

**Biotechnologies for Cancer Diagnostics: Cell
Sorting, Protein Analysis and Imaging of
Cellular Metabolism**

Thesis by
Kiwook Hwang

In Partial Fulfillment of the Requirements for the degree
of
Doctor of Philosophy



CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California
2013
(Defended May 30th, 2013)

To my dearest wife,

Hana Kim

Acknowledgement

First of all, I would like to thank to my advisor, Professor James Heath. With great patience, he has guided and trained me to become a research scientist during my Ph. D. journey. He also granted me academic freedom and incredibly valuable resources, including the organization of Nanosystems Biology Cancer Center (NSBCC) between Caltech, UCLA and the Institute for Systems Biology (ISB). I appreciate him for giving me many chances to collaborate with brilliant senior scientists, including Professor Caius Radu (UCLA), Professor Paul Mischel (UCSD), Professor Raphael Levine (UCLA), Professor Arion Chatziioannou (UCLA), Professor Heather Christofk (UCLA) and Professor Michael Phelps (UCLA), who have helped and advised me during my graduate researches. Also, I would like to acknowledge my thesis committee members, Professor Long Cai, Professor Shu-ou Shan and Professor William Goddard III, who spent much time providing insightful comments and suggestions on my progress report and research proposals. Not to forget, I thank the Samsung Scholarship which has also been supporting my Ph.D. studies.

In Heath lab, I have been fortunately able to collaborate with a number of gifted past and present group members. I would like to thank Dr. Gabe Kwong, Dr. Young Shik Shin and Dr. Jun Wang, with whom I worked on NACS, SCBC and RIMChip projects, respectively. Through them, I have learned practical techniques for the preparation of bio samples and the fabrication of microfluidic chips, as well as the scientific way of designing experiments. I also appreciate Dr. Heather Agnew and Wei Wei in studying the application of RIMChip in clinical systems. Thanks to Dr. Habib Ahmad, Dr. Jen-Kan

Yu, Joey Varghese, Alex Sutherland, Ryan Henning and Blake Farrow for valuable scientific discussions. I appreciate Diane Robinson, Amy Crown, Elyse Garlock and Kevin Kan for processing all paper-work and managing lab facilities properly.

Also, I acknowledge many intelligent co-workers at UCLA, especially Dr. Daniel Braas, Alex Dooraghi, Dr. David Nathanson, Dr. Dean Campbell and Jessica Gu in working with the RIMChip project. Without their kind help and assistance, I could not obtain wonderful results at UCLA.

I am grateful to all of my Korean friends for sharing my personal life at Caltech, especially Dr. Oh-Hoon Kwon, Dr. Chang Ho Sohn, Dr. Christopher Chang, Dongwan Kim, Chung Whan Lee and Jake Kim.

Finally, I appreciate my family who always supported me with love and care. In particular, I thank my beloved wife, Hana Kim. Without her unconditional love and patience during the Ph.D. period, this work would never have existed. I dedicate this thesis to her.

Abstract

This thesis presents the development of chip-based technology for informative *in vitro* cancer diagnostics. In the first part of this thesis, I will present my contribution in the development of a technology called “Nucleic Acid Cell Sorting (NACS)”, based on microarrays composed of nucleic acid encoded peptide major histocompatibility complexes (p/MHC), and the experimental and theoretical methods to detect and analyze secreted proteins from single or few cells.

Secondly, a novel portable platform for imaging of cellular metabolism with radio probes is presented. A microfluidic chip, so called “Radiopharmaceutical Imaging Chip” (RIMChip), combined with a beta-particle imaging camera, is developed to visualize the uptake of radio probes in a small number of cells. Due to its sophisticated design, RIMChip allows robust and user-friendly execution of sensitive and quantitative radio assays. The performance of this platform is validated with adherent and suspension cancer cell lines. This platform is then applied to study the metabolic response of cancer cells under the treatment of drugs. Both cases of mouse lymphoma and human glioblastoma cell lines, the metabolic responses to the drug exposures are observed within a short time (~ 1 hour), and are correlated with the arrest of cell-cycle, or with changes in receptor tyrosine kinase signaling.

The last parts of this thesis present summaries of ongoing projects: development of a new agent as an *in vivo* imaging probe for c-MET, and quantitative monitoring of glycolytic metabolism of primary glioblastoma cells. To develop a new agent for c-MET

imaging, the one-bead-one-compound combinatorial library method is used, coupled with iterative screening. The performance of the agent is quantitatively validated with cell-based fluorescent assays. In the case of monitoring the metabolism of primary glioblastoma cell, by RIMChip, cells were sorting according to their expression levels of oncoprotein, or were treated with different kinds of drugs to study the metabolic heterogeneity of cancer cells or metabolic response of glioblastoma cells to drug treatments, respectively.

TABLE OF CONTENTS

Acknowledgements	iv
Abstract.....	vi
Table of Contents	viii
List of Illustrations and Tables.....	vi
Chapter I: Introduction	1
1.1 Cell Sorting in Cancer Research	1
1.2. Proteomics for Cancer Diagnosis.....	2
1.3. Metabolomics for Cancer Diagnosis.....	3
1.4. Thesis Overview	4
1.5. References.....	7
 Chapter II: Quantifying the performance of a new T cell sorting technique, and the preparation of reagents for quantitative analysis of proteins from macrophages	
2.1. Introduction.....	9
2.2. Experimental Methods	14
2.2.1. Microarray fabrication for T cell sorting	14
2.2.2. Synthesis of DNA-SAC conjugates	15
2.2.3. Preparation of T cells	16
2.2.4. Sorting cells with NACS and conventional microarrays.....	16
2.2.5. Microchip fabrication	17
2.2.6. Preparation of barcode arrays.....	18
2.3. Results and Discussions	20
2.3.1. Comparison between NACS and conventional protein microarray	20
2.3.2. Preparation of DNA-encoded antibodies	22
2.4. Conclusions	24
2.5. References.....	24

Chapter III: Fast Metabolic Response to Drug Intervention through Analysis on a Miniaturized, Highly Integrated Molecular Imaging System

3.1. Introduction.....	28
3.2. Experimental Methods	29
3.2.1. RIMChip design concepts	29
3.2.2. RIMChip fabrication	31
3.2.3. The beta-particle camera (Betabox).....	32
3.2.4. Cell sample preparation, viability, and cell-cycle assays	32
3.2.5. Betabox assay	33
3.2.6. Off-chip radioassay	35
3.2.7. Phosphoprotein assay	35
3.2.8. Data processing	35
3.3. Results.....	36
3.3.1. Design of RIMChip and cell loading.....	36
3.3.2. Validation of RIMChip	41
3.3.3. Kinetic study of drug response with RIMChip.....	42
3.4. Discussions	46
3.5. Conclusions	47
3.6. References	47

Chapter IV: Development and characterization of the capture agents library targeting c-MET for in vivo cancer imaging

4.1. Introduction.....	50
4.2. Experimental Methods	52
4.2.1. Synthesis of PCC for c-MET.	52
4.2.2. Functionalization of capture agents	53
4.2.3. Preparation of cells	55
4.2.4. Confocal experiment for 3D imaging of live cells incubated with capture agents	55
4.2.5. Flow cytometry for the quantification of binding affinity.....	56
4.2.6. In vitro radioassay with RIMChip	56
4.2.7. In vivo treatment and imaging with microPET/CT	57
4.3. Results and Discussions	57

