 910-1 (1965).

53. Herzenberg, L.A. et al. The history and future of the fluorescence activated cell sorter and flow cytometry: a view from Stanford. *Clin Chem* **48**, 1819-27 (2002).

54. Julius, M.H., Masuda, T. & Herzenberg, L.A. Demonstration that antigen-binding cells are precursors of antibody-producing cells after purification with a fluorescence-activated cell sorter. *Proc Natl Acad Sci U S A* **69**, 1934-8 (1972).

55. Czerkinsky, C.C., Nilsson, L.A., Nygren, H., Ouchterlony, O. & Tarkowski, A. A solid-phase enzyme-linked immunospot (ELISPOT) assay for enumeration of specific antibody-secreting cells. *J Immunol Methods* **65**, 109-21 (1983).

56. Bendall, S.C. et al. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* **332**, 687-96 (2011).

57. Sandberg, J.K., Fast, N.M. & Nixon, D.F. Functional heterogeneity of cytokines and cytolytic effector molecules in human CD8+ T lymphocytes. *J Immunol* **167**, 181-7 (2001).

58. Thorsen, T., Maerkl, S.J. & Quake, S.R. Microfluidic large-scale integration. *Science* **298**, 580-4 (2002).

59. Ottesen, E.A., Hong, J.W., Quake, S.R. & Leadbetter, J.R. Microfluidic digital PCR enables multigene analysis of individual environmental bacteria. *Science* **314**, 1464-7 (2006).

60. Lorthongpanich, C. et al. Single-cell DNA-methylation analysis reveals epigenetic chimerism in preimplantation embryos. *Science* **341**, 1110-2 (2013).

61.   Heath, J.R. & Davis, M.E. Nanotechnology and cancer. *Annu Rev Med* **59**, 251-65 (2008).

62.   Zare, R.N. & Kim, S. Microfluidic platforms for single-cell analysis. *Annu Rev Biomed Eng* **12**, 187-201 (2010).

63.   Sun, J. et al. A microfluidic platform for systems pathology: multiparameter single-cell signaling measurements of clinical brain tumor specimens. *Cancer Res* **70**, 6128-38 (2010).

64.   Zhou, J., Wu, Y., Lee, S.K. & Fan, R. High-content single-cell analysis on-chip using a laser microarray scanner. *Lab Chip* **12**, 5025-33 (2012).

65.   Di Carlo, D., Aghdam, N. & Lee, L.P. Single-cell enzyme concentrations, kinetics, and inhibition analysis using high-density hydrodynamic cell isolation arrays. *Anal Chem* **78**, 4925-30 (2006).

66.   Di Carlo, D., Wu, L.Y. & Lee, L.P. Dynamic single cell culture array. *Lab Chip* **6**, 1445-9 (2006).

67.   Faley, S.L. et al. Microfluidic single cell arrays to interrogate signalling dynamics of individual, patient-derived hematopoietic stem cells. *Lab Chip* **9**, 2659-64 (2009).

68.   Faley, S.L., Copland, M., Reboud, J. & Cooper, J.M. Cell chip array for microfluidic proteomics enabling rapid in situ assessment of intracellular protein phosphorylation. *Biomicrofluidics* **5**, 24106 (2011).

69.   Huebner, A. et al. Quantitative detection of protein expression in single cells using droplet microfluidics. *Chem Commun (Camb)*, 1218-20 (2007).

70. Martino, C. et al. Intracellular protein determination using droplet-based immunoassays. *Anal Chem* **83**, 5361-8 (2011).

71. Brouzes, E. et al. Droplet microfluidic technology for single-cell high-throughput screening. *Proc Natl Acad Sci U S A* **106**, 14195-200 (2009).

72. Love, J.C., Ronan, J.L., Grotenbreg, G.M., van der Veen, A.G. & Ploegh, H.L. A microengraving method for rapid selection of single cells producing antigen-specific antibodies. *Nat Biotechnol* **24**, 703-7 (2006).

73. Han, Q., Bradshaw, E.M., Nilsson, B., Hafler, D.A. & Love, J.C. Multidimensional analysis of the frequencies and rates of cytokine secretion from single cells by quantitative microengraving. *Lab Chip* **10**, 1391-400 (2010).

74. Varadarajan, N. et al. Rapid, efficient functional characterization and recovery of HIV-specific human CD8+ T cells using microengraving. *Proc Natl Acad Sci U S A* **109**, 3885-90 (2012).

75. Han, Q. et al. Polyfunctional responses by human T cells result from sequential release of cytokines. *Proc Natl Acad Sci U S A* **109**, 1607-12 (2012).

76. Shin, Y.S. et al. Chemistries for patterning robust DNA microbarcodes enable multiplex assays of cytoplasm proteins from single cancer cells. *Chemphyschem* **11**, 3063-9 (2010).

77. Ahmad, H. et al. A robotics platform for automated batch fabrication of high density, microfluidics-based DNA microarrays, with applications to single cell, multiplex assays of secreted proteins. *Rev Sci Instrum* **82**, 094301 (2011).

78. Lu, Y. et al. High-throughput secretomic analysis of single cells to assess functional cellular heterogeneity. *Anal Chem* **85**, 2548-56 (2013).

79. Fu, A.Y., Spence, C., Scherer, A., Arnold, F.H. & Quake, S.R. A microfabricated fluorescence-activated cell sorter. *Nat Biotechnol* **17**, 1109-11 (1999).

80. Zhuang, X. Nano-imaging with Storm. *Nat Photonics* **3**, 365-367 (2009).

81. Cai, L. Turning single cells into microarrays by super-resolution barcoding. *Brief Funct Genomics* **12**, 75-80 (2013).

82. Shin, Y.S. et al. Protein signaling networks from single cell fluctuations and information theory profiling. *Biophys J* **100**, 2378-86 (2011).

83. Raj, A. & van Oudenaarden, A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**, 216-26 (2008).

84. Paszek, P. et al. Population robustness arising from cellular heterogeneity. *Proc Natl Acad Sci U S A* **107**, 11644-9 (2010).

85. Quintana, E. et al. Phenotypic heterogeneity among tumorigenic melanoma cells from patients that is reversible and not hierarchically organized. *Cancer Cell* **18**, 510-23 (2010).

86. Nathanson, D.A. et al. Targeted therapy resistance mediated by dynamic regulation of extrachromosomal mutant EGFR DNA. *Science* **343**, 72-6 (2014).

C h a p t e r   2

Development of the microchip: single-cell functional proteomics chips for profiling cancer cell signal transduction pathways

2.1 INTRODUCTION

Although signal transduction inhibitors occasionally offer clinical benefit for cancer patients[1], signal flux emanating from oncogenes is often distributed through multiple pathways[2], potentially underlying the failure of most such inhibitors[3]. Measuring signal flux through multiple pathways, in response to signal transduction inhibitors, may help uncover network interactions that contribute to therapeutic resistance, and which are not predicted by analyzing pathways in isolation[4]. The cellular and molecular complexity of a solid tumor microenvironment[5] suggests the need to study signaling in individual cancer cells.

Protein-protein interactions within signaling pathways are often elucidated by assessing the levels of relevant pathway proteins in model and tumor-derived cell lines, and with various genetic and molecular perturbations. Such interactions, and the implied signaling networks, may also be elucidated via quantitative measurements of multiple pathway related proteins within single cells[6]. At the single cell level, inhibitory and activating protein-protein relationships, as well as stochastic (single-cell) fluctuations, are revealed. However, most techniques for profiling signaling pathways[7, 8] require large numbers of cells. Single-cell immunostaining[9] is promising, and some flow cytometry[6] techniques are relevant, as discussed below.

We describe quantitative, multiplex assays of intracellular signaling proteins from single cancer cells using a platform called the single-cell barcode chip (SCBC). The SCBC simple in concept: a single or defined number of cells, is isolated within a 1~2 nanoliter volume microchamber that contains an antibody array[10] for the capture and detection of a panel of proteins. The SCBC design[11] permits on chip culture and on chip lysis of each individual trapped cell.

Intracellular staining flow cytometry can assay up to 11 phosphoproteins from single cells[6]. Our SCBC can profile a panel of up to 20 proteins, but this limit is not fundamental and may be extendable. Depending on the design, several hundred to several thousand single cells per chip can be measured simultaneously, yielding some statistical assessment for each experiment. The SCBC is a relatively simple platform, and only requires a few hundred cells per assay.

In this chapter, we will describe in detail the enabling technologies, the design and fabrication of SCBCs. The protocol of the single cell proteomic assay and the result extraction, calibration, and statistical analysis will also be covered. We will focus on the unique pieces of information disclosed from single cell analysis and how the information endows us with the capacity of employing physicochemical approaches to understand cancer at a systems level.

To illustrate the process, we used the SCBC to study signal transduction in Glioblastoma Multiforme (GBM), a primary malignant brain tumor[12]. GBM has been genetically characterized, yet the nature of signaling pathways downstream of key oncogenic mutations, such as EGFR activating mutation (EGFRvIII) and PTEN tumor

suppressor gene loss associated with Receptor Tyrosine Kinase (RTK)/PI3K signaling, are incompletely understood[13-15]. Single cell experiments may also help resolve the characteristic heterogeneity of GBM.

We interrogated eleven proteins directly or potentially associated with PI3K signaling through three isogenic GBM cell lines: U87 (expressing wild-type p53, mutant PTEN and low levels of wild-type EGFR, no EGFRvIII)[16, 17], U87 EGFRvIII (U87 cells stably expressing EGFRvIII deletion mutant), and U87 EGFRvIII PTEN (U87 cells co-expressing EGFRvIII and PTEN)[18]. Fig. 2.1 diagrams this biology. Each cell line was investigated under conditions of standard cell culture, in response to EGF stimulation, and after erlotinib treatment followed by EGF stimulation. The proteins assayed represented receptor tyrosine kinases (RTKs) and proteins signifying activation of PI3k and MAPK signaling. They were (p-denotes phosphorylation): p-Src, p-mammalian target of rapamycin (p-mTOR), p-p70 ribosomal protein S6 kinase (p-p70S6K), p-glycogen synthase kinase-3 (p-GSK-3α/β), p-p38 mitogen activated protein kinase (p-p38α), p-extracellular regulated kinase (p-ERK), p-c-Jun N-terminal kinase (p-JNK2), p-Platelet Derived Growth Factor Receptor β (p-PDGFRβ), p-Vascular Endothelial Growth Factor Receptor 2 (p-VEGFR2), tumor protein 53 (P53) and total EGFR.

Figure 2.1 The PI3K pathway activated by EGF-stimulated EGFR or the constitutively activated EGFRvIII. All proteins in light blue with central yellow background were assayed. Orange background proteins were expressed in the cell lines U87 EGFRvIII or U87 EGFRvIII PTEN. The oval, yellow background components are the investigated molecular perturbations.
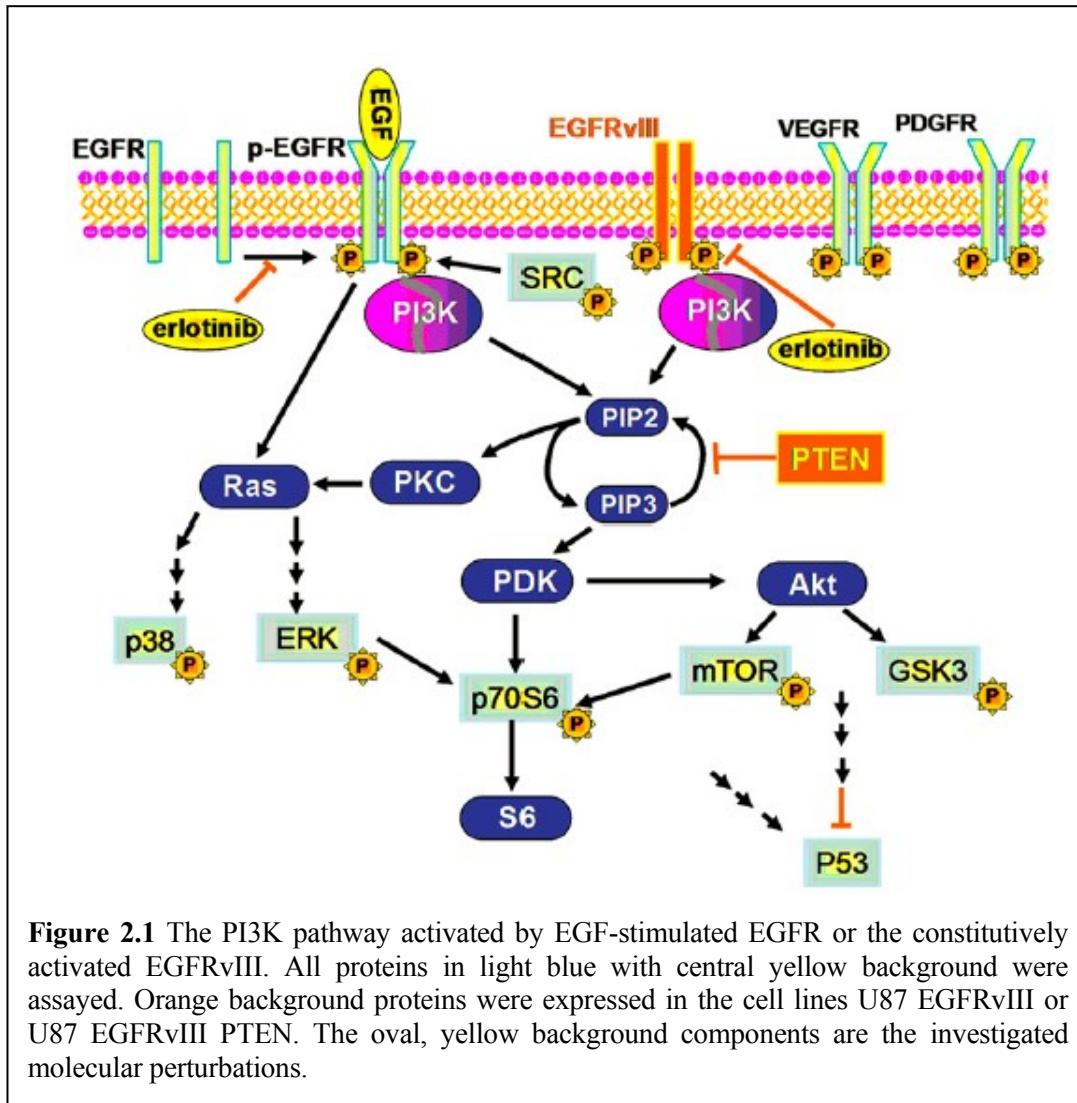
2.2 EXPERIMENTAL METHODS

2.2.1 Enabling technologies: DNA Encoded Antibody Library (DEAL) and DNA barcode microarray

The basic concept of SCBCs is to pattern a many-element capture antibody array in each single-cell microwell so that different proteins are detected at different designated array spots. The key enabling technologies of SCBCs is the miniature antibody arrays. The
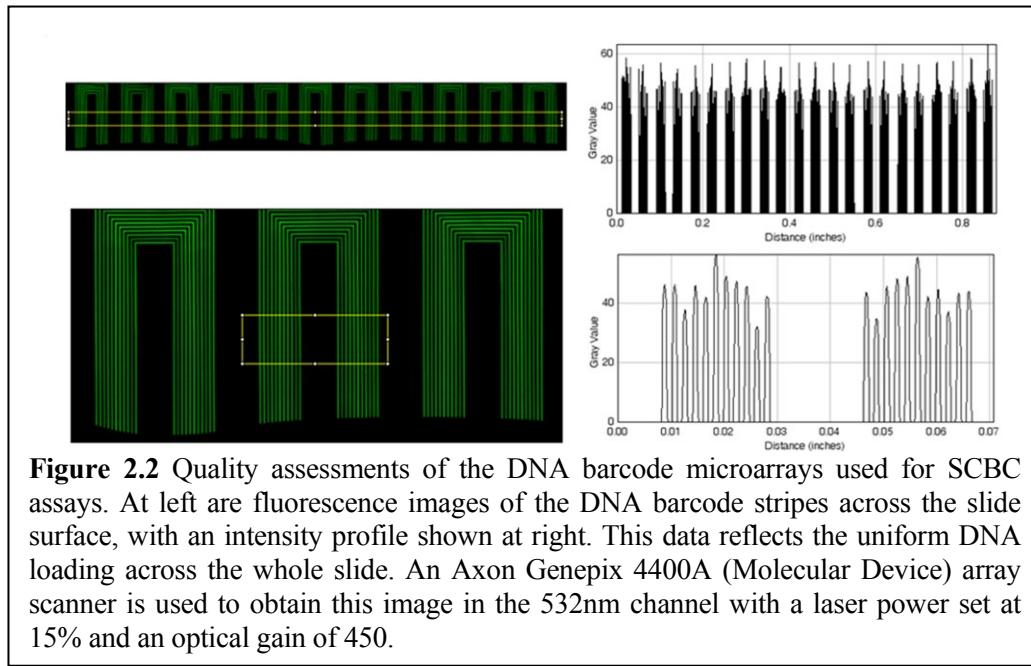
Figure 2.2 Quality assessments of the DNA barcode microarrays used for SCBC assays. At left are fluorescence images of the DNA barcode stripes across the slide surface, with an intensity profile shown at right. This data reflects the uniform DNA loading across the whole slide. An Axon Genepix 4400A (Molecular Device) array scanner is used to obtain this image in the 532nm channel with a laser power set at 15% and an optical gain of 450.

patterning approach for the antibody barcode arrays has unique constraints. Due to the instability of antibody for long term storage, or towards microchip processing conditions, the barcodes are initially patterned as single strand DNA (ssDNA) barcodes, with each barcode stripe having a unique ssDNA label. A cocktail of antibodies labeled with complementary ssDNA' oligomers (DEAL) is used to convert the DNA barcode into an antibody barcode, just prior to running an proteomic assay. The microchamber surface area

for a typical 20-element barcode is between $150 \times 150 \ \mu m^2$ to $1000 \times 150 \ \mu m^2$, implying

each array element needs to be around 10-20 μm wide at a 20-40 μm pitch. Such

dimensions may be read with a standard microarray scanner, but are beyond the resolution

of standard spotting tools. Molecular patterning tools that can approach these dimensions

include molded elastomer stamping, dip pen lithography, and microfluidic flow patterning.

Of these choices, stamping does not permit the required level of multiplexing, while dip

pen does not yield a surface coverage sufficient for stable and sensitive assays. We have

therefore developed microfluidic flow patterning into the method of choice for SCBCs,

including even building robotics systems to automate the task. Specifically, an elastomer

film is molded so that it contains a series of long, serpentine channels. It is adhered to the

top of a glass slide. Solutions containing a different ssDNAs oligomer are flowed through

each channel. After solution evaporation, the molded elastomer is then removed, leaving a

series of 10 to 20 μm wide stripes of different ssDNA oligomers across the glass substrate

(Fig. 2.2). The influence of various barcoding surface chemistries on assay sensitivity are

described in detail in our previous work. The details of microfluidic flow patterning of

DNA barcode microarray and the synthesis of ssDNA oligomers and antibody conjugates

can be found in Appendix A: Supplementary Methods.

2.2.2 Design and fabrication of single-cell proteomic chip

The PDMS microfluidic chip for the single-cell proteomic experiment is comprised of

a two-layer microfluidic network and fabricated by soft lithography[11] (Fig. 2.3 A and B).

Valves isolate the chip into 120 microchambers for cell compartmentalization, cell lysis,

and protein assays (Fig 2.3 C-F). A push-down valve configuration was utilized with a

thick control layer bonded together with a thin flow layer. The molds for the control

layer and flow layer were fabricated with SU8 2010 negative photoresist (~23 μm

thickness) and SPR 220 positive photoresist (~15 μm), respectively. The photoresist

patterns for the flow layer were rounded via thermal treatment. The thick control layer was

Figure 2.3 Single-cell barcode chip design and operation. (A) Design of the SCBC microfluidic chip for on-chip cell lysis and multiplex phosphoprotein detection. V1 to V6 represent valves. V3 isolates chambers from channel and V4 control the diffusion between cell chamber and lysis buffer chamber. V1, V2, V5 and V6 control flow within the microchannels. (B) A photograph of an SCBC. The flow layer (red) and the control valve layer (blue) are delineated with food dyes. (C) A single measurement unit of the SCBC. Single or few cells are isolated in the 2 nL microchamber. Intracellular proteins are assayed by introducing a pre-aliquoted lysis buffer, whereupon the released proteins bind to the antibody microarray within the chamber. (D) A photograph of a single microchamber. (E) Visualization of on-chip lysis buffer diffusion. Red food dye is used to visualize the diffusion process. The chip is sitting on ice and titled to accelerate on-chip diffusion. (F) A drawing of a single cell chamber with critical parts labeled. A cell is isolated in the cell chamber by the valves. The neighboring chamber contains cell lysis buffer. The duplicate DNA barcode copies are converted into an antibody microarray prior to cell loading, counting, and lysis.

molded with 10:1 mixture of GE RTV 615 PDMS prepolymer part A and part B (w/w) and

the flow layer was formed by spin-coating a 20:1 mixture of GE RTV 615 part A and part

B (w/w) on the flow layer mold at 2000rpm for 60 seconds. Both layers were cured at 80°C for 1 hour, whereupon the control layer was cut from its mold and aligned to the flow layer. An additional 60 minutes of thermal treatment at 80°C ensured that the two layers bonded into a monolithic device, which was then peeled from its mold and punched to create appropriate access holes. Finally, the PDMS chip was thermally bonded to the DNA barcode slide to form the working device.

2.2.3 Protocol of single-cell proteomic assays.

The cell determines the copy numbers of a given protein, while the microchamber volume is minimized so those copy numbers are at a detectable level using standard sandwich ELISA immunoassays with fluorescent readouts. At the beginning of the experiment, all SCBC microchannels are blocked with blocking buffer for 60 minutes. A cocktail of all DNA-antibody conjugates is flowed through the channels for 60 minutes, transforming the DNA barcode microarrays into antibody microarrays. Unbound

Figure 2.4 The relationship between on-chip incubation time and obtained intensity: 2 hr incubation is sufficient to capture more than 95% of the released protein following cell lysis.
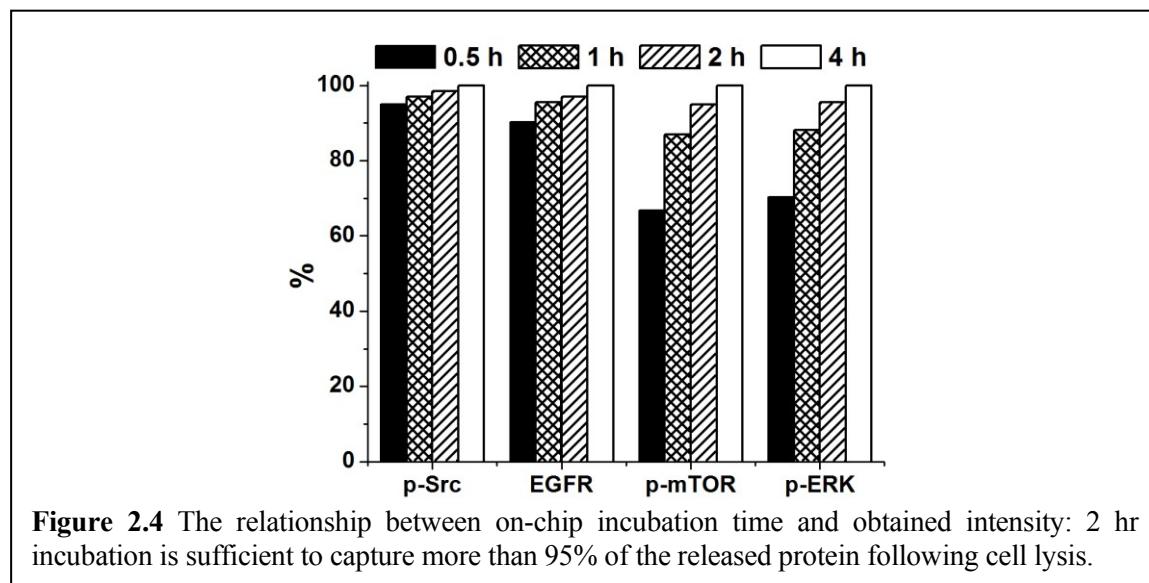
conjugates are removed with washing buffer. Then, 3× lysis buffer is loaded into the lysis buffer chambers and cells are loaded in the cell chamber while keeping the valves between these chambers closed. Upon cell loading, each microchamber contains zero to a few cells, which are counted through the transparent chip under microscope. The valves are then opened to allow on-chip diffusion of lysis buffer to the neighboring cell chambers for 20 min on ice. Cell are lysed via diffusion of lysis buffer which contains phosphatase protease inhibitors. The SCBC was then incubated 20 min on ice and 2 hour at room temperature with gentle shaking to allow capturing the proteins released from the cells.

The two hours incubation used here reaches >95% of maximal intensity for all assays (Fig. 2.4). Cell lysate is quickly removed by washing buffer after incubation. Captured proteins are developed by applying biotinylated detection antibodies and fluorescent dye-labeled streptavidin for visualization. The barcode glass slide is then detached for scanning. A detailed step by step protocol can be found in Appendix A: Supplementary Methods.

2.2.4 Results extraction, calibration and conversion

Once an SCBC assay is complete, an array scanner (Axon Genepix 4400A, Molecular
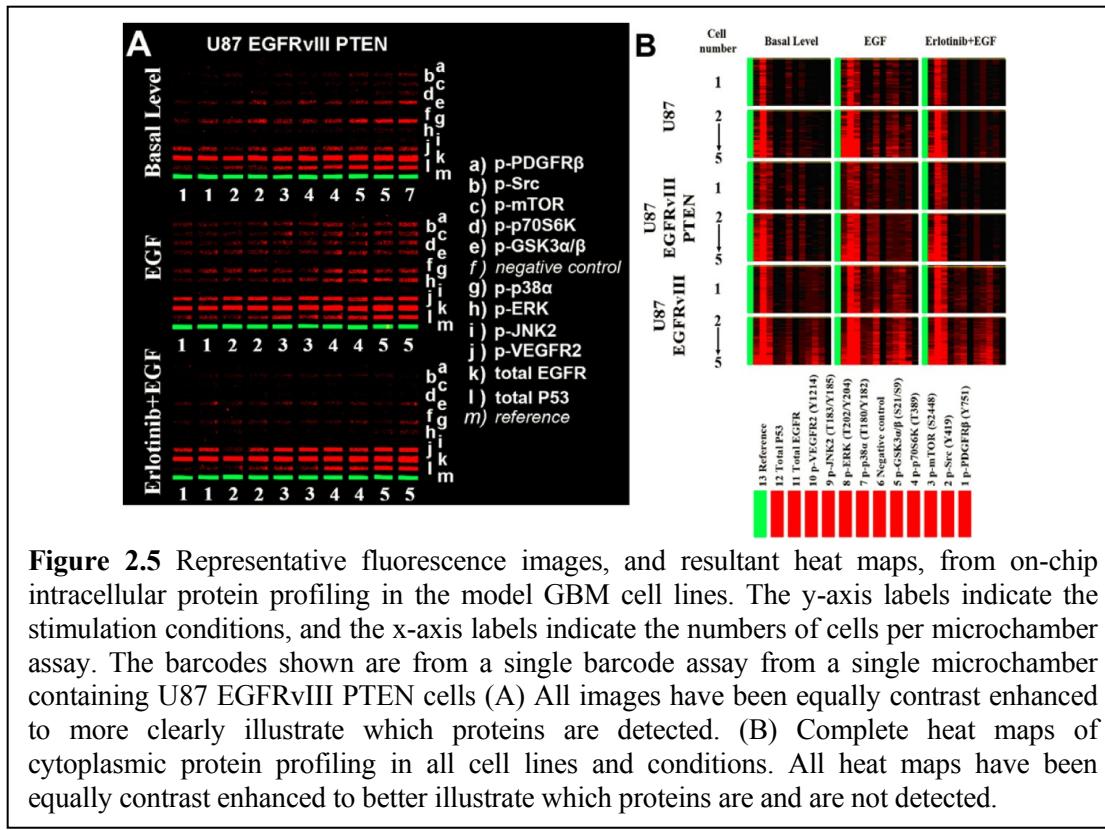
Figure 2.5 Representative fluorescence images, and resultant heat maps, from on-chip intracellular protein profiling in the model GBM cell lines. The y-axis labels indicate the stimulation conditions, and the x-axis labels indicate the numbers of cells per microchamber assay. The barcodes shown are from a single barcode assay from a single microchamber containing U87 EGFRvIII PTEN cells (A) All images have been equally contrast enhanced to more clearly illustrate which proteins are detected. (B) Complete heat maps of cytoplasmic protein profiling in all cell lines and conditions. All heat maps have been equally contrast enhanced to better illustrate which proteins are and are not detected.

Devices) is utilized to digitize the fluorescent levels from the antibody arrays associated

with each microchamber. One array element has a distinct fluorophore to serve as an

alignment marker; this permits individual proteins to be identified via their spatial location

(Fig. 2.5). The digitized data is then loaded into a table (using custom written algorithms).

Each table row corresponds to a specific microchamber address. The columns contain the

numbers and locations of the cells in each microchamber, plus the fluorescence intensities

corresponding to each of the assayed proteins. Those intensities are background normalized

(using 0 cell data) and converted into protein copy numbers using calibrations (Fig. 2.6).

Unlike other single cell proteomic tools, SCBC assays can yield absolute protein level quantification in copy number per cell, which allows clinical studies or investigations in which statistical cell behaviors are compared across a perturbation series or patients. This is achieved by generating calibration curves against standard (recombinant) proteins when available.

The calibration experiments are performed within an SCBC and under exact the same condition as the single-cell proteomic assay described above, except that standard proteins are used, rather than cells. A mixture of standard proteins from the SCBC assayed panel is serially diluted in 1× PBS and flowed into the SCBC microchannels. Fluorescence signals are collected to generate the calibration curves. Because the volume of the microchamber is known, these calibration curves enable a transformation from the fluorescence intensity to

Figure 2.6 Protein calibration test. (A) Fluorescence images used to extract calibration data for the SCBC assays, and representing a serial dilution of assayed proteins. Standard protein are serially diluted in 1× PBS and flwoed into the SCBC microchannels. (B) Calibration curves for several assayed proteins.
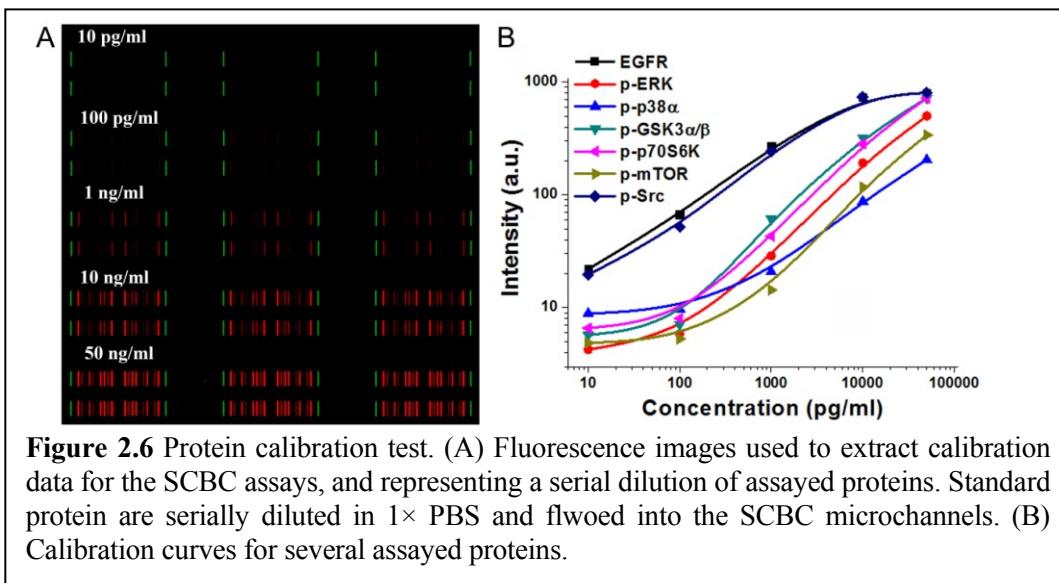
number of molecules for each protein assayed, under the caveat that the standard protein may not be exactly the same as their counterparts from the cells.

2.2.5 Cell culture, stimulation and drug treatment

The human GBM cell lines U87 was purchased from American Tissue Culture Collection. U87 EGFRvIII and U87 EGFRvIII PTEN cells were constructed as previously described[14,18]. Cell lines were routinely maintained in DMEM (ATCC) containing 10% fetal bovine serum in a humidified atmosphere of 5% CO_2, 95% air at 37°C. For EGF stimulation, cells were serum-starved for 24 hours and then stimulated by EGF at 50 ng/mL for 10 min before harvest. For erlotinib treatment, serum-starved cells were treated with 10 μM enlotinib for 24 hours, followed by EGF stimulation (50 ng/mL) for 10 min before harvest. The treated cells were dissociated with trypsin and EDTA and suspended in cold PBS with a concentration of 1000 cell/μL prior to loading to the device.

2.3 RESULTS AND DISCUSSION

2.3.1 Unique information disclosed from single-cell analysis

Single cell proteomics provides the most direct approach for elucidating protein signaling network structure and coordination, and information from such measurements emerges at many levels. An SCBC cell dataset, which is comprised of a statistical number of single-cell assays, yield three types of independent observables. The first observables are averaged level of each assayed protein from single cells, which can be compared against results from assays on bulk cell populations such as Western blotting or sandwich ELISA.

The second and third observables are unique to single-cell multiplex proteomics assays. A measurement of the average level of a protein requires many single-cell measurements. Such measurements, if compiled as a histogram of the frequency of observation versus the measured protein levels, reflect the fluctuations of that protein. Similar to spectroscopy that is the outcome from interactions between matter and radiation energy, single cell protein fluctuations are spectra involving all the protein interactions within a single cell. As a result, they are highly informative toward understanding protein functional activity. A straightforward example is to identify subpopulations of cells by looking at the fluctuation profiles of a single protein marker or a combination of markers (Fig. 2.7)[19]. This information has been intensively used in flow cytometry and flow-activated cell sorting (FACS). A deeper analysis of the protein fluctuation line shapes by a mean field model (See Chapter 3 for details) provides a context for discussing how the average effect of other proteins influences the fluctuations of a specific protein in question. It offers predictive capacity of fluctuation profiles that a widely dispersed fluctuation can

indicate a highly active protein that is involved in multiple functional processes. A narrow, sharp fluctuation, by contrast, represents a protein with limited interactions. The protein fluctuation, namely the spread in copy number of a given protein as measured across each of many otherwise identical single cells, can represent the functional
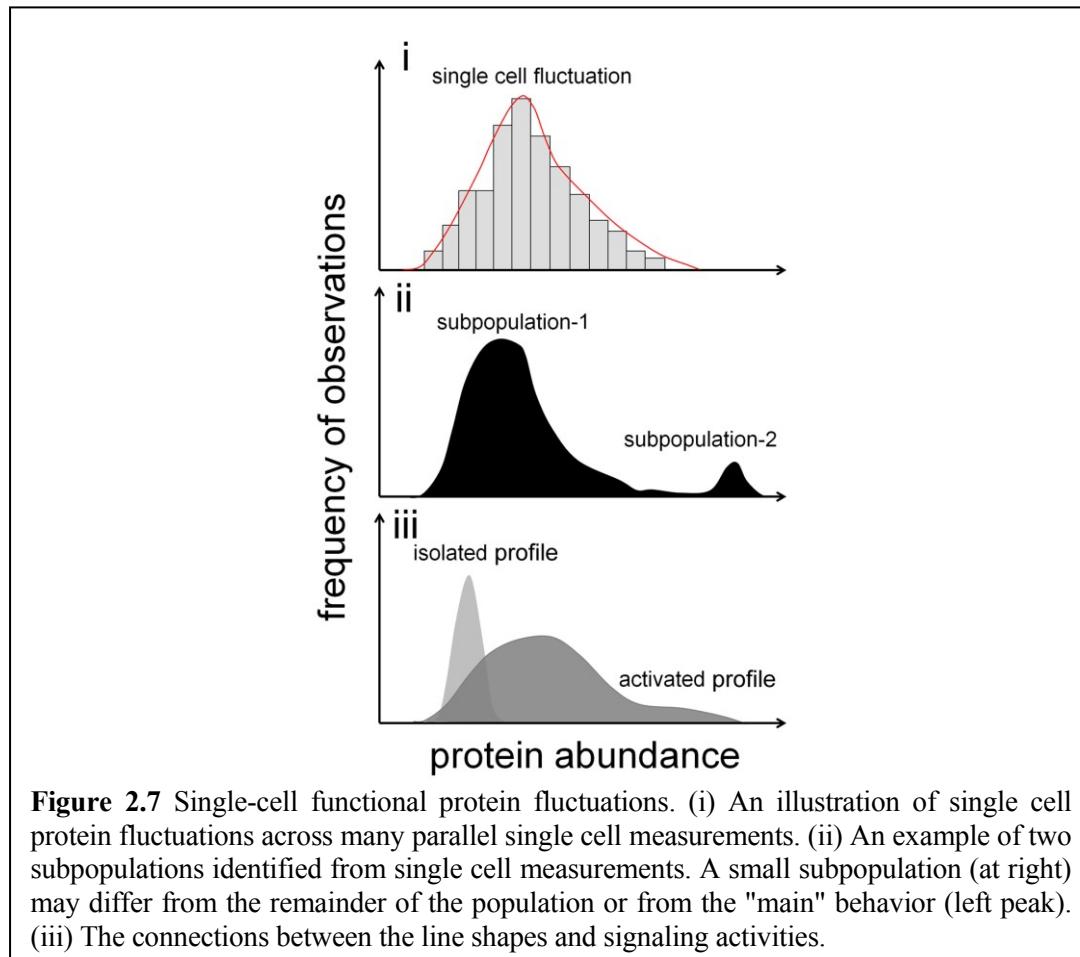
Figure 2.7 Single-cell functional protein fluctuations. (i) An illustration of single cell protein fluctuations across many parallel single cell measurements. (ii) An example of two subpopulations identified from single cell measurements. A small subpopulation (at right) may differ from the remainder of the population or from the "main" behavior (left peak). (iii) The connections between the line shapes and signaling activities.

heterogeneity of that protein, which further contributes the population heterogeneity of those seemingly identical single cells. The concept of a stable population existing in the presence of random fluctuations is reminiscent of many physical systems that are successfully understood using statistical physics. Thus, tools derived from that field can probably be applied to using fluctuations to determine the nature of signaling networks.

The third unique observations are the protein-protein correlations between the various assayed proteins across all the single cell measurements. Correlations and anti-correlations can imply activating and inhibitory interactions. By connecting assayed proteins (network nodes) with different weights of lines to reflect the strength of correlations, a correlation network map can be established and serve as a signature of the signaling network activity. This also means that measurements with higher multiplexing capacity will capture larger number of such interactions, and thus increasingly resolve the associated protein signaling network. In the context of cancer cells, comparing the correlation network of single cancer cells following stimulus, plus a drug treatment with those with stimulus alone could identify how the signaling profiles are affected by the drug. Similarly, comparing the correlation network at initial, drug responsive and drug resistant stages to analyze the evolution of the cancer cell signaling following the drug treatment could yield insight into how tumor cells develop therapeutic resistance to evade the targeted therapy.

2.3.2. Stripping experimental uncertainty out of the biological variation

The protein fluctuations, namely the spread (variance) in copy number of given proteins as measured across each of many otherwise identical single cells, represents the functional heterogeneity of the proteins. However, experimental errors can also contribute the measured variance and thus must be compared against the measured variations for extracting the true biological fluctuations.