4.3.1. Synthesis of capture agents for c-MET.....	57
4.3.2. Cell-based fluorescence imaging assays.....	58
4.4. Conclusions and Future Directions.....	60
4.5. References.....	60
Chapter IV: Imaging of glycolytic metabolism in primary glioblastoma cells with RIMChip	
5.1. Introduction.....	63
5.2. Experimental Methods.....	64
5.2.1. Preparation of cells.....	64
5.2.2. Cell sorting.....	65
5.2.3. In vitro radio assay with RIMChip.....	65
5.3. Results and Discussions.....	66
5.3.1. Imaging of glycolytic metabolism in glioblastoma cells according to their expression level of EGFRvIII.....	66
5.3.2. Imaging of glycolytic metabolism in glioblastoma cells under drug treatment.....	67
5.4. Conclusions and Future Directions.....	68
5.5. References.....	69
Appendix A: Modular nucleic acid assembled p/MHC microarrays for multiplexed sorting of antigen-specific T cells	
A.1. Introduction.....	71
A.2. Experimental Methods.....	74
A.2.1. Microarray fabrication.....	74
A.2.2. Synthesis of DNA-SAC conjugates.....	75
A.2.3. Preparation of T cells.....	76
A.2.4. Sorting cells.....	76
A.3. Results and Discussions.....	78
A.3.1. Design of ssDNA-p/MHC tetramers.....	78
A.3.2. Comparison between NACS and conventional protein microarray.....	80
A.3.3. Specificity of NACS and its detection limit.....	82
A.3.4. Selective release of immobilized T cells with restriction	

endonucleases	84
A.3.5. Sorting of TCR engineered and endogenous primary human T cells by NACS.....	85
A.5. Conclusions.....	88
A.6. References.....	88

Appendix B: Protein signaling networks from single cell fluctuations and
information theory profiling

B.1. Introduction	92
B.2. Experimental Methods	92
B.2.1. Microchip fabrication.....	93
B.2.2. Preparation of barcode arrays	94
B.2.3. Culture and stimulation of THP-1 cells.....	97
B.2.4. On-chip secretion profiling.....	97
B.2.5. Bulk secretion profiling	98
B.2.6. Quantification and statistics.....	98
B.2.7. Data Analysis: Conversion to the number of molecules.....	99
B.2.8. Calculations.....	99
B.2.9. Analysis of experimental and biological variation from SCBC-based single-cell measurement.....	100
B.2.10. Signal-to-noise calculations and experimental error.....	105
B.3. Theoretical methods.....	108
B.3.1. The fluctuations in the secretome.....	108
B.3.2. Theoretical approach	109
B.3.3. Theory of fluctuations.....	110
B.3.4. A quantitative Le Chatelier equation.....	112
B.4. Results and Discussion	114
B.4.1. Computing the covariance matrix	114
B.4.2. The network	115
B.4.3. The composite networks.....	119
B.4.4. The number-based network	121
B.4.5. Antibody perturbations	122
B.5. Conclusions.....	123

B.6. Appendix: Details in Theoretical Methods	124
B.6.1. Introduction to theoretical supplementary methods.....	124
B.6.2. The ensemble: basis for making predictions.....	125
B.6.3. Fluctuations describe the response to small perturbations.....	127
B.6.4. The principle of Le Chatelier.....	128
B.6.5. The equation for the direction of change.....	129
B.6.6. Tiers of the network are eigenvectors of the correlation matrix	131
B.6.7. The spectral representation of the covariance matrix	131
B.6.8. The role of the number of cells in the sample	132
B.6.9. Antibody perturbations	133
B.7. References.....	133

LIST OF ILLUSTRATIONS AND TABLES

<i>Number</i>	<i>Page</i>
Fig. 1.1. Diagnosis of breast cancer with metabolites	3
Fig. 2.1. Self-assembled ssDNA-p/MHC tetramer arrays for multiplexed sorting of antigen-specific cells	11
Fig. 2.2. Design of integrated microchip for single-cell protein secretome analysis	14
Table 2.1. Orthogonal DNA sequences for spatial encoding of p/MHC tetramers	15
Table 2.2. Sequences and terminal functionalization of oligonucleotides	19
Table 2.3. Summary of antibodies used for macrophage experiments	20
Fig. 2.3. Comparison of NACS versus spotted p/MHC arrays	21
Table 2.4. Parameters utilized for the protein assay calibration curve	23
Fig. 2.4. Cross-reactivity check and calibration curves	23
Fig. 3.1. Fabrication of the RIMChip	31
Fig. 3.2. The RIMChip design, operation, and betabox performance	38
Fig. 3.3. Distribution of adherent (liposarcoma) cells within fibronectin-coated cell chambers within channel	39
Fig. 3.4. Effect of surface coating on cell number in microchambers	39
Fig. 3.5. Effect of fibronectin coating on ¹⁸ F-FDG uptake by liposarcoma cells	40
Fig. 3.6. Validations of the RIMChip via genetic and molecular manipulation of glycolytic flux and nucleoside salvage activity in betabox assays, with comparisons against standard assays	42
Fig. 3.7. Betabox assays, correlated with other functional assays, for gauging the response of cancer cells to targeted drugs	44
Fig. 3.8. The kinetics of glucose consumption rate and protein level upon Erlotinib treatment	45
Table 3.1. List of antibodies used for GBM cell proteomic assay	45
Fig. 4.1. c-MET pathway	51
Fig. 4.2. Scheme of capture agent synthesis by in situ click chemistry	53

Fig. 4.3. Scheme of FITC labeling.....	54
Fig. 4.4. Scheme of ^{18}F labeling.....	54
Fig. 4.5. Molecular structure of the 1 st -generation capture agent for targeting human c-MET, conjugated with FITC, IN-CT-1025	57
Fig. 4.6. Expression level of c-MET in different prostate cancer cell lines.....	58
Fig. 4.7. Result of titration experiment with IN-CT-1025.	59
Fig. 5.1. The relation between [^{18}F]FDG uptake and the expression level of EGFRvIII in GBM39 cells.....	67
Fig. 5.5. The relation between [^{18}F]FDG uptake and drug treatment to EGFR+ subpopulations of GBM39 cells.....	68
Fig. A.1. Self-assembled ssDNA-p/MHC tetramer arrays for multiplexed sorting of antigen-specific cells.	73
Table A.1. Orthogonal DNA sequences for spatial encoding of p/MHC tetramers.....	75
Fig. A.2. T cell capture efficiency is optimal when utilizing ssDNA-SAC conjugates to generate NACS p/MHC tetramers.....	80
Fig. A.3. Comparison of NACS versus spotted p/MHC arrays.	81
Fig. A.4. Nucleic acid cell sorting of antigen-specific T cells	83
Fig. A.5. Programmed release of captured T cells by endonuclease cleavage	85
Fig. A.6. NACS sorting of endogenous primary human T cells specific for Epstein-Barr virus and Cytomegalovirus.	87
Fig. B.1. Design of integrated microchip for single-cell protein secretome analysis.....	94
Fig. B.2. Cross-reactivity check and calibration curves.....	95
Table B.1. Sequences and terminal functionalization of oligonucleotides.....	96
Table B.2. Summary of antibodies used for macrophage experiments.....	97
Table B.3. Parameters utilized for the protein assay calibration curve.....	99
Fig. B.3. Experimental and simulation results for extracting the experimental error contribution to the SCBC protein assays	102
Table B.4. Values of parameters used in simulation	103
Fig. B.4. Simulated histograms of average intensity from multiple DNA barcode locations	104

Table B.5. The coefficients of variation for each of the assayed proteins from single-cell experiments.....	105
Fig. B.5. Protein secretion heat maps for different colony sizes of LPS-stimulated macrophages.....	107
Fig. B.6. Fluctuations in the numbers of secreted IL-8 proteins, for all single-cell experiments.....	109
Table B.6. Digital representation of the covariance matrix for 1-cell measurements.....	113
Fig. B.7. The summary network derived from the information theory treatment of the data.....	116
Fig. B.8. PMA and LPS activation and kinetics of protein secretion from activated macrophage cells.....	118
Fig. B.9. Protein-protein interactions via the quantitative Le Chatelier's theorem.....	120
Fig. B.10. The dependence of the dominant eigenvalues of the covariance matrix on the number of cells in the sample.....	121
Fig. B.11. Heat map of the covariance matrix (left) and of the contributions to the first two tiers of the network (right) for measurements on chambers containing 3 cells.....	121
Fig. B.12. Perturbation of protein networks using neutralizing antibodies.....	123

Chapter 1

Introduction

1.1. Cell Sorting in Cancer Research

Cancer is a heterogeneous complex system often composed of different clonal cell subpopulations¹. Increasing reports of intratumor heterogeneity and its augmentation for the selective pressure of tumor microenvironment, and recent achievements in cancer therapeutics requires the investigation and tracking of the subpopulations of cancer cells². Hence, it is important to develop a multiplex and high-throughput technology for sorting cells to diminish the heterogeneity of cancer cells.