The experimental errors mainly include the variation from non-uniform DNA
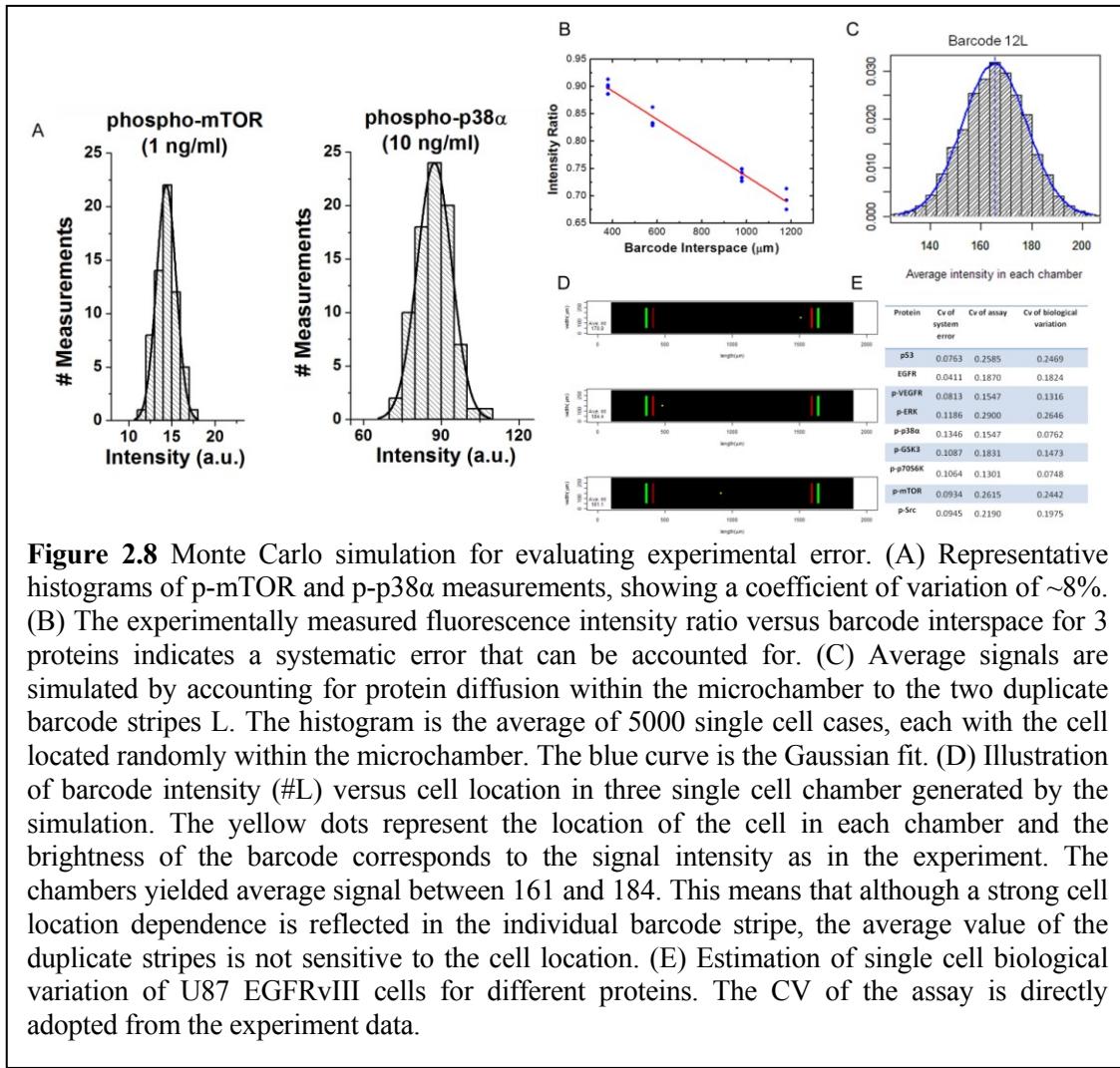
Figure 2.8 Monte Carlo simulation for evaluating experimental error. (A) Representative histograms of p-mTOR and p-p38α measurements, showing a coefficient of variation of ~8%. (B) The experimentally measured fluorescence intensity ratio versus barcode interspace for 3 proteins indicates a systematic error that can be accounted for. (C) Average signals are simulated by accounting for protein diffusion within the microchamber to the two duplicate barcode stripes L. The histogram is the average of 5000 single cell cases, each with the cell located randomly within the microchamber. The blue curve is the Gaussian fit. (D) Illustration of barcode intensity (#L) versus cell location in three single cell chamber generated by the simulation. The yellow dots represent the location of the cell in each chamber and the brightness of the barcode corresponds to the signal intensity as in the experiment. The chambers yielded average signal between 161 and 184. This means that although a strong cell location dependence is reflected in the individual barcode stripe, the average value of the duplicate stripes is not sensitive to the cell location. (E) Estimation of single cell biological variation of U87 EGFRvIII cells for different proteins. The CV of the assay is directly adopted from the experiment data.

barcode patterns which in turn transfers to the final protein signal, and the variation due to

the randomly distributed cell location in the chamber. The former one can be estimated by

the histogram of the fluorescence intensity from the calibration experiment with

recombinant proteins. Since a recombinant protein has fixed concentration over the entire

channel, it represents a uniform protein level without any heterogeneity and location

dependence. As a result, the distribution of the fluorescence intensity of a specific

recombinant protein reflects the detection profile of the DNA barcode. In Fig. 2.8 A, histograms are provided for the protein p-mTOR and p-p38α, and the coefficient of variation (CV) of those histograms is ~8%. These histograms are generated from the calibration measurements that utilized the cocktail of recombinant proteins, and are representative of all of the proteins assayed here over a broad range of concentrations (Fig. 2.8). In general, basically the intensities of all the recombinant proteins at detectable concentration follow a Gaussian distribution with CVs typically lower than 10%.

The cell location within a microchamber, relative to the barcode positions, is another factor for the system error. This error arises because of the competition between antibody/protein binding kinetics, and protein diffusion. In order to minimize this effect, we utilized two sets of barcodes in a chamber and used the averaged signal intensity from two barcodes as the final signal value. Thus, the barcode close to the cell will undergo a higher local protein concentration than its compartment during the course of protein diffusion. Since the cells are randomly distributed in the microchambers, this adds an error to the SCBC system. However, we can record the positions of the cells, as well as the fluorescence intensities from the duplicate barcode arrays. So we can investigate whether this error is systematic. If yes, then it can be accounted for, and its contribution to the experiment error subtracted.

For evaluating experimental errors, a Monte Carlo simulation was carried out by R (R Foundation for Statistical Computing, version 2.10.1). Depending on the processes of the protein release, there are two kinetic scenarios that need to be considered. One is for cytoplasmic and membrane proteins which are fully released from the cell location during

the cell lysis. The other one is for secreted proteins that are gradually secreted by the cell during the incubation. For the former one, we examined the fluorescence signals of the two barcodes within microchambers containing single cells. The intensity ratio of the barcode copies versus the location of the cell can be approximated by a linear relationship (Fig. 2.8 B). The closer the cell is to one barcode, the higher is that barcode signal, while the more distant barcode signal is proportionately weaker. For the latter case, we back-calculate the secretion rate of a single cell for each secreted protein based upon the experiment results. With the exact chamber size and shape, as well as the diffusivity and the secretion rate of a specific protein in hand, the simulation should correctly capture the physics of the measurement since the equations for protein diffusion are known and typical antibody/protein binding kinetics can also be modeled.

To investigate the influence of cell locations on signal variation, we assumed, for the simulation, that each single cell releases a fixed amount of specific amount of a cytoplasmic protein. This assumption removes the effect of biological variation and allows us to focus on the experimental error. The parameters used in the simulation match our experimental environment. The chamber is 2000 μm in length and 100 μm in width, with two sets of DNA barcode M-A and A-M from left to right. Excluding the 200 μm wide valves at each microchamber ends, the effective length of the chamber is 1800 μm. Each barcode stripe is 20 μm wide, and are patterned at 50 μm pitch. For the simulation, a single cell is randomly placed in a chamber.

Fig. 2.8 C is the simulated distributions of fluorescence signal per chamber for 5000 single-cell cases with proteins conjugated to ssDNA oligomers L. The maximum protein

intensity, which is assigned to the case for which a cell sits directly on a given barcode stripe, is set as 200. The barcode variability for a given protein was set as 10%, which is a representative value of the experimental determined uncertainty from the calibration data. The blue curves are the Gaussian fitting with the sample average and sample standard deviation obtained from the simulation. The fits indicate that the average signal per chamber follows a Gaussian distribution after including the randomness of cell locations. This is, then the statistical distribution of the measurements for single cells after accounting for the barcode variations and variations arising from cell locations, but with zero cell heterogeneity, which constitutes the system error.

Fig. 2.8 D shows three representative cases generated from the simulation for the ssDNA L. The average protein signal measured from these chambers ranges from 161 to 184, indicating that a similar average signal (limited by the intrinsic barcode variations) is measured, regardless of cell locations.

By matching the concentration close to the actual level in the experiment, the Monte Carlo simulation can yield the system error of our measurement. Consequently, the single cell biological variation can be calculated by the formula below:

$$CV_{total} = \sqrt{CV_{experiment}^2 + CV_{biological}^2}$$

where the CV_{total} refers to the total CV of the experiment data and $CV_{experiment}$ is dominated by the uncertainty of the barcode variation. An estimation of the biological variation is shown in Fig. 2.8 E where it can be noticed that the biological variation is

dominant in the total variation of the assay and the experimental error is around 10% for this fashion of measurement.

2.3.3 Validation of SCBC technology with conventional methods

Fig. 2.9 shows heat map data from SCBC experiments on U87 EGFRvIII PTEN cells, and from measurements on bulk populations of those cells. Individual microchamber data are shown in Fig. 2.5 B. Fig. 2.9 B shows protein assays measured from a population of EGF stimulated U87 EGFRvIII PTEN cells. These assays used similar cell lysis and assay protocols as the SCBC assays. Comparison across Fig. 2.9 A and B reveals that the bulk

Figure 2.9 SCBC and bulk cell measurements of U87 EGFRvIII PTEN cells. (A) Heat maps of SCBC protein level assays. Each column represents 1 microchamber assay, each row represents a protein. (B) Protein assays from a population of U87 EGFRvIII PTEN cells under EGF stimulation. The contrast of these images has been equally adjusted, and the intensity of EGFR is divided by 5 in the heat maps.

Figure 2.10 Averaged responses for all three cell lines to EGF and erlotinib (eb) + EGF exposures. (A, *Left*) Measured protein expression profiles, in fluorescent intensity units, averaged over the 3-cell measurements (n~20). The signal of EGFR/EGFRvIII is divided by 10. Results are shown as mean ± SD. SD represents the combined experimental error and intrinsic biological variation, but is dominated by the biological variation. (*Right*) Western blot analysis of p-EGFR, p-ERK, p-mTOR, p-Src, p-p70S6K, p-GSK3β and p-Akt expression in U87 EGFRvIII and U87 EGFRvIII PTEN cell lines at the basal, EGF stimulation and eb + EGF treatment states. (B) Heat map of relative expression fold changes of proteins, normalized by unperturbed U87 cells (eb = erlotinib). (C, *Left*) Mean fold change of phosphorylation levels of p-ERK and p-mTOR in different cell lines and conditions, relative to unperturbed U87 cells. (E: EGF; e+E: erlotinib + EGF). (*Right*) Western blot analysis of p-ERK and p-mTOR expression in response to EGF stimulation in U87 EGFRvIII and U87 EGFRvIII PTEN cell lines. These cells were cultured in DMEM medium containing 10% FBS for 24 hours, then in serum free medium for 24 hours or (+) erlotinib (10 μM) treatment in serum free medium, followed by stimulation (+) with EGF (20 ng/ml) for 15 min. Cells were lysed and the listed proteins were detected by western blotting.

assays and SCBC measurements are self-consistent. Comparisons between bulk cell assays, SCBC single cell measurements, and literature results[18, 20-23] were also done to detect distinct phosphorylation states of EGFR under the influence of EGF and erlotinib stimulation. Those results again formed a self consistent data set.

Data, such as is shown in Fig. 2.9, was first averaged to recapitulate measurements of proteins from cell populations for comparison with known biology. It was then more fully analyzed to yield a statistical representation of fluctuations at the single cell level.

Fig. 2.10 A presents the protein abundances (averaged over all 3-cell experiments), measured for each cell line and for all conditions (mean intensities and standard deviations are presented in Appendix B: Supplementary Tables). We compared these SCBC results with literature findings that used conventional bulk cell assays, as well as with our own Western blot assays (Fig. 2.10). In the following discussion, literature citations following the protein names provide validation of our SCBC results.

At basal level, U87 cells (Fig. 2.10 A, *top*) showed low EGFR phosphorylation[23] and modest activation of signaling proteins, including p-Src[22], p-mTOR[24], p-p70S6K[23, 25], p-GSK3α/β[23, 26, 27], p-p38α[24], and p-ERK[18, 22, 26], while p-JNK2 was not detected[23]. U87 EGFRvIII cells (Fig. 2.10 A, *middle*) exhibited increased baseline levels of phosphorylation compared with cells expressing wild-type EGFR, including p-Src[22], p-mTOR, p-p70S6K[25], p-ERK[18] and p-JNK2[28]. In U87 EGFRvIII PTEN cells (Fig. 2.10 A, *bottom*), PTEN coexpression diminished baseline phosphorylation of p-Src, p-mTOR, p-p70S6K[25], p-ERK[18] and p-JNK2 compared with U87 EGFRvIII.

EGF stimulation induced EGFR phosphorylation (Fig. 2.13)[18, 27, 28] and promoted downstream pathway activation in all 3 cell lines, irrespective of PTEN status, including activation of p-p70S6K[25] and p-ERK[18]. The increase of levels of p-ERK in response to EGF stimulation in U87 EGFRvIII and U87 EGFRvIII PTEN cells are demonstrated by the Western blots shown in Fig. 2.10. The level of p-GSK3α/β in response to EGF stimulation was increased in U87[27] and U87 EGFRvIII cells, but remained relatively unchanged in U87 EGFRvIII PTEN cells (consistent with Western blots in Fig. 2.10 A).

Erlotinib inhibition + EGF stimulation diminished phosphorylation of both EGFR and EGFRvIII (Fig. 2.10)[18] relative to EGF stimulation. It led to decreased phosphorylation levels in U87, although those levels are higher than in the unstimulated cells. One previously identified example of this effect is p-p70S6K[14, 18]. Erlotinib + EGF showed little impact on U87 EGFRvIII cells, indicating that PTEN loss confers resistance to EGFR tyrosine kinase inhibitors[14, 21]. The phosphoprotein expression levels decrease, but are above the unstimulated levels. Representative proteins include p-Src[22] and p-p70S6K[14, 18, 25]. Erlotinib significantly diminished phosphorylation levels of p-ERK, p-p70S6K[14, 18], p-mTOR and p-Src only for the U87 EGFRvIII PTEN cells. Those phosphorylation levels are below those observed for unperturbed cells; p-p70S6k and p-mTOR drop to below the detection limit. These results are consistent with previous findings that co-expression of EGFRvIII and PTEN protein by GBM cells is associated with clinical response to EGFR kinase inhibitor therapy[14].

Fig. 2.10 B shows the heat map of relative mean-fold changes in the expression levels of proteins and phosphoproteins for the different cell lines and conditions, normalized by the protein levels measured from unperturbed U87 cells. This plot was calculated as follows. For a microchamber i containing n cells, the fluorescence levels recorded from the two barcode assays for a given protein ρ were averaged to yield $\rho_{i,n}$. The fluorescence intensity for ρ, averaged over all 0-cell measurements, was subtracted as background: $(\rho_{i,n} - \bar{\rho}_0)$. This value was then normalized against the background-subtracted, fluorescence levels of ρ averaged over all n cell measurements for unperturbed U87: $(\rho_{i,n} - \bar{\rho}_0) \cdot (\bar{\rho}_{nU87})^{-1}$. These fold-changes were then averaged over all microchambers

containing 2-5 cells, and combined to produce the heat map of Fig. 2.10 B. This map provides a relative comparison of the pathway activation states in different cell lines and conditions, but it also emphasizes that the phosphorylation of p-ERK (representative of MAPK signaling) exhibits correlation with the phosphorylation of mTOR (PI3K signaling).

Recent work suggests cross talk between the RAS/MAPK and PI3K signaling pathways[3, 29]. Recent work has also uncovered a negative regulatory feedback loop by which mTOR complex 1 signaling through S6K1 suppresses PI3K-mediated activation of MAPK activity, so that inhibition of mTOR signaling through S6K1 can activate MAPK[30, 31]. This implied correlation between PI3K and MAPK signaling can be estimated by comparing the phosphorylation levels of ERK and mTOR in varying genetic contexts that regulate PI3K signaling, and in response to ligand stimulation and/or inhibition. The mean fold changes of p-ERK and p-mTOR in U87 EGFRvIII and U87 EGFRvIII PTEN cell lines are shown in Fig. 2.10 C (left). In U87 EGFRvIII cells, the fold change of p-ERK under basal level, EGF stimulation and erlotinib + EGF treatment are statistically lower than that of p-mTOR (Table 2.3). However, in U87 EGFRvIII PTEN cells, the situation is reversed. Obviously, PTEN expression sensitizes GBM cells to MAPK signaling stimulated by EGF. This preferential activation of MAPK signaling pathways in response to EGF activation in GBM cells containing PTEN was validated by immunoblot analysis (Fig. 2.10 C, *right*), and is consistent with recent findings that minimal levels of ERK signaling are required for optimal EGFRvIII-mediated tumor cell growth in PTEN null glioblastomas[15]. These data demonstrate that SCBC measurements can uncover feedback loops and pathway cross talk in situations where the connectivity is less well defined.

2.3.4 Constructing protein-protein correlation networks to reveal the signaling coordination

Figure 2.11 Graphical representation of the information that is uniquely extracted at the single cell level. All data are collected using the SCBC platform. Representative error bars, based upon the measured standard deviations, are provided. (A) *Relative* abundance levels of P53, recorded for 0, 1, 2, and 3 cell experiments. Note that the lowest abundance levels correlate linearly with numbers of cells in the chamber, and that a small fraction of cells with high p53 levels begin to dominate the measured signal in the 3 cell experiments. (B) Abundance histograms of p-ERK levels measured from single U87 EGFRvIII PTEN cells. The single cell data reveal that the erlotinig+EGF stimulated cells exhibit a much more homogeneous behavior than do the unstimulated cells (biological coefficient of variation of 28% versus 57%). (C-D) Characteristic protein-protein correlations. (C) Plot showing a strong correlation between p-p70S6K and p-mTOR in U87 EGFRvIII cells, regardless of the numbers of cells in the microchamber, for 1-5 cell experiments (note the 0-cell baseline). (D) Plot showing the anti-correlation between p-GSK3α/β and p-ERK in U87 EGFRvIII PTEN cells. The listed % values indicate the fraction of cells in a particular quadrant. Percentages in black font include 1 and 2 cell data; grey font values include 1-4 cell data. Note that the 1 and 2 cell experiments exhibit a similarly high level of anti-correlation, but that the anti-correlation begins to be masked by population effects once the 3 and 4 cell data is included in the scatter plot.

Profiles that reveal the relative importance of the measured biological fluctuations versus the experimental errors are shown in Fig. 2.11 A and B. Two points are relevant for comparing bulk cell assays and single cell measurements. Fig. 2.11 A which plots p53 intensity versus experiment number, for the sets of 1, 2, and 3 cell experiments, illustrates how a small fraction of cells can dominate an assay. Fig. 2.11 B provides histograms of the number of p-ERK molecules detected, versus frequency of detection, for single U87 EGFRvIII PTEN cells under all three conditions. Those histograms may be compared against the averaged p-ERK intensities presented at the bottom of Fig. 2.10 A. According to Fig. 2.10 A, the p-ERK level for the unperturbed cells is only slightly higher than for the EGF + erlotinib exposed cells. However, the coefficient of variation of p-ERK levels is much larger (57%) than in the EGF+erlotinib perturbed cells (28%). This effect, which is not captured in bulk assays, may represent an increased amount of regulation for p-ERK in the EGF+erlotinib perturbed cells[32].

The levels of several proteins associated with PI3K signaling should exhibit coordinated behaviors[6]. A typical protein-protein positive correlation (p-mTOR vs p-p70S6K for unstimulated U87 EGFRvIII cells), and an anti-correlation (p- GSK3α/β vs p-ERK for unstimulated U87 EGFRvIII PTEN) are shown in Supplementary Fig. 2.10 C and D. The positive correlation is independent of the numbers of cells per microchamber assay, while the negative correlation begins to be masked for populations as low as 3 cells. Fig. 2.12 provides 9 SCBC-derived protein correlation networks. The line weight defines the strength of the correlation (see key). We used the Bonferroni method[33], which limits correlations to those that exhibit a p-value ≤ 0.05; correlation coefficients above 0.4, or

below -0.4, are significant. Perturbation by ligand stimulation and/or receptor inhibition reveal new relationships, and the genetic context of those relationships. EGF stimulation of EGFRvIII-expressing GBM cells greatly enhances network connectivity in a way that is very different from what would be expected from simply summing the effects of EGF treatment (U87 + EGF, top middle) and EGFRvIII expression (U87 EGFRvIII, middle left). This represents a clinically and biologically relevant result, since wild type EGFR is always present in EGFRvIII expressing cells[14]. The greatly enhanced network interconnectivity for the EGF stimulated U87 EGFRvIII cells may suggest a mechanism underlying the difficulty of inhibiting downstream signaling in EGFRvIII expressing, PTEN null tumor cells, potentially providing one mechanism for their striking tumorigenicity and their established role in promoting therapeutic resistance. This is consistent with the clinical failure and the lack of p70S6K inhibition observed in EGFRvIII-expressing, PTEN deficient GBM patients treated with erlotinib[14], and suggests that clinically relevant insights may potentially be derived from these types of single cell experiments. Classical genetics is also often used to combine perturbations and phenotypic responses to infer functional relationships between genes[34], but specific interactions are difficult to extract because intermediate interacting partners may contribute combinations of positive and/or negative interactions.

Figure 2.12 Protein correlation maps under different genetic and environmental perturbations. All indicated correlations pass a Bonferroni corrected *p*-value test (*p*=0.05). Underlined proteins are below detection limit.
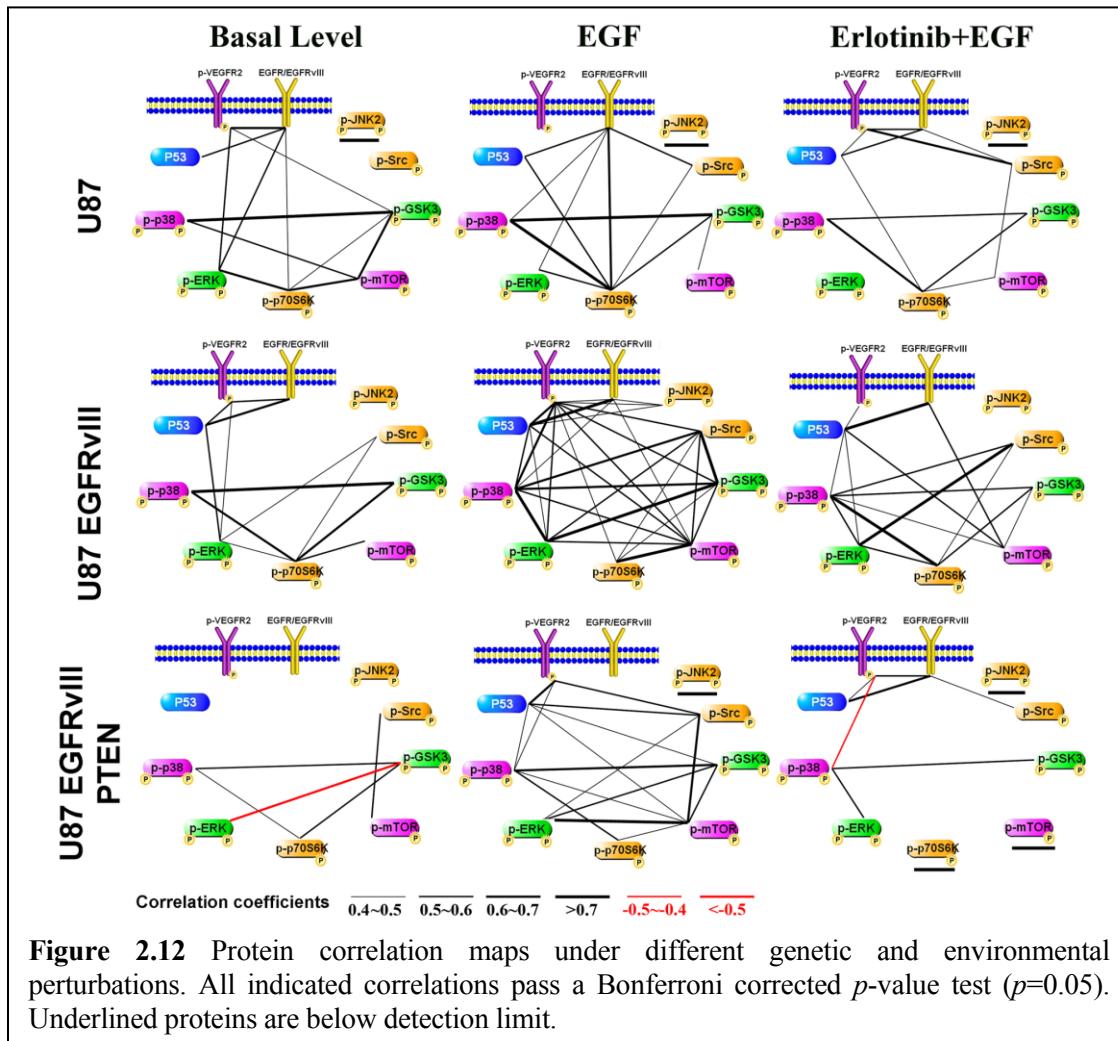
2.4 CONCLUSION

The SCBC provides certain advantages for assaying cytoplasmic proteins. The ability to normalize protein levels to numbers of cells permits for the SCBC data to recapitulate qualitative protein measurements from bulk cell populations, but in a quantitative fashion. One example relates towards interrogating cross talk between the RAS/MAPK and Receptor Tyrosine Kinase (RTK)/PI3K signaling in GBM[3, 29, 30]. Using the SCBC, we found that for U87 EGFRvIII PTEN cells, stimulation with EGF (associated with RTK/PI3K signaling) led to a sharp increase in levels or p-ERK (associated with the RAS/MAPK pathway), a result that was confirmed using Western Blot analysis of the bulk cell lines. Exposure of those same cells to erlotinib + EGF kept the p-ERK levels near the level of unstimulated U87 EGFR vIII PTEN cells.

A second advantage relates to the assessment of the single cell fluctuations, defined by the distribution of the levels of a given protein, measured across many SCBC assays. The measured biological variation that arises from the functional heterogeneity of a genetically identical cell population is significantly higher than the experimental error, and varies across proteins. These fluctuations provide a gauge of the heterogeneity of the cell population, and can be used to predict the thermodynamic stability of specific proteins towards perturbations[32].

The SCBC barcodes could potentially be expanded to 35-40 proteins, depending upon the availability of antibody pairs, but even for just 11 intracellular proteins, the correlation networks extracted from SCBC data already provide interesting parallels with the tumorigenecity and therapeutic resistance of EGFRvIII positive, PTEN null tumors.

Expanding the protein panel will permit a more complete mapping of the connectivity between known GBM signaling pathways, and how that connectivity may be influenced by molecular (i.e., therapeutic) or physical (i.e., hypoxia) perturbations. A further significant challenge will be to extend this platform towards the analysis of clinical specimens, and such work is currently underway.

2.5 REFERENCES

1.   Reynoso, D. & Trent, J.C. Neoadjuvant and adjuvant imatinib treatment in gastrointestinal stromal tumor: current status and recent developments. *Curr Opin Oncol* **22**, 330-5 (2010).

2.   Lemmon, M.A. & Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* **141**, 1117-34 (2010).

3.   She, Q.B. et al. 4E-BP1 is a key effector of the oncogenic activation of the AKT and ERK signaling pathways that integrates their function in tumors. *Cancer Cell* **18**, 39-51 (2010).

4.   Janes, K.A., Reinhardt, H.C. & Yaffe, M.B. Cytokine-induced signaling networks prioritize dynamic range over signal strength. *Cell* **135**, 343-54 (2008).

5.   Marusyk, A. & Polyak, K. Tumor heterogeneity: causes and consequences. *Biochim Biophys Acta* **1805**, 105-17 (2010).

6.   Sachs, K., Perez, O., Pe'er, D., Lauffenburger, D.A. & Nolan, G.P. Causal protein-signaling networks derived from multiparameter single-cell data. *Science* **308**, 523-9 (2005).

7.   Ciaccio, M.F., Wagner, J.P., Chuu, C.P., Lauffenburger, D.A. & Jones, R.B. Systems analysis of EGF receptor signaling dynamics with microwestern arrays. *Nat Methods* **7**, 148-55 (2010).

8.   Du, J. et al. Bead-based profiling of tyrosine kinase phosphorylation identifies SRC as a potential target for glioblastoma therapy. *Nat Biotechnol* **27**, 77-83 (2009).

9. Cheong, R., Wang, C.J. & Levchenko, A. High content cell screening in a microfluidic device. *Mol Cell Proteomics* **8**, 433-42 (2009).

10. Shin, Y.S. et al. Chemistries for patterning robust DNA microbarcodes enable multiplex assays of cytoplasm proteins from single cancer cells. *Chemphyschem* **11**, 3063-9 (2010).

11. Thorsen, T., Maerkl, S.J. & Quake, S.R. Microfluidic large-scale integration. *Science* **298**, 580-4 (2002).

12. Huang, T.T., Sarkaria, S.M., Cloughesy, T.F. & Mischel, P.S. Targeted therapy for malignant glioma patients: lessons learned and the road ahead. *Neurotherapeutics* **6**, 500-12 (2009).

13. Cancer Genome Atlas Research, N. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061-8 (2008).

14. Mellinghoff, I.K. et al. Molecular determinants of the response of glioblastomas to EGFR kinase inhibitors. *N Engl J Med* **353**, 2012-24 (2005).

15. Huang, P.H. et al. Phosphotyrosine signaling analysis of site-specific mutations on EGFRvIII identifies determinants governing glioblastoma cell growth. *Mol Biosyst* **6**, 1227-37 (2010).

16. Furnari, F.B., Lin, H., Huang, H.S. & Cavenee, W.K. Growth suppression of glioma cells by PTEN requires a functional phosphatase catalytic domain. *Proc Natl Acad Sci U S A* **94**, 12479-84 (1997).

17. Bostrom, J. et al. Mutation of the PTEN (MMAC1) tumor suppressor gene in a subset of glioblastomas but not in meningiomas with loss of chromosome arm 10q. *Cancer Res* **58**, 29-33 (1998).

18. Wang, M.Y. et al. Mammalian target of rapamycin inhibition promotes response to epidermal growth factor receptor kinase inhibitors in PTEN-deficient and PTEN-intact glioblastoma cells. *Cancer Res* **66**, 7864-9 (2006).

19. Yuan, T.L., Wulf, G., Burga, L. & Cantley, L.C. Cell-to-cell variability in PI3K protein level regulates PI3K-AKT pathway activity in cell populations. *Curr Biol* **21**, 173-83 (2011).

20. Amos, S., Martin, P.M., Polar, G.A., Parsons, S.J. & Hussaini, I.M. Phorbol 12-myristate 13-acetate induces epidermal growth factor receptor transactivation via protein kinase Cdelta/c-Src pathways in glioblastoma cells. *J Biol Chem* **280**, 7729-38 (2005).

21. Fan, Q.W. et al. A dual phosphoinositide-3-kinase alpha/mTOR inhibitor cooperates with blockade of epidermal growth factor receptor in PTEN-mutant glioma. *Cancer Res* **67**, 7960-5 (2007).

22. Lu, K.V. et al. Fyn and SRC are effectors of oncogenic epidermal growth factor receptor signaling in glioblastoma patients. *Cancer Res* **69**, 6889-98 (2009).

23. Premkumar, D.R., Arnold, B., Jane, E.P. & Pollack, I.F. Synergistic interaction between 17-AAG and phosphatidylinositol 3-kinase inhibition in human malignant glioma cells. *Mol Carcinog* **45**, 47-59 (2006).

24. Aoki, H. et al. Telomere 3' overhang-specific DNA oligonucleotides induce autophagy in malignant glioma cells. *FASEB J* **21**, 2918-30 (2007).

25. Guo, D. et al. The AMPK agonist AICAR inhibits the growth of EGFRvIII-expressing glioblastomas by inhibiting lipogenesis. *Proc Natl Acad Sci U S A* **106**, 12932-7 (2009).

26. Cournoyer, P. & Desrosiers, R.R. Valproic acid enhances protein L-isoaspartyl methyltransferase expression by stimulating extracellular signal-regulated kinase signaling pathway. *Neuropharmacology* **56**, 839-48 (2009).

27. Servidei, T., Riccardi, A., Mozzetti, S., Ferlini, C. & Riccardi, R. Chemoresistant tumor cell lines display altered epidermal growth factor receptor and HER3 signaling and enhanced sensitivity to gefitinib. *Int J Cancer* **123**, 2939-49 (2008).

28. Fromm, J.A., Johnson, S.A. & Johnson, D.L. Epidermal growth factor receptor 1 (EGFR1) and its variant EGFRvIII regulate TATA-binding protein expression through distinct pathways. *Mol Cell Biol* **28**, 6483-95 (2008).

29. Downward, J. Targeting RAS and PI3K in lung cancer. *Nat Med* **14**, 1315-6 (2008).

30. Carracedo, A. et al. Inhibition of mTORC1 leads to MAPK pathway activation through a PI3K-dependent feedback loop in human cancer. *J Clin Invest* **118**, 3065-74 (2008).

31. Kinkade, C.W. et al. Targeting AKT/mTOR and ERK MAPK signaling inhibits hormone-refractory prostate cancer in a preclinical mouse model. *J Clin Invest* **118**, 3051-64 (2008).

32. Shin, Y.S. et al. Protein signaling networks from single cell fluctuations and information theory profiling. *Biophys J* **100**, 2378-86 (2011).

33. Curtin, F. & Schulz, P. Multiple correlations and Bonferroni's correction. *Biol Psychiatry* **44**, 775-7 (1998).

34. Tong, A.H. et al. Global mapping of the yeast genetic interaction network. *Science* **303**, 808-13 (2004).

35. Bailey, R.C., Kwong, G.A., Radu, C.G., Witte, O.N. & Heath, J.R. DNA-encoded antibody libraries: a unified platform for multiplexed cell sorting and detection of genes and proteins. *J Am Chem Soc* **129**, 1959-67 (2007).

36. Huang, B., Wu, H., Kim, S. & Zare, R.N. Coating of poly(dimethylsiloxane) with n-dodecyl-beta-D-maltoside to minimize nonspecific protein adsorption. *Lab Chip* **5**, 1005-7 (2005).

2.6 APPENDIX A: SUPPLEMENTARY METHODS

2.6.1 Synthesis of DNA-Antibody Conjugates

As-received antibodies (Table 2.1) were desalted, buffer exchanged to pH 7.4 PBS and concentrated to 0.5 mg/mL using Zebba protein desalting spin columns (Pierce). Succinimidyl 4-hydrazinonicotinate acetone hydrazone in *N,N*-dimethylformamide (DMF) (SANH, Solulink) was added to the antibodies at variable molar excess of (300:1) of SANH to antibody. Separately, succinimidyl 4-formylbenzoate in DMF (SFB, Solulink) was added at a 16-fold molar excess to 5'aminated 30mer oligomers in PBS. After incubation for 4h at room temperature, excess SANH and SFB were removed and both samples buffered exchanged to pH 6.0 citrate buffer using protein desalting spin columns. A 30-fold excess of derivative DNA was then combined with the antibody and allowed to react overnight at room temperature. Noncoupled DNA was removed using a Pharmacia Superdex 200 gel filtration column (GE) at 0.5 mL/min isocratic flow of PBS. The conjugates were then concentrated to 0.5 mg/mL by Amicon Ultra-4 Centrifugal Filter Unit with Ultracel-10 membrane (Millipore 10kDa) and stored at 4°C. The conjugation yield was determined by Nanodrop (Thermal Scientific). Detailed protocol can be found in the Protein-Oligo Conjugation Kit (Solulink).

Table 2.1 Reagents Used. At top are the Sequences and terminal functionalization of oligonucleotides. All oligonucleotides were synthesized by Integrated DNA Technology (IDT) and purified via high performance liquid chromatography (HPLC). The DNA coding oligomers were pre-tested for orthogonality to ensure that cross-hybridization between non-complementary oligomer strands was negligible (<1% in photon counts).[4] Next are listed the antibodies used for the multiplex protein assays. All antibody pairs except p-EGFR and p-VEGFR2 were purchased from commercial available ELISA kits (R&D systems, DuoSet® Elisa Development Reagents) containing capture antibodies, biotinylated detection antibodies and standard proteins. Capture antibodies bind both phosphoryalted and unphosphorylated proteins. Biotinylated detection

antibodies detect only phosphorylated proteins. p-EGFR antibodies specific for phosphorylation of EGFR at Y1173, Y1068 and Y845 were purchased from R&D systems as capture antibodies and biotin-labeled EGFR was used as detection antibody. p-VEGFR2 (Y1214) capture antibody and biotin-labeled VEGFR2 detection antibody were purchased from Abcam.

Name	DNA Sequence	Melting Point
A	5'- AAA AAA AAA AAA AAT CCT GGA GCT AAG TCC GTA-3'	57.9
A'	5' NH3- AAA AAA AAA ATA CGG ACT TAG CTC CAG GAT-3'	57.2
B	5'-AAA AAA AAA AAA AGC CTC ATT GAA TCA TGC CTA -3'	57.4
B'	5' NH3AAA AAA AAA ATA GGC ATG ATT CAA TGA GGC -3'	55.9
C	5'- AAA AAA AAA AAA AGC ACT CGT CTA CTA TCG CTA -3'	57.6
C'	5' NH3-AAA AAA AAA ATA GCG ATA GTA GAC GAG TGC -3'	56.2
D	5'-AAA AAA AAA AAA AAT GGT CGA GAT GTC AGA GTA -3'	56.5
D'	5' NH3-AAA AAA AAA ATA CTC TGA CAT CTC GAC CAT -3'	55.7
E	5'-AAA AAA AAA AAA AAT GTG AAG TGG CAG TAT CTA -3'	55.7
E'	5' NH3-AAA AAA AAA ATA GAT ACT GCC ACT TCA CAT -3'	54.7
F	5'-AAA AAA AAA AAA AAT CAG GTA AGG TTC ACG GTA -3'	56.9
F'	5' NH3-AAA AAA AAA ATA CCG TGA ACC TTA CCT GAT -3'	56.1
G	5'-AAA AAA AAA AGA GTA GCC TTC CCG AGC ATT-3'	59.3
G'	5' NH3-AAA AAA AAA AAA TGC TCG GGA AGG CTA CTC-3'	58.6
H	5'-AAA AAA AAA AAT TGA CCA AAC TGC GGT GCG-3'	59.9
H'	5' NH3-AAA AAA AAA ACG CAC CGC AGT TTG GTC AAT-3'	60.8
I	5'-AAA AAA AAA ATG CCC TAT TGT TGC GTC GGA-3'	60.1
I'	5' NH3-AAA AAA AAA ATC CGA CGC AAC AAT AGG GCA-3'	60.1
J	5'-AAA AAA AAA ATC TTC TAG TTG TCG AGC AGG-3'	56.5
J'	5' NH3-AAA AAA AAA ACC TGC TCG ACA ACT AGA AGA-3'	57.5
K	5'-AAA AAA AAA ATA ATC TAA TTC TGG TCG CGG-3'	55.4
K'	5' NH3-AAA AAA AAA ACC GCG ACC AGA ATT AGA TTA-3'	56.3
L	5'-AAA AAA AAA AGT GAT TAA GTC TGC TTC GGC-3'	57.2
L'	5' NH3-AAA AAA AAA AGC CGA AGC AGA CTT AAT CAC-3'	57.2
M	5'-Cy3-AAA AAA AAA AGT CGA GGA TTC TGA ACC TGT-3'	57.6
M'	5' NH3-AAA AAA AAA AAC AGG TTC AGA ATC CTC GAC-3'	56.9

DNA label	Antibody	Source
A'	Human p-PDGFRβ (Y751) kit	R&D DYC3096
B'	Human p-Src (Y419) kit	R&D DYC2685

C'	Human p-mTOR (S2448) kit	R&D DYC1665
D'	Human p-p70S6K (T389) kit	R&D DYC896
E'	Human p-GSK3α/β (S21/S9) kit	R&D DYC2630
G'	Human p-p38α (T180/Y182) kit	R&D DYC869
H'	Human p-ERK (T202/Y204) kit	R&D DYC1825
I'	Human p-JNK2 (T183/Y185) kit	R&D DYC2236
K'	Human total EGFR kit	R&D DYC1854
L'	Human total P53 kit	R&D DYC1043
J'	Capture antibody: rabbit anti-human p-VEGFR2 (Y1214)	Abcam ab31480
	Detection antibody: biotin-labeled mouse anti-human VEGFR2	Abcam ab10975
G'	rabbit anti-human EGFR (Y1173)	R&D AF1095
	Detection antibody: biotin-labeled goat anti-human EGFR	R&D BAF231
I'	mouse anti-human EGFR (Y1068)	R&D MAB3570
	Detection antibody: biotin-labeled goat anti-human EGFR	R&D BAF231
K' or L'	rabbit anti-human EGFR (Y845)	R&D AF3394
	Detection antibody: biotin-labeled goat anti-human EGFR	R&D BAF231

2.6.2 Microfluidic flow patterning of DNA barcode microarray

This procedure has been previously described in detail[10], and so only a brief description is provided here. The PDMS elastomer-based microfluidic patterning chips were fabricated via a molding process from a silicon master with photolithographically-defined patterns. The mixture of GE RTV 615 PDMS prepolymer and curing agent (10:1) was stirred, and poured onto the silicon mold, which was pre-treated with trimethyl-chloro-silane (TMCS) vapor to facilitate mold release. The PDMS is then poured on the mold, degassed for 30 min (house vacuum), and then cured at 80°C for 1 h. The solidified PDMS slab was cut from the mold, assess holes drilled and then bonded onto a poly-L-lysine glass slide (VWR). The microfluidic patterning chip contained 13 parallel microchannels patterned

such that they cover a large area (3cm×2cm) of the glass slide for creating the DNA barcode microarray. Each microchannel is approximately 0.5 meters long and 20 micrometers wide.

For creating a DNA barcode array pattern, multiple DNA solutions (one for each barcode stripe), are each diluted in a mixture of DMSO and deionized water (v/v=1:2) with a final DNA concentration of 267 μM. These solutions are each flowed into a specific microfluidic channel. The solution-filled chip was placed in a desiccator to allow solvent (DMSO and water) to evaporate completely through the gas-permeable PDMS, leaving the DNA molecules behind. This evaporation process took 3-5 days to complete. Last, the PDMS elastomer was removed from the glass slide, and the barcode-patterned DNA was fixed to the glass surface by thermal treatment at 80°C for 4 hours. Residue crystals were readily removed by rapidly dipping the slide in deionized water. Each DNA barcode chip needs validation before bonding to bonding the single cell assay chip. A small area near edge was validated to check the DNA loading and uniformity. It was blocked with 1% Bovine Serum Albumin (BSA)/PBS for 1 hr, and then hybridized with fluorescent Cy3-labeled complementary DNA for 1 hr. After washing three times with 1% BSA/PBS and PBS, the slide was dried by nitrogen gun and scanned by Axon Genepix 4400A . Under laser power of 15% and gain of 450, fluorescence intensity above 10,000 was acceptable for cytoplasmic protein detection at the single cell level.

2.6.3 DNA spot microarrays for conjugate validation, cross-reactivity check and other bulk protein measurements.

The DNA-1° antibody conjugate validation, check of cross-reactivity among antibodies and the measurement of EGFR phosphorylation states were performed on DNA spot microarrays printed by Institute of Systems Biology (Seattle, WA) using the same DNA oligos as those in microfluidic DNA barcode patterning[35]. For device assembly, a 12-well PDMS slab was bonded to the glass slide with DNA spot microarrays, as shown in Fig. 2.13. Each well contains repeated 13 DNA spot microarrays. The diameter of each spot is 150 μm, and distance between neighboring spots is 400 μm.