As the current state of the art, fluorescent activated cell sorting (FACS) has been widely used since its invention in early 1970s³ due to its multiplexity and ability to sort small subpopulations (cells with frequency around 0.03%). However, its sorting speed cannot be over one million cells per hour, and it requires expensive instrumentation as well as a well-trained operator. Furthermore, rare cells, such as circulating tumor cells, cannot be sorted by FACS because of their extremely low frequency.

With the recent advance in microfluidics technology, powerful techniques have been developed to overcome the limitation of FACS. High-throughput cell sorting has been achieved by Wolff and co-workers, who described the 'on-chip' system allowing sorting of 12,000 cells per second at 100-fold enrichment⁴, and by Simmonet and Groisman, whose microfluidic system is reported to sort 17,000 cells per second at 83-fold enrichment with 40% of recovery⁵. In the case of rare cells, Toner and co-workers

have managed to sort circulating tumor cells from peripheral blood by using a microfluidic chip with a large number of microposts⁶.

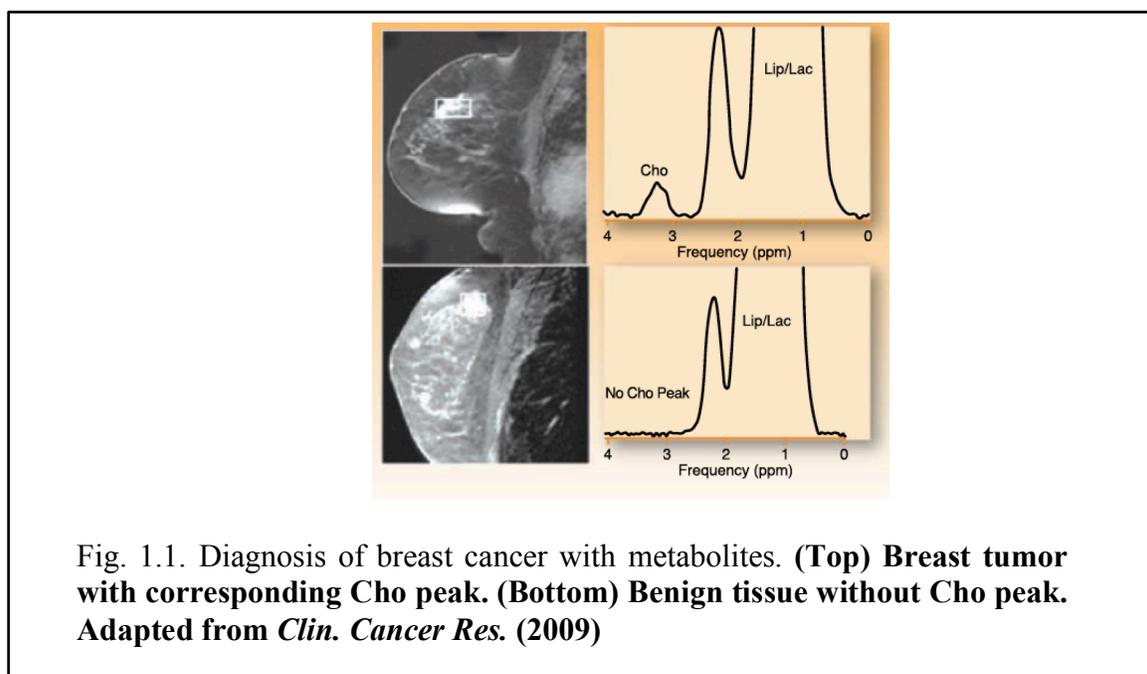
1.2. Proteomics for Cancer Diagnosis

In contrast to DNA, which is subject to one major form of modification as methylation, proteins can be altered in many ways after their translation, such as by phosphorylation, acetylation, and glycosylation, which can bring a functional shift that potentially affects disease development, progression and response to clinical therapies. Therefore, a pressing need exists for improved and innovative technologies to profile oncoproteins for better understanding and treatment of cancer.

Conventionally, there are two approaches to proteomics for cancer diagnosis. Serum-based proteomics has been focused on for its non-invasiveness and opportunity to reduce time for assaying. As in the case of pancreatic cancer, with antibody arrays, signatures of sera distinguish cancer patients from healthy controls⁷. Serum proteomics is also used in monitoring breast cancer patients to detect early local recurrence or metastatic deposits⁸. In addition to the serum-based proteomics, tissue-based proteomics gives opportunities to improve diagnostic sensitivity and disease classification. As Kashani-Sabet and co-workers reported, tissue proteomics through immunohistochemistry using five markers increases the sensitivity and specificity of the assay to 91% and 95%, respectively, for the diagnosis of melanoma⁹. In the case of lung cancer, Ring et al. described a five-marker diagnostic assay that distinguishes adenocarcinoma from squamous cell carcinoma¹⁰.

1.3. Metabolomics for Cancer Diagnosis

Besides the abnormal expression and post-modification of oncoproteins, cancer cells can be identified for their specific profiles of metabolites associated with the alteration of a particular metabolic pathway¹¹. In the case of the breast cancer, compared with healthy tissue or benign tumors, breast biopsy samples shows reliably elevated total choline-containing compounds (tCho), low glycerophosphocoline and low glucose through NMR studies^{12,13}. Similar to breast cancer, prostate cancer represents particular metabolic profile, especially high level of tCho, phosphocholine, lactate and alanine¹⁴.



As a form of *in vivo* metabolomics, positron emission tomography (PET) imaging has been widely used in clinics to diagnose cancer before the introduction of metabolomics. Glucose, among many others, is the most important metabolite in PET imaging due to the “Warburg effect”¹⁵, the phenomenon in cancer cells that cellular energy is generated through aerobic glycolysis instead of mitochondrial oxidative

phosphorylation. Although the exact mechanism is not yet well understood at present, most cancer cells uptake a high level of glucose compared with their normal counterpart, and their glucose uptake can be visualized by the measurement of radioactivity from fluorodeoxyglucose (FDG), a radioactive glucose analog¹⁶.

1.4. Thesis Overview

This thesis presents the development of biotechnologies for *in vitro* cancer diagnostics. In Chapter 2, I will introduce my contribution to the development of cell sorting technology based on DNA-encoded p/MHC complexes for antigen-specific T cells. Through the development of genetic engineering, T cells are able to recognize cells presenting particular antigens, including cancer-specific antigens. Exploiting this characteristic of T cells selecting target populations of cells, researchers has been developed a new way of cancer treatment, so called T cell immunotherapy. By the injection of engineered and activated cancer-specific T cells to patients, the regression of tumor has been reported in subsets of patients with metastatic cancers. However, it was not easy to characterize the antigen-specific T cells during the treatment due to their small populations. To achieve highly multiplexed T cell detection, Chapter 2 will introduce the technique named as “Nucleic Acid Cell Sorting (NACS)”. By employing p/MHC tetramers site-specifically conjugated with single-stranded DNA oligomers (ssDNA), specific target T cells from cellular suspensions can be captured and immobilized via hybridization to a complementary-printed substrate. The immobilization of T cells is optimized by engineering streptavidin to have cysteine residue at a specific site for the formation of ssDNA-p/MHC tetramer guaranteeing the superiority of NACS over conventional spotted arrays. With restriction enzymes and sophisticatedly designed ssDNA, the selective release of sorted T

cells is also investigated. Finally, the capability of NACS as a clinical application is studied by monitoring a cancer-specific T cell population from a melanoma patient undergoing T cell immunotherapy. Chapter 2 has been taken in part from *J. Am. Chem. Soc.*, **2009**, *131*(28), 9695–9703.

In order to clinical treatments for cancer cells, it is highly important to understand their protein-protein signaling network. With the development of single-cell barcode chip (SCBC), the measurement of multiple proteins in single-cells became available, but more studies were needed for the quantitative understanding of protein-protein network to predict the effect of perturbation, especially induced by targeted cancer therapy. In Chapter 2, I will also explain briefly about my contribution to develop experimental and theoretical methods for measuring protein secretion level of single or few cells and analyzing the results quantitatively. As a model system, mimicking innate immune response of human macrophages to Gram-negative bacteria, macrophages stimulated with lipopolysaccharide are employed. The secreted proteins are measured with SCBC to generate a covariance matrix for the analysis through a quantitative version of Le Chatelier principle, derived from information theory. Importantly, through the quantitative analysis, it is possible to rank the contributions of proteins to the entire network as well as to generate the diagram of protein-protein correlation network. Lastly, without actual experiments, the role of perturbation in the network can be theoretically predictable from the fluctuation of protein secretions. This prediction is then validated with the experiments with cells treated neutralizing antibodies to eliminate the secretion of certain proteins, which shows great accordance with the experimental result. Chapter 2 has also been taken in part from *Biophys. J.*, **2011**, *100*(10), 2378-2386.

In addition to proteins, metabolites can be used to diagnose cancer. Due to their high proliferation rate, in general, cancer cells uptake huge amount of glucose compared to their normal counterpart. For the monitoring of metabolites in small number of cells, in Chapter 3, a new platform, composed of a beta-particle camera and a microfluidic chip, will be introduced. The microfluidic chip, so called “Radiopharmaceutical Imaging Chip(RIMChip)” is sophisticatedly designed for the visualization of radio probes inside of cells with a small sample size (10 to 100 cells). The performance of RIMChip is validated with the uptake of [F-18]fluorodeoxyglucose (FDG), a glucose analog, in adherent and suspension cell lines. This chip is then utilized to study the response kinetics of lymphoma and glioblastoma cell lines under gemcitabine and erlotinib treatment, respectively. Within short time (~ 1 hour), the decrease of FDG uptake in those treated cells is observed, correlated to the arrest of cell-cycle (gemcitabine treatment to lymphoma cells) or the decreased signaling in epidermal growth factor receptor (EGFR) pathway (erlotinib treatment to glioblastoma cells) confirmed by separate cell-based assays. Chapter 3 has been taken in part from *J. Nucl. Med.*, **2013**, *in press*. By applying technique, clinical researches have been also proceeding to understand the glycolytic metabolism in primary glioblastoma cells. The results of ongoing study, imaging of glucose uptake in glioblastoma cells according to their expression level of oncoprotein (EGFRvIII) or in gliolastoma cells under different types of drug treatments, are briefly presented in Chapter 5.

Chapter 4 describes the ongoing project to develop novel agents for in vivo imaging of c-MET, a receptor tyrosine kinase largely amplified in prostate cancer cells.

The protein-catalyzed capture agents, PCCs, are synthesized from artificial amino acids by iteratively screening one-bead-one-compound combinatorial library method. Compared to commercial antibodies, PCC is inexpensive and stable for long-term storage at room temperature. The 1st generation of PCC for c-MET is developed and its performance is validated with cell-based fluorescent imaging assays for the quantitative comparison with commercial anti c-MET antibody.

Lastly, for reference, *J. Am. Chem. Soc.*, **2009**, *131*(28), 9695–9703 and *Biophys. J.*, **2011**, *100*(10), 2378–2386 are added as Appendix A and B.

1.5. References

1. Fisher, R., Pusztai, L. & Swanton, C. Cancer heterogeneity: implications for targeted therapeutics. *Br J Cancer* **108**, 479–485 (2013).
2. Barteneva, N. S., Ketman, K., Fasler-Kan, E., Potashnikova, D. & Vorobjev, I. A. Cell sorting in cancer research-Diminishing degree of cell heterogeneity. *BBA-Reviews on Cancer* **1836**, 105–122 (2013).
3. Sklar, L. A. & Finney, D. A. Analysis of ligand-receptor interactions with the fluorescence activated cell sorter. *Cytometry* **3**, 161–165 (1982).
4. Wolff, A. *et al.* Integrating advanced functionality in a microfabricated high-throughput fluorescent-activated cell sorter. *Lab Chip* **3**, 22–27 (2003).
5. Simonnet, C. & Groisman, A. High-throughput and high-resolution flow cytometry in molded microfluidic devices. *Anal. Chem.* **78**, 5653–5663 (2006).
6. Nagrath, S. *et al.* Isolation of rare circulating tumour cells in cancer patients by microchip technology. *Nature* **450**, 1235–1239 (2007).
7. Orchekowski, R. *et al.* Antibody microarray profiling reveals individual and combined serum proteins associated with pancreatic cancer. *Cancer Res* **65**, 11193–11202 (2005).

8. Carlsson, A. *et al.* Serum protein profiling of systemic lupus erythematosus and systemic sclerosis using recombinant antibody microarrays. *Mol. Cell Proteomics* **10**, M110.005033 (2011).
9. Kashani-Sabet, M. *et al.* A multi-marker assay to distinguish malignant melanomas from benign nevi. *P Natl Acad Sci USA* **106**, 6268–6272 (2009).
10. Ring, B. Z. *et al.* A novel five-antibody immunohistochemical test for subclassification of lung carcinoma. *Mod. Pathol.* **22**, 1032–1043 (2009).
11. Spratlin, J. L., Serkova, N. J. & Eckhardt, S. G. Clinical applications of metabolomics in oncology: a review. *Clin Cancer Res* **15**, 431–440 (2009).
12. Bathen, T. F. *et al.* MR-determined metabolic phenotype of breast cancer in prediction of lymphatic spread, grade, and hormone status. *Breast Cancer Res Treat* **104**, 181–189 (2007).
13. Sitter, B. *et al.* Comparison of HR MAS MR spectroscopic profiles of breast cancer tissue with clinical parameters. *NMR Biomed* **19**, 30–40 (2006).
14. Swanson, M. G. *et al.* Quantitative analysis of prostate metabolites using ¹H HR-MAS spectroscopy. *Magn Reson Med* **55**, 1257–1264 (2006).
15. Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. *Science* **324**, 1029–1033 (2009).
16. Phelps, M. E. Positron emission tomography provides molecular imaging of biological processes. *P Natl Acad Sci USA* **97**, 9226–9233 (2000).

Chapter 2

Quantifying the performance of a new T cell sorting technique, and the preparation of reagents for quantitative analysis of proteins from macrophages

2.1. Introduction

In this chapter, I will briefly present the background of two projects in which I took participation at the early years of my graduate studies and give explanations about my contribution on each projects: the quantification of a novel sorting method for tumor antigen specific T cells and the preparation of reagents for the quantitative analysis of protein signaling network in macrophages.

T cells, which play a central role in an adaptive immune response, recognize different antigens via the interaction between T cell receptor (TCR) and antigenic peptides bound to Major Histocompatibility Complex (p/MHC) on the surface of antigen presenting cells (APC). In order to detect antigen-specific T cells, four enzymatically biotinylated p/MHC monomers are coupled with a streptavidin molecule primed by chromophore to form a p/MHC tetramer, and the tetramer enables its counter part T cell to be detected via flow cytometry¹.