1) Conjugate validation. a) Block each well with 1% BSA for 1 h. b) Add 50 μl of conjugate (5 μg/ml) into the well and incubate at room temperature for 1 h. c) Aspirate each well and wash with 1% BSA, repeating the process two times for a total of 3 washes. d) Add 50 μl of standard protein in PBS and incubate at room temperature for 1 h. Meanwhile, add 50 μl of PBS in another well as the negative control. e) Wash each well with 1% BSA for three times. f) Add 50 μl of biotin-labeled detection antibody and incubate at room temperature for 1 h. g) Wash each well with 1% BSA for three times. h) Add a mixture of Cy5 fluorescent dye-labeled streptavidin and Cy3-labeled reference complementary ssDNA and incubate at room temperature for 1 h. i) Wash each well with 1% BSA for three times and PBS for two times. j) The PDMS device is removed from the glass slide, and is immediately dipped 3 times each in the following solutions in order, 1X PBS, 0.5X PBS, deionized Millipore H_2O, and finally dried with a nitrogen gun. k) The slide was scanned by Axon Genepix 4400A. 5 μm of resolution is selected. Two color channels (the green Cy3 and the red Cy5 channel) are turned on to collect fluorescence signals.

2) Cross-reactivity check of antibodies. The experimental procedure is similar with the conjugate validation. In order to validate the cross-reactivity of antibodies, all conjugates and detection antibodies are added in each well, but only one standard protein is added in each well.

3) Measurement of EGFR phosphorylation states. The experimental procedure is similar with the conjugate validation, except using cell lysate instead of standard proteins. The treated cells are washed with cold PBS to remove residual media. Cell lysis buffer (1X) is then added and the cells are incubated on ice for 10 minutes. The cell extract is then collected and spun at 14,000 g in a microcentrifuge at $4^{o}C$ for 10 minutes, and the supernatant is removed for use. p-EGFR/ssDNA conjugates specific for phosphorylation of EGFR at Y1173, Y1068 and Y845 are used as capture antibodies and biotin-labeled EGFR is used as detection antibody.

Figure 2.13 Spot-based microarray platform utilized for multiplex protein assays from large numbers of cells, and application in monitoring EGFR phosphorylation and validation of conjugate cross-reactivity. The spot-based platform is a variant of ELISA. The sandwich immunoassays are identical to the DEAL-based assays used within the SCBC. (A) Layout of DNA spot microarrays. The inset shows a basic unit consisting of 13 DNA spots (#A to #M, comprised of the same ssDNA oligomers used for the DNA barcodes). Spot diameters are 150 μm. The basic unit is repeated over 6 columns and 50 rows to make a DNA spot array containing 60X50 spots. (B) Photograph of a spot-based chip for multiplexed protein measurement from large numbers of cells, and for antibody cross-reactivity studies. This 12-well PDMS slab is bonded to the glass slide with the pre-printed DNA spot microarrays. Each well has a volume of 50 μl and contains a complete set of 13 DNA spots for simultaneously assaying the entire panel of proteins. (C) Spot-based EGFR phosphorylation measurements in U87, U87 EGFRvIII and U87 EGFRvIII PTEN cells, under the conditions of standard culture, EGF stimulation, and erlotinib (eb) + EGF stimulation. The same numbers of cells were lysed in each condition for comparison. As a control, EGFR standard protein was used to validate the specificity of the phospho-specific antibodies. (D) Cross-reactivity of panel of assayed proteins. A pin-spotted microarray was used for the DEAL-based protein detection. Each row represents a different condition. For each row, the full cocktail of DEAL conjugates (for the 11 assayed proteins) was used, with one standard protein was tested per row. Red spots are signal from the standard proteins and the green spots alignment reference signals from Cy3-labeled DNA sequence #M'.

2.6.4 Step by step protocol of single-cell proteomic assay

The use of an SCBC was as follows:

1) Blocking: All microfluidic channels were first blocked with the blocking buffer (3% w/v Bovine Serum Albumin + 0.1% N-Dodecyl-β-D-maltoside in 1X PBS) for 60 min. N-Dodecyl-β-D-maltoside was reported to minimize non-specific protein adsorption on PDMS[36].

2) Forming capture antibody microarrays: A solution containing all DNA-antibody conjugates was flowed through the assay channels of the SCBCs for 60 min, which transformed the DNA barcode microarrays into antibody microarrays enabling the subsequent surface-bound immunoassay. The unbound conjugates were removed by flowing the washing buffer (3% BSA + 1% phosphatase inhibitor) for 10 min. The DNA-antibody conjugate solution (100 μl) was prepared by mixing all synthesized conjugates in 3% BSA with a final concentration of 10 μg/mL of each conjugate.

3) On-chip cell lysis and intracellular protein measurement: The concentrated cell lysis buffer (4X) was prepared by mixing Cell Lysis Buffer (Cell Signaling, containing 20 mM Tris-HCl, 150 mM NaCl, 1 mM Na_2EDTA, 1 mM EGTA, 1% Triton, 2.5 mM sodium pyrophosphate, 1 mM -glycerophosphate, 1 mM Na_3VO_4, and 1 g/ml leupeptin), Complete Protease Inhibitor Cocktail (Roche) and Phosphatase Inhibitor Cocktail 2 (Sigma). This lysis buffer can efficiently extract nuclear, membrane associated, and cytoplasmic proteins. The lysis buffer was loaded into the lysis buffer chambers (Fig. 2.3) while the valve 4 (Fig. 2.3 A) was kept closed by applying 18-20 psi constant pressure. After cell treatment, the treated cells were dissociated with trypsin and EDTA, centrifuged at $4^{\circ}C$ and

suspended in cold PBS with a concentration of 1000 cell/µL. Then, cells were loaded into the pre-chilled microfluidic chips, and valve 3 (Fig. 2.3 A) were closed to compartmentalize cells, which converted the 8 channels into 120 isolated microchambers. The cells on the chip were recorded by a CCD camera for cell counting at 4°C. Subsequently, the chip was sitting on the ice and the valve 4 was opened for on-chip diffusion of lysis buffer to the neighboring cell chambers (Fig 2.3 D-F). The chip was tilted to accelerate the lysis buffer diffusion. After 20 minutes of lysis buffer diffusion and on-chip cell lysis, the valve 4 was closed and the chip was incubated on ice for 20 more minutes with shaking to complete the on-chip cell lysis. The chip was incubated at room temperature with shaking for 2 hours to allow capture of target proteins by antibody microarrays within the microchambers. Afterwards, the unbound cell lysate was quickly removed by flowing the washing buffer for 10 min.

4) Applying detection antibodies: A mixture of biotin-labeled detection antibodies was flowed into the SCBC for 60 min at room temperature to complete the DEAL assay. The detection antibody solution contained biotinylated detection antibodies at ~3 µg/mL (or specified in the insert of the ELISA kit; the concentration varies from lot to lot) prepared in washing buffer. Then, unbound detection antibodies in the SCBC were removed by flowing the washing buffer for 10 min.

5) Fluorescence probes: Cy5 fluorescent dye-labeled streptavidin (eBioscience, 2µg/mL) and the reference, Cy3-labeled complementary ssDNA (DNA code

M/M', 25 nM), were mixed together and were then flowed into the SCBC

for 60 min. Afterwards, 3% BSA/PBS was flowed for 20 min to remove

unbound Steptavidin-Cy5, and 1X Phosphate Buffered Saline Tween-20

(PBST) was flowed for 40 min as the final wash step.

6) Rinse: The PDMS chip device was removed from the DNA-patterned glass
   slide, and was immediately dipped 3 times each in the following solutions in
   order, PBST, 1X PBS, 0.5X PBS, deionized Millipore H_2O, and finally dried
   with a nitrogen gun.

7) Optical readout: The slide was scanned by Axon Genepix 4400A (Molecular
   Devices). The finest resolution (2.5 μm) was selected. Two color channels
   (the green Cy3 and the red Cy5 channel) were turned on to collect
   fluorescence signals.

2.6.5 Protocols for immunoblot assays

Western blotting was as previously described. Cultured cells in 60 mm dish were

lysed and homogenized using buffer containing 10 mmol/L Tris, pH 7.4, 100 mmol/L

NaCl, 1 mmol/L EDTA, 1 mmol/L EGTA, 1 mmol/L NaF, 20 mmol/L $Na_4P_2O_7$, 2 mmol/L

Na_3VO_4, 0.1% sodium dodecyl sulfate, 0.5% sodium deoxycholate, 1% Triton X-100, 10%

glycerol, 10 μg/mL leupeptin, 60 μg/mL aprotinin, and 1 mmol/L phenylmethanesulfonyl

fluoride. Equal amounts of protein extracts were separated by using 8% or 10% SDS-

PAGE, and then transferred to a polyvinylidene difluoride membrane (Bio-Rad

Laboratories Inc, Hercules, CA). After blocking for 1 hour in a Tris-buffered saline

containing 0.1% Tween 20 and 5% nonfat milk, the membrane was probed with various

primary antibodies, followed by secondary antibodies conjugated to horseradish peroxidase. The immunoreactivity was revealed by use of an ECL kit (Amersham Biosciences Co, Piscataway, NJ). The antibodies were obtained from Cell Signaling (p-Akt Ser473, Akt, p-GSK3β Ser9, GSK3β, PTEN, phospho-EGFR), Sigma (α-actin) and Upstate (EGFR/EGFRvIII cocktail antibody).

2.7 APPENDIX B: SUPPLEMENTARY TABLES

Table 2.2 Mean intensity, standard deviation and coefficient of variation (CV) of standard proteins in calibration experiments. (bottom rows) Parameters utilized for protein assay calibration curves. The calibration curves were well fit by a four-parameter Morgan-Mercer-Flodin (MMF) model, which is a common model for sigmoidal or S-shaped growth. Based on this model, the fluorescence intensity can be translated into concentration (pg/mL). $y = \dfrac{ab + cx^d}{b + x^d}$

	EGFR	p-ERK	p-p38α	p-GSK3α/β	p-p70S6K	p-mTOR	p-Src
50 ng/ml	803.86±15.27	500.16±46.12	204.45±18.92	720.68±22.72	719.11±17.12	339.45±26.08	804.87±14.74
CV	0.02	0.09	0.09	0.03	0.02	0.08	0.02
10 ng/ml	733.32±8.43	191.47±11.93	87.40±6.64	321.67±14.40	283.54±14.35	117.64±6.01	736.16±18.34
CV	0.01	0.06	0.08	0.04	0.05	0.05	0.02
1 ng/ml	269.06±8.83	28.84±3.25	20.80±3.61	61.07±4.62	43.17±3.4	14.35±1.13	243.66±10.37
CV	0.03	0.11	0.17	0.08	0.08	0.08	0.04
100 pg/ml	67.14±4.3	5.86±0.93	9.74±1.44	7.09±0.85	8.02±0.88	5.29±0.59	52.16±4.66
CV	0.06	0.16	0.15	0.12	0.11	0.11	0.09
10 pg/ml	22.01±2.14	4.25±0.65	8.86±0.9	5.71±0.91	6.59±0.98	4.87±0.57	19.58±1.62
CV	0.10	0.15	0.10	0.16	0.15	0.12	0.08
0	12.45±1.5	4.34±0.71	7.2±0.95	5.11±0.69	6.87±0.59	4.43±0.37	8.97±0.86
CV	0.12	0.16	0.13	0.13	0.09	0.08	0.10
Parameters for Protein Assay Calibration Curves							
	a	b	c	d	r		
EGFR	1.230	667.38	1179.21	0.758	0.9999		
p-ERK	3.864	28142.75	884.84	0.970	0.9999		
p-p38α	7.911	11874.57	382.78	0.876	0.9999		
p-GSK3α/β	3.283	7972.09	1234.7	0.861	0.9999		
p-p70S6K	5.808	26921.69	1218.11	0.976	0.9999		
p-mTOR	4.280	101832	562.39	1.103	0.9999		
p-Src	10.618	1161.08	1173.15	0.821	0.9999		

Table 2.3 Mean intensity and standard deviation (std, red font) for proteins assayed from U87, U87 EGFRvIII and U87 EGFRvIII PTEN cells, as a function of # of cells, and under the different conditions. At the bottom (blue font) is presented the mean fold change and standard error of protein and phosphoprotein levels. For calculating the mean fold change values, unperturbed U87 was set as the base line, and mean fold changes of protein levels in the cell lines and conditions were calculated. This mean fold change was calculated as follows: For a microchamber i containing n cells, the fluorescence levels recorded from the two barcode assays for a given protein ρ were averaged to yield $\rho_{i,n}$. The fluorescence intensity for ρ, averaged over all 0-cell measurements, was subtracted as background: $(\rho_{i,n} - \overline{\rho}_0)$. This value was then normalized against the background-

subtracted, fluorescence levels of ρ averaged over all n cell measurements for unperturbed U87:

$$(\rho_{i,n} - \bar{\rho}_0) \cdot (\bar{\rho}_{n,U87})^{-1} - 1.$$ These values for the fold-change in the levels of each assayed protein for each microchamber were then averaged over all microchambers containing 2-5 cells. Standard error of median (SEM) is calculated accordingly. Number of measurements is also listed.

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
0	403.81	2.24	2.29	14.87	2.52	1.20	11.48	3.14	0.65	2.29	1.77	3.04	1.85	**U87**
1	382.03	19.80	384.80	19.94	3.49	3.78	22.47	2.84	16.48	5.42	5.01	7.51	2.65	
2	396.52	28.69	544.72	26.87	3.88	6.11	24.01	2.98	18.19	10.01	7.28	10.38	3.63	
3	423.98	40.12	595.11	28.03	4.03	7.10	27.17	3.06	18.82	12.95	7.70	13.83	3.16	
4	402.42	64.10	664.33	32.98	4.80	10.99	30.06	3.59	18.55	16.21	7.87	16.90	1.93	
5	373.08	92.42	706.23	58.71	5.63	13.76	31.76	3.20	22.43	18.16	10.15	13.87	3.19	

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
0	415.70	1.28	3.61	24.41	2.39	0.83	7.22	2.65	0.73	1.31	1.56	1.24	0.47	**U87 + EGF**
1	420.95	28.64	418.24	46.37	3.23	7.20	31.43	2.59	29.34	31.99	8.08	15.44	2.53	
2	424.75	48.68	578.54	51.67	3.55	10.61	33.95	2.57	31.12	40.95	10.65	25.94	3.55	
3	414.15	73.76	726.12	67.83	4.89	13.63	39.64	2.37	34.41	45.93	16.56	28.78	3.46	
4	413.53	118.79	773.29	110.58	8.35	15.94	44.71	2.94	43.29	50.62	19.06	32.03	2.83	
5	424.47	136.49	780.69	126.49	8.63	21.94	52.65	3.93	56.73	60.23	26.76	35.21	3.19	

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
0	384.05	2.11	3.13	29.50	2.20	1.34	10.18	0.88	0.55	2.76	1.70	2.14	0.80	**U87 + EGF + eb**
1	384.42	18.67	374.74	45.24	5.53	8.64	19.45	3.73	12.02	22.08	10.51	13.21	5.56	
2	381.35	25.94	495.02	51.43	5.32	10.16	21.87	4.02	12.81	26.20	12.08	15.90	5.14	
3	384.71	39.22	566.42	56.64	5.57	11.42	22.58	3.74	13.43	26.16	13.15	18.84	5.29	
4	386.49	54.69	683.23	66.43	5.35	12.32	22.77	3.24	14.44	27.33	15.18	20.87	5.12	
5	380.24	87.64	733.72	79.67	5.16	13.39	23.37	4.16	16.80	35.69	17.91	22.45	6.09	

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
0	30.30	1.32	0.76	5.64	1.57	0.74	2.37	3.57	0.51	0.90	0.22	1.52	1.15	U87 *std*
1	38.66	14.51	123.07	7.29	1.36	1.82	10.41	1.33	5.47	2.65	2.54	7.18	1.40	

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
2	51.31	12.62	96.37	13.87	1.48	2.88	6.75	1.35	5.36	5.19	4.88	9.10	1.72	
3	57.36	20.20	118.85	12.29	1.31	4.22	9.88	1.47	9.00	4.15	4.28	8.34	1.75	
4	35.00	39.63	97.98	16.35	1.00	4.36	6.69	1.58	7.55	11.40	5.14	10.84	1.07	
5	35.53	50.58	101.77	25.57	1.24	9.03	10.58	1.79	6.46	11.33	7.12	6.87	1.92	

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
0	22.30	0.63	0.83	5.19	0.81	0.29	1.40	2.29	0.24	0.68	0.13	0.89	0.24	U87 + EGF *std*
1	42.69	12.96	145.74	15.27	1.41	4.91	13.76	1.48	11.19	9.84	4.43	8.51	1.33	
2	48.00	34.25	116.05	18.47	1.61	3.06	11.74	1.63	9.89	9.85	4.66	12.94	2.42	
3	32.29	28.70	71.26	32.32	2.21	4.13	7.43	1.32	8.52	9.00	10.12	7.05	1.50	
4	18.47	43.52	32.99	16.56	2.87	4.75	15.57	1.27	8.40	11.44	7.10	9.16	2.22	
5	38.37	41.30	33.32	46.91	0.89	15.81	16.73	2.26	21.09	17.43	10.40	7.34	2.54	

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
0	14.50	1.06	0.93	6.78	0.49	0.22	2.86	0.40	0.37	1.09	0.12	0.44	0.29	U87 + EGF + eb *(std)*
1	13.10	5.26	70.17	9.73	0.94	1.35	2.71	1.15	2.40	3.08	2.63	4.51	1.00	
2	10.21	10.18	79.79	11.04	0.76	1.32	2.94	1.17	1.75	3.18	2.29	4.33	0.48	
3	10.50	25.71	94.81	10.24	1.00	1.68	2.37	1.06	2.22	2.63	3.49	4.92	0.41	
4	11.34	25.70	72.01	6.13	1.16	1.41	2.67	0.86	2.48	6.37	2.82	6.69	0.72	
5	11.26	16.17	21.01	7.68	1.72	0.91	3.03	0.92	2.97	5.85	7.69	3.73	1.15	

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
0	378.74	3.05	3.40	26.61	3.03	1.72	13.04	1.29	1.21	3.34	1.68	2.71	1.07	U87 EGF RVIII
1	388.54	21.66	478.11	42.73	9.52	10.69	29.73	4.75	16.71	23.13	26.61	19.04	5.27	
2	386.66	27.32	567.97	52.53	12.53	14.39	30.08	4.94	17.91	26.31	35.86	21.50	5.28	
3	382.84	43.24	688.07	66.40	14.62	20.55	36.38	4.38	21.24	29.40	50.41	24.61	4.89	
4	387.14	57.77	735.43	76.10	15.73	23.22	41.96	5.06	24.66	36.88	60.61	34.79	5.98	
5	373.04	88.48	802.16	94.07	20.86	28.07	52.42	5.44	30.10	45.04	73.66	42.85	5.18	

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
0	397.04	3.03	4.36	24.97	2.82	2.22	13.49	1.01	0.93	3.73	1.81	2.89	1.10	U87 EGF RVIII + EGF
1	413.40	19.25	476.95	59.57	15.50	26.92	47.99	2.66	25.87	39.32	43.89	22.05	5.01	
2	411.70	27.59	579.62	81.44	24.98	30.33	61.07	2.77	29.64	46.46	58.61	28.30	7.42	
3	416.12	43.10	721.14	122.76	29.63	37.57	71.37	3.16	34.59	54.16	75.14	33.90	7.88	
4	421.36	63.40	762.14	144.66	33.76	41.02	79.86	3.37	40.71	59.80	89.14	41.76	8.78	
5	398.36	83.97	801.47	164.69	36.45	49.82	95.95	3.20	51.10	67.18	103.44	50.34	9.11	

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
0	384.74	2.75	3.95	26.24	2.81	1.65	11.84	1.43	1.35	2.69	1.68	3.19	1.05	U87 EGF RVIII
1	399.74	17.94	453.10	61.89	12.45	24.27	29.58	4.05	20.64	26.95	39.78	15.23	1.77	
2	400.97	30.28	567.72	64.22	12.49	31.19	36.65	4.17	27.59	36.97	50.72	22.34	3.08	

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
3	411.64	42.38	673.86	78.75	18.90	36.67	43.97	4.60	35.59	42.62	61.40	23.60	3.56	+ EGF + eb
4	414.72	67.17	724.37	85.05	13.69	43.11	53.92	5.08	38.90	56.30	73.36	33.72	5.49	
5	414.24	87.63	783.60	106.46	20.06	47.88	63.99	5.14	47.02	60.55	85.44	38.61	4.64	

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
0	16.35	1.47	0.78	7.77	0.90	0.70	4.44	0.90	0.42	2.28	0.25	0.59	0.69	U87 EGFRVIII *(std)*
1	13.32	5.60	89.41	6.61	3.59	3.10	4.60	1.19	3.06	3.01	6.96	4.17	2.06	
2	9.22	7.64	83.71	10.94	3.51	4.71	7.00	2.07	2.99	4.50	8.36	4.64	1.78	
3	16.46	9.58	36.77	12.41	3.27	4.57	6.39	1.48	3.66	4.66	7.60	4.85	1.55	
4	19.57	6.24	54.94	17.45	6.38	5.43	5.20	1.84	4.14	5.16	8.15	7.97	2.17	
5	12.94	16.61	13.30	14.15	5.20	3.75	9.95	1.67	3.26	5.58	8.89	12.37	1.09	

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
0	21.09	1.35	0.78	6.59	0.73	0.88	3.19	0.30	0.43	1.86	0.11	1.17	0.70	U87 EGFRVIII + EGF *(std)*
1	33.42	5.13	84.21	17.17	6.39	6.09	14.42	1.57	7.54	10.47	10.76	9.44	2.61	
2	24.88	5.66	53.66	21.92	9.04	7.21	19.65	1.66	10.65	17.12	15.11	17.02	5.72	
3	25.84	6.22	46.98	50.42	7.44	8.05	20.51	1.87	9.30	14.76	20.83	18.29	6.57	
4	28.09	8.78	42.57	46.16	12.30	8.45	26.53	2.33	15.34	13.77	21.48	18.16	6.23	
5	26.02	20.46	68.48	56.14	5.66	7.59	29.04	1.37	18.14	19.27	37.54	28.27	10.53	

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
0	20.46	0.81	1.16	8.90	0.68	0.75	2.93	0.73	0.59	1.34	0.15	0.74	0.36	U87 EGFRVIII + EGF + eb *(std)*
1	11.74	5.03	49.82	15.18	8.12	7.17	7.58	2.80	7.78	7.34	6.76	7.33	1.52	
2	13.41	7.39	56.69	8.32	9.91	5.18	9.58	2.55	9.48	10.18	11.16	7.99	2.90	
3	21.42	9.16	41.18	18.06	10.82	12.24	15.43	3.07	11.44	9.91	15.22	8.22	2.28	
4	19.28	6.26	37.52	20.58	7.75	16.65	21.53	3.57	14.95	21.57	21.22	13.80	3.44	
5	18.90	16.79	50.15	36.58	5.99	14.13	15.25	3.25	13.60	21.34	19.09	13.68	5.17	

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
0	384.61	3.25	2.88	20.10	3.44	1.10	12.88	0.87	0.96	1.41	1.73	3.47	0.36	U87 EGFRVIII PTEN
1	381.18	27.33	283.56	37.96	5.83	7.53	30.75	2.89	27.92	7.15	7.32	11.83	4.41	
2	369.80	36.31	363.47	42.77	5.58	8.36	34.88	3.26	30.67	7.33	7.55	12.98	4.03	
3	368.64	56.97	466.70	46.03	6.21	10.66	35.75	2.73	31.46	7.52	9.04	14.48	3.82	
4	383.82	68.22	545.53	52.54	6.74	12.67	39.44	2.71	32.61	9.10	10.76	14.48	4.27	
5	382.19	75.71	574.20	54.77	6.96	13.37	41.22	2.88	33.99	9.13	11.93	17.07	3.13	

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
0	385.52	2.86	3.87	26.30	2.72	1.53	9.63	0.80	0.72	2.58	1.74	1.90	0.58	U87 EGFRVIII PTEN + EGF
1	412.33	25.18	262.04	54.41	2.71	24.29	32.12	2.01	23.84	19.58	15.64	23.02	2.24	
2	400.03	34.14	360.20	69.42	3.84	38.11	39.96	1.63	28.92	22.94	24.77	29.28	1.79	
3	395.77	49.89	472.01	73.00	4.74	41.88	43.11	2.43	30.30	25.24	27.60	31.28	1.90	
4	398.65	66.69	596.30	81.81	4.69	43.17	44.88	2.59	31.01	27.00	27.42	32.67	2.31	
5	390.22	90.98	623.85	85.15	6.44	46.63	47.75	2.39	32.39	31.21	30.47	33.48	2.39	

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
0	369.50	1.75	4.97	23.63	2.50	1.05	8.19	2.84	1.19	2.43	1.62	1.85	0.85	U87 EGFRVIII PTEN + EGF + eb
1	372.30	13.11	326.30	48.22	2.92	3.61	16.24	2.22	16.72	3.06	2.76	8.55	1.81	
2	371.09	27.80	415.36	57.37	3.17	5.88	17.46	2.82	19.43	3.31	3.88	10.83	2.75	
3	372.73	50.64	494.14	65.81	3.17	7.02	17.99	2.88	21.24	3.94	3.52	12.55	1.94	
4	363.36	61.77	573.10	66.83	4.26	8.84	20.16	3.19	22.26	4.11	3.52	13.05	2.78	
5	365.16	91.42	641.41	73.24	5.86	14.10	22.46	3.92	25.08	4.76	4.67	16.52	3.45	

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
0	20.64	1.09	0.80	2.70	1.19	0.51	2.29	0.52	0.89	0.71	0.14	0.72	0.25	U87 EGFRVIII PTEN *(std)*
1	47.40	11.84	61.25	13.69	2.92	6.38	9.14	1.54	11.50	2.99	3.31	5.43	2.90	
2	45.83	14.26	135.31	15.15	2.23	6.85	12.06	1.78	15.94	3.09	3.91	7.61	1.94	
3	55.28	21.83	108.94	16.52	1.62	5.25	12.18	1.70	9.85	3.30	3.84	6.17	2.89	
4	52.66	26.28	103.55	10.65	2.51	9.18	13.24	1.43	10.44	4.45	6.83	5.92	2.34	
5	49.54	22.53	137.77	14.42	2.32	7.13	11.18	1.64	12.87	3.19	4.32	6.29	2.60	

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
0	21.54	0.90	1.79	5.88	1.03	1.02	1.72	0.51	0.42	1.29	0.14	0.95	0.33	U87 EGFRVIII PTEN + EGF *(std)*
1	16.96	5.31	54.66	13.53	1.43	9.91	10.23	1.19	7.98	3.09	5.83	5.21	1.50	
2	23.34	7.01	106.16	11.03	1.33	11.46	5.57	1.01	5.73	2.98	8.03	8.10	1.26	
3	16.90	16.34	77.90	12.78	2.22	11.49	5.16	1.65	7.67	3.63	7.30	6.68	1.05	
4	13.22	12.60	77.81	13.57	0.96	8.72	7.80	2.00	7.60	4.86	8.45	8.13	1.48	
5	23.63	19.13	39.10	16.58	1.66	12.37	3.93	1.53	6.16	5.53	6.59	8.37	1.25	

cell	ref	P53	EGFR	p-VEGFR	p-JNK2	p-ERK	p-p38	control	p-GSK3	p-p70S6K	p-mTOR	p-Src	p-PDGFR	
0	25.19	0.94	3.42	6.46	0.77	0.47	2.23	2.14	1.54	0.77	0.20	0.81	0.53	U87 EGFRVIII PTEN + EGF + eb *(std)*
1	23.68	7.16	59.41	11.59	1.16	2.06	4.56	0.93	3.19	1.66	1.07	3.88	1.07	
2	33.84	13.34	73.28	16.67	1.30	4.60	5.27	1.05	3.71	0.87	2.28	4.04	1.53	
3	23.88	10.81	62.81	15.33	1.25	5.13	6.02	0.67	3.60	1.76	1.04	3.05	1.02	
4	32.82	14.57	65.15	12.98	1.48	4.58	6.64	1.01	5.87	1.67	1.13	4.61	1.70	
5	33.56	22.56	44.35	12.00	2.16	9.65	6.41	1.01	5.02	1.44	1.59	6.96	1.62	

Mean Fold Change of Protein & Phosphoprotein Levels												
Mean fold change	Reference	P53	EGFR	p-VEGFR	**p-ERK**	p-p38	p-GSK3	p-p70S6K	**p-mTOR**	p-Src	# of Measurements
U87 (EGF)	0.06	0.78	0.12	1.94	**0.89**	1.12	1.05	3.34	**1.35**	1.93	42
U87 (Erlotinib+EGF)	-0.05	-0.07	-0.04	0.89	**0.55**	-0.20	-0.28	1.38	**0.97**	0.65	52
U87 EGFRvIII	-0.04	-0.05	0.10	1.32	**1.56**	0.54	0.12	1.70	**6.69**	1.63	55
U87 EGFRvIII (EGF)	0.03	-0.03	0.13	4.60	**4.09**	2.76	0.93	3.74	**11.05**	2.31	80

U87 EGFRvIII (Erlotinib+EGF)	0.02	0.02	0.09	2.21	**4.31**	1.15	0.83	2.97	**8.90**	1.43	74
U87 EGFRvIII PTEN	-0.04	0.08	-0.24	0.48	**0.27**	0.53	0.64	-0.40	**0.21**	0.16	83
U87 EGFRvIII PTEN (EGF)	-0.01	0.15	-0.24	2.22	**5.28**	1.19	0.61	1.25	**3.15**	2.10	76
U87 EGFRvIII PTEN (Erlotinib+EGF)	-0.07	0.04	-0.19	1.53	**-0.04**	-0.31	0.09	-0.87	**-0.63**	0.12	63

Standard Error of Protein & Phosphoprotein Levels

Standard Error of Median (SEM)	Reference	P53	EGFR	p-VEGFR	**p-ERK**	p-p38	p-GSK3	p-p70S6K	**p-mTOR**	p-Src
U87 (EGF)	0.03	0.21	0.04	0.59	**0.23**	0.22	0.16	0.48	**0.38**	0.53
U87 (Erlotinib+EGF)	0.02	0.10	0.03	0.32	**0.15**	0.07	0.04	0.22	**0.23**	0.24
U87 EGFRvIII	0.02	0.07	0.03	0.38	**0.25**	0.14	0.06	0.25	**0.88**	0.36
U87 EGFRvIII (EGF)	0.02	0.06	0.02	0.78	**0.40**	0.29	0.12	0.40	**1.17**	0.47
U87 EGFRvIII (Erlotinib+EGF)	0.02	0.07	0.02	0.44	**0.46**	0.18	0.12	0.33	**0.97**	0.31
U87 EGFRvIII PTEN	0.02	0.09	0.03	0.24	**0.16**	0.13	0.10	0.06	**0.15**	0.17
U87 EGFRvIII PTEN (EGF)	0.02	0.08	0.03	0.45	**0.55**	0.15	0.08	0.17	**0.42**	0.41
U87 EGFRvIII PTEN (Erlotinib+EGF)	0.02	0.09	0.04	1.16	**0.20**	0.10	0.06	0.04	**0.08**	0.31

Table 2.4 Correlation coefficients of cytoplasmic proteins measured by SCBC in all cell lines and conditions. Correlation coefficients were calculated based on 1-2 cells. The cut-off for correlation and anti-correlation is 0.4 and -0.4, respectively, and red font is used to indicate the above-threshold correlations, which was drawn in the correlation maps (shown in red in the following tables). The number of measurements is shown after the table. The proteins shown in blue are below the detection limit.

	p-JNK2	p53	EGFR	p-VEGFR	p-ERK	p-p38	p-GSK3	p-p70S6K	p-mTOR	p-Src	U87 (63 measurements)
p-JNK2	1.00	0.05	0.24	0.14	0.17	0.08	0.09	0.16	0.14	0.41	
p53		1.00	0.50	0.34	0.16	0.05	0.02	0.04	-0.07	0.10	
EGFR			1.00	0.65	0.57	0.20	0.27	0.44	0.14	0.07	
p-VEGFR				1.00	0.53	0.35	0.49	0.39	0.25	-0.18	
p-ERK					1.00	0.32	0.40	0.70	0.37	0.03	
p-p38						1.00	0.72	0.32	0.56	0.00	
p-GSK3							1.00	0.44	0.61	-0.19	
p-p70S6K								1.00	0.61	0.10	
p-mTOR									1.00	-0.01	
p-Src										1.00	

	p-JNK2	p53	EGFR	p-VEGFR	p-ERK	p-p38	p-GSK3	p-p70S6K	p-mTOR	p-Src	U87 + EGF (54 measurements)
p-JNK2	1.00	0.16	0.26	0.16	-0.09	0.23	0.24	0.02	0.11	0.19	
p53		1.00	0.50	0.05	0.39	0.37	0.24	0.51	0.23	0.31	
EGFR			1.00	-0.12	0.44	0.46	0.20	0.61	0.03	0.51	
p-VEGFR				1.00	-0.04	-0.13	0.04	-0.11	0.32	0.03	
p-ERK					1.00	0.31	0.12	0.52	0.12	0.40	
p-p38						1.00	0.80	0.71	0.39	0.32	
p-GSK3							1.00	0.54	0.46	0.27	
p-p70S6K								1.00	0.35	0.47	
p-mTOR									1.00	0.34	
p-Src										1.00	

	p-JNK2	p53	EGFR	p-VEGFR	p-ERK	p-p38	p-GSK3	p-p70S6K	p-mTOR	p-Src	U87 + EGF + eb (62 measurements)
p-JNK2	1.00	0.15	0.28	0.36	0.29	0.00	-0.11	-0.04	0.20	0.46	
p53		1.00	0.58	0.41	0.37	0.33	0.10	0.42	0.27	0.30	
EGFR			1.00	0.60	0.34	0.22	0.06	0.27	0.21	0.50	
p-VEGFR				1.00	0.36	0.16	-0.13	0.22	0.36	0.66	
p-ERK					1.00	0.27	0.07	0.35	0.29	0.34	
p-p38						1.00	0.59	0.62	0.20	0.13	
p-GSK3							1.00	0.51	0.16	-0.06	
p-p70S6K								1.00	0.46	0.22	
p-mTOR									1.00	0.47	
p-Src										1.00	

	p-JNK2	p53	EGFR	p-VEGFR	p-ERK	p-p38	p-GSK3	p-p70S6K	p-mTOR	p-Src	U87 EGFRvIII (68 measurements)
p-JNK2	1.00	0.10	0.08	0.18	0.33	0.17	0.16	0.22	0.39	0.23	
p53		1.00	0.65	0.51	0.50	0.09	0.19	0.30	0.19	0.35	
EGFR			1.00	0.55	0.39	-0.04	0.06	0.21	0.18	0.24	
p-VEGFR				1.00	0.41	0.15	0.30	0.26	0.39	0.37	
p-ERK					1.00	0.17	0.27	0.43	0.40	0.41	
p-p38						1.00	0.74	0.66	0.18	0.32	
p-GSK3							1.00	0.64	0.31	0.32	
p-p70S6K								1.00	0.52	0.44	
p-mTOR									1.00	0.36	
p-Src										1.00	
	p-JNK2	p53	EGFR	p-VEGFR	p-ERK	p-p38	p-GSK3	p-p70S6K	p-mTOR	p-Src	U87 EGFRvIII + EGF (72 measurements)
p-JNK2	1.00	0.46	0.42	0.48	0.20	0.32	0.31	0.18	0.23	0.19	
p53		1.00	0.71	0.68	0.43	0.57	0.28	0.40	0.60	0.35	
EGFR			1.00	0.47	0.34	0.41	0.12	0.11	0.46	0.14	
p-VEGFR				1.00	0.58	0.66	0.56	0.47	0.57	0.54	
p-ERK					1.00	0.73	0.74	0.36	0.54	0.59	
p-p38						1.00	0.61	0.34	0.60	0.68	
p-GSK3							1.00	0.49	0.52	0.71	
p-p70S6K								1.00	0.71	0.41	
p-mTOR									1.00	0.51	
p-Src										1.00	
	p-JNK2	p53	EGFR	p-VEGFR	p-ERK	p-p38	p-GSK3	p-p70S6K	p-mTOR	p-Src	U87 EGFRvIII + EGF + eb (58 measurements)
p-JNK2	1.00	-0.03	0.08	0.02	-0.22	0.04	0.20	-0.20	0.14	-0.09	
p53		1.00	0.72	0.43	0.47	0.44	0.30	0.40	0.53	0.36	
EGFR			1.00	0.17	0.29	0.32	0.24	0.27	0.57	0.31	
p-VEGFR				1.00	0.33	0.16	0.22	0.11	0.21	0.22	
p-ERK					1.00	0.59	0.28	0.59	0.31	0.77	
p-p38						1.00	0.56	0.77	0.42	0.50	
p-GSK3							1.00	0.59	0.45	0.14	
p-p70S6K								1.00	0.30	0.46	
p-mTOR									1.00	0.39	
p-Src										1.00	
	p-JNK2	p53	EGFR	p-VEGFR	p-ERK	p-p38	p-GSK3	p-p70S6K	p-mTOR	p-Src	U87 EGFRvIII PTEN (72 measurements)
p-JNK2	1.00	0.19	0.03	0.22	-0.18	-0.12	0.07	0.01	-0.16	-0.13	
p53		1.00	0.31	0.31	0.16	-0.01	-0.08	-0.18	0.22	-0.01	
EGFR			1.00	0.09	-0.06	-0.03	0.13	-0.10	0.15	-0.05	

	p-JNK2	p53	EGFR	p-VEGFR	p-ERK	p-p38	p-GSK3	p-p70S6K	p-mTOR	p-Src	
p-VEGFR				1.00	0.01	0.00	0.28	0.28	-0.07	-0.33	
p-ERK					1.00	-0.13	-0.51	-0.08	0.11	0.04	
p-p38						1.00	0.54	0.42	0.15	0.26	
p-GSK3							1.00	0.55	0.17	0.14	
p-p70S6K								1.00	0.26	0.12	
p-mTOR									1.00	0.54	
p-Src										1.00	
	p-JNK2	p53	EGFR	p-VEGFR	p-ERK	p-p38	p-GSK3	p-p70S6K	p-mTOR	p-Src	U87 EGFRvIII PTEN + EGF (82 measurements)
p-JNK2	1.00	0.34	0.16	0.27	0.30	0.36	0.39	0.23	0.34	0.29	
p53		1.00	0.32	0.61	0.32	0.44	0.44	0.29	0.46	0.57	
EGFR			1.00	0.33	0.26	0.25	0.03	0.18	0.28	0.13	
p-VEGFR				1.00	0.14	0.42	0.29	0.34	0.30	0.57	
p-ERK					1.00	0.36	0.53	0.29	0.75	0.48	
p-p38						1.00	0.67	0.52	0.48	0.31	
p-GSK3							1.00	0.38	0.54	0.28	
p-p70S6K								1.00	0.41	0.30	
p-mTOR									1.00	0.65	
p-Src										1.00	
	p-JNK2	p53	EGFR	p-VEGFR	p-ERK	p-p38	p-GSK3	p-p70S6K	p-mTOR	p-Src	U87 EGFRvIII PTEN + EGF + eb (73 measurements)
p-JNK2	1.00	0.24	0.10	-0.02	0.01	0.29	0.24	0.10	0.07	0.04	
p53		1.00	0.62	0.46	0.27	0.07	0.15	0.09	0.15	0.33	
EGFR			1.00	0.54	-0.03	-0.14	0.11	-0.11	0.06	0.49	
p-VEGFR				1.00	-0.32	-0.47	-0.08	-0.15	-0.04	0.33	
p-ERK					1.00	0.52	0.25	0.32	0.43	-0.02	
p-p38						1.00	0.50	0.39	0.40	-0.10	
p-GSK3							1.00	0.21	0.40	-0.13	
p-p70S6K								1.00	0.22	-0.04	
p-mTOR									1.00	0.01	
p-Src										1.00	

C h a p t e r   3

Applications in fundamental cancer biology: hypoxia induces a phase transition within a kinase signaling network in cancer cells

3.1 INTRODUCTION

In most solid organ cancers, increased interstitial pressure, vascular constriction, abnormal leaky blood vessels, and edema result in a hypoxic microenvironment, particularly in the center of the tumor[1-5]. Hypoxia, in part by stabilizing the hypoxia inducible transcription factor (HIF), can increase the biological aggressiveness of tumors, promoting glycolysis, cellular proliferation, and angiogenesis. It can also make tumors less responsive to many therapies[6-9].

Signaling through mTOR is often a critical component of the hypoxic response[10-13]. Amplification and activating mutations of receptor tyrosine kinases, mutation of phosphoinositide 3-kinase (PI3K) and its regulatory subunits, and loss of the phosphatase and tensin homolog (PTEN) tumor suppressor protein can lead to elevated growth factor–independent activation of mTOR signaling[10, 14]. The hypoxic microenvironment indirectly regulates mTOR, in part by regulating intracellular ATP levels[15], to promote tumor cell growth and proliferation. This can occur via activation of hypoxia-inducible factor-1α (HIF-1α) dependent glycolysis, and by stimulating angiogenesis[16]. Most models of mTOR signaling in cancer assume a continuous relationship between the level of growth factor receptor pathway signaling, and/or ATP and nutrient levels, and the degree of mTORC1 activation. However, most signaling cascades actually behave as excitable devices with

built in excitability thresholds, enabling them to integrate diverse temporal and spatial inputs to produce specific signaling responses[17]. It is not known how physical perturbations like altering pO_2 can influence the excitability of signaling networks, and whether such effects yield continuous or discrete transitions. This question is important because if mTOR signaling becomes uninhibitable at levels of hypoxia that are frequently reached within the center of a tumor, a potentially targetable mechanism of drug resistance can be identified.

We set out to study how varying pO_2 from 21% (ambient) to 1% (hypoxia) influences mTORC1 and HIF-1α signaling within model GBM cancer cells that exhibit persistent mTORC1 activation[18, 19]. We used the Single Cell Barcode Chip (SCBC)[20, 21] to investigate U87 EGFRvIII cells (which are GBM cells that stably express the epidermal growth factor receptor activating mutation (EGFRvIII)). The SCBC is an integrated microfluidics platform[22] designed for the quantification of a panel of functional proteins from statistical numbers of single cells[21]. The panel, which was designed to capture key aspects of both HIF-1α and mTORC1 signaling[9, 12], included 3 secreted proteins (vascular endothelial growth factor (VEGF), Interleukin-6 (IL-6) and Matrix Metalloprotease-1 (MMP1)), and 1 cytoplasmic protein (HIF-1α) and 3 cytoplasmic phosphoproteins (phospho(p)-mTOR, p-extracellular signal-regulated kinase 1 (p-ERK1), p-P70 ribosomal protein S6 kinase (p-P70S6K)). Advantages of these proteins is the availability of high quality antibody pairs for our assays, and the fact that they are produced by single cells at a level that allows us to accurately convert single cell fluorescence signal into copy numbers per cell detected.

An SCBC cell data set, which is comprised of a statistical number of single cell assays, yields three types of independent observables. The first are averaged levels of each assayed

protein from single cells. The second are the protein fluctuations, which are histograms

of the observation frequency versus the measured protein levels. The third observations are

the correlations between the various assayed proteins. The last two observables are unique

to single cell multiplex proteomics assays, and all three observation types are employed to

understand how changes in pO_2 influence mTORC1 and HIF-1α signaling. We provide the

interpretation in three stages of increasing level of detail, where the last stage is a theory

with predictive capabilities. We first discuss a mean field, qualitative model that provides a

context for discussing how the average effect of other proteins influences the fluctuations

of a specific protein in question. The experimentally measured fluctuations, when

interpreted within this model, point towards a pO_2-dependent deregulation of mTORC1

signaling, and imply that mTORC1 signaling will be difficult to inhibit near 1.5% pO_2.

This picture is shown to be correct through the use of the ATP-competitive mTOR inhibitor

PP242[23] on both the GBM cell lines, as well as a neurosphere culture model grown from a

human-derived GBM xenograft tumor that also expresses the EGFRvIII mutation. We then

attempt to understand the pO_2-dependent deregulation of mTORC1 in two more detailed

ways. We first present a steady state kinetic model to capture the relationships between O_2,

p-mTOR, HIF-1α, and PP242. The kinetic approach indicates that there is a switch in

mTORC1 signaling near 1.5% pO_2, and that there is a value of pO_2 near 1.5% for which

mTOR is un-inhibitable. Finally we discuss a quantitative version of the Le Chatelier's

principle that relies upon the single cell proteomics assays as input[24], and, unlike the mean

field model, allows for the explicit treatment of protein-protein correlations. The theory is

validated by using it to predict the effect of changes of pO_2 on the mean numbers of the

assayed proteins. This prediction fails between 2 and 1.5% pO_2, which implies that changing pO_2 through this range is a strong perturbation to the cells. The theory then shows that the deregulation of mTORC1 signaling is associated with a phase transition in the signaling network. The implication is that, near 1.5% pO_2, the network switches from one set of protein-protein interactions to another. At the switching point, the network is unstable, and the coordinated signaling between mTOR and its effector proteins is lost.

3.2 EXPERIMENTAL METHODS

3.2.1 Microchip design and fabrication

The SCBC platform (Fig. 3.1) contains 240, 1.7 nanoliter volume microchambers. Each microchamber has an upper assaying compartment that contains a 9-element DNA barcode. A second compartment, separated by a valve, serves as a lysis buffer reservoir. Eight elements of the barcode are converted to a miniature antibody array for assaying a panel of proteins by loading a DNA-antibody conjugate cocktail, while one element provides an alignment marker (Fig. 3.2).

Figure 3.1 The design the of single cell proteomic chip. The SCBC microfluidic chip is designed for enabling on-chip cell culture, cell lysis and multiplex protein detection. V1-V6 represents inputs for activating specific valves. V3 isolated the chamber from the microchannels and V4 controls the diffusion between the cell compartments and the lysis buffer reservoirs. The cartoon illustrates the immunoassays for detecting functional proteins from cell lysate. The optical images show a cell viability check by trypan blue diffusion.

DNA barcode arrays for the fabrication of SCBCs are flow patterned using molded polydimethylsiloxane (PDMS) microfluidics templates. This procedure has been previously described in Chapter 2. The PDMS microfluidic chip for the single cell assay was fabricated by two-layer soft lithography (Fig. 3.1). The fabrication of these chips has been previously described in Chapter 2, and with a few specific differences, as described here. The channel surface of the as-fabricated PDMS chip was coated with collagen type 1(BD Biosciences, 0.1mg/ml in DI water) before thermally bonded to the DNA barcode slide to form the working device. The collagen coating promoted cell adherence during the on-chip cell culture.

3.2.2 Experiment setups and procedures

A custom-build hypoxia setup is used for providing control oxygen environment with real time pO_2 monitoring. Cells are loaded from an upstream inlet into the SCBC and distributed randomly among the microchambers. Controlling the cell loading density enables about half of the microchambers to contain a single cell while others may be empty or contain 2 or more cells. After cell loading and counting, the microchip is incubated in a controlled O_2 environment for 7 hours (Fig. 3.2 A), followed by an on-chip cell lysis (Fig 3.2 B). An O_2 sensor (0.1% accuracy) measured both the level and the equilibration rate of the pO_2. Secreted proteins are captured during incubation, and intracellular proteins are captured following lysis. A detection antibody cocktail and the fluorescent probes are loaded afterward to complete the on-chip immunoassay, which is read with a Genepix array scanner (See Fig. 3.2 B for detailed execution scheme) The incubation time was chosen to ensure cell viability at all pO_2 explored (Fig. 3.1), and to enable capture of sufficient

89

numbers of secreted proteins. Refer to Chapter 2 for a step by step procedure of the

single cell proteomic assay.

3.2.3 Cell lines and reagents
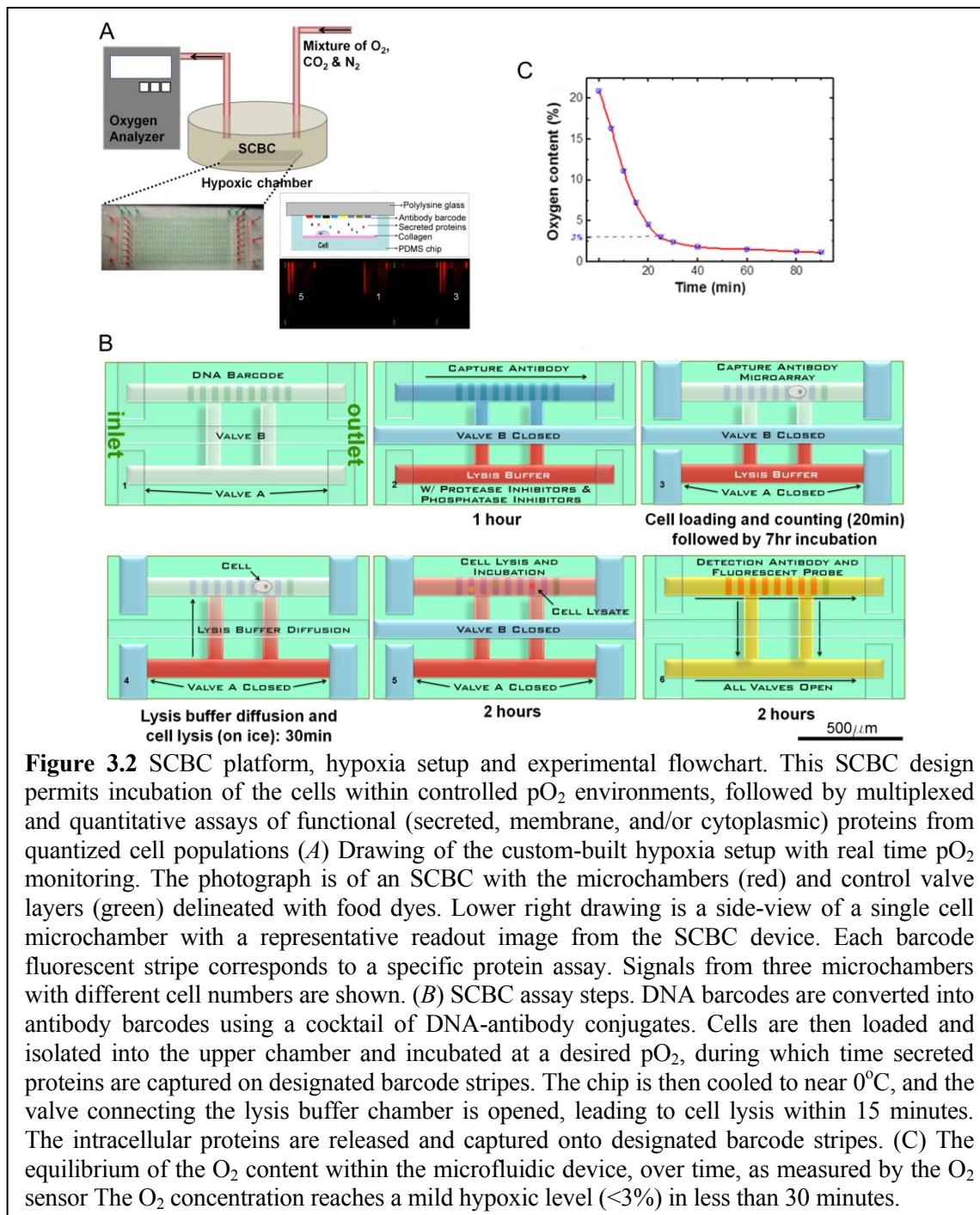
Figure 3.2 SCBC platform, hypoxia setup and experimental flowchart. This SCBC design permits incubation of the cells within controlled pO_2 environments, followed by multiplexed and quantitative assays of functional (secreted, membrane, and/or cytoplasmic) proteins from quantized cell populations (*A*) Drawing of the custom-built hypoxia setup with real time pO_2 monitoring. The photograph is of an SCBC with the microchambers (red) and control valve layers (green) delineated with food dyes. Lower right drawing is a side-view of a single cell microchamber with a representative readout image from the SCBC device. Each barcode fluorescent stripe corresponds to a specific protein assay. Signals from three microchambers with different cell numbers are shown. (*B*) SCBC assay steps. DNA barcodes are converted into antibody barcodes using a cocktail of DNA-antibody conjugates. Cells are then loaded and isolated into the upper chamber and incubated at a desired pO_2, during which time secreted proteins are captured on designated barcode stripes. The chip is then cooled to near $0^{\circ}C$, and the valve connecting the lysis buffer chamber is opened, leading to cell lysis within 15 minutes. The intracellular proteins are released and captured onto designated barcode stripes. (C) The equilibrium of the O_2 content within the microfluidic device, over time, as measured by the O_2 sensor The O_2 concentration reaches a mild hypoxic level (<3%) in less than 30 minutes.

U87 EGFRvIII cells were constructed as previously described[25] and routinely maintained in Dulbecco's Modified Eagle's Medium (DMEM, American Type Culture Collection) containing 10% fetal bovine serum (FBS) in humidified atmosphere of 5% CO_2 and 95% air at 37°C. GBM39 human glioblastoma cells were generated as previously described[26] and maintained in NeuroCult®-XF Proliferation Medium (STEMCELL Technologies, Inc.) containing 20ng/mL epidermal growth factor (EGF, Sigma) and fibroblast growth factor (FGF, Sigma) and 1µg/mL Heparin (Sigma) in humidified atmosphere of 5% CO_2 and 95% air at 37°C. The DNA and antibody reagents are listed in the *SI Appendix,* Table S1. The DNA-antibody conjugates were synthesized as described in the Chapter 2 and validated with standard proteins by DNA spot microarray before use[27].

3.2.4 Protein assays on bulk cell culture

The validation of the DNA-antibody conjugates involved separate calibrations for each of the different immunoassays, as well as quantitating the cross-reactivity between those immunoassays (Fig. 3.3). All bulk protein assays in this study started with spotted DNA microarrays that were obtained from the Institute for Systems Biology (Seattle, Washington). The spotted arrays and the flow patterned barcode arrays utilized the same DNA oligomer pairs (Appendix A: Supplementary Tables) for each detected protein. The description of the microwell-based multiplexed immunoassays from statistical numbers of cells followed the protocols described in Chapter 2.

For mTOR kinase inhibition bulk assay, U87 EGFRvIII cells were cultured in DMEM with 1% FBS at a density of 150,000 cells/mL and at O_2 levels controlled to be 21%, 3%,

2%, 1.5% or 1% for 7 hours, with or without addition of 3µM of the mTOR kinase

inhibitor PP242 (2-(4-amino-1-isopropyl-1H-pyrazolo[3,4-d]pyrimidin-3-yl)-1H-indol-5-

ol, Sigma-Aldrich). GBM39 neurosphere cells were dissociated with TriplE (Invitrogen) to

form single cell suspension and then were cultured in laminin (Sigma) pre-coated dishes

with NeuroCult®-XF Proliferation Medium at a density of 150,000 cells/mL and at various

conditions identical to U87 EGFRvIII cells above. Following incubation, the treated cells

were then washed by cold PBS to remove residual media. A mixture of Cell lysis buffer

Figure 3.3 Antibody microarray utilized for the multiplex protein assays from statistical numbers of cells, as well as for cross-reactivity checks of the assayed protein panel. Layout (left-top) and photograph (left-bottom) of the microarray. The layout inset shows a basic 13-spot repeat unit. Locations #A to #M are spotted with the same ssDNA oligomers that are used for DNA barcode SCBC microarrays. Spot diameters are 150 µm. The basic unit is repeated over 60 columns and 50 rows. The 12-well PDMS slab (left-bottom) is bonded to the array-spotted glass slide. Each well has a volume of 50 µl and contains 5 complete sets of 13 DNA spots for simultaneously assaying the entire panel of proteins. At right is presented the cross-reactivity data for the panel of assayed proteins. For each row in the figure, the full cocktail of antibody-DNA conjugates (for the 7 assayed proteins, except p-EGFR due to lack of standard protein) was used. One standard protein was tested per row. The red spots are signal from the standard proteins and the green spots are alignment reference signals from Cy3-labeled DNA sequence complementary to spot #M. The p-EGFR antibody does not cross-react with other proteins in the panel.

(Cell Signaling, containing 20mM Tris-HCl, 150mM NaCl, 1mM Na_2EDTA, 1mM EGTA,

1% Triton, 2.5mM sodium pyrophosphate, 1mM β-glycerophosphate, 1mM Na_3VO_4 and

1µg/ml leupeptin), Complete Protease Inhibitor (Roche) and Phosphatase Inhibitor Cocktail

2 (Sigma) was added, and the mixture was stored on ice for 10 minutes. The cell extract

was then collected and spun at 14,000g at 4°C for 10 minutes. The resulting supernatant

was re-centrifuged to remove remaining cell debris. The cell lysate and media were then

added into corresponding wells for profiling secreted and intracellular proteins.

3.2.5 Protein calibration for bulk and SCBC assays

1. Calibration curves for bulk protein measurement: The assay for generating

calibration curves was performed under conditions identical to the mTOR kinase inhibition

assay described earlier, except that standard proteins were used instead of cell lysate or

Figure 3.4 Calibrations of the assayed protein panel. The calibrations at left were carried out using SCBC barcodes within an SCBC, and under conditions similar to those single cell experiments, except that standard proteins were used. The calibration at right utilized the microwell format of the spotted arrays, and provides calibration data for the bulk cell assays. These curves are utilized to transform the fluorescence intensities into concentrations (or copy numbers) of proteins. The error bars represent the standard deviations of the measurements. Two dagger symbols, color coded for each curve, are drawn to delineate the upper and lower range of protein levels that were detected within the experiments. For both the SCBC and bulk assays, these ranges are generally within the linear response regime of the calibration curves. Note that p-EGFR was not included here due to lack of standard protein.

medium. A cocktail of standard proteins was serially diluted in 1X PBS and added into

different wells. Fluorescence signals were collected and plotted versus protein

concentrations (Fig. 3.4).

2. Calibration curves for SCBC measurement: These calibrations were performed within an SCBC and under exactly the same condition as the single cell proteomic assay described above, except that standard proteins were utilized, rather than cells. A mixture of standard proteins from the SCBC assayed panel was serially diluted in 1X PBS and flowed into the SCBC microchannels. Fluorescence signals were collected to generate the calibration curves (Fig. 3.4). Since the volume of the microchambers is known, these calibration curves enable a transformation from the fluorescence intensity to the number of molecules for each protein assayed, under the caveat that the standard proteins may not be exactly the same as their counterparts from the GBM cells.

3.3 PHYSICAL APPROACHES

3.3.1 A mean-field model for understanding protein fluctuations

Protein fluctuations can be highly informative toward understanding protein functional activity. Quantifying and linking the fluctuation profiles of each signaling node with different external perturbations will yield vast amount of information regarding the signaling transduction mechanism, stimuli effectiveness or drug efficacy. A mean-field model is applied here to provide a context for discussing how the average effect of other proteins influences the fluctuations of a specific protein in question. It offers predictive capacity of fluctuation profiles that a widely dispersed fluctuation can indicate a highly active protein that is involved in multiple functional processes. A narrow, sharp fluctuation, by contrast, represents a protein with limited interactions.

To illustrate this point, we carried out Monte Carlo simulations to generate histograms for a hypothetical functional protein at several degrees of activity. The simulation is designed to capture the protein fluctuation profile for different degrees of protein activity and is programmed by R (r-project.org). The protein is assumed to be able to participate in up to four independent functional processes. Each process requires a range of protein copy numbers represented by a Gaussian distribution and an associated probability that it is active in any given single cell. If the protein does not carry out any process, its concentration is set as an inactive baseline. The more active the functional protein is, the more processes it will participate in (up to four in this simulation).

To be specific, the hypothesized processes and their required concentrations and fraction active values are listed in Table 3.1.

In the Fig. 3.6, 1 active process means that protein is confined to be involved only in process-1 or doing nothing; 2 active processes represents that the protein can access both process-1 and process-2 and so on. Finally 4 active processes indicate the protein is able to participate into all four parallel functional processes listed above. The Gaussian distributions for representing the required range of protein copy number have been set to have a fixed coefficient of variation (CV) as 0.15. 200 single cell events are generated for each case and the corresponding histogram is plotted in Fig. 3.6 C. The averaged required protein copy number and fraction active value for each process are arbitrarily chosen for calculation convenience. They can be altered freely without affecting the final conclusion of the simulation.

Table 3.1 Parameters used for the Monte Carlo simulation of the fluctuation profile of a hypothetical protein.

Process	Averaged required protein copy #	Distribution type	Fraction active
Baseline	100	Gaussian	
Process-1	150	Gaussian	0.7
Process-2	200	Gaussian	0.3
Process-3	250	Gaussian	0.5
Process-4	300	Gaussian	0.8

3.3.2 Single-cell ensemble, a basis for making predictions

Here we describe a physically motivated approach based upon the maximum entropy formalism[28], which is being increasingly used in biology[29-38]. However, we use entropy not as a statistical measure of dispersion, but as a physical quantity[39, 40]. This allows us to apply a thermodynamic-like approach and to derive a quantitative Le Chatelier's principle[41].

The system we consider is many independent replicas of a compartment containing a single cell in a nutrient solution at thermal equilibrium. Because the system is not large, different replicas of it can differ in the number, N_i, of functional proteins of kind i. We seek to represent these fluctuations by taking the different replicas as different samples from an ensemble of single cell compartments where the mean number $\bar{N}_i$ of proteins of kind i over the ensemble is given. Another given quantity is the energy, (and volume that we do not indicate explicitly). We now seek the most probable distribution of protein numbers in different compartments. The solution is well known because if many compartments are measured, then the required distribution is the one whose entropy is maximal. In textbooks of statistical mechanics this search for the most probable distribution is sometime called the Boltzmann approach. It is possible to show [42] that this approach does not require the system to be macroscopic in size. It is sufficient if we measure enough replicas so that the distribution of proteins does not significantly change as we add more measurements. If each replica is macroscopic, the fluctuations will be small and rare. Repeated measurements will give the same results. If each replica is small we can observe the fluctuations, which is the unique information available to SCBCs.

The key point is that even if the fluctuations are not small it is possible to make predictions. The probability of a system in a particular composition can be shown to be given by a joint probability density function

$$P(N_1, N_2,..) = \exp\left\{\beta\left(\sum_i \mu_i N_i - E\right)\right\}\middle/ \Xi \tag{3.1}$$

This straightforward result is perhaps misleading in its simplicity. It is most directly derived by the method of Lagrange undetermined multipliers. The numerical value of these multipliers is determined at the final stage by imposing the condition that the distribution (Eq. 3.1) reproduces the given values of the means. There are as many multipliers as conditions.

β is the Lagrange multiplier that is determined by the mean value of the energy and, as usual, is related to the temperature T as $\beta = 1/kT$ where k is Boltzmann's constant. The μ_i's are the chemical potentials as introduced in the thermodynamics of systems of more than one component [41, 43]. The Lagrange multipliers that correspond to the given (mean) number of species i are known as the Planck potentials and denoted as α_i. It is often more convenient to work with μ_i, $\alpha_i = \beta\mu_i$. If our system were macroscopic in size we would call μ_i 'the chemical potential of protein i'. For convenience we retain the designation 'potential' because, as we shall show, μ_i retains essential properties of the chemical potential even when fluctuations are finite. Ξ is a function of all the Lagrange multipliers and its role is to insure that the sum of the probability over all possible compositions yields one.

There are at least two points where important details are not revealed by the notation used in Eq. S1. Both are relevant in what follows. First is the condition that the numerical values of the chemical potentials are determined by the given mean numbers, the $\bar{N}_i$'s, of the proteins. Strictly speaking, we should write the chemical potentials as functions of the $\bar{N}_i$'s. The other point arises when we want to treat the actual numbers N_i's of the different proteins as continuous variables. This is needed, for example, to compute averages, normalize the distribution (Eq. 3.1), etc. The integration for each protein is over $dN/N!$ where $N!$, the factorial of N, arises to account for the Gibb's paradox. Therefore, as a function of the continuous variable N the distribution for one protein is

$$P(N) \propto \left(Q^N / N!\right) \exp(-\beta \mu N) \tag{3.2}$$

Here Q is the factor that arises by summing over all the internal states of the protein that are occupied at the temperature T.

3.3.3 Fluctuations describe the response to small perturbations

We show that by measuring the fluctuations in the unperturbed system we can predict how the system responds to <u>small</u> perturbations [41]. Proof: Say that we make a small change in the value of the chemical potential μ_i from its current equilibrium value to some new value $\mu_i + \delta \mu_i$. We do so isothermally. This change in μ_i potentially changes the equilibrium mean concentration of all species from $\bar{N}_j$ to $\bar{N}_j + \delta \bar{N}_j$, for all j. To compute the change in concentrations, we need to consider the change in the ensemble as

represented by Eq. 3.1. In the algebraic developments in Eq. 3.4 below we make use of

the definition of the mean concentration

$$\bar{N}_j = \sum N_j P(N_1, N_2,..)$$
(3.3)

The summation in Eq. 3.3 is over all the possible compositions, each weighted by its

probability $P(N_1, N_2,..)$ computed as the distribution of maximal entropy. The same

meaning for the summation is used also in Eq. 3.4 below. We denote this averaging by an

over bar. From Eq. 3.1, the variation of the distribution that occurs when a particular

chemical   potential   is   changed   by   a   small   amount   is

$\delta P(N_1, N_2,..) = \beta \delta \mu_i \, N_i P(N_1, N_2,..)$. Note that it is in using this lowest term in the

Taylor series that we assume that the change is small. It follows that on the average the

proteins respond to the change as:

$$\begin{aligned}
\delta \bar{N}_j &= \sum N_j \delta P(N_1, N_2,..) \\
&= \sum \left(N_j - \bar{N}_j\right)\delta P(N_1, N_2,..) \\
&= \beta \delta \mu_i \sum \left(N_j - \bar{N}_j\right) N_i P(N_1, N_2,..) \\
&= \beta \delta \mu_i \sum \left(N_j - \bar{N}_j\right)\left(N_i - \bar{N}_i\right) P(N_1, N_2,..) \\
&= \beta \delta \mu_i \overline{\left(N_j - \bar{N}_j\right)\left(N_i - \bar{N}_i\right)}
\end{aligned}$$
(3.4)

Note that the conservation of normalization implies that the average change in the

probability must be zero, $0 = \sum \delta P(N_1, N_2,..)$ and we have used this result in the

derivation above. In the last line in Eq. 3.4 we have avoided writing the summation over all

compositions by the use of the over bar to designate an average over the probability

$P(N_1, N_2,..)$, which is the notation introduced in Eq. 3.3.

As a special case of Eq. 3.4, for a small change in $\delta\mu_i$, it's readily to reach that

$$\delta\bar{N}_i = \beta\delta\mu_i\overline{(N_i-\bar{N}_i)(N_i-\bar{N}_i)} = \beta\delta\mu_i\overline{(N_i-\bar{N}_i)}^2 .$$ Because the variance is positive, a

change in the mean copy number of protein i when its own potential is changed from μ_i to

$\mu_i+\delta\mu_i$ is always in the same direction (positive or negative) as $\delta\mu_i$ itself. It is in this sense

that we refer to μ_i as the potential of the protein i.

The key point that carries into the general case is that, to linear order in the

perturbation, the change in the mean number of proteins due to a perturbation can be

computed as an average over the unperturbed distribution of copy numbers. The change in

the mean is proportional to the variance of the distribution of fluctuations. Therefore, the

lesser the fluctuations (i.e., the narrower the histogram), the more resilient to change is the

distribution[24]. This is wholly consistent with the conclusion derived from the mean-field

model discussed above.

3.3.4 A quantitative version of the principle of Le Chatelier

Taylor theorem states that, in the leading order, the change of a function is the sum of

the changes. Therefore the expression for an isothermal variation in all the chemical

potentials leads to a change of the distribution of the form:

$$\delta P(N_1, N_2, ..) = \beta \sum_i N_i P(N_1, N_2, ..) \delta\mu_i \qquad\qquad (3.5)$$

The summation in Eq. 3.5 is an ordinary sum over the finite number S of signaling proteins, $i = 1, 2, .., S$. Then we have the general equation of change that is an extended form of Eq. S4 valid for all possible small isothermal changes in the chemical potentials

$$\delta \bar{N}_j = \beta \sum_i \overline{\left(N_j - \bar{N}_j \right)\left(N_i - \bar{N}_i \right)} \, \delta \mu_i \tag{3.6}$$

This is the result of the quantitative Le Chatelier's principle.

The principle in its simplistic statement claims that the system responds to a perturbation in a direction that restores equilibrium. For example, when the temperature of a heat bath is increased, the mean energy of an immersed system goes up so that the distribution remains canonical. The proof for our case starts from Eq. 3.3. When the chemical potential of protein i is changed, for an ensemble at maximal entropy the mean value of protein j changes by

$$\frac{\partial \bar{N}_j}{\partial \mu_i} = \sum N_j \frac{\partial P\left(N_1, N_2, .. \right)}{\partial \mu_i} \tag{3.7}$$

where, as emphasized in Eq. 3.3, the distribution $P\left(N_1, N_2, .. \right)$ is not arbitrary, but is the one of maximal entropy as exhibited in Eq. 3.1. Eq. 3.4 is recovered when the derivative in Eq. 3.7 is evaluated. It appears trivial, but it is not without meaning. What we have proven is that computing a small change in the distribution $P\left(N_1, N_2, .. \right)$ when a particular chemical potential is changed from the value μ_i to a new value $\mu_i + \delta \mu_i$ is the same as computing the derivative of the distribution $P\left(N_1, N_2, .. \right)$ at the point where the value of the chemical potential is μ_i. Then the change in the distribution is

$\left(\partial P\left(N_1, N_2,..\right)/\partial \mu_i\right)\delta \mu_i$. Of course, this is what differential calculus is about. Yet the result is not pure mathematics. It shows that the new distribution is a distribution of maximal entropy of the functional form Eq. 3.1 as otherwise the result will not hold. It says that a small change in the chemical potential μ_i, and no other change, leads to a new distribution which is also one of maximal entropy.

Typically we do not see the theorem of Le Chatelier stated as in Eq. 3.6. This is because of the practical point that the number fluctuations are typically not easy to observe in a macroscopic system. Here however we deal with functional proteins released by a single cell in which the distribution is clearly observed and the covariance can be computed from the experimental data as long as that the number of replicas is not small.

The (symmetric) square matrix $\overline{\left(N_j - \bar{N}_j\right)\left(N_i - \bar{N}_i\right)}$ is the covariance matrix of the copy number fluctuations in the steady state concentrations, the $\bar{N}_j$'s. The covariance matrix has the dimensions of S by S where S is the number of signaling molecules that take part. In practice we have to compromise on this definition meaning that S is the number of signaling molecules that can be detected. If an important protein is not detected, then the network that we infer will be incomplete and the predictions may fail. However, if they do, this is informative, because it signals the presence of new constraints needed to capture reproducible behavior in the system, which had not been taken into account.

3.4 RESULTS AND DISCUSSION

3.4.1 Single-cell data collected from SCBCs

Figure 3.5 Representative SCBC data collected from U87 EGFRvIII cell lines. (*A*) Scatter plots of assayed protein levels measured from U87 EGFRvIII single cells at 21% and 1% pO$_2$. The averaged fluorescence intensity with standard error of the mean (SEM) are overlaid for each protein. Statistical uniqueness is evaluated by two-tailed student's t test assuming unequal variance (*NS: not significant; *P<0.05; **P<0.005; ***P<0.0005*). (*B*) Western blotting results for several of the cytoplasmic proteins from U87 EGFRvIII cells assayed at 21% and 1% pO$_2$. (*C*) Scatter plots of the assayed levels of p-P70S6K and p-mTOR at 21% pO$_2$ for individual microchambers containing 1, 2 or 3 cells, indicating the statistical uniqueness of data sets representing different quantized cell populations. (*D*) Complete single cell dataset of secreted and cytoplasmic proteins profiling from individual SCBC microchambers containing single or few cells at different pO$_2$ values, presented as a heat map. Columns containing 1-, 2- and 3-cell data are indicated by the color code along the top row. The heat maps were generated by Cluster and Treeview (EisenSoftware).

Figure 3.6 Measured single cell fluctuations for four cytoplasmic proteins as a function of pO$_2$, and a simulation of fluctuations for a hypothetical protein. (*A*) Single cell fluctuation profiles for HIF-1α at various pO$_2$. (*B*) Single cell fluctuations for p-mTOR, p-P70S6K and p-ERK1 at various pO$_2$. Note that these fluctuations exhibit a sharpening at 1.5% pO$_2$. (*C*) Single cell fluctuation profiles from a Monte Carlo simulation that assumes a hypothetical protein participates in varying numbers of functional processes. Note the comparison of this simulation to the measured fluctuations of HIF-1α.

For the U87 EGFRvIII cell line, we collected single cell data at 21%, 3%, 2%, 1.5% and 1% pO$_2$. At each condition, we assayed ~100 single cells, 60 zero-cell chambers, and 50 two-cell chambers. After background subtraction, scatter plots of the single-cell proteomic data (Fig. 3.5 A) can be compared against bulk cell population protein assays using western blotting or sandwich ELISA (Fig. 3.5 B and D). The statistical uniqueness of 1-cell data was established via comparison against 2-cell data (Fig. 3.5 C). For each protein measured, the fluorescence intensity is converted into copy numbers detected using

calibration data (Fig. 3.4) that relied on standard proteins. For a given protein, a

histogram of copy number versus frequency of observation reflects the fluctuations of that
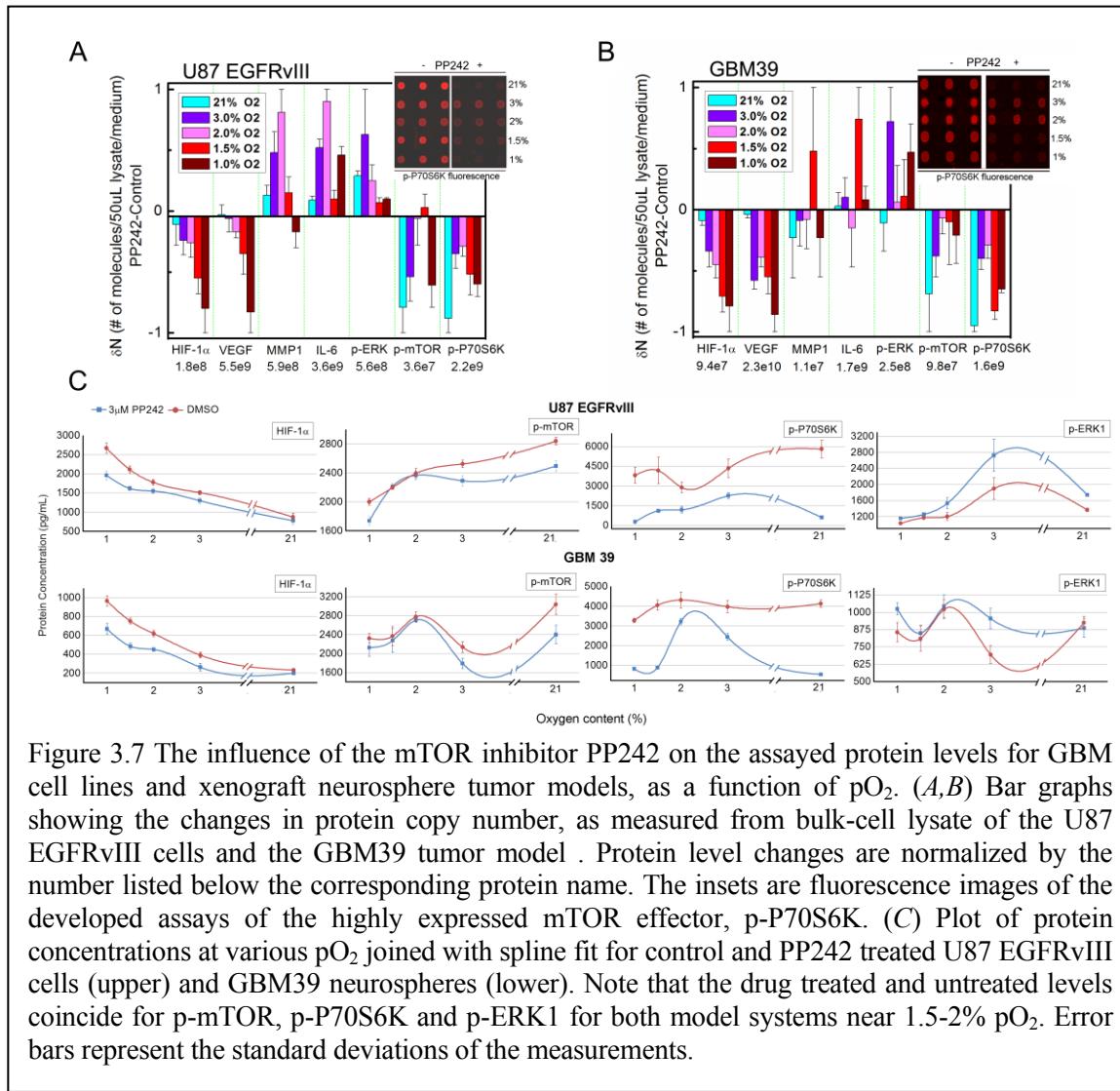
protein.

Figure 3.7 The influence of the mTOR inhibitor PP242 on the assayed protein levels for GBM cell lines and xenograft neurosphere tumor models, as a function of pO_2. (*A,B*) Bar graphs showing the changes in protein copy number, as measured from bulk-cell lysate of the U87 EGFRvIII cells and the GBM39 tumor model . Protein level changes are normalized by the number listed below the corresponding protein name. The insets are fluorescence images of the developed assays of the highly expressed mTOR effector, p-P70S6K. (*C*) Plot of protein concentrations at various pO_2 joined with spline fit for control and PP242 treated U87 EGFRvIII cells (upper) and GBM39 neurospheres (lower). Note that the drug treated and untreated levels coincide for p-mTOR, p-P70S6K and p-ERK1 for both model systems near 1.5-2% pO_2. Error bars represent the standard deviations of the measurements.

3.4.2 Protein fluctuations reveals a deregulation in mTORC1 signaling near 1.5% oxygen partial pressure (pO_2).

Fig. 3.6 A and B show the single cell fluctuations for the four cytoplasmic proteins at different pO_2 values. HIF-1α has a unique profile compared to the phosphoproteins related to mTORC1 signaling (mTORC1 designates mTOR and its effectors, which include P70S6K and ERK1[12]). As pO_2 decreases, the HIF-1α fluctuations evolve from a narrow and peaked histogram into a widely dispersed profile, with the average shifting to higher copy numbers. By contrast, the three phosphoprotein fluctuations exhibit broad widths at 21%, 3%, 2%, and 1% pO_2, but are sharply peaked at 1.5% pO_2 (Fig. 3.6). This has implications for a signaling network transition.

Protein fluctuations can be highly informative toward understanding protein functional activity. A widely dispersed fluctuation can indicate a highly active protein that is involved in multiple functional processes. A narrow, sharp fluctuation, by contrast, represents a protein with limited interactions. To illustrate this point, we carried out Monte Carlo simulations to generate histograms for a hypothetical functional protein at several degrees of activity. The protein was assumed to participate in up to four independent functional processes. Each process required a range of protein copy numbers, and had an associated probability that it was active in any given single cell (See Materials and Methods for detail). The simulated histograms (Fig. 3.6 C) reveal that the fluctuations are increasingly dispersed, as the number of potentially active functional processes increases. This plot effectively emulates the fluctuations of HIF-1α as pO_2 is lowered (Fig. 3.6 A). The implication is that HIF-1α is increasingly activated as the cells transition from normoxia to hypoxia. This conclusion may be drawn by simply inspecting the fluctuation profiles of HIF-1α, but it is also in strong agreement with the literature[8, 9].

By analogy with the above discussion of HIF-1α, we hypothesized that the phosphoproteins associated with mTORC1 signaling (Fig. 3.6 B) become isolated from cell signaling processes around 1.5% pO_2. This has implications, because mTORC1 is considered an important drug target in GBM (and other) tumors. The decoupling of mTORC1 from its effector proteins within this hypoxic window could account for a level of resistance to mTOR kinase inhibitors. We tested this prediction by assaying for the effects of the mTOR inhibitor PP242[23] on the phosphorylation levels of mTOR, P70S6K, and ERK1, as a function of pO_2, on bulk U87 EGFRvIII cell cultures, since those cells were the ones analyzed using the SCBC platform. We also tested our prediction on a tumor model by similarly analyzing neurosphere cultures derived from the human origin GBM39 xenograft[26]. This model also carried the EGFRvIII mutation. GBM neurospheres can provide realistic tumor models relative to cell lines[44], and have even been shown to exhibit stem-like behaviors under hypoxic stress[45]. Inhibition of mTOR by PP242 leads to down-regulation of the phosphorylation of both mTOR and P70S6K, and increased phosphorylation of ERK1, due to the activation of a negative feedback loop downstream of mTORC1 that targets the PI3K pathway[46]. The protein assays used here were multiplexed sandwich ELISA immunoassays from statistical numbers of cells based upon a published technique[27]. The cells were assayed in the presence of a 3 μM solution of PP242, or a DMSO control, under varying pO_2. As shown in Fig. 3.7, the results clearly support the prediction. We found that mTORC1 signaling is inhibited by PP242 for both U87 EGFRvIII cells and for the GBM39 neurosphere cultures at 21%, 3%, and 1% pO_2, but is not inhibited between 2–1.5% pO_2.

3.4.3 Steady-State kinetic model identifies a switch in mTORC1 signaling near 1.5% pO$_2$

We now look towards achieving a better mechanistic understanding of the behavior of mTOR signaling near 1.5% pO$_2$ via a steady state kinetic model.

As master regulators of hypoxic GBM cells, HIF-1α and mTORC1 act in an integrated way[12]. Our data suggest that their interplay is critical for the signaling network transition. PP242, as an ATP competitive inhibitor, can directly inhibit mTORC1 activity. mTORC1 activity will also be inhibited by HIF-1α dependent transcriptional regulation, which can occur through REDD1 (regulated in development and DNA damage responses 1) or BINP3 (BCL2/adenovirus E1B 19 kDa protein-interacting protein 3), when exposed to hypoxia[47-49]. Furthermore, our measurements (Fig. 3.7), and other reports[50], indicate that the HIF-1α expression level can be suppressed by addition of PP242 under hypoxia. Thus, since HIF-1α can repress mTORC1, suppression of HIF-1α could potentially promote mTORC1 activity. This effect may compete against PP242 direct inhibition of mTORC1 during the course of hypoxia, thus providing a potential mechanistic explanation of the undruggability of mTORC1 signaling between 1.5–2% pO$_2$. This is summarized by the network hypothesis illustrated in Fig. 3.8 A. The network is a greatly simplified version of what is known from the literature, but we are able to work with it here because, as a steady-state kinetic model, it only requires that that the flux into and out of a particular protein channel equal a constant value, for a given set of physical conditions. Thus, we are accounting for the net influence of the network components on each other, but not necessarily the direct influence. The network of Fig. 3.8 A has the

nuance that the indicated protein-protein and protein-molecule interactions are not necessarily linear relationships.

We first address the relationship between HIF-1α and pO_2. Under normoxic conditions, HIF-1α is hydroxylated at conserved proline residues by HIF prolyl-hydroxylases (PHD1-3), allowing its further ubiquitination and degradation[25]. But HIF-1α is stabilized in hypoxic environments because HIF prolyl-hydroxylase utilizes oxygen as a co-substrate, and so is inhibited at a low pO_2[51]. It has also been reported that the formation of HIF-1α is directly regulated by NF-κB[52], while the activation of NF-κB is strongly related to the production of reactive oxygen species (ROS)[53]. Therefore, it is reasonable to assume that the production of HIF-1α is partially related to oxygen concentration.