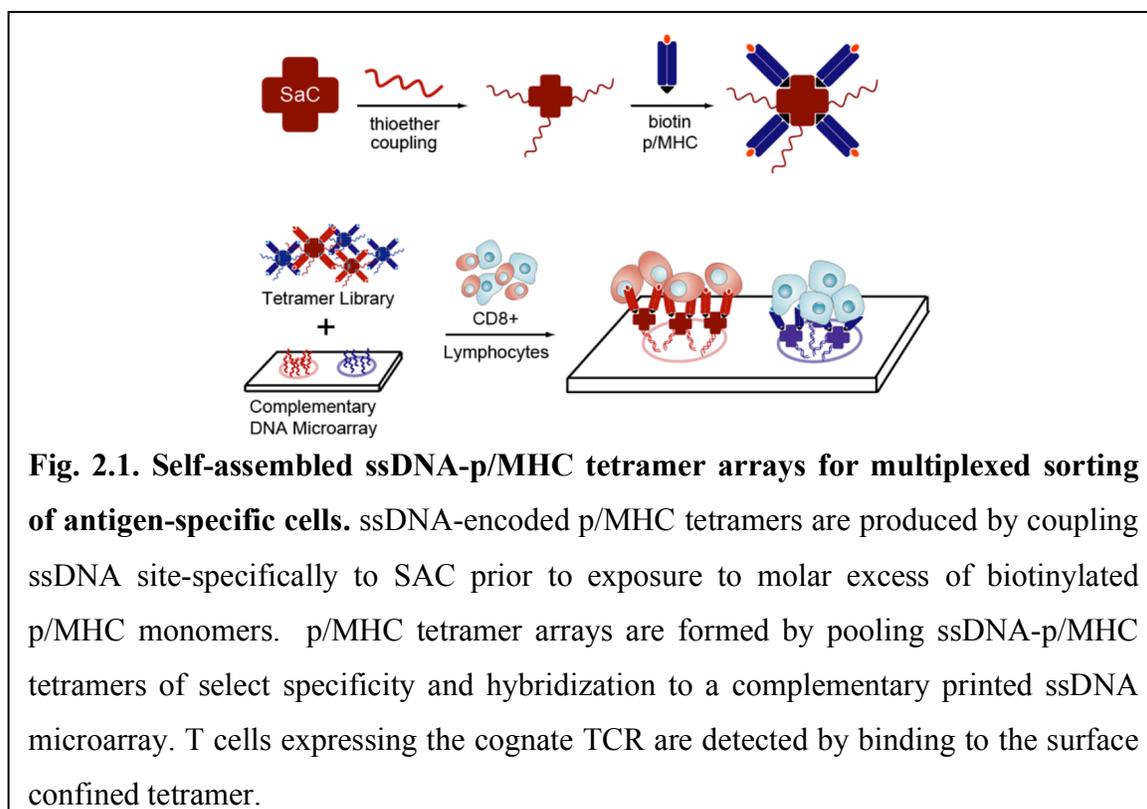
The p/MHC tetramer based flow cytometry has inevitable restrictions. At first, each antigen-specificity requires unique optical dye molecule having little spectral overlap with others. Secondly, the existence of detection limit in flow cytometry, serial

flow cytometry detection for multiplexing study is restricted by the sample size. Due to the existence of diverse optical characters in each cell population, moreover, the matter of compensation always follows. Although researchers used polychromatic flow cytometry utilizing quantum dots to resolve those problems²⁻⁴, high cost, long sample preparation time, and complex color compensation still hinder further study.

To avoid such problems, several researchers have reported antigen-specific cell sorting by microarrays, which contains directly printed p/MHC tetramers on a supporting substrates⁵⁻⁸. When a population of cells is applied on the array, only target antigen-specific T cells bind to the region having p/MHC tetramers. Because this method uses the location of p/MHC tetramers to separate a target antigen-specific T cell from the variety of other cells instead of the chromophore's emission spectrum, multiplexing is only restricted the number of reagent able to fit on a substrate.

In order to increase the capture efficiency, orientation of p/MHC tetramer should be preserved in a way to react approaching cells and the tetramer must not be denatured through the microarray fabrication. Analogous protein arrays, made via antibody adsorption to unmodified and derivatized surfaces, have shown that the adsorbed antibodies could fully react because of surface-induced effects such as protein denaturation and orientational change to inactive. As the result, amount of functional antibody is decreased, immobilization occurs heterogeneously, and required concentrations of antibody is increased to compensate for the protein loss⁹⁻¹¹. To avoid these problems, several mild chemistries have been studied for protein immobilization¹²⁻¹⁸, but often the surface that meets the demands of application requires a high level of

technical expertise and/or is limited in accessibility. Therefore, the development of a new technology to fabricate p/MHC array in an easy way is important.



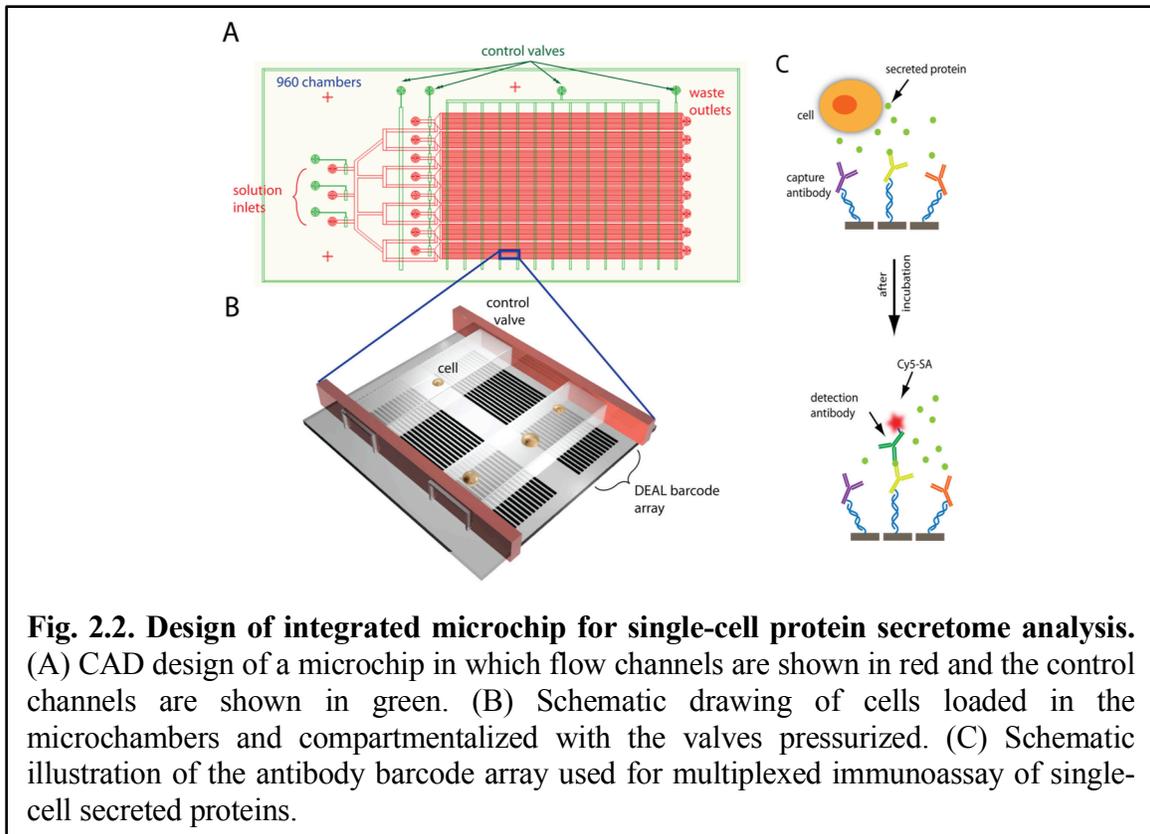
In Heath lab, a new p/MHC arrays conjugated with nucleic acid for multiplexed antigen-specific lymphocytes sorting is developed. Instead of direct printing, each specific p/MHC tetramer is conjugated to unique sequence of ssDNA designed to be orthogonal to other sequences. On the glass surface where the complementary DNA sequences are printed, the ssDNA-p/MHC tetramer conjugates are self-assembled by DNA hybridization, then used as a p/MHC array to sort mixed population of antigen-specific T cells (**Fig. 2.1.**). This method is called “Nucleic Acid Cell Sorting (NACS).” Because NACS employs DNA as a linkage molecule, fabrication of p/MHC arrays is simple and highly modular¹⁹⁻²². Traditional DNA microarray technologies are also available to make cDNA glass substrates. In order to produce ssDNA-p/MHC tetramer

conjugates having maximum capacity to immobilize cells, cysteine-engineered SA (SAC) is used as a scaffold. NACS arrays have better performance than traditional p/MHC arrays made by direct printing, and the specificity, multiplexing and sensitivity of NACS are studied. Selective detachment after cell sorting by NACS and its application to primary human T cells is also introduced. The detailed results are shown in Appendix A, as well as in *J. Am. Chem. Soc.*, **2009**, *131*(28), 9695–9703.