Thus, at steady state of HIF-1α ([HIF-1] for short),

$$\frac{d[\text{HIF-1}]}{dt} = -k_{\text{HIF}}[O_2][\text{HIF-1}] + (P_{\text{HIF}} + k'_{\text{HIF}}[O_2]) = 0 \tag{3.8}$$

where k_{HIF} is the rate constant for HIF-1α hydroxylation. The term $P_{\text{HIF}} + k'_{\text{HIF}}[O_2]$ represents the formation rate of HIF-1α that is partially dependent upon pO_2.

From Eq. 3.8, the steady-state HIF-1α concentration, which fits well with the experimental data (Fig. 3.8 B), is given by:

$$[\text{HIF-1}] = \frac{P_{\text{HIF}} + k'_{\text{HIF}}[O_2]}{k_{\text{HIF}}[O_2]} \tag{3.9}$$

The level of p-mTOR in non-drug treated sample at a fixed pO_2 is written as:

$$\frac{d[\text{p-mTOR}]}{dt} = k_{\text{phos}}[\text{mTOR}] - k_{\text{dephos}}[\text{p-mTOR}] = 0 \tag{3.10}$$

where k_{phos} is the phosphorylation constant of mTOR by phosphorylases or kinases and k_{dephos} is the de-phosphorylation constant of p-mTOR by phosphatases in a cell. Thus, as expected, the steady-state concentration of p-mTOR for non-drug treated sample is simply proportional to total mTOR concentration

$$[\text{p-mTOR}] = (k_{\text{phos}} / k_{\text{dephos}})[\text{mTOR}] \qquad (3.11)$$

Introduction of PP242 inhibits mTORC1 activity and down regulates mTOR S2448 phosphorylation[54], which is equivalent to reducing k_{phos} to a smaller value, $k_{\text{phos}}^{\text{PP242}}$. Thus we have

$$[\text{p-mTOR}]' / [\text{p-mTOR}] = k_{\text{phos}}^{\text{PP242}} / k_{\text{phos}} = \gamma_{\text{PP242}}^{\text{p-mTOR}} \quad (3.12)$$

Since the oxygen dependent mTORC1 regulation and the drug inhibition of mTOR are independent processes, we take the inhibition constant $\gamma_{\text{PP242}}^{\text{p-mTOR}}$ as independent of pO_2, and so it can be readily calculated from experiment data.

We now turn towards quantifying the dependence of [p-mTOR] on [HIF-1] for non-drug treated samples (Fig. 3.8 C). These two proteins exhibit a clear (inverse) linear relationship, and so the plot is fitted using:

$$[\text{p-mTOR}] = -\gamma_{\text{HIF}}^{\text{p-mTOR}} \cdot [\text{HIF-1}] + [\text{p-mTOR}]^0 \qquad (3.13)$$

where $[\text{p-mTOR}]^0$ is the extrapolated concentration of p-mTOR at zero concentration of HIF-1α, and $\gamma_{\text{HIF}}^{\text{p-mTOR}}$ captures the strength of the inhibition of p-mTOR by HIF-1α dependent transcriptional regulations.

Eq. 3.13 describes the increase in p-mTOR with decreasing HIF-1α. PP242 further inhibits HIF-1α, which thus tends to promote p-mTOR.

$$\text{Promotion}^{\text{p-mTOR}} = \gamma_{\text{HIF}}^{\text{p-mTOR}} \cdot \Delta[\text{HIF-1}]_{\text{PP242}} \qquad (3.14)$$

where $\Delta[\text{HIF-1}]_{\text{PP242}} = \left|[\text{HIF-1}]_{\text{PP242}} - [\text{HIF-1}]\right|$ is the inhibitory influence of the mTOR inhibitor PP242 on HIF-1α. To determine this value, we note that, for our experiments, PP242 concentration was either 0 (the DMSO control) or 3μM, while HIF-1α (through the influence of hypoxia) varied from around 800 to 2800 pg/ml. Over range, the suppression of HIF-1α by PP242 appears to quadratic with respect to HIF-1α concentration for non-drug treated samples (Fig. 3.8 D).

$$\Delta[\text{HIF-1}]_{\text{PP242}} = \gamma_{\text{PP242}}^{\text{HIF}}[\text{HIF-1}]^2 \qquad (3.15)$$

As a result, according to the proposed mechanism, the p-mTOR concentration at different pO_2 for PP242 treated cells is given by

$$[\text{p-mTOR}]_{PP242} = [\text{p-mTOR}] - \text{Inhibition}^{\text{p-mTOR}} + \text{Promotion}^{\text{p-mTOR}} \quad (3.16)$$

where $\text{Inhibition}^{\text{p-mTOR}} = ([\text{p-mTOR}] - [\text{p-mTOR}]') = (1 - \gamma_{\text{pp242}}^{\text{p-mTOR}})[\text{p-mTOR}]$ represents the direct inhibition on mTORC1 by PP242.

We combined steady state chemical kinetic analysis with the fitting of data from calibrated microwell-based sandwich ELISA assays on proteins collected from lysed U87 EGFRvIII cells (Fig. 3.7 A and C; Appendix A: Supplementary Tables). The parameters used in the model are extracted from fits to the experimental data (Fig. 3.8 B–D), and are provided in the table below. With these relationships, it is straightforward to calculate the

p-mTOR concentrations from only the pO_2 values, the presence or absence of PP242,

and the fitted parameters through Eq. 3.16 (Fig. 3.8 E).

Table 3.2 Parameters used in the steady-state kinetic model.

parameters	values
P_{HIF}	1877.8 pg/mL·s
k_{HIF}	100 s^{-1}
$k'_{HIF}{}'$	83580 pg/mL·s
$[\text{p-mTOR}]^0$	3242 pg/mL
$\gamma^{\text{p-mTOR}}_{HIF}$	0.47
$\gamma^{\text{p-mTOR}}_{PP242}$	0.88
γ^{HIF}_{PP242}	9.8e-5 mL/pg

In Fig. 3.8 B–D we present the relationships between HIF-1α and pO_2, p-mTOR and

HIF-1α, and the influence of PP242 on HIF-1α. With these relationships in hand, we can

calculate the dependence of the p-mTOR level on pO_2 using, as input, only the measured

pO_2 values, the presence or absence of PP242, and the fitted parameters (Fig. 3.8 E;

Table 3.2). This result is of interest in three ways. First, the kinetic model accurately

captures the p-mTOR levels in the absence of PP242 inhibition, for all values of pO_2.

Second, it predicts a pO_2 level for which p-mTOR is not influenced by PP242. For the

parameters fitted here this level is near 1.25% pO_2, but can be shifted to slightly higher

pO_2 levels by altering some of the fitting parameters, while keeping them within their

statistical margins of error. However, for any of the fitted parameters, the kinetic model

also predicts PP242 inhibition of mTOR at pO_2 levels above the crossing point, and

PP242 activation of mTOR below the crossing point. This is clearly not observed

experimentally. The implication is that new regulators of mTOR, not included in the
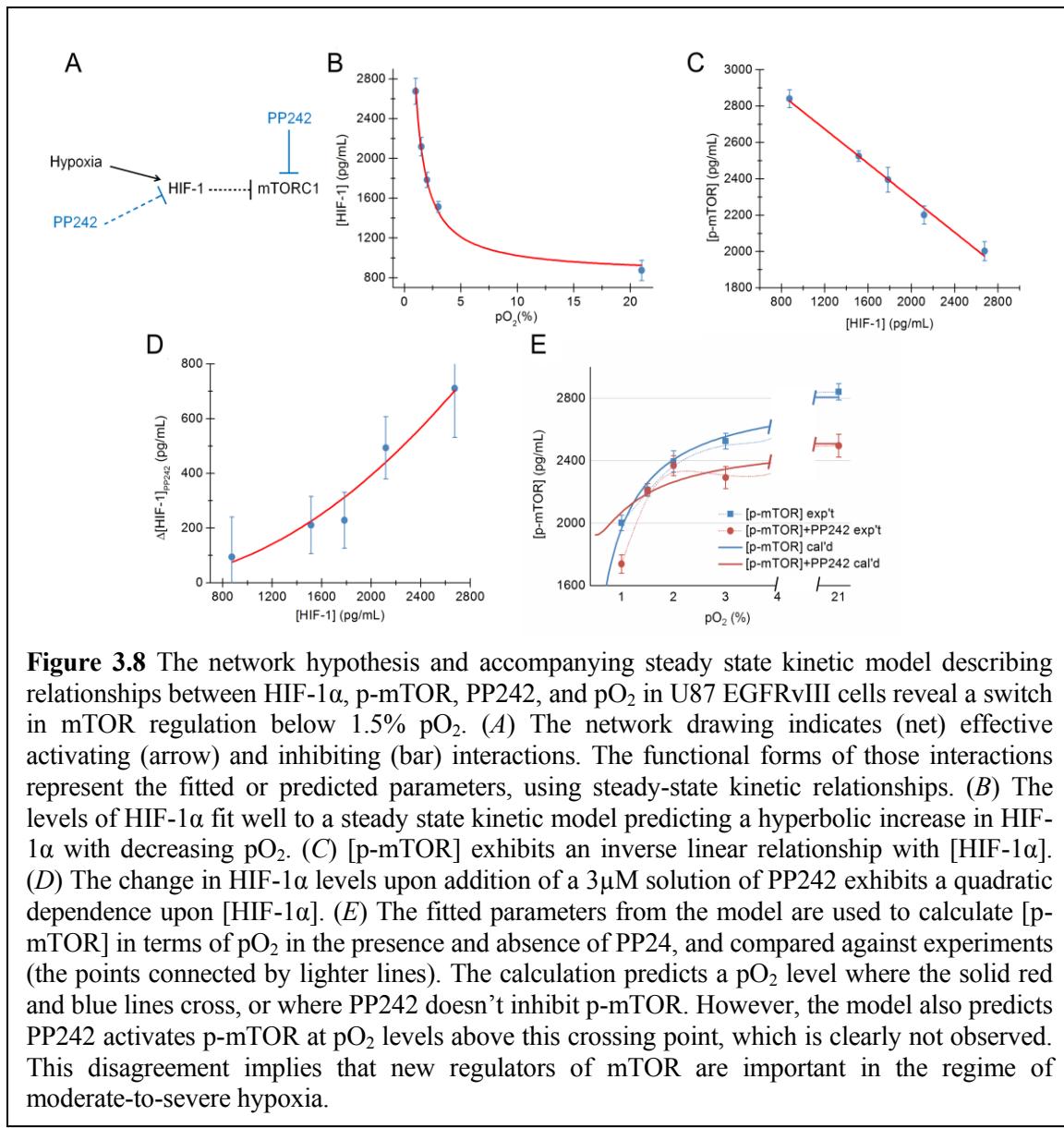
Figure 3.8 The network hypothesis and accompanying steady state kinetic model describing relationships between HIF-1α, p-mTOR, PP242, and pO_2 in U87 EGFRvIII cells reveal a switch in mTOR regulation below 1.5% pO_2. (*A*) The network drawing indicates (net) effective activating (arrow) and inhibiting (bar) interactions. The functional forms of those interactions represent the fitted or predicted parameters, using steady-state kinetic relationships. (*B*) The levels of HIF-1α fit well to a steady state kinetic model predicting a hyperbolic increase in HIF-1α with decreasing pO_2. (*C*) [p-mTOR] exhibits an inverse linear relationship with [HIF-1α]. (*D*) The change in HIF-1α levels upon addition of a 3μM solution of PP242 exhibits a quadratic dependence upon [HIF-1α]. (*E*) The fitted parameters from the model are used to calculate [p-mTOR] in terms of pO_2 in the presence and absence of PP24, and compared against experiments (the points connected by lighter lines). The calculation predicts a pO_2 level where the solid red and blue lines cross, or where PP242 doesn't inhibit p-mTOR. However, the model also predicts PP242 activates p-mTOR at pO_2 levels above this crossing point, which is clearly not observed. This disagreement implies that new regulators of mTOR are important in the regime of moderate-to-severe hypoxia.

model of Fig. 3.8 A, would have to be invoked to account for the observed behavior at very low pO_2 levels.

The influence of hypoxia and PP242 on the GBM39 model exhibits many similarities to that observed for U87 EGRFvIII cells, but only certain aspects of the kinetic model translate to that system. For example, HIF-1α exhibits a clear hyperbolic dependence upon decreasing pO_2 in both models, but the other relationships are not as clear for GBM39. This
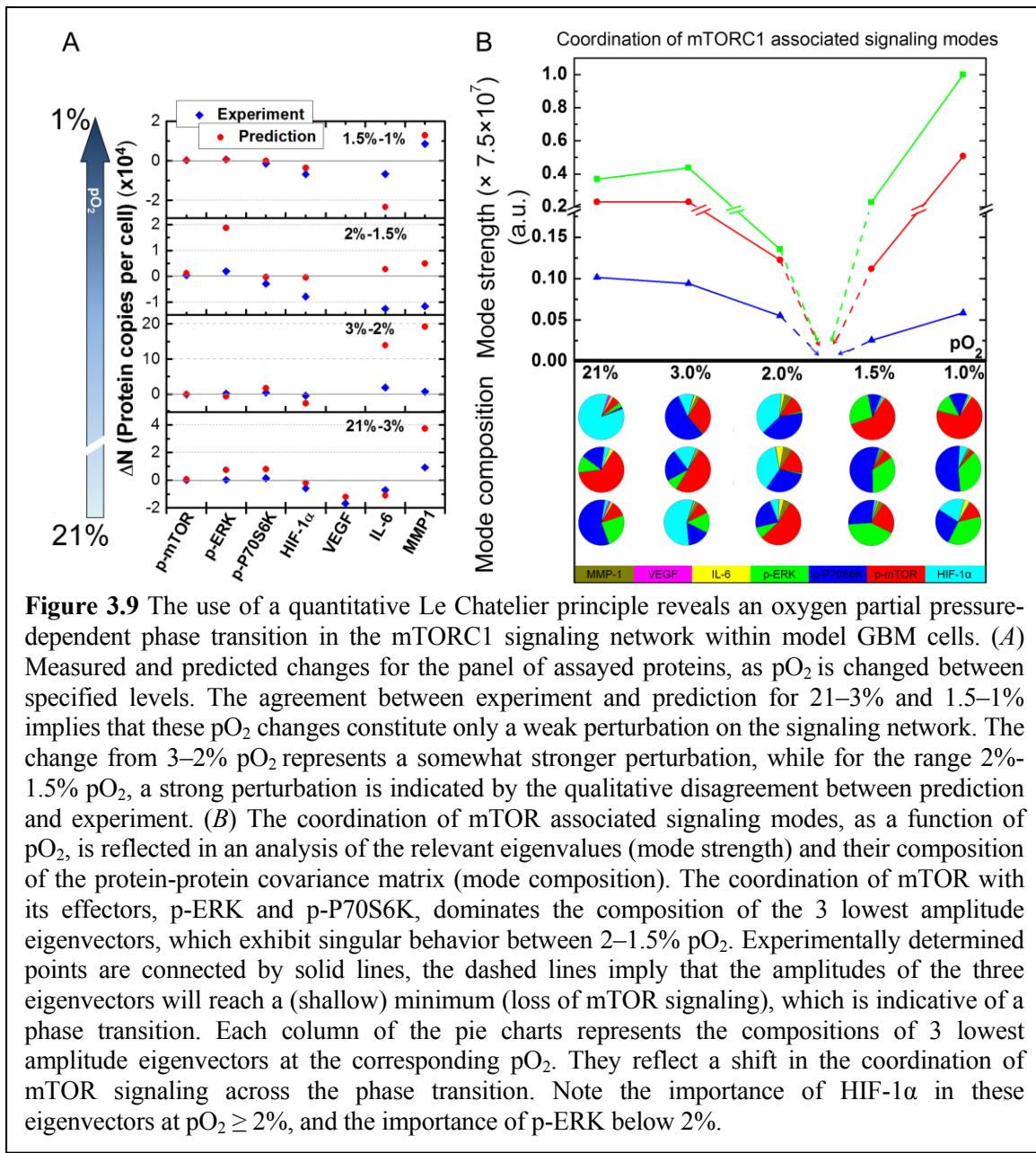
Figure 3.9 The use of a quantitative Le Chatelier principle reveals an oxygen partial pressure-dependent phase transition in the mTORC1 signaling network within model GBM cells. (*A*) Measured and predicted changes for the panel of assayed proteins, as pO_2 is changed between specified levels. The agreement between experiment and prediction for 21–3% and 1.5–1% implies that these pO_2 changes constitute only a weak perturbation on the signaling network. The change from 3–2% pO_2 represents a somewhat stronger perturbation, while for the range 2%-1.5% pO_2, a strong perturbation is indicated by the qualitative disagreement between prediction and experiment. (*B*) The coordination of mTOR associated signaling modes, as a function of pO_2, is reflected in an analysis of the relevant eigenvalues (mode strength) and their composition of the protein-protein covariance matrix (mode composition). The coordination of mTOR with its effectors, p-ERK and p-P70S6K, dominates the composition of the 3 lowest amplitude eigenvectors, which exhibit singular behavior between 2–1.5% pO_2. Experimentally determined points are connected by solid lines, the dashed lines imply that the amplitudes of the three eigenvectors will reach a (shallow) minimum (loss of mTOR signaling), which is indicative of a phase transition. Each column of the pie charts represents the compositions of 3 lowest amplitude eigenvectors at the corresponding pO_2. They reflect a shift in the coordination of mTOR signaling across the phase transition. Note the importance of HIF-1α in these eigenvectors at $pO_2 \geq 2\%$, and the importance of p-ERK below 2%.

is not surprising, given that the GBM39 protein assays are sampling a neurosphere model

of a tumor, which is comprised of a heterogeneous mixture of cellular phenotypes.

The kinetic model gives some mechanistic insight into the switch in mTORC1, partly

through its failure to predict the influence of PP242 below 1.5% pO_2. This failure

presumably arises because certain protein-protein interactions are neglected in this range.

Those interactions are implicit in the protein fluctuations. Thus, we turn to the quantitative Le Chatelier's principle, because it explicitly recognizes individual protein-protein correlations, and the predictive nature of this theory may help shed light on the uninhibitability of mTORC1 between 1.5–2% pO_2.

3.4.4 Quantitative Le Chatelier's Principle identifies a phase transition in mTORC1 signaling between 2% and 1.5% pO_2

For the Le Chatelier approach, the goal is to understand whether a change in pO_2 constitutes a strong or a weak perturbation to the U87 EGFRvIII cells. We previously reported on the development and validation of this approach[24]. In that earlier work, we used the theory to predict how the levels of a panel of secreted proteins from a human macrophage cell line, stimulated with lipopolysaccharide to emulate gram(-) bacteria, would respond to the addition of neutralizing antibodies. The theory requires single cell data as input, and can predict how the levels of certain proteins will respond to a weak perturbation. A strong perturbation is implied when the theoretical prediction and the experimental measurement are in strong disagreement.

For the theory, we first use the measured data to compute the mean number $\overline{N_i}$ of molecules for each protein i per cell, and the mean of the joint numbers of proteins i and j, $\overline{N_i N_j}$. Thereby we compute the covariance matrix Σ, which is a symmetric $P \times P$ matrix where P is the size of the protein panel assayed, and the matrix elements Σ_{ij} represent the covariance between proteins i and j (Appendix A: Supplementary Tables). Given the protein-protein covariance matrix Σ, we write the quantitative Le Chatelier's principle as

the matrix equation $\Delta \overline{N} = \beta \Sigma \Delta \mu$, where $\Delta \mu$ is a column vector whose P components give the change in the chemical potentials of the P proteins due to the change in external conditions. $\beta = 1/k_B T$, where T is the temperature and k_B is Boltzmann's constant (Theoretic details can be found in Chapter 3.3). This relates the change $\Delta \overline{N}$ in the mean number of molecules of each protein to external perturbations, such as O_2 pressure changes, or addition of a drug. Applying this approach to the single cell data, we found that the state of the signaling network at 3% pO_2 was only weakly perturbed from that at 21% pO_2 (Fig. 3.9 A). The change between 3 and 2% pO_2 was a stronger perturbation (we correctly predict the signs of the changes in protein levels, but the predicted levels for proteins IL6 and MMP1 deviate significantly from experiment). We could not predict the measured changes between 2% and 1.5% pO_2. We could, however, describe the changes between 1.5% and 1% pO_2. We do not show a prediction for VEGF at low pO_2 because, in this range, VEGF appears decoupled from the other proteins (Fig. 3.10).

Based upon these observations we hypothesized that the states corresponding to ~2–21% pO_2 represented one phase of the signaling network, while those between 1–1.5% pO_2 represented a second phase, with a phase transition occurring in between. We tested this hypothesis by analyzing the protein-protein covariance matrix to view the coordination of mTORC1 signaling, as pO_2 was varied. This approach goes beyond measuring specific protein-protein pairwise interactions, because it accounts for all of the proteins that are simultaneously assayed from each single cell. For the analysis, the eigenvalues (Fig. 3.9 B) of covariance matrix describe the amplitude of the coordinated protein-protein interaction modes, while the eigenvectors (Fig. 3.9 C) describe the composition of those modes. Such

an analysis draws from the Gibbs phase rule[55, 56], which states that, at a phase transition, a degree of freedom is lost for each coexisting phase. Consider the water liquid/solid phase transition. Away from the transition, temperature can be readily varied by warming or cooling, but at the transition when ice and water coexist, it is not possible to change the temperature.

The nature of the hypoxia-induced transition is that, at the phase transition, the signaling network undergoes a switch in connectivity during which the functional phosphoproteins related to mTORC1 signaling are isolated and inactivated. This is reflected in how the fluctuations of Fig. 3.6 B sharpen at 1.5% pO_2, but more rigorously in Fig. 3.9 B and C. Above 2% pO_2, these eigenvectors capture 75–95% of the covariance, and hence signaling network coordination, between the proteins HIF-1α, p-P70S6K, and p-mTOR; below 1.5% pO_2, they capture 80–100% of the covariance between p-P70S6K, p-mTOR, and p-ERK1. The amplitudes of these eigenvectors are strongly influenced by pO_2 and they each point to a minimum between 1.5% and 2% pO_2 (Fig. 3.9 B). Because the cell is a finite system, the minimum will likely not be sharp. This eigenvalue singularity indicates a loss of degrees of freedom (or the loss of mTORC1 signaling coordination) and thereby points to the existence of a phase transition associated with mTORC1 signaling between 1.5% and 2% pO_2. Recall the quantitative Le Chatelier's principle $\Delta \overline{\boldsymbol{N}} = \beta \boldsymbol{\Sigma} \Delta \boldsymbol{\mu}$ where the vector $\Delta \overline{\boldsymbol{N}}$ of change in protein numbers has P components. The matrix equation tells us that we can identify P linearly independent ways in which an external perturbation can influence the response of the proteins within the network. If the matrix $\boldsymbol{\Sigma}$ is singular (i.e. it has one or more zero eigenvalues), there are fewer independently

allowable variations. This is the loss of degrees of freedom. This analysis leads to the surprising prediction that mTORC1 signaling will be intrinsically uncontrollable in the U87 EGFRvIII cells between 1.5% and 2% pO$_2$, but may be influenced at higher or lower pO$_2$

Figure 3.10 The protein compositions of the 7 eigenvectors of the covariance matrix, as a function of pO$_2$. Note that below 2% pO$_2$, VEGF is almost completely decoupled from the other proteins, as indicated by its dominance within the 1st (highest amplitude) eigenvector.
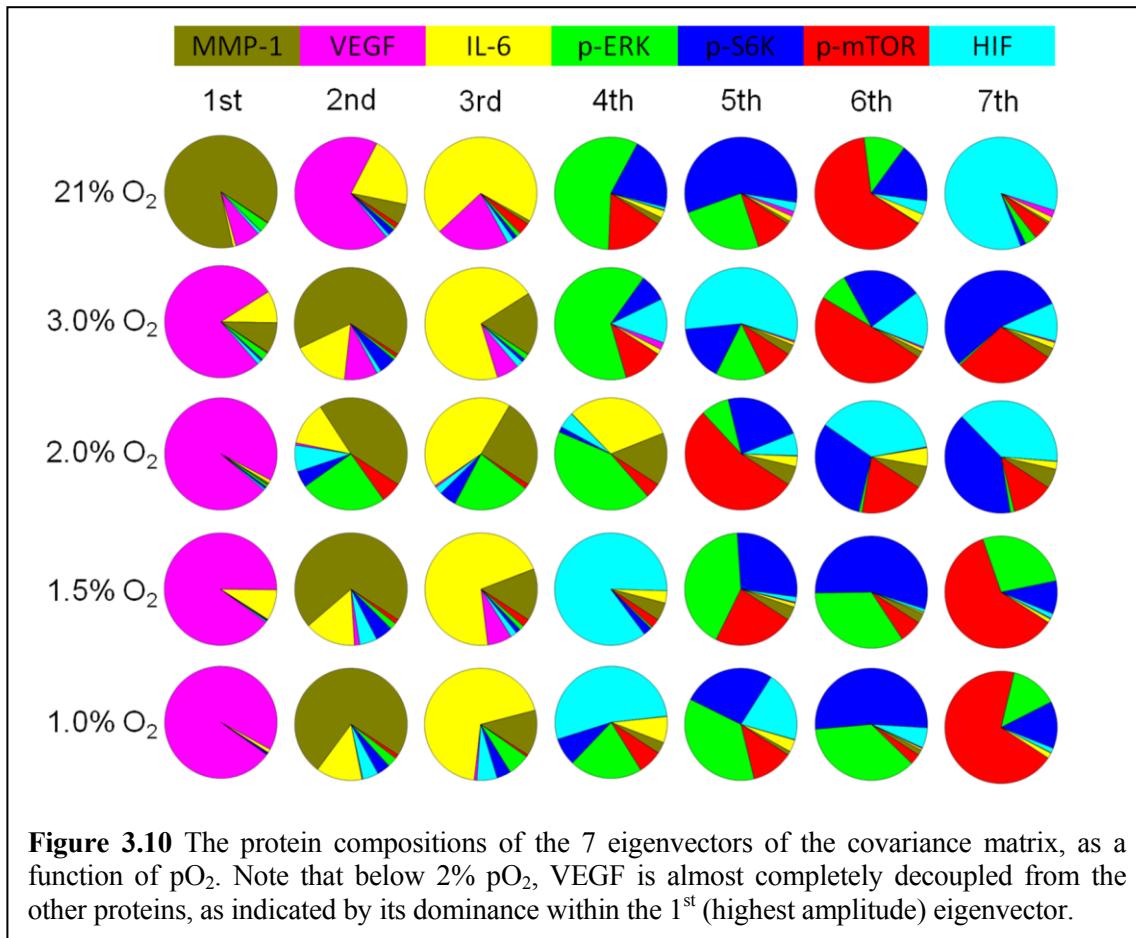
values. The proof follows from the near zero eigenvalues of the covariance matrix; the associated eigenvectors are those localized on the phosphoproteins associated with mTORC1 signaling. Near the transition, even large changes in the chemical potentials of p-mTOR and its effector proteins p-ERK and p-70S6K result in very small changes in their mean numbers.

The hypoxia induced phase transition is a multi-dimensional transition that behaves in a complementary manner to a regular transition of the inverse relation $\Delta\boldsymbol{\mu} = \beta^{-1}\boldsymbol{\Sigma}^{-1}\Delta\overline{N}$. The latter implies that near a phase coexistence where $\boldsymbol{\Sigma}^{-1}$ has a low eigenvalue, large changes of the number of molecules (extensive variables) will barely influence the chemical potential (the conjugated intensive variables). This bears an analogy to the liquid/solid transition of water where finite changes of the internal energy (the extensive variable) via the addition of heat do not alter the temperature (the conjugated intensive variable). Given that intensive and extensive variables come in conjugate pairs and are interchangeable through Legendre transforms[56], both transition manners can be appreciated.

3.5 CONCLUSION

We found that in model GBM cell lines and in a mouse GBM xenograft neurosphere model, the change in mTOR signaling from normoxia to hypoxia involves a discontinuous transition between two phases - i.e. changing pO_2 induces a switch in mTORC1 signaling. These results point to a fundamentally different approach towards understanding and predicting certain cellular behaviors. They may also provide a clue towards understanding the clinical failure of mTOR inhibitors on GBM tumors[57, 58]. Our measurements were guided by the existing biological literature, but our concern was not with capturing the detailed biomolecular interactions within the cells, but rather on understanding how the state of the signaling network is influenced by physical (pO_2) or molecular (therapeutic) perturbations. The approach is driven by new experimental tools for quantitating the levels of a panel of functional proteins from single cells, while the theory is grounded in well-established physico-chemical principles.

3.6 REFERENCES

1. Provenzano, P.P. et al. Enzymatic targeting of the stroma ablates physical barriers to treatment of pancreatic ductal adenocarcinoma. *Cancer Cell* **21**, 418-29 (2012).

2. Goel, S. et al. Normalization of the vasculature for treatment of cancer and other diseases. *Physiol Rev* **91**, 1071-121 (2011).

3. Bertout, J.A., Patel, S.A. & Simon, M.C. The impact of O2 availability on human cancer. *Nat Rev Cancer* **8**, 967-75 (2008).

4. Brown, J.M. & Wilson, W.R. Exploiting tumour hypoxia in cancer treatment. *Nat Rev Cancer* **4**, 437-47 (2004).

5. Collingridge, D.R., Piepmeier, J.M., Rockwell, S. & Knisely, J.P. Polarographic measurements of oxygen tension in human glioma and surrounding peritumoural brain tissue. *Radiother Oncol* **53**, 127-31 (1999).

6. Harris, A.L. Hypoxia--a key regulatory factor in tumour growth. *Nat Rev Cancer* **2**, 38-47 (2002).

7. Tannock, I.F. Conventional cancer therapy: promise broken or promise delayed? *Lancet* **351 Suppl 2**, SII9-16 (1998).

8. Semenza, G.L. Oxygen sensing, homeostasis, and disease. *N Engl J Med* **365**, 537-47 (2011).

9. Semenza, G.L. Targeting HIF-1 for cancer therapy. *Nat Rev Cancer* **3**, 721-32 (2003).

10. Guertin, D.A. & Sabatini, D.M. The pharmacology of mTOR inhibition. *Sci Signal* **2**, pe24 (2009).

11. Menon, S. et al. Chronic activation of mTOR complex 1 is sufficient to cause hepatocellular carcinoma in mice. *Sci Signal* **5**, ra24 (2012).

12. Wouters, B.G. & Koritzinsky, M. Hypoxia signalling through mTOR and the unfolded protein response in cancer. *Nat Rev Cancer* **8**, 851-64 (2008).

13. Fan, Q.W. et al. A dual PI3 kinase/mTOR inhibitor reveals emergent efficacy in glioma. *Cancer Cell* **9**, 341-9 (2006).

14. Akhavan, D., Cloughesy, T.F. & Mischel, P.S. mTOR signaling in glioblastoma: lessons learned from bench to bedside. *Neuro Oncol* **12**, 882-9 (2010).

15. Houghton, P.J. et al. Initial testing (stage 1) of the mTOR kinase inhibitor AZD8055 by the pediatric preclinical testing program. *Pediatr Blood Cancer* **58**, 191-9 (2012).

16. Majumder, P.K. et al. mTOR inhibition reverses Akt-dependent prostate intraepithelial neoplasia through regulation of apoptotic and HIF-1-dependent pathways. *Nat Med* **10**, 594-601 (2004).

17. Kholodenko, B., Yaffe, M.B. & Kolch, W. Computational approaches for analyzing information flow in biological networks. *Sci Signal* **5**, re1 (2012).

18. Wang, M.Y. et al. Mammalian target of rapamycin inhibition promotes response to epidermal growth factor receptor kinase inhibitors in PTEN-deficient and PTEN-intact glioblastoma cells. *Cancer Res* **66**, 7864-9 (2006).

19. Tanaka, K. et al. Oncogenic EGFR signaling activates an mTORC2-NF-kappaB pathway that promotes chemotherapy resistance. *Cancer Discov* **1**, 524-38 (2011).

20. Ma, C. et al. A clinical microchip for evaluation of single immune cells reveals high functional heterogeneity in phenotypically similar T cells. *Nat Med* **17**, 738-43 (2011).

21. Shi, Q. et al. Single-cell proteomic chip for profiling intracellular signaling pathways in single tumor cells. *Proc Natl Acad Sci U S A* **109**, 419-24 (2012).

22. Quake, S.R. & Scherer, A. From micro- to nanofabrication with soft materials. *Science* **290**, 1536-40 (2000).

23. Apsel, B. et al. Targeted polypharmacology: discovery of dual inhibitors of tyrosine and phosphoinositide kinases. *Nat Chem Biol* **4**, 691-9 (2008).

24. Shin, Y.S. et al. Protein signaling networks from single cell fluctuations and information theory profiling. *Biophys J* **100**, 2378-86 (2011).

25. Maxwell, P.H. et al. The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature* **399**, 271-5 (1999).

26. Sarkaria, J.N. et al. Use of an orthotopic xenograft model for assessing the effect of epidermal growth factor receptor amplification on glioblastoma radiation response. *Clin Cancer Res* **12**, 2264-71 (2006).

27. Bailey, R.C., Kwong, G.A., Radu, C.G., Witte, O.N. & Heath, J.R. DNA-encoded antibody libraries: a unified platform for multiplexed cell sorting and detection of genes and proteins. *J Am Chem Soc* **129**, 1959-67 (2007).

28. Levine, R.D., Tribus, M. The Maximum Entropy Formalism (MIT Press, Cambridge, MA, 1979).

29. Nemenman, I. Fluctuation-dissipation theorem and models of learning. *Neural Comput* **17**, 2006-33 (2005).

30. Varshavsky, R., Gottlieb, A., Linial, M. & Horn, D. Novel unsupervised feature filtering of biological data. *Bioinformatics* **22**, e507-13 (2006).

31. Lezon, T.R., Banavar, J.R., Cieplak, M., Maritan, A. & Fedoroff, N.V. Using the principle of entropy maximization to infer genetic interaction networks from gene expression patterns. *Proc Natl Acad Sci U S A* **103**, 19033-8 (2006).

32. Slonim, N., Atwal, G.S., Tkacik, G. & Bialek, W. Information-based clustering. *Proc Natl Acad Sci U S A* **102**, 18297-302 (2005).

33. Ziv, E., Middendorf, M. & Wiggins, C.H. Information-theoretic approach to network modularity. *Phys Rev E Stat Nonlin Soft Matter Phys* **71**, 046117 (2005).

34. Schneidman, E., Berry, M.J., 2nd, Segev, R. & Bialek, W. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* **440**, 1007-12 (2006).

35. Mora, T., Walczak, A.M., Bialek, W. & Callan, C.G., Jr. Maximum entropy models for antibody diversity. *Proc Natl Acad Sci U S A* **107**, 5405-10 (2010).

36. Remacle, F., Kravchenko-Balasha, N., Levitzki, A. & Levine, R.D. Information-theoretic analysis of phenotype changes in early stages of carcinogenesis. *Proc Natl Acad Sci U S A* **107**, 10324-9 (2010).

37. Margolin, A.A., Wang, K., Califano, A. & Nemenman, I. Multivariate dependence and genetic networks inference. *IET Syst Biol* **4**, 428-40 (2010).

38. Graeber, T.G. et al. Maximal entropy inference of oncogenicity from phosphorylation signaling. *Proc Natl Acad Sci U S A* **107**, 6112-7 (2010).

39. Levine, R.D.a.B., R. B. Energy disposal and energy consumption in elementary chemical-reactions - information theoretic approach. *Acc Chem Res* **7**, 393-400 (1974).

40. Levine, R.D. Statistical dynamics. In Theory of Reactive Collisions (CRC, Boca Raton, FL, Boca Raton, FL, 1985).

41. Callen, H.B. Thermodynamics and an Introduction to Thermostatics (Wiley, NY, 1985).

42. Levine, R.D. How large is 'large' for a thermodynamic-like behavior. *Physica E* **9**, 591-599 (2001).

43. Mayer, J.E. & Mayer, M.G. Statistical mechanics (Wiley, New York, 1966).

44. Joo, K.M. et al. Patient-specific orthotopic glioblastoma xenograft models recapitulate the histopathology and biology of human glioblastomas in situ. *Cell Rep* **3**, 260-73 (2013).

45. Bar, E.E., Lin, A., Mahairaki, V., Matsui, W. & Eberhart, C.G. Hypoxia increases the expression of stem-cell markers and promotes clonogenicity in glioblastoma neurospheres. *Am J Pathol* **177**, 1491-502 (2010).

46. Carracedo, A. et al. Inhibition of mTORC1 leads to MAPK pathway activation through a PI3K-dependent feedback loop in human cancer. *J Clin Invest* **118**, 3065-74 (2008).

47. DeYoung, M.P., Horak, P., Sofer, A., Sgroi, D. & Ellisen, L.W. Hypoxia regulates TSC1/2-mTOR signaling and tumor suppression through REDD1-mediated 14-3-3 shuttling. *Genes Dev* **22**, 239-51 (2008).

48. Horak, P. et al. Negative feedback control of HIF-1 through REDD1-regulated ROS suppresses tumorigenesis. *Proc Natl Acad Sci U S A* **107**, 4675-80 (2010).

49. Li, Y. et al. Bnip3 mediates the hypoxia-induced inhibition on mammalian target of rapamycin by interacting with Rheb. *J Biol Chem* **282**, 35803-13 (2007).

50. Chen, H. et al. mTOR activates hypoxia-inducible factor-1alpha and inhibits neuronal apoptosis in the developing rat brain during the early phase after hypoxia-ischemia. *Neurosci Lett* **507**, 118-23 (2012).

51. Semenza, G.L. Hydroxylation of HIF-1: oxygen sensing at the molecular level. *Physiology (Bethesda)* **19**, 176-82 (2004).

52. van Uden, P., Kenneth, N.S. & Rocha, S. Regulation of hypoxia-inducible factor-1alpha by NF-kappaB. *Biochem J* **412**, 477-84 (2008).

53. Michiels, C., Minet, E., Mottet, D. & Raes, M. Regulation of gene expression by oxygen: NF-kappaB and HIF-1, two extremes. *Free Radic Biol Med* **33**, 1231-42 (2002).

54. Rosner, M., Siegel, N., Valli, A., Fuchs, C. & Hengstschlager, M. mTOR phosphorylated at S2448 binds to raptor and rictor. *Amino Acids* **38**, 223-8 (2010).

55. Gibbs, J.W. On the equilibrium of heterogeneous substrates. *Trans Conn Acad Arts and Sci* **3**, 343-524 (1874).

56. Alberty, R.A. Calculation of thermodynamic properties of species of biochemical reactants using the inverse Legendre transform. *J Phys Chem B* **109**, 9132-9 (2005).

57. Cloughesy, T.F. et al. Antitumor activity of rapamycin in a Phase I trial for patients with recurrent PTEN-deficient glioblastoma. *PLoS Med* **5**, e8 (2008).

58.   Chang, S.M. et al. Phase II study of CCI-779 in patients with recurrent glioblastoma multiforme. *Invest New Drugs* **23**, 357-61 (2005).

3.7 APPENDIX A: SUPPLEMENTARY TABLES

Table 3.3 Reagents Used in this study. The upper table provides the sequences of the oligonucleotides used in the protein immunoassays. All oligonucleotides were synthesized by Integrated DNA Technology (IDT) and purified via high performance liquid chromatography (HPLC). The DNA coding oligomers were pre-tested for orthogonality to ensure that cross-hybridization between non-complementary oligomer strands was negligible (<1% in photon counts). Below the oligonucleotides is a list of the antibodies and standard proteins used for the multiplex protein assay.

Name	DNA Sequence	Melting Point
B	5'-AAA AAA AAA AGC CTC ATT GAA TCA TGC CTA -3'	57.4
B'	5' NH3-AAA AAA AAA ATA GGC ATG ATT CAA TGA GGC -3'	55.9
C	5'- AAA AAA AAA AGC ACT CGT CTA CTA TCG CTA -3'	57.6
C'	5' NH3-AAA AAA AAA ATA GCG ATA GTA GAC GAG TGC -3'	56.2
D	5'- AAA AAA AAA AAT GGT CGA GAT GTC AGA GTA -3'	56.5
D'	5' NH3-AAA AAA AAA ATA CTC TGA CAT CTC GAC CAT -3'	55.7
E	5'- AAA AAA AAA AAT GTG AAG TGG CAG TAT CTA -3'	55.7
E'	5' NH3-AAA AAA AAA ATA GAT ACT GCC ACT TCA CAT -3'	54.7
G	5'-AAA AAA AAA AGA GTA GCC TTC CCG AGC ATT-3'	59.3
G'	5' NH3-AAA AAA AAA AAA TGC TCG GGA AGG CTA CTC-3'	58.6
H	5'-AAA AAA AAA AAT TGA CCA AAC TGC GGT GCG-3'	59.9
H'	5' NH3-AAA AAA AAA ACG CAC CGC AGT TTG GTC AAT-3'	60.8
K	5'-AAA AAA AAA ATA ATC TAA TTC TGG TCG CGG-3'	55.4
K'	5' NH3-AAA AAA AAA ACC GCG ACC AGA ATT AGA TTA-3'	56.3
L	5'-AAA AAA AAA AGT GAT TAA GTC TGC TTC GGC-3'	57.2
L'	5' NH3-AAA AAA AAA AGC CGA AGC AGA CTT AAT CAC-3'	57.2
M	5'-AAA AAA AAA AGT CGA GGA TTC TGA ACC TGT-3'	57.6
M'	5' Cy3-AAA AAA AAA AAC AGG TTC AGA ATC CTC GAC-3'	56.9
DNA label	**Antibody**	**Source**
B'	Capture: Human VEGF Antibody	R&D MAB293
	Detection: Human VEGF 165 Biotinylated Antibody	R&D BAF293
	Standard: Recombinant Human VEGF 165	R&D 293-VE
C'	Human p-mTOR (S2448) DuoSet® ELISA kit	R&D DYC1665
D'	Human p-P70S6Kinase (T389) DuoSet® ELISA kit	R&D DYC896
E'	Capture: Human IL-6 Antibody	R&D MAB206
	Detection: Human IL-6 Biotinylated Antibody	R&D BAF206
	Standard: Recombinant Human IL-6	R&D 206-IL
G'	Capture: Rabbit anti-human phospho-EGF R (Y1173)	R&D AF1095
	Detection: Biotinylated Goat anti-human EGF R	R&D BAF231
H'	Human p-ERK1 (T202/Y204) DuoSet® ELISA kit	R&D DYC1825
K'	Capture: Human MMP-1 Antibody	R&D AF901
	Detection: Human MMP-1 Biotinylated Antibody	R&D BAF901
	Standard: Recombinant Human MMP-1	R&D 901-MP
L'	Human total HIF-1α DuoSet® ELISA kit	R&D DYC1935

Table 3.4 Data and parameters used in protein calibration curves. Mean intensities and standard deviations of standard proteins and fitting parameters for bulk (upper half) and SCBC (lower half) calibration curves. The calibration curves for bulk protein measurements were fit by fourth order polynomials and those for SCBC protein measurements were well fit by a four parameter Morgan-Mercer-Flodin (MMF) model.