Protein-signaling pathways play important roles in tissue processes ranging from tumorigenesis to wound healing. Elucidation of these signaling pathways is challenging, in large part²³⁻²⁷, because of the heterogeneous nature of tissues⁴. Such heterogeneity makes it difficult to separate cell-autonomous alterations in function from alterations that are triggered via paracrine signaling, and it can mask the cellular origins of paracrine signaling. Intracellular signaling pathways can be resolved via multiplex protein measurements at the single-cell level²⁸. For secreted protein signaling, there are additional experimental challenges. Intracellular staining flow cytometry (ICS-FC) requires the use of protein transport inhibitors which can influence the measurements³. In addition, the largest number of cytokines simultaneously assayed in single-cells by ICS-FC is only 5². Finally, certain biological perturbations, such as the influence of one cell on another, are difficult to decipher using ICS-FC. Other methods, such as multiplex fluorospot assays^{23-27,29}, have even more significant limitations.

In Heath lab, an experimental/theoretical approach designed to unravel the coordinated relationships between secreted proteins, and to understand how molecular and cellular perturbations can influence those relationships. The starting points are single,

lipopolysaccharide (LPS)-stimulated, human macrophage cells^{4,30}. LPS stimulation activates the Toll-like Receptor-4 (TLR-4), and emulates the innate immune response to Gram-negative bacteria. We characterize the secretome, at the single-cell level, through the use of a microchip platform in which single, stimulated macrophage cells are isolated into 3 nanoliter (nl) volume microchambers, with ~1000 microchambers per chip. Each microchamber permits duplicate assays for each of a dozen proteins that are secreted over the course of a several-hour incubation period following cell stimulation. The barcode assays are developed using detection antibodies and fluorescent labels, and then converted into numbers of molecules detected (**Fig. 2.2.**). We demonstrate that the observed spread in protein levels is dominated by the cellular behaviors (the biological fluctuations), rather than the experimental error. These fluctuations are utilized to compute a covariance matrix linking the different proteins. This matrix is analyzed at both coarse and fine levels to extract the protein-protein interactions. We demonstrate that our system has the stability properties requisite for the application of a quantitative version of a Le Chatelier-like principle, which permits a description of the response of the system to a perturbation. This is a prediction in the strict thermodynamic sense. The fluctuations, as assessed from the multiplexed protein assays from unperturbed single-cells, are used to predict the results when the cells are perturbed by the presence of other cells, or through molecular (antibody) perturbations. The detailed results are shown in Appendix B, as well as in *Biophys. J.*, **2011**, *100(10)*, 2378-2386.



2.2. Experimental Methods

2.2.1. Microarray fabrication for T cell sorting

All DNAs were acquired from IDT. By the facility at the Institute for Systems Biology (ISB, Seattle), DNA strands were printed on amine coated glass slides (GAPSII, Corning) in the way to form 12x12 arrays having alternative rows of A, B, and C spots with a SMPXB15 pin (Arrayit). Sequence of each strand and its counterparts is written at the following **Table 2.1**.

Table 2.1. Orthogonal DNA sequences for spatial encoding of p/MHC tetramers

Name	Sequence*
A	5' - AAA AAA AAA AAA AAT CCT GGA GCT AAG TCC GTA AAA AAA AAA AAT CCT GGA GCT AAG TCC GTA AAA AAA AAA AAA A
A'	5' - NH ₂ - AAA AAA AAA ATA CGG ACT TAG CTC CAG GAT
B	5' - AAA AAA AAA AAA AGC CTC ATT GAA TCA TGC CTA AAA AAA AAA AGC CTC ATT GAA TCA TGC CTA AAA AAA AAA AAA A
B'	5' - NH ₂ - AAA AAA AAA ATA GGC ATG ATT CAA TGA GGC
C	5' - AAA AAA AAA AAA AGC ACT CGT CTA CTA TCG CTA AAA AAA AAA AGC ACT CGT CTA CTA TCG CTA AAA AAA AAA AAA A
C'	5' - NH ₂ - AAA AAA AAA ATA GCG ATA GTA GAC GAG TGC

* All sequences to be conjugated to SAC (A', B', C') were designed with a polyA linker followed by a 20mer hybridization region. The 5' amine is required for the attachment of the hetero-bifunctional maleimide derivative MHPH. Sequences printed on glass substrates (A, B, C) were designed with two hybridization regions separated by polyAs. This was designed to facilitate electrostatic adsorption to amine glass substrates.

2.2.2. Synthesis of DNA-SAC conjugates

In order to express SAC, pET-3a plasmid, as a kind gift from Takeshi Sano (Harvard Medical School), was used. SAC was expressed according to the previously published protocol³¹, and buffer exchanged to PBS with 5 mM Tris(2-carboxyethyl) phosphine hydrochloride (Solulink), using zeba desalting columns (Pierce). MHPH (3-N-Maleimido-6-hydraziniumpyridine hydrochloride, Solulink) was added to SAC at a molar excess of 300:1. Meanwhile, SFB (Succinimidyl 4-formylbenzamide, Solulink) in DMF was added to 5' aminated DNA oligos at a molar excess of 40:1. In order to finish reactions, these mixtures were incubated at room temperature (RT) for 3-4 hours. Then,

both mixture were buffer exchanged to pH 6.0 buffer solution (50 mM sodium citrate, 150 mM NaCl), and SFB-conjugated DNA oligos were added to MHPH-conjugated SAC at a molar excess of 20:1, and incubated overnight at RT. With Phamacia Superdex 200 gel filtration column at 0.5 ml/min isocratic flow of PBS, SAC-DNA oligo conjugates was purified, and concentrated by 10K mwco concentration filters (Millipore).

2.2.3. Preparation of T cells

In order to make Jurkat^{a-MART-1} cells, MSGV1-F5AfT2AB retroviral vector expressing the F5 MART-1 TCR, obtained from Steven A. Rosenberg and Richard Morgan (Surgery Branch, National Cancer Institute Bethesda, MD), was used to infect Jurkat cells. Cells were culture with RPMI 1640 (ATCC) supplemented with 10% Fetal Bovine Serum (ATCC).

2.2.4. Sorting cells with NACS and conventional microarrays

The HLA-A*0201 restricted MHC class I monomers loaded with MART-126-35 (ELAGIGILTV) were made in house with previous published protocols³², and lipophilic cell membrane staining dyes DiO, DiD, and DiL were obtained from Invitrogen.

To prevent non-specific cell binding, 1mg/ml PEG-NHS ester (Sunbio) in PBS was applied on microarray slides for 2 hours at RT. p/MHC monomers were combined in a 4:1 molar excess with ssDNA-SAC at 37°C for 20 min to form ssDNA-p/MHC tetramers. With the tetramers in 200µl media, DNA microarrays was hybridized for 1 hour at 37°C, and washed with 3% FBS in PBS. T cells (10⁶ cells/ml) were incubated on the array for 30 min at 37°C, and washed with the same media.

To compare the performance of sorting with conventional microarray techniques, SuperEpoxy(covalent) and SuperProtein(hydrophobic) slides were obtained from Arrayit, GAPSII(electrostatic) slide was purchased from Corning, and Hydrogel(hydrophilic) slide was acquired from Xantec. According to manufacturer's protocol, fluorescent MART-1 tetramers were immobilized on each slide, and the result was analyzed with ImageJ and Origin. All cell capture images were obtained by bright field (Nikon Eclipse TE2000) and/or confocal microscopy (Nikon E800).