Calibration curves for bulk protein measurement							
	p-ERK	p-mTOR	p-S6K	HIF-1α	MMP-1	VEGF	IL-6
50ng/ml	6433± 93.2	5216±	5758.2±	9586± 99.5	11224±	40209±	29293±

		208.5	558.8		1020.7	2289.7	1273.5
10ng/ml	1608± 110.5	800.6± 41.7	1185± 61.2	1945.5± 78.7	1535.9± 78.3	6015± 1201.3	3525.4± 133.4
5ng/ml	743.8± 44.3	409.2± 69.8	456.7± 32.7	805.5± 33.2	660± 31.8	2417.8± 141.6	1250.5± 50.6
1ng/ml	149.2± 13	51.4± 11.2	105.4± 16.4	127.6± 24.4	129.7± 10.6	452±16.1	248.7± 25.2
500pg/ml	63.5± 12.5	29.5± 3.5	72.3± 7	53.3±6.7	74.8± 4.5	255.7± 32.7	132.8± 13.5
100pg/ml	20.6±1.6	13.8±0.4	37±7.3	11±1.1	21.2±2.5	79.6± 17.4	45±4.9
10pg/ml	17.2± 2	12.4± 0.5	34.7± 4.2	7±0.8	8.6±2.4	66.3± 23.2	30.4± 7.7

$$y = a + bx + cx^2 + dx^3 + ex^4$$

	a	b	c	d	e	r
p-ERK	9.04	0.13	3.71E-6	-7.5E-11	0	0.9999
p-mTOR	12.1	2.35E-2	1.8E-5	-1.47E-9	2.29E-14	0.9999
p-S6K	32.6	7.39E-2	-2.87E-7	5.47E-10	-1.05E-14	0.9999
HIF-1α	1.59	0.11	1.19E-5	-3.92E-10	3.72E-15	0.9999
MMP-1	9.03	0.123	-5.56E-7	4.24E-10	-7.45E-15	0.9999
VEGF	69.6	0.328	3.09E-5	-4.28E-10	0	0.9999
IL-6	24.7	0.226	-6.84E-6	2.33E-9	-4.11E-14	0.9999

Calibration curves for SCBC protein measurement						
	p-ERK	p-mTOR	p-S6K	HIF-1α	MMP-1	VEGF	IL-6
50ng/ml	189.5± 7.7	170.6± 8.1	215.7± 5.1	248.6± 15.2	218.7± 14.8	254± 16.6	251.7± 4.3
10ng/ml	81.3± 3.1	65.9± 3.5	85.1± 4.3	97.9± 5.4	83.5± 4	182.3± 6.8	240.7± 14.9
1ng/ml	11.3± 0.63	12± 0.64	13± 1.02	16.8± 1.1	23.1± 1.6	70.7± 3.01	105.2± 6.05
100pg/ml	2.76± 0.15	6.31± 0.77	2.41± 0.26	4.3±0.28	8.4±1.05	18.7±2.3	14.3± 1.31
10pg/ml	2.28± 0.24	6.09± 0.64	1.98± 0.29	2.62± 0.28	5.1±0.47	5.1±0.33	8.12±0.81
0	2.2± 0.24	6.08± 0.72	1.5± 0.17	2.05± 0.18	4.11±0.53	1.29±0.14	7.9±0.77

$$y = \frac{ab + cx^d}{b + x^d}$$

	a	b	c	d	r
p-ERK	2.134	28.005	267.22	1.077	0.9999
p-mTOR	5.997	41.482	261.08	1.105	0.9999
p-S6K	1.518	31.53	370.82	0.965	0.9999
HIF-1α	2.27	31.667	478.25	0.901	0.9999
MMP-1	4.162	227.94	4267	0.637	0.9999
VEGF	1.239	3.552	316.86	0.68	0.9999
IL-6	8.124	1.524	253.24	1.443	0.9999

Table 3.5 Mean intensities and standard deviations for the SCBC protein assays from U87 EGFRvIII cells as a function of number of cells, and under different oxygen contents.

cell	ref	p-EGFR	p-ERK	p-mTOR	p-S6K	MMP-1	HIF-1α	VEGF	IL-6	Copy #
Normoxia condition at 21% O_2, 7 hours incubation										
1	81.4± 5.9	154.6±49.5	12.2±4	18.8±9.1	13.5±4.3	33.8± 16.8	16.9±4.2	50.1±16	67.7± 38.6	113
2	81.2± 5.1	190.6±38.6	16.4± 6.4	28.5±19.4	17.3±7.6	38.2± 14.1	24.7±7.1	105.5±24.1	123.8 ±50.2	43

3	82.1±6.2	204.1±59.6	23.7±9.5	38.4±19.7	23.5±9	44.8±28.9	28.7±8.1	158.2±22.7	179.3±28.2	40
Hypoxia condition at 3.0% O₂, 7 hours incubation										
1	80.6±4	150.9±52.6	11.9±3.6	18.4±7.1	12.7±2.8	30.8±9.9	24.9±7.4	85.4±32.5	79.6±34.7	77
2	80.6±4.8	184.8±53.4	17±8.04	27.4±12.8	17.3±6.9	34.7±14.9	31±8.8	149.6±30.7	141.9±58.9	61
3	80.9±4.2	199.5±46.4	23.2±9.6	37.8±18.9	21.5±6.9	39.5±11.4	37.3±8.2	183.2±11.6	192.9±51.5	23
Hypoxia condition at 2.0% O₂, 7 hours incubation										
1	80±4.4	161.4±51.6	11.2±3.6	19.2±6.3	9.2±2.3	27.3±5.1	30.1±4.4	116.8±36.8	43.9±18.9	86
2	80±3.7	186.8±53.5	15.4±4.8	27.4±11.7	12.2±3.9	44.4±48.4	40.7±6.6	161.5±42.6	96±28.7	96
3	80.7±3.4	206.3±35.2	20.4±5.7	37.8±15.3	15.7±3.7	58.4±62	45.8±7.9	196.2±33.6	154.6±35.4	48
Hypoxia condition at 1.5% O₂, 7 hours incubation										
1	80.7±6.1	150.6±49.1	10.4±1.5	18.4±4.6	11.3±3.1	31.1±15	40.5±7.9	116.8±47.2	63.3±41	92
2	80.4±3.7	180.9±42.6	15.1±5.25	25.8±13.1	18±3.5	34.9±21.7	50.4±10.1	170.4±39	118.4±52.3	77
3	81.3±4.1	201.1±35.1	20.5±11.5	36.2±15.1	19.1±9.8	39.±19.3	61.9±11.1	204.5±16.1	167±56.7	34
Hypoxia condition at 1.0% O₂, 7 hours incubation										
1	81±5.1	153.8±57.1	10.2±3.4	18±5.9	12.3±5.1	28.2±9.5	47.5±11.6	108.6±28.5	84.3±36.2	157
2	80.3±5	181.9±58.7	14.7±5.9	26.1±11	17.4±8.3	31.5±8.2	62.9±16.1	174.1±32.3	136.4±51.9	101
3	80.8±4.9	199.5±41.7	20.7±9.1	36.3±15.1	19.7±6.6	35.9±13.1	75±17.3	209±20.5	188.2±40.1	58

Table 3.6 Fluorescence intensities and change in number of molecules of the mTOR kinase inhibition assay on U87 EFGRvIII bulk cell populations. The mean intensities and standard deviations (SD) are shown with or without addition of PP242, and at different O₂. The molecular weights (in kDa) used in the calibrations are 120, 289, 44, 70, 52, 20 and 20.3 for HIF-1α, p-mTOR(Ser2448), p-ERK1(T202/Y204), p-P70S6K(T389), MMP-1, VEGF and IL-6, respectively.

Protein	+ PP242 (3μM) -				Change in # of molecules (PP242+) - (PP242-)	
	mean	SD	mean	SD	mean	SD
U87 EGFRvIII normoxia condition at 21% O₂, 7 hours incubation						
HIF-1α	94.9	13.4	107.0	13.3	-2.35E+07	3.7E+07

p-mTOR	160.8	6.3	191.7	4.8	-3.59E+07	9.34E+06
p-ERK1	247.0	4.4	194.1	6.5	2.56E+08	3.8E+07
p-P70S6K	78.1	4.7	547.6	77.9	-2.23E+09	3.0E+08
MMP-1	437.3	7.2	415.2	10.6	9.63E+07	5.6E+07
VEGF	4858.6	189.5	4944.9	178.1	-1.70E+08	5.2+08
IL-6	525.9	10.9	468.9	13.1	3.72E+08	1.1E+08
U87 EGFRvIII hypoxia condition at 3.0% O_2, 7 hours incubation						
HIF-1α	164.9	12.4	194.8	7.7	-5.29E+07	2.62E+07
p-mTOR	143.3	5.9	163.3	4.3	-2.43E+07	9.06E+06
p-ERK1	390.2	58.9	269.4	38.2	5.65E+08	3.29E+08
p-P70S6K	204.3	18.2	389.8	65.9	-8.98E+08	3.13E+08
MMP-1	546.2	23.4	463.6	17.2	3.50E+08	1.23E+08
VEGF	7726.9	208.1	7980.9	360.4	-4.30E+08	7.07E+08
IL-6	894.5	41.9	542	21.8	2.13E+09	2.75E+08
U87 EGFRvIII hypoxia condition at 2.0% O_2, 7 hours incubation						
HIF-1α	200.8	9.9	234.4	11.3	-5.73E+07	2.58E+07
p-mTOR	149.7	5.4	152	5.8	-2.79E+06	9.75E+06
p-ERK1	217.1	19.3	171	13.7	2.25E+08	1.16E+08
p-P70S6K	121.6	18.4	256.4	34.5	-7.29E+08	2.07E+08
MMP-1	540.3	18.6	402.1	26	5.94E+08	1.39E+08
VEGF	7467.2	155.4	8137.6	110.4	-1.14E+09	3.26E+08
IL-6	1066.9	69.1	452.7	31.7	3.64E+09	4.19E+08
U87 EGFRvIII hypoxia condition at 1.5% O_2, 7 hours incubation						
HIF-1α	210.8	9.6	285.4	14.4	-1.24E+08	2.87E+07
p-mTOR	136.9	3.2	135.7	2.4	1.47E+06	5.13E+06
p-ERK1	178.3	5.8	166.4	3.1	5.82E+07	3.25E+07
p-P70S6K	115.9	7.6	373	92	-1.31E+09	4.32E+08
MMP-1	337.4	16.8	313.8	11.8	1.07E+08	9.28E+07
VEGF	8414.4	461.1	9851.6	495.1	-2.33E+09	1.10E+09
IL-6	528	28.2	468.1	26.2	3.91E+08	3.01E+08
U87 EGFRvIII hypoxia condition at 1.0% O_2, 7 hours incubation						
HIF-1α	261.8	18.4	375.2	21.7	-1.78E+08	4.48E+07
p-mTOR	99.8	4.3	119.7	3.8	-2.74E+07	7.99E+06
p-ERK1	164.7	1.96	147.5	1.5	8.51E+07	1.22E+07
p-P70S6K	54.3	7.9	338.8	53.8	-1.52E+09	2.61E+08
MMP-1	289	12.9	316.3	17.1	-1.24E+08	9.76E+07
VEGF	7567.1	179.1	10991.9	683.7	-5.54E+09	1.11E+09
IL-6	1102.3	45.2	762.7	29.5	1.87E+09	2.91E+08

Table 3.7 Protein-protein covariance matrices at various pO_2 levels from U87 EGFRvIII single cell data. The covariance matrices are symmetric, so only the upper half parts are shown.

21% pO_2	p-mTOR	p-ERK1	p-p70S6K	HIF-1α	VEGF	IL-6	MMP1
p-mTOR	**2.52E+07**	2.11E+07	1.09E+07	1.19E+05	3.98E+07	2.29E+07	8.11E+06
p-ERK1		**9.84E+07**	2.37E+07	2.61E+06	-1.26E+07	-3.75E+06	1.24E+08
p-p70S6K			**3.77E+07**	6.46E+05	5.66E+07	1.98E+07	-9.52E+05
HIF-1α				**8.82E+06**	3.66E+07	3.36E+06	3.82E+07
VEGF					**2.17E+09**	5.68E+08	1.27E+08
IL-6						**4.32E+08**	-1.39E+07
MMP1							**3.74E+09**
3% pO_2	p-mTOR	p-ERK1	p-p70S6K	HIF-1α	VEGF	IL-6	MMP1
p-mTOR	**1.74E+07**	9.48E+06	6.62E+06	3.54E+06	1.96E+07	4.00E+06	1.06E+07
p-ERK1		**7.69E+07**	6.28E+06	9.49E+06	1.57E+08	2.27E+07	9.07E+06
p-p70S6K			**1.48E+07**	7.96E+06	3.47E+07	1.90E+07	4.55E+07
HIF-1α				**3.42E+07**	8.72E+07	5.73E+06	2.23E+07
VEGF					**5.49E+09**	6.09E+08	5.49E+08
IL-6						**4.08E+08**	1.42E+08
MMP1							**6.75E+08**

2% pO$_2$	p-mTOR	p-ERK1	p-p70S6K	HIF-1α	VEGF	IL-6	MMP1
p-mTOR	**1.19E+07**	1.08E+07	2.85E+06	1.54E+06	6.55E+07	3.65E+06	1.09E+07
p-ERK1		**7.94E+07**	1.03E+07	7.63E+06	5.18E+08	2.41E+07	3.15E+07
p-p70S6K			**1.07E+07**	5.76E+06	3.73E+08	1.23E+07	8.22E+06
HIF-1α				**1.16E+07**	2.96E+08	7.32E+06	1.98E+07
VEGF					**6.13E+10**	6.33E+08	2.87E+08
IL-6						**7.95E+07**	9.79E+06
MMP1							**1.17E+08**
1.5% pO$_2$	p-mTOR	p-ERK1	p-p70S6K	HIF-1α	VEGF	IL-6	MMP1
p-mTOR	**6.04E+06**	6.27E+06	5.36E+06	9.68E+05	1.51E+08	3.67E+06	3.10E+07
p-ERK1		**1.37E+07**	6.18E+06	2.83E+06	2.73E+07	8.39E+05	4.34E+07
p-p70S6K			**1.86E+07**	7.04E+06	1.73E+08	2.92E+07	1.07E+08
HIF-1α				**5.27E+07**	2.48E+08	6.03E+07	1.01E+08
VEGF					**6.51E+10**	6.37E+09	2.89E+07
IL-6						**1.25E+09**	2.02E+08
MMP1							**1.52E+09**
1% pO$_2$	p-mTOR	p-ERK1	p-p70S6K	HIF-1α	VEGF	IL-6	MMP1
p-mTOR	**1.03E+07**	1.67E+07	1.27E+07	7.24E+06	3.75E+06	1.47E+06	9.11E+06
p-ERK1		**6.96E+07**	1.96E+07	2.06E+07	3.56E+07	2.19E+07	2.09E+07
p-p70S6K			**5.37E+07**	6.61E+06	4.52E+08	2.28E+07	3.21E+07
HIF-1α				**1.19E+08**	1.14E+08	2.02E+07	3.48E+07
VEGF					**8.66E+10**	9.64E+08	9.25E+07
IL-6						**3.14E+08**	7.93E+07
MMP1							**7.37E+08**

C h a p t e r 4

Applications in preclinical cancer research: collective behaviors in signaling coordination—signaling modes—identifies adaptive network dynamics and defines effective targeted therapy strategies

4.1 INTRODUCTION

For almost all patients with advanced cancer, targeted therapies yield modest clinical benefit, as tumors rapidly become resistant[1]. Resistance to any single therapy can occur when drug-resistant tumor cell subpopulations expand to drive recurrence[2] in a process akin to Darwinian-type evolution under the selection pressure of the drug. The timescale of the appearance of resistance may be governed by several factors, one of which is long term cell-cycle selection of the resistant subpopulation. Deep sequencing of clinical tumor samples can potentially detect that rare cell subpopulation, and thus help guide the selection of a therapy that includes a second drug that forestalls resistance by targeting that population[3-6].

An alternative resistance mechanism is one in which cancer cells targeted by the inhibitor adapt, altering their protein signaling networks so as to maintain the signal flux through those networks that is required for tumor maintenance and growth[7-9]. In this mechanism, the drug is a perturbation that shifts the cancer cells from one steady state to another. Such resistance can develop quickly, and would not be detectable by sequencing. Instead, the challenge is to measure the structure and response dynamics of the protein signaling networks that are influenced by the drug[10, 11]. Anticipating how those networks will respond to the drug might then provide insights for identifying effective therapy

combinations.

To test this idea, we exploited a clinically relevant model of acquired cancer drug resistance, with the goal of understanding the general nature of the resistance, and of identifying combinations of targeted therapies for effective treatment. GBM39 is a human-derived model of glioblastoma multiforme (GBM) that is maintained by serial transplantation in xenografts[12]. It retains tumor heterogeneity, an invasive growth capacity, and a drug response profile that are representative of clinical behavior[13]. GBM39 expresses high levels of the epidermal growth factor receptor (EGFR) variant(v)III oncogene, which sensitizes tumor cells to the mechanistic target of rapamycin (mTOR) kinase inhibitor CC214-2[14]. Here, we determined the structure of the hyperactivated phosphoprotein networks, including those associated with mTOR signaling, using multiplex assays of phosphoproteins from statistical numbers of single cancer cells[15, 16] that were untreated, responding to CC214-2, and resistant to CC214-2. The evolution of that structure, between the untreated and responsive states, provides guidance for selecting targeted therapy combinations that can successfully arrest tumor growth. It also provides guidance for identifying those therapies and therapy combinations that will not be effective.

4.2 EXPERIMENTAL METHODS

4.2.1 Establishment of *in vivo* mouse xenograft model recapitulating the clinical scenario of acquired resistance

Figure 4.1 Characterization of GBM 39 *in vivo* mouse model. (a) Tumor growth curve for control (sample size n=11), responsive (n=7) and resistant (n=7) xenografts (variations expressed as s.d.), (b) ^{18}F-FDG PET scanned images, PET-CT, and CT scanned images for the three conditions (Right, n=4 for each condition). The arrow indicates the localization of the tumor. (c) Immunohistochemistry results for the three conditions (n= 8, 6, 6 respectively for vehicle, responsive and resistant groups). *$P<0.05$ (Student's T-test); scale bar: 100 μm.

Mice bearing GBM39 flank xenografts treated for 19 days with CC214-2 (100mg/kg, once every two days by gavage), demonstrated significant inhibition of glucose uptake as measured by ^{18}F-fluorodeoxyglucose (FDG) positron emission tomography (PET) (Figs. 4.1 a, b and 4.2), reduced mTOR Complex(C)1 and mTORC2 signaling, reduced cellular proliferation as measured by Ki67 staining (Fig. 4.1 c, Appendix A: Supplementary Figures and Appendix B: Supplementary Table) and a reduction in tumor volume relative to
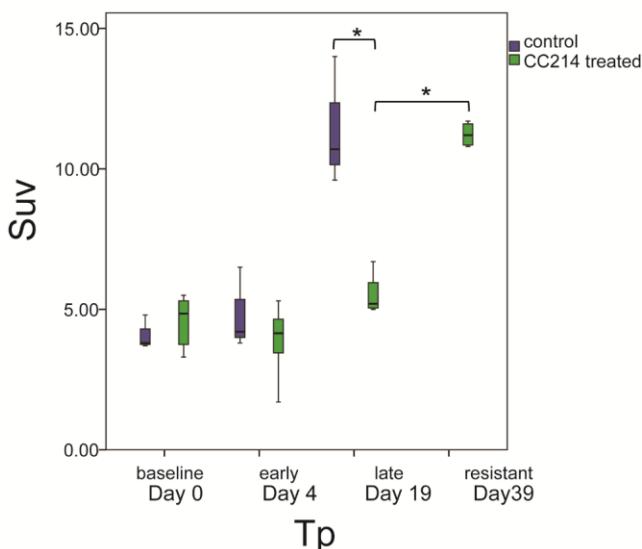
Figure 4.2 ^{18}F-FDG Positron Emission Tomography (PET) on GBM39 xenografts. Statistical analysis of the Standardized Uptake Value (Suv) registered by imaging; *P<0.05 (Student's T-test; n= 4 for each time point; data represents the average of 4 independent measurements for each time point with variations expressed as s.d.); Tp= Time points.

untreated mice (Fig. 4.1 a). Minor tumor cell death was observed. By day 27, rapid tumor regrowth was appreciated with concomitant increases in glucose uptake, mTOR signaling, cellular proliferation and tumor volume, thus modeling the clinical scenario of acquired resistance (Fig. 4.1 a–c).

GBM39 primary neurospheres were provided by Prof. David James (UCSF, San Francisco, U.S.A.) and authenticated by luciferase reporter expression before the beginning of the in vivo experiments. GBM39 cells were tested for pathogens, including mycoplasma, by IDEXX RADIL and all tests results were negative. GBM39 cells were cultured in NeuroCult (StemCell Technologies) supplemented with Heparin (1 μg/mL), Epidermal Growth Factor (EGF, 20 ng/mL) and Fibroblast Growth Factor (FGF, 20 ng/mL; SIGMA) and 100 U/mL penicillin and streptomycin (Gibco) in a humidified 5% CO_2 (vol/vol) incubator, at 37°C. CC214-1 and CC214-2 were provided by Celgene Corporation (San Diego, U.S.A.)[17]. GBM39 flank xenografts were obtained in full compliance with the

UCLA-Division of Laboratory Animal Medicine (DLAM) regulation and with the UCSD-Institution of Animal Care and Use Committee (IACUC) regulations. GBM39 cells were resuspended in PBS (Cellgro) plus Matrigel (BD Biosciences), 1:1 solution (vol/vol), at 1×10^7 cells/ml density. One million of GBM39 cells were injected in the flank of each 4 weeks old female athymic mouse. Tumor sizes were measured using automated caliper. For the drug treatments, CC214-2 was administered by oral gavage, 100mg/kg, once every two days, in a suspension containing 0.5% carboxymethylcellulose (Sigma), 0.25% Teewn-80 (Sigma) in nanopure water.

4.2.2 MicroPET/CT characterizations

Four mice for each group were anesthetized with isoflurane (2% in 100% oxygen), warmed and injected with 20μCi [F^{18}]-FDG. After an uptake period of 60 minutes, mice were placed in a dedicated imaging chamber designed for use for the CT and both PET systems. Data were acquired using an Inveon scanner (Siemens Preclinical Solutions), a Gensiys4 (Sofie Biosciences, Culver City, CA) and a MicroCAT II CT (microCAT; Imtek Inc.) instrument. Acquisition of PET images was performed for 10 min on each scanner followed by 8 min CT acquisition[18].

PET and CT Images were analyzed using OsiriX Imaging Software (version 3.8; OsiriX). MicroCT and PET images were reviewed blinded to detect tumor burden. Consecutive 2-dimensional regions of interest (ROI) were drawn on tumor on coronal and axial to detect the maximum FDG uptake. These regions encompassed the entire metabolically active tumor. Display of representative images was done according the

shown color scale proportional to tissue concentration, with red being the highest and lower values in yellow, green, and blue (Figs. 4.1 b and 4.2). Student's T-test was used to assess statistical significance. The variation between groups was similar and expressed as standard deviation.

4.2.3 Immunohistochemistry (IHC) and immunoblotting

For IHC assays, Paraffin embedded GBM39 xenografts blocks were sectioned at the UCLA Pathology Histology and Tissue Core Facility and at the UCSD Histology and Immunohistochemistry core followed by immunohistochemistry stains performed as described in Mellinghoff *et al.*[19]. Three images at 40x magnification per IHC slide were captured using DP 26 camera mounted on an Olympus BX43 microscope. Quantitative analysis of the IHC stained slides was performed with Microsuite Five software (Olympus; Figs. 4.1 c and 4.14 and Appendix B: Supplementary Tables). In the IHC quantification, the following number of xenografts was considered for each group: n= 8 for controls, n= 6 for CC214-2 responsive group, n= 6 for CC214-2 resistant group, n= 4 for Dasatinib group, n= 4 for U0126 group, n= 4 for U0126 plus Dasatinib group, n= 2 for each combination with CC214-2, n= 2 for each drug removed group. Student's T-test was used to assess statistical significance.

Western blot (Figs. 4.3 a, 4.7 a and 4.11 b) was done loading 10 μg of protein lysates. Lysates were collected in RIPA buffer (Boston BioProducts) with addiction of protease plus phosphatase inhibitor cocktail, 10μL/mL each (Thermo Scientific). Gradient 4-15% pre-casted gels were used for the electrophoretic protein separation in mono-dimension (Bio-Rad). Proteins were transferred on nitrocellulose membranes using Trans-blot Turbo

Transfer system (Bio-Rad). Blots were then blocked in Tris-buffered saline, 0.1%

Tween20 (vol/vol) and 5% BSA (Fischer Scientific, vol/vol) for 1 hour. The primary

antibodies were incubated overnight, at 4ºC. After washing, the membranes were incubated

with secondary HRP conjugated antibodies for 1 hour at room temperature. West Femto

Trial kit (Thermo Scientific) was used to develop the immunoreactivities.

4.2.4 Preparing single-cell suspension from solid tumors.

Freshly resected xenografts were finely cut in sterile conditions and digested for 3

hours at 37°C, under constant rotation (200 rpm), in a solution containing 1.5% BSA

(g/mL, Gemini), 0.3% collagenase type 2 (g/mL, Worthington), 0.3% collagenase type 4

(g/mL, Worthington) and 10 μg/mL DNAse I enzyme (Sigma). Single cell suspensions

were then filtered with a 40 μm cell strainer and pellets were treated for 2 minutes with 3

mL ACK buffer (Lonza). Solutions were neutralized with DMEM medium (Gibco) and cell

viability assessed by trypan blue exclusion. Frozen stocks were made re-suspending cell

Figure 4.3 Cell preparation and validation for the SCBC test. (a) Western blot results for the cells that were sorted, plated and harvested at three different time points: 0, 12, 24 hours. The fact that the levels of key functional proteins do not change over time validates the cell plating step for the test. (b) EGFR expression level of sorted U87/EGFR cells and GBM39 cells. After sorting, cells were stained with PE-conjugated rat anti-human EGFR antibodies. It's readily to see that GBM39 cells are smaller in size and more heterogeneous in terms of EGFR expression on the cell surface.

pellets in Bambanker (Wako) and storing cryovials at – 80°C (Fig. 4.11 a).

4.2.5 Magnetic-activated cell sorting (MACS) and plating

A solid tumor is composed of many different types of cells including immune cells, stromal cells, and cancer cells. In order to analyze the main characteristics of the cancer cells, it is required to sort out cancer cells, even specific subset of the cancer cells of a tumor sample. There are several challenges for cell sorting. First, the technique should be very specific, robust, and reproducible. Second, it should be able to handle small amount of sample. Third, cells after sorting should be healthy in order to provide representative information of the tumor. We chose to use magnetic bead-activated cell sorting (MACS) targeting EGFR cell surface marker. EGFR is one of the major cell surface markers for GBM which is over expressed in approximately 50–60% of glioblastoma (GBM) tumors[20]. Cell sorting was carried out with Human EGF R/ErbB1+ Cancer Cells PlusCellect kit from R&D systems (Catalog # PLS1095) and followed the manufacturer's protocol. Typical cell number available from the mouse tissue sample ranges from 500,000 to 1,000,000 and the yield of the EGFR+ cell sorting is sample specific with a range of 60–70% for most of our cases. Cell variability after sorting is another critical issue here since primary cells are normally fragile. A short time (2 hours) incubation step on a laminin pre-coated petri dish was introduced to enable only healthy cells to attach to the plate and subsequently to be transferred to the SCBC test. The cell viability was greater than 95% after employing the surface plating step. Immunoblot analyses (Fig. 4.3 a) also confirm that the additional short time culturing does not induce significant physiological changes in the primary cells. After 2 hours of incubation, dead suspended cells were removed by aspirating the media. Cells

that attached to the surface were trypinized and re-suspended in the cell media (NeuroCult®-XF Proliferation Medium, STEMCELL Technologies, Inc.) at 1,000 cells/µL for loading to the SCBC.
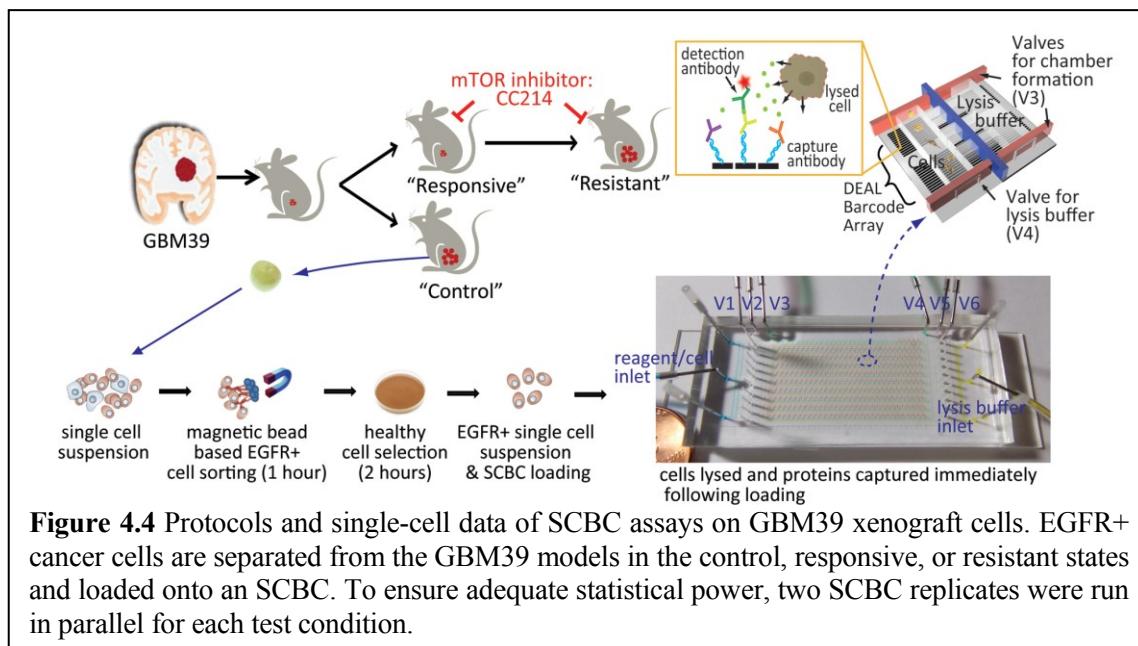
Figure 4.4 Protocols and single-cell data of SCBC assays on GBM39 xenograft cells. EGFR+ cancer cells are separated from the GBM39 models in the control, responsive, or resistant states and loaded onto an SCBC. To ensure adequate statistical power, two SCBC replicates were run in parallel for each test condition.

4.2.6 Microchip design, fabrication and experimental procedures.

We utilized the single cell barcode chip (SCBC) platform to quantify the levels and correlative interactions of 9 proteins and phosphoproteins[15, 16] from statistical numbers of single EGFR+ tumor cells (Figs. 4.4 and 4.5). GBM39 cells from harvested tumor were processed into a single cell suspension. The EGFR+ cells (which include the 70% EGFRvIII+ subset) were sorted and applied to the SCBC platform (Fig. 4.4). An SCBC contains 320 1.5 nanoliter volume microchambers, each designed for cell lysis, and each equipped with a full antibody array. Following cell lysis, the proteins are captured on the antibody array, which is developed so that specific protein levels are encoded as a fluorescence signal on particular array spots. A one-chip data set includes the digitized

fluorescence signals from 9 proteins measured from each of ~100 1-cell microchambers (Fig. 4.11 a) and 100 0-cell microchambers (providing a measurement of signal background). Two SCBCs were typically used for each condition studied.

The fabrication of SCBC, the protein panel validation and calibration and the experimental procedures of the single-cell proteomic assay follow the same protocol described in Chapter 2.

Figure 4.5 Design and working scheme of SCBC. (a) Picture of SCBC. (b) Picture of one chamber unit: valves for chamber formation (red), valve for lysis buffer control (green), cell chamber compartment (blue), and lysis buffer (yellow) reservoir are delineated by food dyes. (c) Optical image of a single cell chamber.
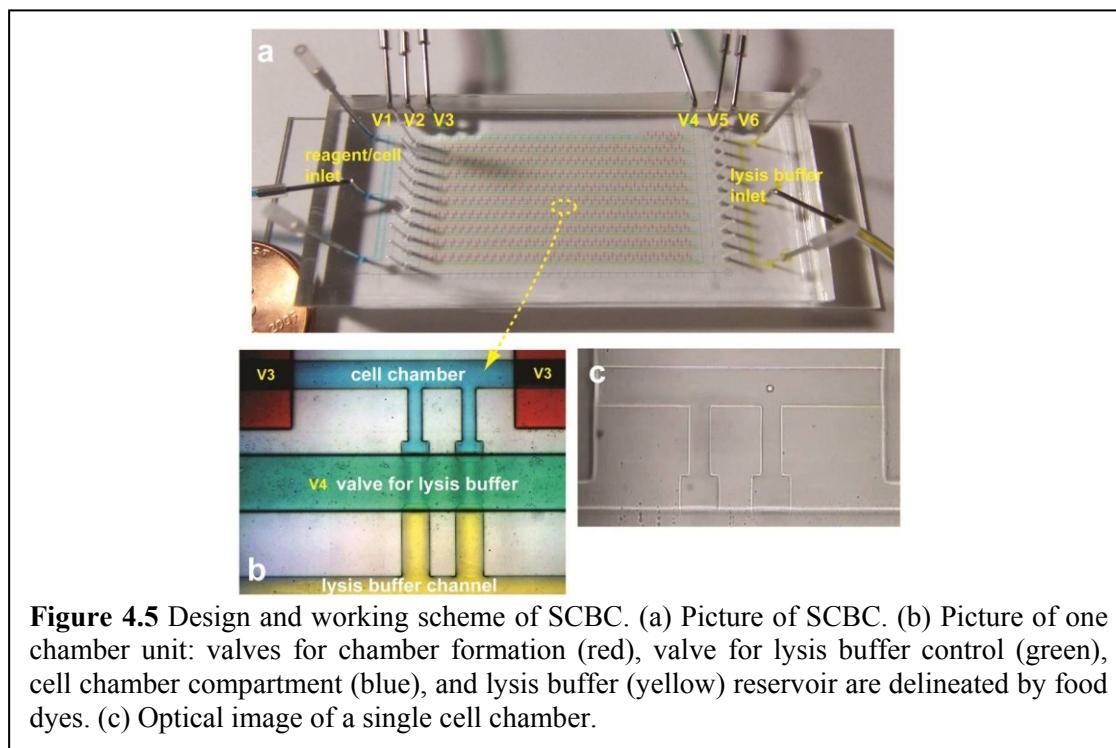
4.2.7 DEAL based cell capturing and viability test

EGFR + and EGFRvIII + cells were sorted from GBM39 neurospheres by DEAL technology[21]. DEAL arrays were blocked with 1% BSA in PBS solution (g/mL) for 30 minutes, washed in PBS and deionized water and incubated with oligo-Cetuximab (Bristol-Myers) conjugate for 30 minutes, at 37°C. Single cell preparation of GBM39 cells,

suspended in the culture medium, were applied to the array for 40 minutes, on ice. The

DEAL arrays were then washed with BSA 0.1% in PBS (g/mL), a solution 1:2 of trypan

blue in PBS (vol/vol) was applied on the captured cells that were then covered with a cover

slip. Dead cells were visualized and counted as the number of trypan blue positive cells,

with a Nikon Eclipse TS100 scope (Fig 4.7).

Fifteen thousand of GBM39 cells were seeded in 12 well plates and, after 24 hours,

Figure 4.6 Calibration curves of the assayed protein panel. The calibrations were carried out using SCBC under the same condition to those single cell experiments, except that standard proteins were used. The calibration curve for p53 couldn't be generated due to the non-specific binding of recombinant standard protein.
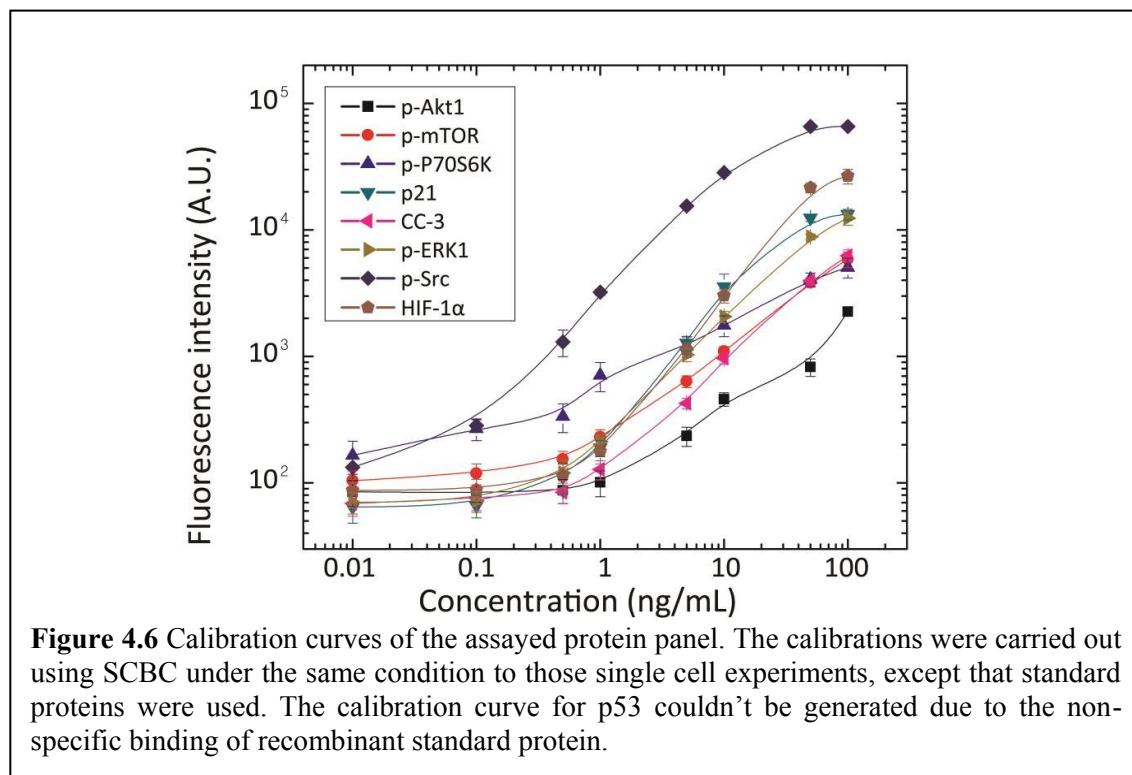
treated with CC214-1 2 μM, U0126 5 μM, Dasatinib 100 nM, for 24 hours. Cell viability

and cell death were evaluated using Bio-Rad TC-20 cell counter. Representative images of

the cells were taken using a Nikon Eclipse TS100 scope equipped with Canon S51S camera

(Fig. 4.8). Student's T-test was used to assess statistical significance. The variation between

the sample sets was similar and expressed as standard deviation.

4.2.8 Single Nucleotide Polymorphism (SNPs) analysis

Two controls and two CC214-2 resistant xenografts were used for the DNA extractions, using the Qiagen QIAamp DNA Mini kit protocol. Few microgram (1.5-2 µg) of DNA from control and resistant samples were analyzed by Affymetrix SNP 6.0 array
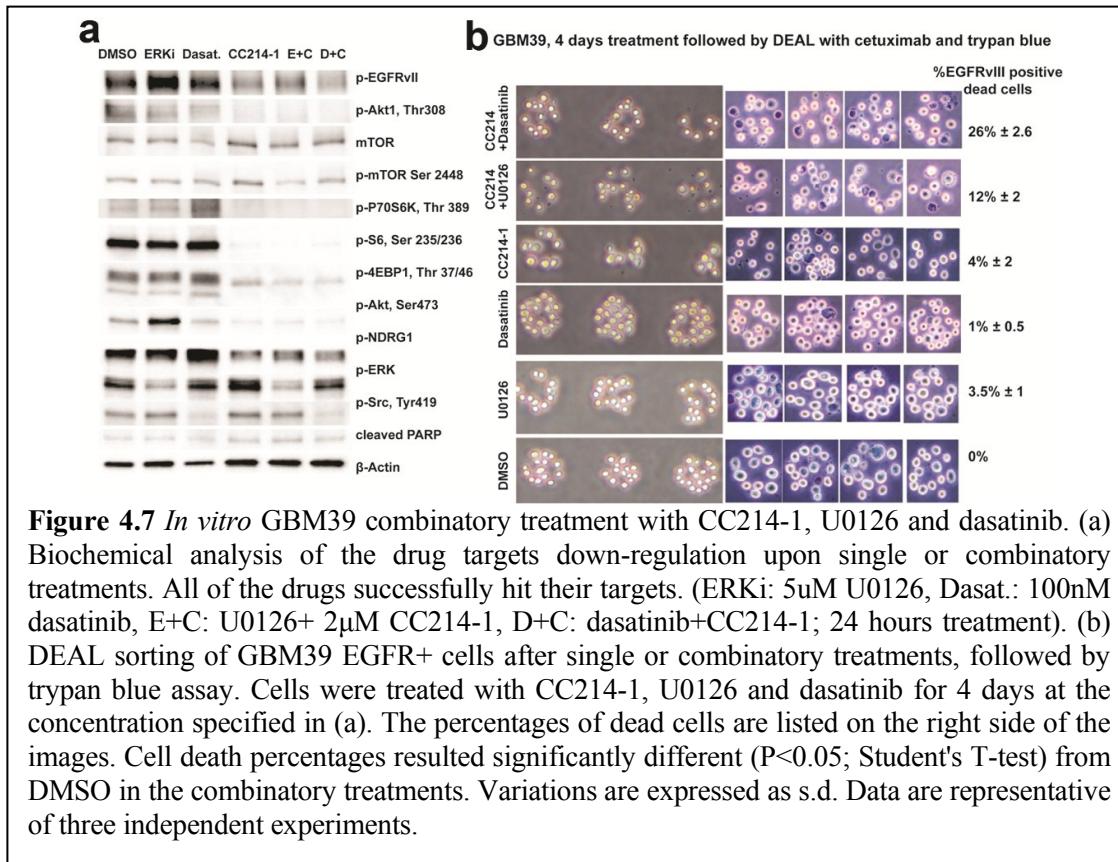
Figure 4.7 *In vitro* GBM39 combinatory treatment with CC214-1, U0126 and dasatinib. (a) Biochemical analysis of the drug targets down-regulation upon single or combinatory treatments. All of the drugs successfully hit their targets. (ERKi: 5uM U0126, Dasat.: 100nM dasatinib, E+C: U0126+ 2µM CC214-1, D+C: dasatinib+CC214-1; 24 hours treatment). (b) DEAL sorting of GBM39 EGFR+ cells after single or combinatory treatments, followed by trypan blue assay. Cells were treated with CC214-1, U0126 and dasatinib for 4 days at the concentration specified in (a). The percentages of dead cells are listed on the right side of the images. Cell death percentages resulted significantly different (P<0.05; Student's T-test) from DMSO in the combinatory treatments. Variations are expressed as s.d. Data are representative of three independent experiments.

(200kb filter, 50 markers) at the Clinical Microarray Core, University of California Los Angeles. The DNA sequences used for the Genome-wide Affymetrix SNP6.0 array sequence analysis have been deposited to Gene Expression Omnibus (GEO). The accession code is GSE53042.

4.2.9 *In vivo* drug treatment
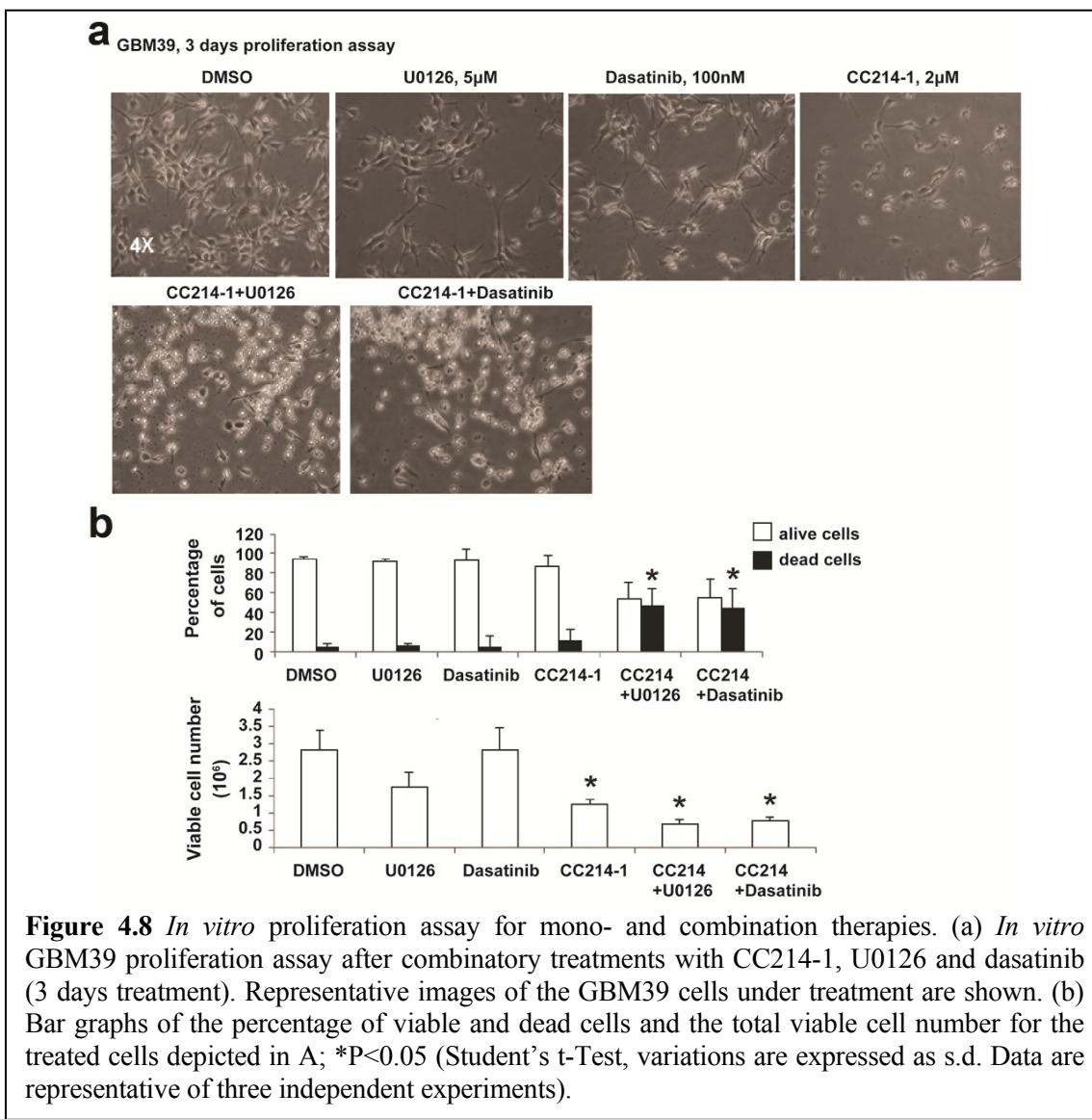
Figure 4.8 *In vitro* proliferation assay for mono- and combination therapies. (a) *In vitro* GBM39 proliferation assay after combinatory treatments with CC214-1, U0126 and dasatinib (3 days treatment). Representative images of the GBM39 cells under treatment are shown. (b) Bar graphs of the percentage of viable and dead cells and the total viable cell number for the treated cells depicted in A; *P<0.05 (Student's t-Test, variations are expressed as s.d. Data are representative of three independent experiments).

For the drug treatments, CC214-2 was administered by oral gavage, 100mg/kg, once every two days, in a suspension containing 0.5% carboxymethylcellulose (Sigma), 0.25% Teewn-80 (Sigma) in nanopore water. Dasatinib (Selleckchem) was administered by oral gavage, 30mg/kg, once every two days, dissolved in the CC214-2 suspension. U0126 (Selleckchem) was administered by intra-peritoneal injection, 25 μmol/kg, once every two days, in a suspension containing 40% DMSO (vol/vol, Fisher) in PBS (Cellgro). The injection of 1mL saline 0.9% NaCl (Baxter) was used if signs of weight loss were

registered. Mice were euthanized when tumors reached 15 mm diameter. To determine

the number of animals requested, we carried out power calculations using STATA software

(version 8), performed the Monte Carlo simulation command (simpower) and determined

the sample size to detect a significant difference in our tumor size comparison study. To

ensure statistical significance of drug effects, we used a sufficient, but not excessive,

sample size: for vehicle treated controls, CC214-2 responsive and CC214-2 resistant mice n

= 11, 7, 7 respectively for each group; for Dasatinib, U0126, either alone or in

combination, and for the combinatory treatments with CC214-2, n = 4 for each group.

4.3 PHYSICAL APPROACHES

4.3.1 Collective behaviors in signaling coordination: singling modes hypothesis

The typical topological structure of the signaling networks is comprised of many different wires interconnecting various signaling proteins (nodes). Each wire represents a protein-protein pairwise interaction between two signaling nodes (Fig. 4.9). The structure
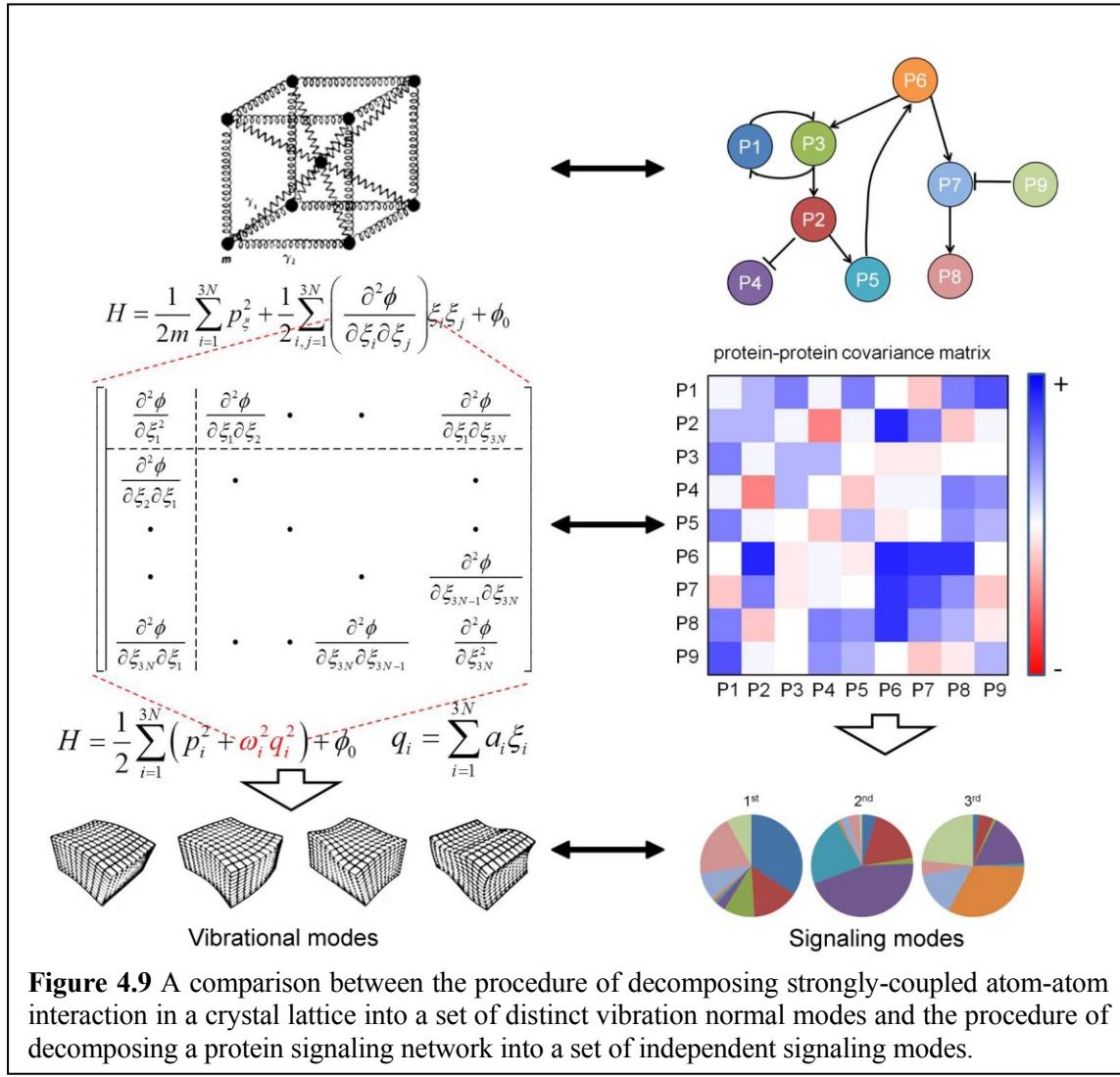

Figure 4.9 A comparison between the procedure of decomposing strongly-coupled atom-atom interaction in a crystal lattice into a set of distinct vibration normal modes and the procedure of decomposing a protein signaling network into a set of independent signaling modes.

can be very complicated if many signaling protein are involved, which makes it hard to

quantitatively trace and predict the exact outcome from a signaling input. Physical and computational models are therefore needed to structure the deluge of data and to make them accessible to meaningful interrogation and analysis[22].

By analogy with the study of atomic interactions in a crystal lattice where the atom-atom pairwise interactions can be greatly simplified by diagonalizing the system potential matrix to decompose the strongly coupled atomic interactions into a series of independent normal vibrational modes (Fig. 4.9), a simplified picture of signaling network coordination can also be achieved by diagonalizing protein-protein correlation or covariance matrices to decompose the pairwise interactions into a set of distinct linear combinations of signaling proteins (i.e., independent signaling modes). The protein-protein covariance matrices are the strength measure of the protein interactions. The composition of each mode is a linear combination of a panel of proteins that are the key nodes of the signaling network under study (i.e., eigenvectors) and the strength of each mode is the corresponding eigenvalue (Fig. 4.9). Interrogating how the strength and composition of the dominative signaling modes evolve with the tumor progression (treatment naïve, drug responding and resistant) and/or respond to external perturbations (different drug treatments) will lead a conceptual advance of viewing and simplifying the protein signaling network and provides the basis of developing physico-chemical theory for single cell proteomics with predictive capacity.

4.3.2 Principal component analysis (PCA)

An optimal tool for extracting the dominative signaling modes from the single cell proteomic dataset is principal component analysis (PCA). In this chapter, PCA is carried

out carried out for single cell data from tumors at all three stages, as well as the three clustered subgroups identified in control sample. Each column of the dataset is mean-centered and divided by the standard deviation to form a standardized dataset first. A normalized PCA (coded in R) is used to peel off layer after layer of systematic co-variations from the data, in terms of principal components (PCs). Different principal components typically capture different parts of the cell responses—for example, one might correlate best with cell division, whereas another correlates best with cell death[22]. The correlations between functional protein levels and PCs are calculated to quantify the dominative protein pattern of the signaling network coordination and its response to external perturbations such as drug treatment. The square of a protein projection on a PC defines the contribution of that protein to that PC. The contributions of each functional protein to the first three PCs for control sample are calculated and shown in the pie charts of Fig. 4.13. In the subgroup with most mTOR activity, different groups of functional protein preferentially occupy different PCs and become orthogonal to one another (Figs. 4.12 c and 4.13), which implies potentially independent signaling modes are active within the same cellular subpopulations.

4.3.3 Quantifying the functional heterogeneity

Cancer cells are stable populations existing in the presence of large heterogeneity. The population is stable exactly because it is heterogeneous. Single cell proteomic measurements capture the protein fluctuations, while simultaneously providing a measure of the stability of a tumor that is comprised of those cells, and providing a bridge to statistical physics models with predictive capacity. The heterogeneity can be quantified in

many different ways. In this chapter, we apply agglomerative hierarchical clustering (AHC) analysis that is applied by XLSTAT software (Addinsoft) on the single cell data extracted from control, responsive and resistant tumors respectively (Fig. 4.13). The proximity among single cell observations is measured by the dissimilarity coefficients of Euclidian distance. Ward's minimum variance method is employed as a strategy to calculate the dissimilarity in order to minimize the total within-cluster variance and thus keep each clustered group as homogeneous as possible[23]. The truncation level is determined automatically by the software based upon the entropy and tries to create homogeneous groups. The calculated dissimilarity coefficients are used as indices for quantifying the functional heterogeneity of the tumor. As we can see in Fig. 4.12 b, this heterogeneity index correlates pretty well with the tumor progression during the course of CC214-2 treatment.

4.3.4 Partial least square (PLS) modeling of immunohistochemical data

To demonstrate the relationships between single cell proteomic data and tumor response to different therapy combinations, we use a PLS model to ask whether an IHC tissue analysis of the mouse models explored in Fig. 4.14 b could yield a deeper understanding of the SCBC data analysis.

The digitized IHC data under different drug combinations was used to establish the explanatory metrics **X** (independent metrics) (Figs. 4.10 and 4.15). Prior to establishing the corresponding dependent metrics **Y**, two characteristic terms that can be directly extracted from the tumor growth curves were introduced. One is the transitory growth rate (TGR) at

the time of sacrifice, which is defined as the average percentage of tumor volume change per day of the last three time points available on the growth curves, except for cases C+U DR, C+D DR and V2 where last two time points were used instead of three. The other term is the cell cycle measure (time constant τ) that can be extracted by fitting the growth curves (Fig. 4.15 a) with the exponential growth function. The tumor volume in the growth curves was modeled by the exponential growth function $V_t = V_0 + \varepsilon + V_0' e^{t/\tau}$, where V_t is the normalized tumor volume over time. V_0 can be understood as the portion of the solid tumor that was not engaged in tumor growth and V_0' was the portion that was engaged in the tumor growth[24]. $\varepsilon \sim N(0, \sigma)$ is an error term from a normal distribution with mean zero and SD σ. At the initial state (t=0), $V_t = V_0 + V_0' = 1 - \varepsilon$. The application of effective drug combinations can significantly shut down the tumor progression and thus lead to large values of the cell cycle measure τ as expected (Fig. 4.10). Please note that the functional form used here is an oversimplified model without separately considering the initial/consistent cell death caused by the drugs or any phenotypical switch and emergence of resistance during the treatment. Only the single metric τ was used to roughly assess the drug effect. But it still captures the essential information to disclose the independent signaling modes as discussed in the paper.

The PLS model was constructed in XLSTAT software (Addinsoft) according to the following iterative formulae[25]:

$$\mathbf{w_a} = \text{first eigenvector of } (\mathbf{E'_{i-1}F_{i-1}F'_{i-1}E_{i-1}})$$

$$\mathbf{t_i} = \mathbf{E_{i-1}w_i}$$

$$\mathbf{E_i} = \mathbf{E_{i-1}} - \mathbf{t_i}\mathbf{p_i'}$$

$$\mathbf{F_i} = \mathbf{F_{i-1}} - \mathbf{t_i}\mathbf{q_i'}$$

$$\mathbf{p_i'} = (\mathbf{t_i't_t})^{-1}\mathbf{t_i'E_{i-1}}$$

$$\mathbf{q_i'} = (\mathbf{t_i't_t})^{-1}\mathbf{t_i'F_{i-1}}$$

where the $\mathbf{E_i}$ represents the residue of the ith principal component of the explanatory metrics with the score vector $\mathbf{t_i}$, loading vector $\mathbf{p_i}$ while $\mathbf{F_i}$ represents the residue of the ith principal component of the dependent metrics with score vector ti and loading vector $\mathbf{q_i}$. $\mathbf{w_i}$ is the loading weight that strikes a balance between modeling $\mathbf{X}$ and modeling $\mathbf{Y}$. The prime represents the matrix transpose. The residue matrices $\mathbf{E_0}$ and $\mathbf{F_0}$ just contain the mean-centered $\mathbf{X}$- and $\mathbf{Y}$- variables. The regression coefficient matrix that leads out the functional form between $\mathbf{X}$ and $\mathbf{Y}$ can be calculated as $\mathbf{B_h} = \mathbf{W_h}(\mathbf{P_h'W_h})^{-1}\mathbf{Q_h'}$, where h is the number of principal components used in the model.

Eight observations (V1, C, D, U, C+D, D+U, C+U and C-R) were employed to establish the calibration phase of the model. The stability and predictive quality of the model were assessed by calculating Q^2cum index that involves the predicted residual sums of squares (PRESS) statistic and sum of squares of error (SSE) for a model with one less component[26].

$$Q^2cum(h) = 1 - \prod_{j=1}^{h} \frac{\sum_{k=1}^{n} PRESS_{kj}}{\sum_{k=1}^{n} SSE_{k(j-1)}}$$

PRESS statistic requires a leave-one-out cross-validation. The Q^2cum index measures the global contribution of the h first principal components to the predictive capacity of the model. As a result, the optimal number of principal components used in the model can be determined with respect to this index. The first two PCs yielded the highest

				Explanatory Metrics								Dependent Metrics	
	Tunel	Ki-67	p-Src	p-ERK1	HIF-1α	p-P70S6K	p-S6	p-4EBP1	p-AKT	p-NDRG1		TGR at sacrifice (%)	Cell cycle measure (τ) (day)
V1	0.17	38.95	14.93	17.52	14.23	12.24	22.05	13.49	11.07	19.57		19	6.61
C	1.46	25.96	19.98	11.49	12.71	8.78	15.90	9.47	7.51	11.94		7.7	8.3
D	0.22	40.37	15.46	17.24	13.89	11.24	18.98	13.56	10.67	23.15		8.6	6.56
U	0.18	41.41	13.00	7.68	14.76	10.71	17.19	14.21	9.76	22.91		16	6.02
U+D	0.29	40.32	8.10	6.81	10.89	10.55	17.19	12.93	11.30	22.65		12.7	7.01
C+D	1.75	29.30	9.26	13.75	13.08	8.68	11.76	8.36	5.82	5.06		7.1	17.98
C+U	1.23	27.53	16.77	11.11	13.09	9.03	10.56	9.09	5.88	4.72		5.28	15.99
C-R	0.17	40.20	19.32	25.89	16.52	10.29	23.19	20.22	16.90	26.47		18.5	5.84
C+D DR	0.07	49.91	18.61	15.05	10.59	12.62	18.55	17.60	10.20	21.37		33	2.73
C+U DR	0.10	52.60	14.57	15.35	14.52	10.93	19.59	12.52	9.69	21.70		15.7	5.45
V2	0.13	43.44	14.82	17.13	13.40	11.50	18.34	12.86	11.02	21.29		16.59	6.58

Figure 4.10 Explanatory and dependent matrices used for PLS modeling. The explanatory matrix is composed of numerical IHC data for samples treated under different drug combinations, including V1 for vehicle, C for CC214-2 treated, D for dasatinib treated, U for U0126 treated, U+D for combinatory treatment of U0126 and dasatinib, C+D for CC214-2 and dasatinib, C+U for CC214-2 and U0126, C-R for resistant stage of CC214-2 treated sample, C+D DR for C+D treated samples after drug removal, C+U DR for C+U treated samples after drug removal and V2 for the control sample in the first CC214-2 only mouse experiment. The values of TGR at sacrifice and cell cycle measure are directly extracted from the corresponding growth curves under different drug treatments. The orange part of the table represents the calibration phase of the model and blue part represents the prediction phase.

Q^2cum index for this model and thus were employed in the subsequent calculations.

In the prediction phase, the established model was used to predict the TGR at sacrifice and cell cycle measure τ for the observations C+D DR, C+U DR and V2. The predicted values were compared against the observed values extracted from the growth curves and

shown in Fig. 4.16 a. The good match between observation and prediction further

validates the model reliability and stability.

4.4 RESULTS AND DISCUSSION

4.4.1 Single-cell proteomic analysis of three drug treatment stages

Figure 4.11 Single-cell proteomic data for three drug treatment stages (a) SCBC data represented as one-dimensional scatter plots (n=133, 62, 143 for control, responsive resistant groups) for each protein at each test condition. The averaged fluorescence intensity with standard error of the mean (SEM) is overlaid for each protein. Statistical uniqueness is evaluated by two-tailed Mann-Whitney test for pairwise comparison (black stars) and Kruskal-Wallis ANOVA for comparison among three groups (blue stars, NS, not significant; *P < 0.05; **P < 0.005; ***P < 0.0005). (b) Immunoblots of various proteins from bulk assays for the control and resistant states of GBM39.

The   single cell proteomic data is shown in Fig. 4.11. Comparing the SCBC

measurements with immunohistochemical (IHC) and immunoblot analyses performed on

bulk tumor samples confirmed that CC214-2 significantly suppressed mTORC1 and

mTORC2 signaling, which became reactivated during acquired resistance (Figs. 4.1 c and
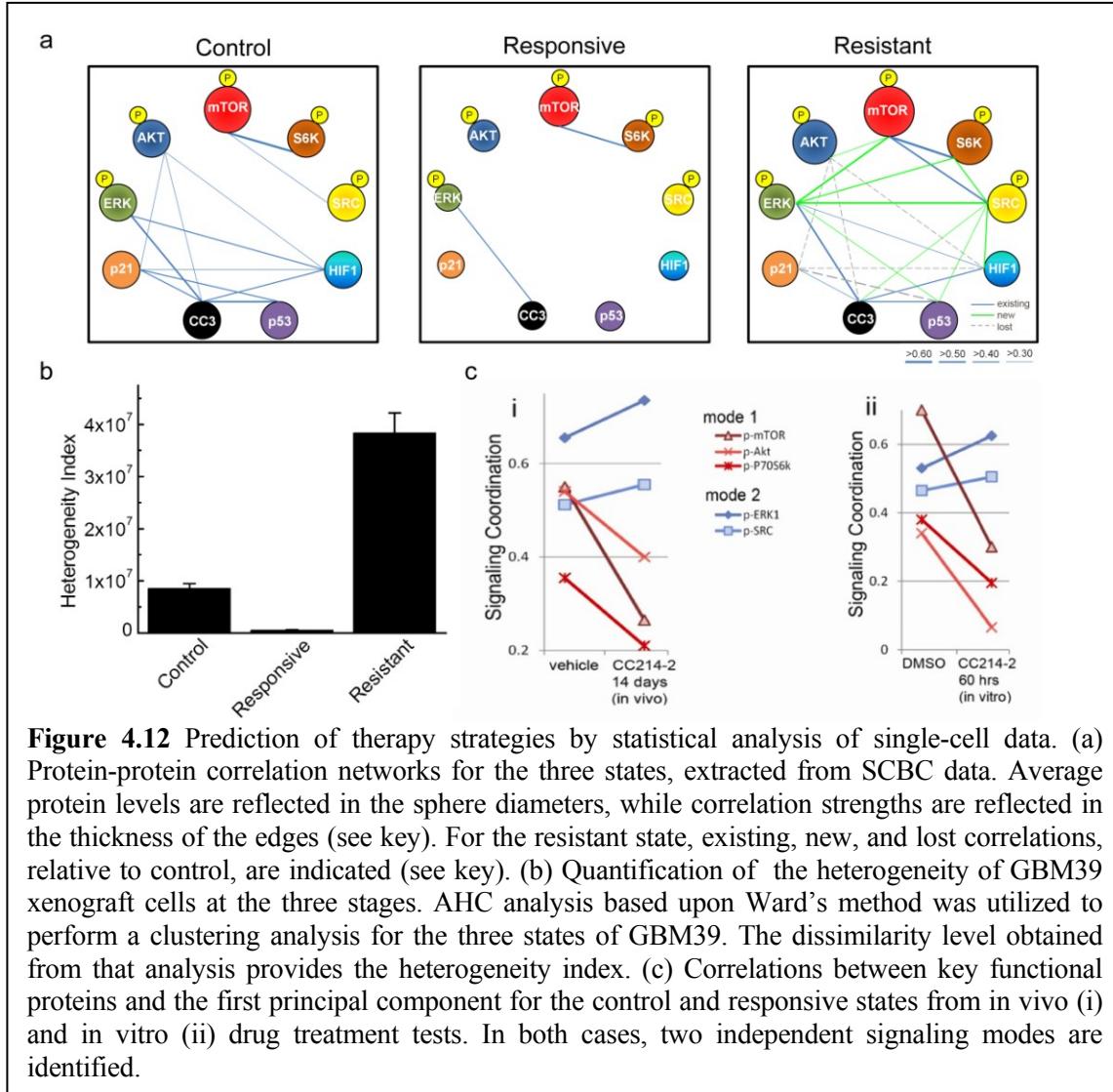
Figure 4.12 Prediction of therapy strategies by statistical analysis of single-cell data. (a) Protein-protein correlation networks for the three states, extracted from SCBC data. Average protein levels are reflected in the sphere diameters, while correlation strengths are reflected in the thickness of the edges (see key). For the resistant state, existing, new, and lost correlations, relative to control, are indicated (see key). (b) Quantification of the heterogeneity of GBM39 xenograft cells at the three stages. AHC analysis based upon Ward's method was utilized to perform a clustering analysis for the three states of GBM39. The dissimilarity level obtained from that analysis provides the heterogeneity index. (c) Correlations between key functional proteins and the first principal component for the control and responsive states from in vivo (i) and in vitro (ii) drug treatment tests. In both cases, two independent signaling modes are identified.

4.11 b and Appendix B: Supplementary Tables).

The SCBC data set enabled statistical analyses of the functional protein correlations in

tumor cells for all three states. This analysis provides insight into how changes in signaling

coordination, rather than just protein levels, might be implicated in acquired resistance.

During response, CC214-2 profoundly diminished the levels of most proteins (Fig. 4.11 b

and the sphere sizes in Fig. 4.12 a), as well as the protein signaling coordination (the

loss of network edges of Fig. 4.12 a). This is reflected in a near 10-fold drop in the

functional heterogeneity of the cell population (Fig. 4.12 b). The functional heterogeneity

index, defined as the dissimilarity value in the agglomerative hierarchical clustering (AHC)

analysis based upon Ward's minimum variance method (See Chapter 4.3.3), is a metric of

the dispersion of the functional protein levels across all single cell assays[23]. Note the

appearance of 9 new interactions (green edges in Fig. 4.12 a), 8 of which are associated

with phospho(p)- extracellular-signal-regulated kinases 1 (p-ERK1) and p-proto-oncogene

tyrosine-protein kinase (p-Src). These results point to the possibility that a gain of function

through ERK/Src might be leading to CC214-2 resistance by promoting downstream

mTOR signaling[27-29]. Additionally, acquired resistance was associated with a sharp

increase in functional heterogeneity, an outcome inconsistent with clonal selection of a

resistant subpopulation, suggesting that mTOR kinase inhibitor resistance may be mediated

by signaling network adaptation.

4.4.2 Signaling modes extraction by PCA predicts effective therapy strategies

To clarify how the protein coordination was altered by mTOR kinase inhibition, we performed principal component analysis (PCA), using the two-dimensional matrix of measured protein-protein covariances as input. The PCA analysis of the protein-protein covariance matrix reflects protein signaling coordination, rather than absolute protein

Figure 4.13 Clustering and PCA of single-cell data for control and responsive tumor samples. Agglomerative Hierarchical Clustering (AHC) based upon Ward's method was used for the clustering analysis. (a) Three major subpopulations with different functional phenotypes were observed for control sample. The subpopulation with magenta color shows the highest mTOR activity. (b) The AHC and PCA analysis for responsive sample. The greatly elevated p-ERK1/p-Src partition in PC1 implies that ERK/Src signaling, that has less prominent regulatory activity in control (PC3), takes over the signaling in responsive stage due to the repression of the previously dominant regulatory proteins associated with mTORC1/C2.

Figure 4.14 *In vivo* validation of 7 mono- or combination therapies. (a) In vivo test results for the 7 mono- or combination therapies based upon the predictions from the SCBC data analysis. All 7 predictions proved correct. P values derived from Student's T-test (variations expressed as s.d.; n= 11 for controls, n= 6 for CC214-2 group, n= 4 for Dasatinib group, n= 4 for U0126 group, n= 4 for each combinatory treatment group) (b) IHC images of drug targets for the combinatory treatments of CC214-2 and ERK and/or Src inhibitors, attesting the successful down-regulation of all of the drug targets (digitized results and statistical analysis shown in the Appendix B: Supplementary Tables). Scale bar: 100μm.

levels. The pie charts at the top of the plot represent the composition of the 3 top principal components (PCs) of the full single cell dataset. The 1st PC contains many contributions, reflecting the highly interconnected correlation network of Fig. 4.12 a. However, the 2nd and 3rd PCs are dominated by mTORC1 signaling (p-mTOR and p-P70S6K) and ERK/Src signaling, respectively. PCA analysis of the subpopulation most characterized by mTOR

signaling (purple) is striking. PC1 (capturing 26% of the total variance) is dominated by

mTORC1 signaling (p-mTOR and p-P70S6K); PC2 (21% of the variance) is dominated by

mTORC2 signaling (p-Akt S473) and PC3 (16% of the variance) is dominated by p-ERK1
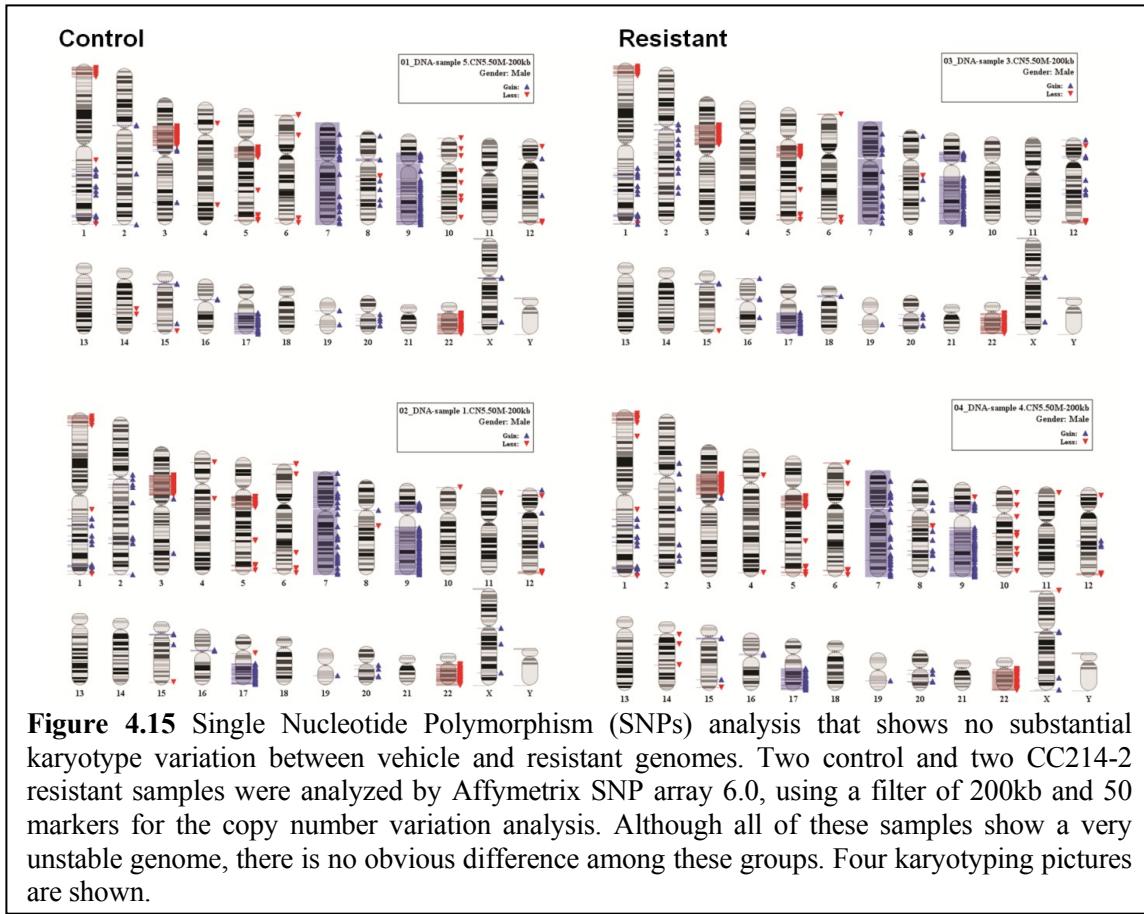
Figure 4.15 Single Nucleotide Polymorphism (SNPs) analysis that shows no substantial karyotype variation between vehicle and resistant genomes. Two control and two CC214-2 resistant samples were analyzed by Affymetrix SNP array 6.0, using a filter of 200kb and 50 markers for the copy number variation analysis. Although all of these samples show a very unstable genome, there is no obvious difference among these groups. Four karyotyping pictures are shown.

and p-Src signaling. This implies potentially independent signaling modes are active

within the same cellular subpopulations.

We then calculated the correlations between the assayed proteins and the first principal

component (PC1) in response to CC214-2. PC1 captures the most essential feature of the

signaling network, and so this analysis estimates the influence of a given protein on

signaling coordination. For the control tumor, PC1 is populated by p-ERK1, p-Src, p-Akt1,

p-mTOR, and p-P70S6 kinase (p-P70S6K, Fig. 4.12 ci). In response to CC214-2, these

5 proteins split into 2 groups, or modes (a term we use to imply collective behavior). The

influence of p-Akt1, p-mTOR, and p-P70S6K (red mode of Fig. 4.12 ci) on signaling

coordination, relative to p-ERK1 and p-Src (blue mode), is diminished. Among the proteins

whose levels are altered by treatment with CC214-2, the signaling coordination associated

with mTORC1/C2 signaling was repressed while the signaling coordination associated with

MAPK/ERK and Src signaling were increased. This observation suggested that the latter

may have gained the ability to maintain signal flux to key downstream mTOR effectors

(Fig. 4.13). These data suggest that targeting any one of the 5 proteins is unlikely to exhibit

a strong effect, as would targeting two proteins from the same mode. However,

simultaneously targeting 1 protein from each mode would be predicted to constitute an

effective therapy. Correspondingly, we treated mice implanted with GBM39 using

combinations of CC214-2 (C), dasatinib (D, Src inhibitor), and U0126 (U, MEK/ERK

inhibitor) to test four therapies or therapy combinations expected to be ineffective, and 3

therapy combinations expected to be effective.

4.4.3 *In vivo* validation of predicted therapy strategies

Consistent with our model, treatment with C, D, U or D + U could not induce long

term tumor growth inhibition (Figs. 4.7, 4.8 and 4.14 a). In contrast, combining C with

either D and/or U completely suppressed tumor growth in vivo, with no adverse effects of

either the tumor or the treatment. Treatment was stopped after 47 days with no sign of

recurrence (Fig. 4.14 a). Removal of combination therapy resulted in rapid tumor regrowth

(Fig. 4.14 a). This was concomitant with reactivation of signaling pathways (Fig. 4.14 b,

Appendix A: Supplementary Figures and Appendix B: Supplementary Tables). These results demonstrate that in a therapeutically representative in vivo model of one of the most aggressive and treatment resistant of all human cancers, GBM39[12], long-term disease remissions can be induced and sustained if independent signaling modes are sufficiently inhibited.

4.4.4 *In vitro* perturbation identifies the fast network adaptation mechanism

We next addressed the question of whether the resistance to CC214-2 that appears to be mediated by ERK/Src signaling arose from a clonal subpopulation of tumor cells, or from an adaptation of the signaling networks within those same tumor cells that responded to CC214-2. For that experiment, we treated the EGFR+ cells resected from a GBM39 control with 2μM solutions of CC214-1 (in vitro use) for 60 hours. We reasoned that 60 hours of in vitro treatment was sufficient for the tumor cells to establish a new steady state. The treatment did not induce significant cell death, and the GBM39 tumor cells do not divide during the 60 hour exposure to CC214-1 (Fig. 4.8). The results, shown in Fig. 4.12 c (ii), are consistent with the in vivo observations, and clearly indicate that resistance arises from adaptation of the cancer cells to the drug, rather than from the emergence of a subpopulation of drug-resistant cells. Single Nucleotide Polymorphism (SNP) analysis of the vehicle and resistant genomes detected no substantial karyotype variation between the two, providing further evidence of an adaptive mechanism (Fig. 4.15). The in vitro recapitulation of in vivo observations implies such an analysis for identifying effective therapy combinations could potentially be carried out on biopsied tissue from GBM patients, within a clinically relevant time-scale.

4.4.5 PLS modeling on tissue analysis independently confirms the signaling modes hypothesis.

We used a Partial Least Squares (PLS) model to ask whether an IHC tissue analysis of the mouse models explored in Fig. 4.14 b could yield a deeper understanding of the SCBC

Figure 4.16 PLS modeling confirms the independent signaling modes as effective combination therapies. (a) Validation of the PLS modeling: The calibration phase of the model was constructed by using part of observations (orange part). The first two PCs were used to perform a leave-one-out cross-validation to assess the model stability. The established model was then employed to predict the TGR at sacrifice and cell cycle metric τ for the remaining observations (blue part). The prediction from the model matches well with the observed values extracted from the growth curves in Figure 4a, showing the validity of the model. (b) The correlations of the IHC assayed proteins, as well as the functional observations, with PC1 and PC2. TGR strongly correlates with oncogenic kinases such as p-P70S6K, p-Akt1 and p-4EBP1. p-Src and p-ERK1 largely dominate PC2, constituting an independent signaling mode that accounts for the resistance of mTOR kinase inhibitor. (c) PLS modeling shows that, for effective drug combinations, it projects qualitatively differently from linear superposition of individual drugs, which in turn implies that synergistic drug combinations do not simply act in a linearly additive manner.

data analysis of Figs. 4.12 c and 4.13. (See Chapter 4.3.4). Such an analysis can demonstrate biological relationships between different classes of information[30, 31]. The approach is similar to PCA, but seeks to identify those PCs of one data set (the IHC analysis) that can best predict a second data set (functional observations of the tumor). Quantitative IHC assays of a panel of functional proteins assayed from resected tumor tissues from the control model and for all 7 tested therapy combinations, at the time points of sacrifice, were loaded into an explanatory matrix (Figs. 4.10 and 4.16). The dependent matrix was constructed from two characteristic parameters of tumor growth under different therapies in Fig. 4.14 a. One is transitory growth rate (TGR) at sacrifice, defined as the average percentage of tumor volume change per day of the last three time points measured before the sacrifice. The other one is a cell cycle metric that is the time constant τ extracted by fitting the growth curves with the exponential growth function (see Chapter 4.3.4 and Fig. 4.10 for detail).