2.2.5. Microchip fabrication.

The Single-cell barcode chips (SCBC) were assembled from a DNA barcode microarray glass slide and a PDMS slab containing a microfluidic circuit^{22,33}. The DNA barcode array was created with microchannel-guided flow patterning technique³³. Each barcode was comprised of thirteen stripes of uniquely designed ssDNA molecules. PDMS microfluidic chip was fabricated using a two-layer soft lithography approach³⁴. The control layer was molded from a SU8 2010 negative photoresist (~ 20 μm in thickness) silicon master using a mixture of GE RTV 615 PDMS prepolymer part A and part B (5:1). The flow layer was fabricated by spin-casting the pre-polymer of GE RTV 615 PDMS part A and part B (20:1) onto a SPR 220 positive photoresist master at ~2000 rpm for 1minute. The SPR 220 mold was ~ 18 mm in height after rounding via thermal treatment. The control layer PDMS chip was then carefully aligned and placed onto the flow layer, which was still situated on its silicon master mold, and an additional 60 min thermal treatment at 80 °C was performed to enable bonding. Afterward, this two-layer PDMS chip was cut off and access holes drilled. In order to improve the biocompatibility of PDMS, we performed

a solvent extraction step, which removes uncrosslinked oligomers, solvent and residues of the curing agent through serial extractions/washes of PDMS with several solvents^{35,36}. We noticed that this step significantly improves the biocompatibility and the reproducible protein detection. Finally, the microfluidic-containing PDMS slab was thermally bonded onto the barcode-patterned glass slide to give a fully assembled microchip.

2.2.6. Preparation of barcode arrays

The barcode array initially consists of 13 uniquely designed DNA strands labeled in order as A through M. Prior to loading cells, a cocktail containing all capture antibodies conjugated to different complementary DNA strands (A'-L') is flowed through the chambers, thus transforming, via DNA hybridization, the DNA barcode into an antibody array. These dozen proteins that comprised the panel used here were encoded by the DNA strands A through L, respectively. The DNA oligomer sequences and the antibody pairs used are listed in **Table 2.2.** and **Table 2.3.** DNA-encoded antibodies were synthesized as previously described¹⁸ and quantified with BCA kit (Pierce) according to the manufacturer's guide.

Table 2.2. Sequences and terminal functionalization of oligonucleotides*.

Name	Sequence
A	5'- AAA AAA AAA AAA AGT CCT CGC TTC GTC TAT GAG-3'
A'	5' NH3-AAA AAA AAA ACT CAT AGA CGA AGC GAG GAC-3'
B	5'-AAA AAA AAA AAA AGC CTC ATT GAA TCA TGC CTA -3'
B'	5' NH3-AAA AAA AAA ATA GGC ATG ATT CAA TGA GGC -3'
C	5'- AAA AAA AAA AAA AGC ACT CGT CTA CTA TCG CTA -3'
C'	5' NH3-AAA AAA AAA ATA GCG ATA GTA GAC GAG TGC -3'
D	5'-AAA AAA AAA AAA AAT GGT CGA GAT GTC AGA GTA -3'
D'	5' NH3-AAA AAA AAA ATA CTC TGA CAT CTC GAC CAT -3'
E	5'-AAA AAA AAA AAA AAT GTG AAG TGG CAG TAT CTA -3'
E'	5' NH3-AAA AAA AAA ATA GAT ACT GCC ACT TCA CAT -3'
F	5'-AAA AAA AAA AAA AAT CAG GTA AGG TTC ACG GTA -3'
F'	5' NH3-AAA AAA AAA ATA CCG TGA ACC TTA CCT GAT -3'
G	5'-AAA AAA AAA AGA GTA GCC TTC CCG AGC ATT-3'
G'	5' NH3-AAA AAA AAA AAA TGC TCG GGA AGG CTA CTC-3'
H	5'-AAA AAA AAA AAT TGA CCA AAC TGC GGT GCG-3'
H'	5' NH3-AAA AAA AAA ACG CAC CGC AGT TTG GTC AAT-3'
I	5'-AAA AAA AAA ATG CCC TAT TGT TGC GTC GGA-3'
I'	5' NH3-AAA AAA AAA ATC CGA CGC AAC AAT AGG GCA-3'
J	5'-AAA AAA AAA ATC TTC TAG TTG TCG AGC AGG-3'
J'	5' NH3-AAA AAA AAA ACC TGC TCG ACA ACT AGA AGA-3'
K	5'-AAA AAA AAA ATA ATC TAA TTC TGG TCG CGG-3'
K'	5' NH3-AAA AAA AAA ACC GCG ACC AGA ATT AGA TTA-3'
L	5'-AAA AAA AAA AGT GAT TAA GTC TGC TTC GGC-3'
L'	5' NH3-AAA AAA AAA AGC CGA AGC AGA CTT AAT CAC-3'
M	5'-AAA AAA AAA AGT CGA GGA TTC TGA ACC TGT-3'
M'	5' Cy3-AAA AAA AAA AAC AGG TTC AGA ATC CTC GAC-3'

* All oligonucleotides were synthesized by Integrated DNA Technology (IDT) and purified via high-performance liquid chromatography (HPLC).

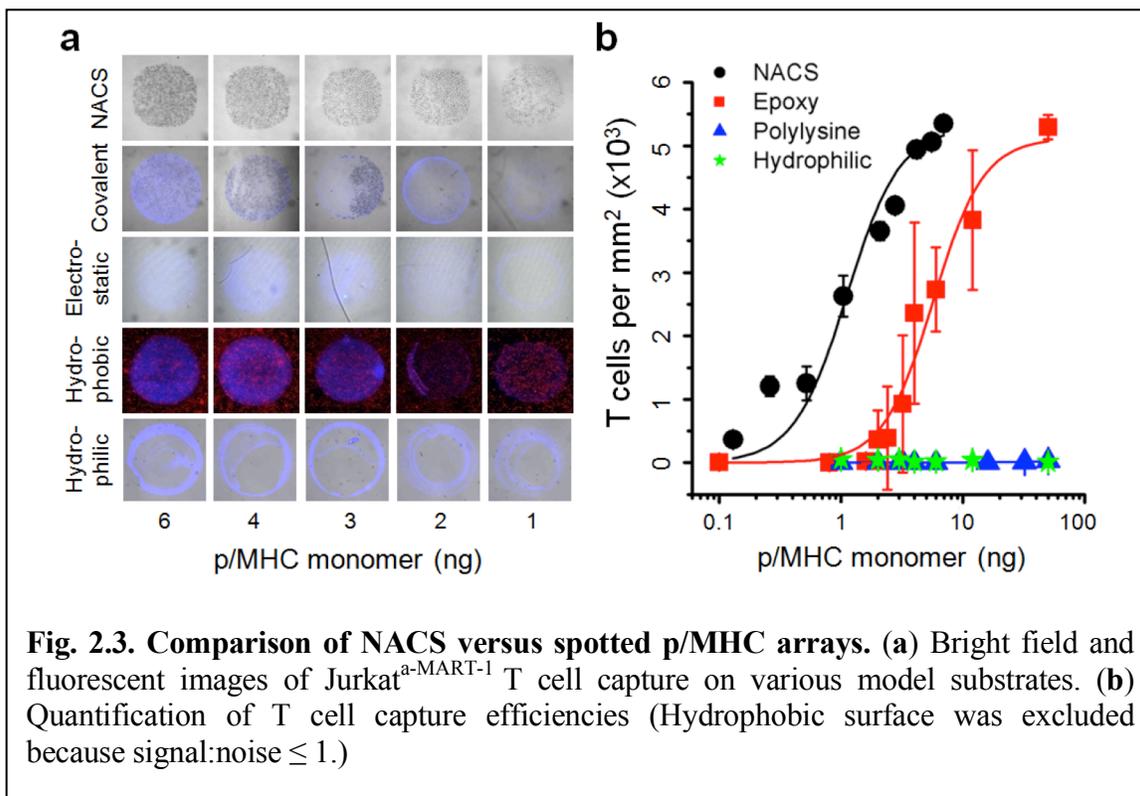
Table 2.3. Summary of antibodies used for macrophage experiments