The PLS model was developed using a subset of the treatment and treatment combinations (Figs 4.10 and 4.16 a, orange shaded data sets), and then validated through predictions of the remaining measurements (Figs 4.10 and 4.16 a, blue shaded data sets). The agreement between predictions and observations supports the validity of the model. In Fig. 4.16 b we plot the correlations of the assayed proteins, as well as the functional observations, with the x- and y-axis as the 1st and 2nd PCs, respectively. The mTORC1/C2 associated proteins that constitute mode 1 of Fig. 4.12 c lie principally along PC1, as does the TGR. This means that the TGR is largely predicted using just the IHC measurements of the mTORC1/C2 proteins. Importantly, the TGR correlates with the phosphorylation of

ribosomal protein S6 (p-S6), the eukaryotic translation initiation factor 4E-binding protein 1 (p-4EBP1) and the oncogenic protein kinase Akt (p-Akt), which are key factors regulating protein translation and cellular proliferation. The cell cycle measure τ is anti-correlated with the TGR and the mTORC1/C2 associated proteins, but also lies largely along PC1. These relationships indicate that mTORC1/C2 and their effectors are the primary drivers that account for tumor growth. The functional proteins p-ERK1 and p-Src constitute a second group largely aligned along PC2, with only weak relationships to proteins in the first group. This is consistent with a resistance mechanism associated with MAPK/ERK1 signaling and Src signaling. It also provides independent confirmation of the two signaling modes that are pointed to in the single cell analysis.

The PLS model also allows for a comparison of the therapy combinations, with each combination represented by its respective IHC data set (Appendix B: Supplementary Tables). Examination of the projections of C, U and C+U (Fig. 4.16 c) on the first two PCs indicate that C+U projects qualitatively differently from the linear superposition of single-input C and U, as does the therapy combination C+D, indicating cooperative effect of the two drugs that are non-linear. The therapeutically ineffective combination of D+U, however, is very close to the linear superposition of D and U.

4.5 CONCLUSION

Our data provide evidence that tumor cells can respond to a therapy by adopting a new steady state that restores the tumor growth characteristics temporarily disrupted by the therapy. This resistance mechanism is pre-existing; the same cells that respond to the therapy also adapt and develop resistance to it. For the human-derived GBM39 model explored here, a single cell proteomics analysis of the phosphoprotein signaling networks associated with tumor growth can resolve the independent signaling modes that drive tumor growth in both the untreated and drug resistant states. Such analyses can be rapidly carried out using untreated tumor biopsies, and so may represent a new approach for guiding the selection of targeted combination therapies that can anticipate resistance, and thereby lead to the induction of sustained long term disease remission.

4.6 REFERENCES

1. Masui, K. et al. A tale of two approaches: complementary mechanisms of cytotoxic and targeted therapy resistance may inform next-generation cancer treatments. *Carcinogenesis* **34**, 725-38 (2013).

2. Bozic, I. et al. Evolutionary dynamics of cancer in response to targeted combination therapy. *Elife* **2**, e00747 (2013).

3. Al-Lazikani, B., Banerji, U. & Workman, P. Combinatorial drug therapy for cancer in the post-genomic era. *Nat Biotechnol* **30**, 679-92 (2012).

4. Wacker, S.A., Houghtaling, B.R., Elemento, O. & Kapoor, T.M. Using transcriptome sequencing to identify mechanisms of drug action and resistance. *Nat Chem Biol* **8**, 235-7 (2012).

5. Cancer Genome Atlas Research, N. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061-8 (2008).

6. Parsons, D.W. et al. An integrated genomic analysis of human glioblastoma multiforme. *Science* **321**, 1807-12 (2008).

7. Lee, M.J. et al. Sequential application of anticancer drugs enhances cell death by rewiring apoptotic signaling networks. *Cell* **149**, 780-94 (2012).

8. Krakstad, C. & Chekenya, M. Survival signalling and apoptosis resistance in glioblastomas: opportunities for targeted therapeutics. *Mol Cancer* **9**, 135 (2010).

9. Elkabets, M. et al. mTORC1 inhibition is required for sensitivity to PI3K p110alpha inhibitors in PIK3CA-mutant breast cancer. *Sci Transl Med* **5**, 196ra99 (2013).

10. Yaffe, M.B. The scientific drunk and the lamppost: massive sequencing efforts in cancer discovery and treatment. *Sci Signal* **6**, pe13 (2013).

11. Cloughesy, T.F. & Mischel, P.S. New strategies in the molecular targeting of glioblastoma: how do you hit a moving target? *Clin Cancer Res* **17**, 6-11 (2011).

12. Sarkaria, J.N. et al. Identification of molecular characteristics correlated with glioblastoma sensitivity to EGFR kinase inhibition through use of an intracranial xenograft test panel. *Mol Cancer Ther* **6**, 1167-74 (2007).

13. Giannini, C. et al. Patient tumor EGFR and PDGFRA gene amplifications retained in an invasive intracranial xenograft model of glioblastoma multiforme. *Neuro Oncol* **7**, 164-76 (2005).

14. Gini, B. et al. The mTOR kinase inhibitors, CC214-1 and CC214-2, preferentially block the growth of EGFRvIII-activated glioblastomas. *Clin Cancer Res* **19**, 5722-32 (2013).

15. Wei, W. et al. Hypoxia induces a phase transition within a kinase signaling network in cancer cells. *Proc Natl Acad Sci U S A* **110**, E1352-60 (2013).

16. Shi, Q. et al. Single-cell proteomic chip for profiling intracellular signaling pathways in single tumor cells. *Proc Natl Acad Sci U S A* **109**, 419-24 (2012).

17. Mortensen, D.S. et al. Use of core modification in the discovery of CC214-2, an orally available, selective inhibitor of mTOR kinase. *Bioorg Med Chem Lett* **23**, 1588-91 (2013).

18. Fueger, B.J. et al. Impact of animal handling on the results of 18F-FDG PET studies in mice. *J Nucl Med* **47**, 999-1006 (2006).

19. Mellinghoff, I.K. et al. Molecular determinants of the response of glioblastomas to EGFR kinase inhibitors. *N Engl J Med* **353**, 2012-24 (2005).

20. Heimberger, A.B., Suki, D., Yang, D., Shi, W. & Aldape, K. The natural history of EGFR and EGFRvIII in glioblastoma patients. *J Transl Med* **3**, 38 (2005).

21. Bailey, R.C., Kwong, G.A., Radu, C.G., Witte, O.N. & Heath, J.R. DNA-encoded antibody libraries: a unified platform for multiplexed cell sorting and detection of genes and proteins. *J Am Chem Soc* **129**, 1959-67 (2007).

22. Kholodenko, B., Yaffe, M.B. & Kolch, W. Computational approaches for analyzing information flow in biological networks. *Sci Signal* **5**, re1 (2012).

23. Ward, J.H. Hierarchical Grouping to Optimize an Objective Function. *J Am Stat Assoc* **58**, 236-244 (1963).

24. Novak, B., Pataki, Z., Ciliberto, A. & Tyson, J.J. Mathematical model of the cell division cycle of fission yeast. *Chaos* **11**, 277-286 (2001).

25. Martens, H. & Martens, M. Multivariate analysis of quality : an introduction (J. Wiley, Chichester England ; New York, 2001).

26. Allen, D.M. The Relationship between Variable Selection and Data Augmentation and a Method for Prediction. *Technometrics* **16**, 125-127 (1974).

27. Sunayama, J. et al. Crosstalk between the PI3K/mTOR and MEK/ERK pathways involved in the maintenance of self-renewal and tumorigenicity of glioblastoma stem-like cells. *Stem Cells* **28**, 1930-9 (2010).

28. Mendoza, M.C., Er, E.E. & Blenis, J. The Ras-ERK and PI3K-mTOR pathways: cross-talk and compensation. *Trends Biochem Sci* **36**, 320-8 (2011).

29. Carracedo, A. et al. Inhibition of mTORC1 leads to MAPK pathway activation through a PI3K-dependent feedback loop in human cancer. *J Clin Invest* **118**, 3065-74 (2008).

30. Janes, K.A. et al. A systems model of signaling identifies a molecular basis set for cytokine-induced apoptosis. *Science* **310**, 1646-53 (2005).

31. Janes, K.A. & Yaffe, M.B. Data-driven modelling of signal-transduction networks. *Nat Rev Mol Cell Biol* **7**, 820-8 (2006).

4.7 APPENDIX A: SUPPLEMENTARY FIGURES

Figure 4.17 IHC on GBM39 xenograft samples. Statistical analysis of the quantitative IHC graphs for the mTOR, Erk and Src biomarkers and biological response markers (quantitative values mentioned in Table S2; C: CC214-2 responsive samples; R: CC214-2 resistant samples; D: dasatinib; U: U0126; stop: xenografts collected after interruption of the treatments); *P< 0.05; **P<0.005; ***P≤0.0005, N.S., Not Significant; (Student's T-test, values represent the average of three independent section's fields).

Figure 4.18 IHC on GBM39 xenograft samples. Statistical analysis of the quantitative IHC graphs for additional mTOR biomarkers (quantitative values mentioned in Table S2; C: CC214-2 responsive samples; R: CC214-2 resistant samples; D: dasatinib; U: U0126; stop: xenografts collected after interruption of the treatments); *P< 0.05; **P<0.005; ***P≤0.0005, N.S., Not Significant; (Student's T-test, values represent the average of three independent section's fields).

4.8 APPENDIX B SUPPLEMENTARY TABLES

Table 4.1 Reagents Used. The upper part of the table provides the sequences of the oligonucleotides used in the protein immunoassays. All oligonucleotides were synthesized by Integrated DNA Technology (IDT) and purified via high performance liquid chromatography (HPLC). The DNA coding oligomers were pre-tested for orthogonality to ensure that cross-hybridization between non-complementary oligomer strands was negligible (<1% in photon counts). Below the oligonucleotides is a list of the antibodies and standard proteins used for the SCBC multiplexed protein assay as well as the immunoblotting and the immunohistochemistry.

Name	DNA Sequence	Melting Point
B	5'-AAA AAA AAA AAA AGC CTC ATT GAA TCA TGC CTA -3'	57.4
B'	5' NH3AAA AAA AAA ATA GGC ATG ATT CAA TGA GGC -3'	55.9
C	5'- AAA AAA AAA AAA AGC ACT CGT CTA CTA TCG CTA -3'	57.6
C'	5' NH3-AAA AAA AAA ATA GCG ATA GTA GAC GAG TGC -3'	56.2
D	5'-AAA AAA AAA AAA AAT GGT CGA GAT GTC AGA GTA -3'	56.5
D'	5' NH3-AAA AAA AAA ATA CTC TGA CAT CTC GAC CAT -3'	55.7
E	5'-AAA AAA AAA AAA AAT GTG AAG TGG CAG TAT CTA -3'	55.7
E'	5' NH3-AAA AAA AAA ATA GAT ACT GCC ACT TCA CAT -3'	54.7
F	5'-AAA AAA AAA AAA AAT CAG GTA AGG TTC ACG GTA -3'	56.9
F'	5' NH3-AAA AAA AAA ATA CCG TGA ACC TTA CCT GAT -3'	56.1
G	5'-AAA AAA AAA AGA GTA GCC TTC CCG AGC ATT-3'	59.3
G'	5' NH3-AAA AAA AAA AAA TGC TCG GGA AGG CTA CTC-3'	58.6
H	5'-AAA AAA AAA AAT TGA CCA AAC TGC GGT GCG-3'	59.9
H'	5' NH3-AAA AAA AAA ACG CAC CGC AGT TTG GTC AAT-3'	60.8
K	5'-AAA AAA AAA ATA ATC TAA TTC TGG TCG CGG-3'	55.4
K'	5' NH3-AAA AAA AAA ACC GCG ACC AGA ATT AGA TTA-3'	56.3
L	5'-AAA AAA AAA AGT GAT TAA GTC TGC TTC GGC-3'	57.2
L'	5' NH3-AAA AAA AAA AGC CGA AGC AGA CTT AAT CAC-3'	57.2
M	5'-AAA AAA AAA AGT CGA GGA TTC TGA ACC TGT-3'	57.6
M'	5' Cy3-AAA AAA AAA AAC AGG TTC AGA ATC CTC GAC-3'	56.9
DNA Label	**Antibody for Conjugation in SCBC**	**Source**
B'	Human/Mouse Phospho-Akt1 (S473) DuoSet® IC ELISA kit	R&D DYC2289B
C'	Human Phospho-TOR (S2448) DuoSet® IC ELISA kit	R&D DYC1665
D'	Phospho-p70 S6 Kinase (T389) DuoSet® IC ELISA kit	R&D DYC896
E'	Human Total p21/CIP1/CDKN1A DuoSet® IC ELISA kit	R&D DYC1047
F'	Human Total p53 DuoSet® IC ELISA kit	R&D DYC1043
G'	Human/Mouse Cleaved Caspase-3 (Asp175) DuoSet® IC ELISA kit	R&D DYC835

H'	Human/Mouse/Rat Phospho-ERK1 (T202/Y204) DuoSet® IC ELISA kit	R&D DYC1825
K'	Human Phospho-Src (Y419) DuoSet® IC ELISA kit	R&D DYC2685
L'	Human total HIF-1α DuoSet® IC ELISA kit	R&D DYC1935
Catalog #	**Antibody for Immunoblotting and Immunohistochemistry**	**Source**
4060	p-Akt Ser473 (D9E)	Cell Signaling
9275	p-Akt Thr308	Cell Signaling
4857	p-S6 Ser235-236 (91B2)	Cell Signaling
9205	p-P70S6K Thr389	Cell Signaling
2855	p-4E-BP1 Thr37-46 (236B4)	Cell Signaling
4370	p-ERK Thr202-204 (D13.14.4E)	Cell Signaling
2971	p-mTOR Ser2448	Cell Signaling
2976	p-mTOR Ser2448 (49F9)	Cell Signaling
2972	mTOR	Cell Signaling
2101	p-Src Tyr416	Cell Signaling
3217	p-NDRG1 Thr346	Cell Signaling
5482	p-NDRG1 Thr346 (D98G11)	Cell Signaling
2280	raptor (24C12)	Cell Signaling
06-847	EGFR	Millipore
36-9700	p-EGFR Tyr1086	Invitrogen
2997	p-PRAS40 Thr246 (C77D7)	Cell Signaling
441100G	p-PRAS40 Thr246	Invitrogen
9541	cleaved PARP Asp214	Cell Signaling
ABM-2052	PTEN (6H2.1)	Cascade Bioscience
VP-RM04	Ki67 (SP6)	Vectorlabs
SMC-184D	HIF-1α	StressMarq
16314-015	TUNEL	Invitrogen
NB-600-501	Actin (AC15)	Novus Biologicals

Table 4.2. IHC quantification of the GBM39 xenograft stains. Statistical significant differences (P<0.05, student's T-test) vs. vehicle samples highlighted in green (C= CC214-2 responsive xenografts; R= CC214-2 resistant xenografts; D= Dasatinib; U= U0126; C+D= CC214-2 + Dasatinib; C+U= CC214-2 + U0126; DR= Drug Removed). Values represent the average of three independent section's fields.

	Ki-67 (%)	Tunel (%)	p-Src	p-ERK	HIF1a	p-mTOR
vehicle	43.44	0.13	14.82	17.13	13.40	11.25
C	25.96	1.46	19.98	11.49	12.71	7.25
R	42.69	0.14	19.32	26.43	16.23	9.70
D	40.37	0.22	15.46	17.24	13.89	12.74
U	42.26	0.23	13.38	7.53	14.25	11.92
U+D	41.80	0.26	7.75	6.96	10.91	14.56
C+D	29.30	1.75	9.26	13.75	13.08	10.02
C+D DR	49.91	0.07	18.61	15.05	10.59	12.90
C+U	27.53	1.23	16.77	11.11	13.09	9.13
C+U DR	52.60	0.10	14.57	15.35	14.52	12.12
C+D+U	22.05	1.61	9.45	10.13	10.88	13.74
C+D+U DR	54.84	0.17	18.60	17.95	10.78	12.99

	p-S6K	p-S6	p-4EBP1	p-AKT	p-NDRG1	p-PRAS40
vehicle	11.50	18.34	12.86	11.02	21.29	15.42
C	8.78	15.90	9.48	7.51	11.94	9.12
R	9.98	23.31	19.28	14.86	25.01	16.03
D	11.24	18.98	13.56	10.67	23.15	14.17
U	10.60	17.23	13.74	10.39	23.45	14.01
U+D	10.39	16.91	12.53	10.76	21.53	13.70
C+D	8.68	11.76	8.36	5.82	5.06	7.30
C+D DR	12.62	18.55	17.60	10.20	21.37	12.61
C+U	9.03	10.56	9.09	5.88	4.72	8.47
C+U DR	10.93	19.59	12.52	9.69	21.70	13.97
C+D+U	8.55	11.24	9.73	6.34	5.69	9.08
C+D+U DR	11.60	21.78	13.22	11.70	20.53	11.75

C h a p t e r   5

Translating single-cell functional proteomics into the clinic

5.1 INTRODUCTION

There exists a sufficient pharmacy to treat many patients with advanced cancers, such as GBM. Resolving the functional heterogeneity to determine the signaling network coordination within an individual patient's tumor can potentially inform an effective therapy strategy to treat the patient[1]. In Chapter 4, we have shown that functional proteomics assays, executed on statistical number of single cells in patient derived models that we perturb with the drugs we would use in the clinic to treat the patient, yield deep insight towards identifying independent drug targets within the tumor, which in turn inform therapy combinations that are highly effective in treating the tumors in GBM mouse model. We have also shown that, such analysis can be performed ex vivo on tumor biopsy samples within in a clinically relevant time scale. It therefore provides the rationale to extend our current work into the clinic, which will enable us to interrogate GBM tumor samples in a way that could potentially yield a straightforward, rapid interpretation to give therapeutic guidance to the attending physician.

Our approach centers on resolving the heterogeneity within a GBM tumor biopsy sample at the level of the activated, functional protein signaling networks that are associated with the aberrant nature of the tumor. However, such translation is not straightforward. The first challenge is that primary cells directly from tissues contain significantly lower copy number of a given functional protein than do cultured cells. Thus,

assay sensitivity is an important factor here, which demands improvement on engineering design and surface chemistries of the microchip. A reliable measurement/decision protocol is also required to streamline the single-cell proteomic assays and to ensure a robust and highly reproducible data collection and analysis, so that meaningful comparison can be made between datasets collected across time points, patient samples and assay conditions. Additionally, a challenge with any clinical study that attempts to match patients with appropriate therapies and therapy combinations, and for which the disease is highly heterogeneous, is that we don't know, prior to analysis, what drugs will be required, and so we need to design a trial that can potentially accommodate multiple drugs from different manufacturers and select the appropriate dose smartly.

In this chapter, we first discuss the advances in engineering and surface chemistry that address the technical challenges for translating SCBCs into the clinic. We further present some preliminary data collected from a pediatric GBM patient bearing an EGFR amplified tumor to demonstrate the workflow of the clinical translation.

5.2 EXPERIMENTAL METHODS

5.2.1 Surface chemistry optimization for clinical applications

Primary cells are smaller in size and have a higher dynamic range in terms of oncogenic signaling compared to the genetic modified cell lines. An antibody microarray with high sensitivity is therefore critical in this context, which in turn requires a higher DNA loading for patterning the ssDNA microarray. To increase the loading and overall uniformity, we developed a method based upon covalent binding between ssDNA and poly-L-lysine (PLL) in place of the original evaporation method (See Chapter 2). An

Figure 5.1 The reaction scheme for covalent DNA patterning. Accessible α-amine groups present on the amine-terminated ssDNA and ε-amines on lysine residues react with NHS-esters and form amide bonds. A covalent amide bond is formed when the NHS-ester crosslinking agent reacts with a primary amine releasing N -hydroxysuccinimide
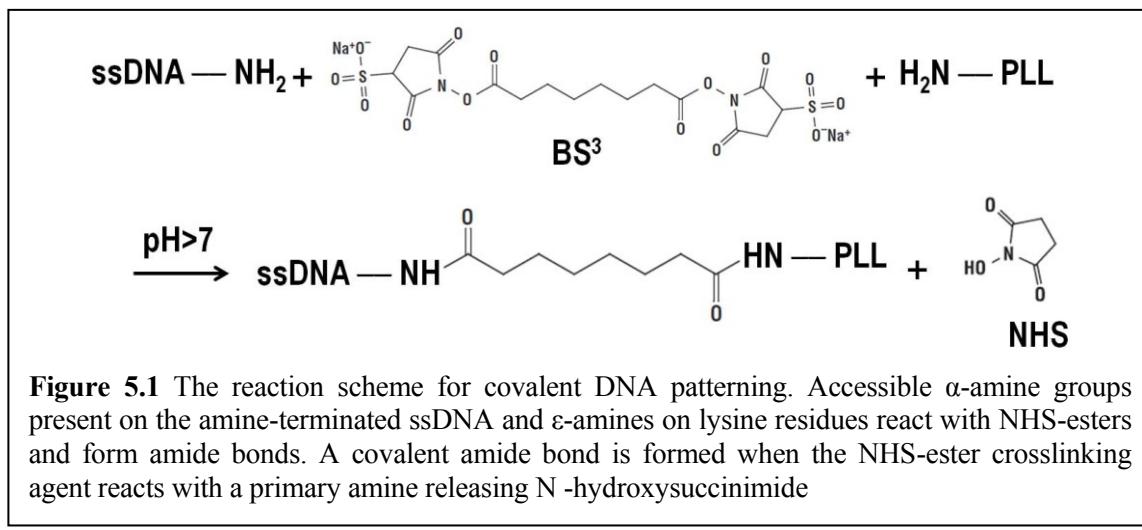
additional PLL coating step is also included (Fig. 5.1).

Specifically, after bonding of PDMS device to the PLL slide, 0.1% PLL solution (Sigma Aldrich) is flowed through the microchannels followed by air blow drying. Then a library of amine modified ssDNAs, diluted in a mixture of DMSO and deionized water (v/v=3:2) with a final concentration of 300μM and mixed with 2mM BS3 solution (a linker molecule that contains an amine-reactive N-hydroxysulfosuccinimide (NHS) ester at each

Figure 5.2. The surface chemistry of DNA barcode patterning, and its importance for quantitative single-cell protein immunoassays. **a.** The microfluidic flow patterning template used to prepare barcodes on poly-L-lysine(PLL)-coated glass slides. i. The elastomer flow patterning mold contains 1 channel for each barcode stripe – a mold for a 20 element barcode is drawn. The channels meander across the glass surface, and are on the order of 1 meter long and 10-20μm wide, depending upon the design. ii. ssDNA oligomers are initially patterned, and the quality of those DNA barcodes is assessed by hybridizing each strand with a complementary, dye-labeled ssDNA' oligomer. iii. The digitized fluorescence micrograph shows the uniformity of a 10-element barcode, across the region indicated by the yellow bar in ii. **b.** Digitized fluorescence data reflects the DNA loading of 20μm wide barcode stripes, based on various patterning chemistries. "O" and "X" mean that the indicated chemistry was or was not used, respectively; CV (coefficient of variation) values through the entire slide were listed, showing the loading uniformity from various patterning strategies. **c.** Calibration data for the protein p-ERK, measured using the various chemistries. Note that surface chemistry improvements yield more than a 10-fold increase in assay sensitivity, enabling single cell assays of both highly challenging primary cells (the GBM patient sample), and model cell lines.

end of an 8-carbon spacer arm, v/v=1:1), is flowed into each of the microfluidic channels. The solution-filled chip is then placed in a sealed petri-dish with controlled moisture for 90 minutes to immobilize amine-terminated ssDNAs to the PLL surface. After incubation, the PDMS elastomer is removed from the glass slide in water containing 0.02% SDS followed with intensive washing in 0.02% SDS in water.

Fig. 5.2 shows how different surface chemistries will affect the DNA barcode patterning with respect to DNA loading and overall uniformity. Comparing with other methods, covalent binding method gives out highest loading and best over uniformity (lowest CV, Fig. 5.2 b). It also shorten the original 3-5 days process by evaporation method to only 1 day. The high DNA loading achieved plus other surface chemistry optimizations

such as matching best antibody ELISA pairs and using brighter dyes finally transfer to a more than 10-fold increase in assay sensitivity and more than 50-fold in signal to noise ratio (Fig. 5.2 c).

5.2.2 High throughput solutions

Single-cell functional proteomic microchips are the diagnostic workhorse for clinical applications. The nature of the single-cell biology is that statistical numbers of single cells

Figure 5.3 Illustration of the two-layer cross-stripe DNA microarray construction. (a, *top*) Schematic illustration of the chemical patterning to produce the high density cross-stripe ssDNA arrays. Stripes of DNA are first patterned onto a PLL coated glass slide using a flow patterning mold . A second flow patterning step, orientated perpendicular to the first, is used to create unique addresses at the intersections of the cross stripes. The color coded DNA oligomers illustrate the patterning/hybridization sequence to produce the final array. (a, *bottom*) Validation of the cross-stripe barcode microarray. Each square unit of fluorescent spots represents many copies of a 3×3 array. The plot at right provides the fluorescence intensity profile of the vertical line through the two square unit. (b) The microfluidic flow patterning template used to prepare high density 10μm × 10μm cross-stripe DNA array throughout the entire slide by covalent method. It will generate more than 240,000 array spots in the whole slide after patterning.

must be analyzed for any given assay to generate a meaningful result[2]. Although our first microchip prototype contained only 120 microchambers[3], the SCBC has been consistently developed and optimized since then and now it contains 320 microchambers per chip. Two chips have to be run in parallel each time to ensure enough statistics, which is not yet

optimal. Another valve-free SCBC design[4] developed for assessing cell-cell communications has around 10,000 microchambers and allows around 1,000 single cell assays per chip (Fig. 5.3 a). But it requires subtle operations and is less robust, which is not acceptable for analyzing precious clinical materials. A high throughput proteomic microchip that is robust and easy to handle is therefore desired for the clinical translation.

The density of integration of SCBCs depends on the density of the antibody microarray and thus the density of ssDNA barcode. The two layer DNA patterning strategy is developed to achieve a significantly high density DNA microarray. Previously, we started with an n-element stripe-structured ssDNA microarray (Fig.5.2 a) that was further converted to an n-element antibody microarray by DNA hybridizations. In the new patterning approach, a second set of ssDNA are flow patterned at the right angle to the first set which yield an addressable n×m array, where n and m are the numbers of microchannels utilized for the two flow patterning steps (Fig. 5.3). This approach has been successfully demonstrated as a 3×3 array with 20μm × 20μm feature size in the central area of a standard microscope glass slide by traditional evaporation method[4]. A higher density array with 10μm × 10μm feature size throughout the entire slide by covalent binding method is still under developing with promising preliminary results (Fig. 5.3 b).

Two layer DNA flow patterning yields microarray slides that are similar to commercial products, but with much extremely higher density and a much smaller feature size (10μm vs. 150μm). This in turn provides us great flexibility to design the SCBCs with much more microchambers per chip to ensure sufficient statistics. The smaller chamber volume, as a byproduct, will further increase the assay sensitivity. The multiplexity can

also be easily expanded from assaying 9–20 functional proteins per chamber to 16–36

functional proteins.

5.2.3 The protocol and workflow of analyzing patient biopsy samples
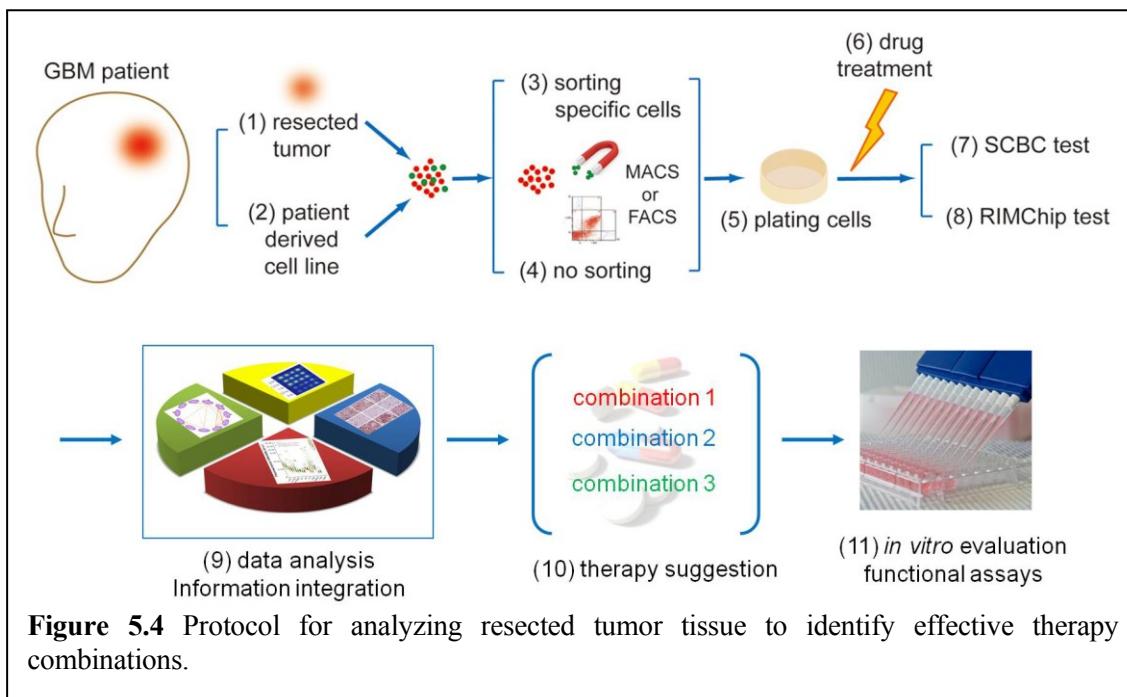
Figure 5.4 Protocol for analyzing resected tumor tissue to identify effective therapy combinations.

Fig. 5.4 illustrates the general workflow from collecting patient's tissue biopsy samples

to single-cell proteomics test and finally down to give out therapeutic guidance. Briefly,

immediately following tissue resection, we use standard protocols to dissociate the solid

tissue into single cell suspension (See Chapter 4.2 for detail). We select EGFR+ cells from

the population, which apparently permits capture of other relevant cellular subsets from the

tumor. Those cells are extracted using either MACS or FACS, and then cultured for a short

period of time (2 hours), which allows for the removal of inviable cells. The remaining

cells are loaded into an SCBC (or Betabox) and immediately analyzed. For investigating

drug perturbations, the EGFR+ cells are cultured for a set of time in the presence of a

targeted inhibitor at a relevant dosing level, prior to analysis. The time between the tumor resection and SCBC or Betabox assay completion is about one day.

The Betabox[5] is a combined microchip/Beta particle camera device which is an *in vitro* equivalent of [^{18}F]FDG Positron Emission Tomography (PET) scans[6]. It only requires about 100 cells per assay and can simultaneously test up to 4 different conditions. Betabox assays yield metabolic assessments of drug target engagement, drug responses, as well as pharmacokinetic information within a few hours.

Once the SCBC assay is complete, a genepix array scanner is utilized to digitize the fluorescent signal. A custom algorithm is applied to transfer the scanned image into a data table for analyzing the signaling network, and to reveal protein signaling modes (See Chapter 4.4.2 for detail). Comparing single cell proteomics data of the EGFR+ tumor cells with and without drug treatment and integrating information from Betabox and other pathological/molecular characterizations will help identify the effective drug combinations to treat the tumor.

The period from surgical resection to completion of all the experiments is about 2-3 days. Two additional days are required for data analysis. By the end of the fifth day, the analyzed data is discussed with the attending physician(s), with suggestions for effective therapy emerging from those conversations. These predictions will be first evaluated *in vitro* through a set of functional assays such as cell viability and proliferation tests. The patient can potentially start the therapy as soon as a week after the surgery.

5.3 RESULTS AND DISCUSSION

5.3.1 Fast signaling adaptation of a pediatric GBM tumor to lapatinib: a patient sample case study

Figure 5.5 Single-cell scatter plots and correlation networks on pediatric GBM cells. (a) SCBC data represented as one dimensional scatter plot for vehicle and Lapatinib treated sample. The averaged fluorescence intensity with standard error of the mean (SEM) is overlaid for each protein. Statistical uniqueness is evaluated by two-tail Mann-Whitney test for pairwise comparison(*P<0.05; **P<0.005;***P<0.0005, NS: not significant). (b) Protein-protein correlation networks for control and Lapatinib treated samples, extracted from SCBC data. The correlation strengths are reflected in the thickness of the edges (see keys). For Lapatinib treated sample, new correlations, relative to control, emerge between p-ERK1 and core mTOR effectors.

The resected GBM tumor was pathologically characterized as EGFR amplified. As a result, Lapatinib, as an FDA approved EGFR inhibitor, was suggested by the physicians in the first place to treat the patient. We then utilized SCBC platform to quantify the levels and correlative interactions of 10 proteins and phosphoproteins from statistical numbers of single EGFR+ tumor cells in both treatment naïve and 48 hours Lapatinib perturbed[7] (with a clinically relevant dosage, 2000nM) stages. We tested p-mTOR, p-P70S6K and p-4EBP1

associated with mTORC1 signaling; p-Akt1, p-NDRG associated with mTORC2 signaling; p-ERK1, p-Src and three p53 related functional proteins. Both mTORC1/C2 associated proteins and p-ERK/p-Src are downstream effectors of EGF Receptor tyrosine kinase signaling[3] (Fig. 5.5). The results, however suggested that Lapatinib neither induced a considerable cell killing nor shut down the oncogenic signaling as it is supposed to, although it has successfully hit the target by evaluating the phosphorylation level of two

Figure 5.6 Prediction of therapy strategies by statistical analysis of the single-cell data. (a) Correlations between key functional proteins and the first principal component (PC) for control and Lapatinib perturbed samples. Most of the mTOR effectors still dominate the signaling coordination in the drug treated sample. The influence of p-ERK1 on the signaling coordination significantly rise up in drug treated stage compared to control. The composition of the first three PCs for control sample reveal the decoupling between mTOR signaling and ERK/Src signaling. (b) The down regulated p-EGFR phosphorylation levels indicate Lapatinib has successfully engaged its target. (c) Quantification of the heterogeneity of the cancer cells at control and Lapatinib perturbed stages. The heterogeneity index is obtained from calculating the dissimilarity level of the AHC analysis (Ward's method, see chapter 4 for detail). (d) The *in vitro* validation shows that adapted tumor cells is more resistant to Lapatinib treatment.

related EGFR phosphorylation sites Tyr1173 and Tyr1068 (Fig. 5.6 b)[8, 9].

The expression level and spread of fluctuation of p-ERK1 greatly up-regulated in Lapatinib treated sample suggested a potential gain of function of MAPK/ERK signaling (Fig 5.5 a). By comparing the correlation networks of the tumor cells with and without drug treatment, it's readily to see that the drug actually triggered p-ERK1 to take over the signaling coordination and at the same time, the interactions among the core mTORC1/C2 effectors became even stronger (Fig 5.5 b). In other words, the cancer cells quickly adapted to EGFR inhibition by activating ERK/Src signaling and mTOR signaling, suggesting that mTOR and ERK/Src signaling might provide two independent druggable pathways, which is in accord with our findings in the GBM 39 mouse model (See Chapter 4 for detail). The calculation of the influence of functional proteins on signaling coordination, as well as the population heterogeneity, further confirmed that Lapatinib did seem to be a bad player here (Fig. 5.6 a and c). The analysis of the composition of the first 3 principal components also implied the decoupling between mTOR signaling and ERK/Src signaling is present in this tumor (Fig. 5.6 a).

Integrating the information above, we predicted that the inhibition of EGFR by Lapatinib would not be likely to suppress the tumor growth. In contrast, the adapted tumor cells may progress even faster due to the highly elevated mTOR signaling which drives the tumor to grow. We validated our prediction *in vitro* by assessing the cell viability and proliferation after 2 days Lapatinib treatment for both fresh tumor cells and adapted tumor cells (2000nM Lapatinib 2-day pretreated cells). We started with two sets of petri-dishes (3 repeats in each condition) with identical cell number representing fresh or adapted tumor cells respectively. After 2-day Lapatinib treatment at 2000nM o.p.d., slight cell death was

observed in fresh tumor samples while a cell growth was observed in the adapted tumor samples, which is wholly consistent with our predictions (Fig 5.6 d).

5.3.2 Prediction on effective therapy combination

The single cell analysis on treatment naïve samples showed a highly activated mTOR signaling and thus suggested that mTOR is a reasonable drug target for treating this tumor.

Figure 5.7 Validation of therapy predictions via SCBC and functional tests. (a) SCBC data represented as one dimensional scatter plot for control and drug combination, XL-765 plus Trametinib (XL765+T) treated sample. The averaged fluorescence intensity with standard error of the mean (SEM) is overlaid for each protein. Statistical uniqueness is evaluated by two-tail Mann-Whitney test for pairwise comparison(*P<0.05; **P<0.005;***P<0.0005, NS: not significant). (b) Protein-protein correlation networks for control and Lapatinib treated samples, extracted from SCBC data. The correlation strengths are reflected in the thickness of the edges (see keys). For XL-765+T treated sample, the interaction among core mTOR effectors has been diminished. (c) Quantification of the heterogeneity of the cancer cells at control and XL-765+T treated stages. (d) Cell viability assay for control and XL-765+T treated samples. The drug combination successfully induces considerable cell death.

The data on Lapatinib perturbed samples further implied that ERK signaling may emerge to serve as a resistant mechanism through fast signaling network adaptation. As a result, we suggested that a combination of mTOR inhibitor (XL-765)[10] and MEK/ERK inhibitor (Trametinib)[11] could potentially repress the oncogenic signaling and halt the tumor progression. We tested our prediction via *in vitro* cell viability assay as well as single cell proteomics assay. The results were encouraging (Fig 5.7).

Analysis on the signaling coordination reveals that the combination of XL-765 (1μM) and Trametinib (100nM) has successfully weakened the interactions among the core mTOR effectors without triggering ERK signaling to compensate. Although p-Src shows slightly increased interactions with mTOR effectors, the correlations are not strong. The functional heterogeneity of drug treated samples is also significantly reduced (Fig. 5.7 b and c). Cell viability assays shows that the predicted drug combination (XL-765+Trametinib) can induce a considerable cell death during the course of treatment, which in part validates our prediction (Fig. 5.7 d). The single-cell analysis also implies that a clinically available drug that can hammer mTOR signaling harder would be desired.

5.4 CONCLUSION

Our data clearly illustrates the point that, via that single-cell functional proteomics analysis, we can rapidly determine the signaling network coordination within an individual patient's tumor and gain new insight into combining existing therapies together in much more effective ways. That knowledge can then be used to design a combinatorial therapy to treat the patient. The results and protocol discussed here are preliminary. Additional tests for intracranial xenograft models and GBM patient samples are required to further refine the measurement protocols as well as the analytical approaches to integrate the information from different sources to reach a robust decision strategy.

Drug dosage and dosing schedule are also important factors that need further investigation. It has been shown that varied dosing schedules determined based upon the information such as maximum tolerated dose and pharmacokinetic processes of the drugs have great impact on the dynamics of the acquired resistance to Erlotinit in EGFR-mutant lung cancer. High-dose pulses with low-dose continuous therapy impede the development of resistance to the maximum extent compared with constant dose therapy[12]. It has implications in our scenario as well. If the fast network adaption induced therapy resistance is also a reversible dynamic switch, as we believe so, matching effective drug combination with smart dosing strategy for individual patient could potentially yield stronger synergetic effect with respect to tumor suppression and tumor killing.

5.5 REFERENCES

1. Cloughesy, T.F. & Mischel, P.S. New strategies in the molecular targeting of glioblastoma: how do you hit a moving target? *Clin Cancer Res* **17**, 6-11 (2011).

2. Welch, C.M., Elliott, H., Danuser, G. & Hahn, K.M. Imaging the coordination of multiple signalling activities in living cells. *Nat Rev Mol Cell Biol* **12**, 749-56 (2011).

3. Shi, Q. et al. Single-cell proteomic chip for profiling intracellular signaling pathways in single tumor cells. *Proc Natl Acad Sci U S A* **109**, 419-24 (2012).

4. Wang, J. et al. Quantitating cell-cell interaction functions with applications to glioblastoma multiforme cancer cells. *Nano Lett* **12**, 6101-6 (2012).

5. Wang, J. et al. Fast metabolic response to drug intervention through analysis on a miniaturized, highly integrated molecular imaging system. *J Nucl Med* **54**, 1820-4 (2013).

6. Dooraghi, A.A. et al. Betabox: a beta particle imaging system based on a position sensitive avalanche photodiode. *Phys Med Biol* **58**, 3739-53 (2013).

7. Vivanco, I. et al. Differential sensitivity of glioma- versus lung cancer-specific EGFR mutations to EGFR kinase inhibitors. *Cancer Discov* **2**, 458-71 (2012).

8. Zwick, E., Hackel, P.O., Prenzel, N. & Ullrich, A. The EGF receptor as central transducer of heterologous signalling systems. *Trends Pharmacol Sci* **20**, 408-12 (1999).

9.   Rojas, M., Yao, S. & Lin, Y.Z. Controlling epidermal growth factor (EGF)-stimulated Ras activation in intact cells by a cell-permeable peptide mimicking phosphorylated EGF receptor. *J Biol Chem* **271**, 27456-61 (1996).

10.   Prasad, G. et al. Inhibition of PI3K/mTOR pathways in glioblastoma and implications for combination therapy with temozolomide. *Neuro Oncol* **13**, 384-92 (2011).

11.   Flaherty, K.T. et al. Improved survival with MEK inhibition in BRAF-mutated melanoma. *N Engl J Med* **367**, 107-14 (2012).

12.   Foo, J., Chmielecki, J., Pao, W. & Michor, F. Effects of pharmacokinetic processes and varied dosing schedules on the dynamics of acquired resistance to erlotinib in EGFR-mutant lung cancer. *J Thorac Oncol* **7**, 1583-93 (2012).

Published materials included in the thesis

1. Shi, Q. et al. Single-cell proteomic chip for profiling intracellular signaling pathways in single tumor cells. *Proc Natl Acad Sci U S A* **109**, 419-24 (2012). (Included in Chapter 2, http://www.pnas.org/content/109/2/419.short)

2. Wei, W. et al. Hypoxia induces a phase transition within a kinase signaling network in cancer cells. *Proc Natl Acad Sci U S A* **110**, E1352-60 (2013). (Included in Chapter 3, http://www.pnas.org/content/110/15/E1352.short)