DNA label	primary antibody (vendor)	secondary antibody (vendor)
A'	mouse anti-hu IL-2 (BD Biosciences)	biotin-labeled mouse anti-hu IL-2 (BD Biosciences)
B'	mouse anti-hu MCP-1 (eBioscience)	biotin-labeled armenian hamster anti-hu MCP-1 (eBioscience)
C'	rat anti-hu IL-6 (eBioscience)	biotin-labeled rat anti-hu IL-6 (eBioscience)
D'	rat anti-hu GMCSF (Biolegend)	biotin-labeled rat anti-hu GMCSF (Biolegend)
E'	goat anti-hu MIF(R&D systems)	biotin-labeled goat anti-hu MIF(R&D systems)
F'	mouse anti-hu IFN- γ (eBioscience)	biotin-labeled mouse anti-hu IFN- γ (eBioscience)
G'	mouse anti-hu VEGF (R&D systems)	biotin-labeled goat anti-hu VEGF (R&D systems)
H'	mouse anti-hu IL-1 β (eBioscience)	biotin-labeled mouse anti-hu IL-1 β (eBioscience)
I'	rat anti-hu IL-10 (eBioscience)	biotin-labeled rat anti-hu IL-10 (eBioscience)
J'	mouse anti-hu IL-8 (R&D systems)	biotin-labeled mouse anti-hu IL-8 (R&D systems)
K'	mouse anti-hu MMP9 (R&D systems)	biotin-labeled goat anti-hu MMP9 (R&D systems)
L'	mouse anti-hu TNF- α (eBioscience)	biotin-labeled mouse anti-hu TNF- α (eBioscience)

2.3. Results and Discussions

2.3.1. Comparison between NACS and conventional protein microarray

For the comparison study between NACS and conventional microarrays based on direct protein spotting method, I chose various substrates to cover different spectrum of commonly used strategies (covalent, electrostatic, hydrophobic, and hydrophilic adsorption) for protein immobilization. According to manufacturer's instruction, serial dilutions of MART-1 SA-PE tetramer (HLA-A2.1 MHC molecules loading melanoma epitope MART-1₂₆₋₃₅) were directly printed. Onto these microarray, Jurkat^{a-MART-1} T cells (Jurkat cells transduced with the F5 MART-1 TCR³⁷ specific for peptide epitope MART-1₂₆₋₃₅) were applied. Collected images of immobilized cells on the various substrates (**Fig. 2.3.a**) and quantified analysis (**Fig. 2.3.b**) were obtained.



Compared to conventional microarrays with identical concentration of p/MHC tetramers, NACS was superior. Electrostatic and hydrophilic immobilization have little/no captured T cells, and hydrophobic immobilization has huge noise. Even though covalent immobilization succeeded to capture T cells, it had intra-spot and inter-spot heterogeneity while NACS gave homogeneous result at the same concentration of tetramers. Furthermore, as shown in the quantified analysis, only with a fifth of materials required in covalent immobilization, NACS could capture equivalent T cell. (p/MHC monomer at half max $\equiv K_{1/2} = 1.1$ ng for NACS and 5.7 ng for covalent immobilization).

Two factors can be reasons of superior capturing efficiency in NACS. At first, linked by DNA-cDNA hybridization, p/MHC tetramers of NACS have great orientational freedom while those of conventional microarrays hardly have other options except just

absorbed on the surface. This may increase the density of functional p/MHC tetramers and result in reduced $K_{1/2}$. Secondly, NACS does not disrupt hydration state of arrayed proteins, which can be modified during production and storage process of conventional protein microarray to decrease reproducibility^{11,16,21}. Because p/MHC tetramer arrays are self-assembled in solution immediately prior to an experiment and only stable DNA oligos are printed on the surface when the chip is stored, NACS may show intra-spot and inter-spot homogeneity and guarantee the array reproducibility.

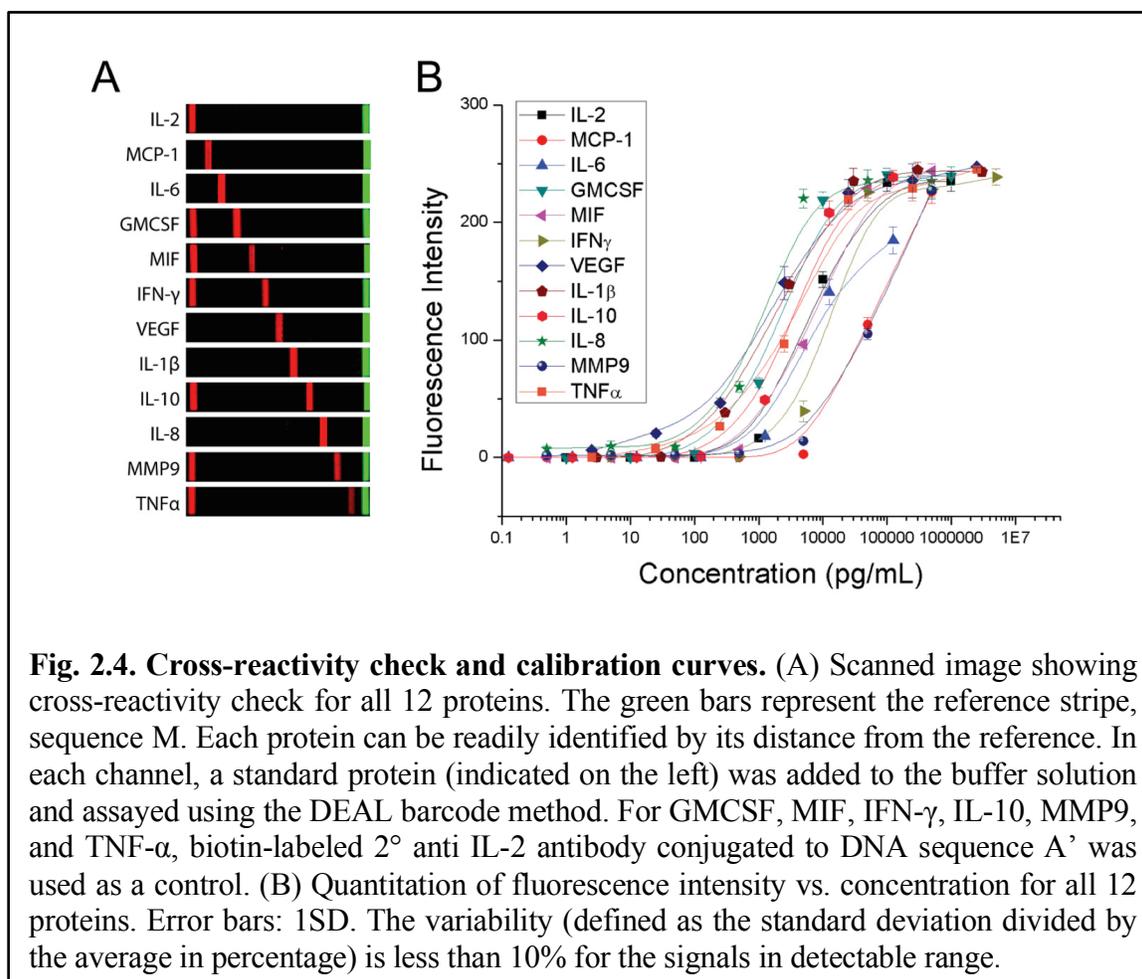
2.3.2. Preparation of DNA-encoded antibodies

In the case of protein measurement with SCBC, all the DNA-encoded antibodies should be quantified with BCA kit and calibrated with recombinant proteins before using in on-chip experiment. Their cross-reactivity also must be checked. Without using valves of SCBC, I verified the DNA conjugated antibodies (**Fig. 2.4.a**), and made calibration curves of each of them with serial dilutions of recombinant solutions (**Fig. 2.4.b**). The four parameter logistic model is used for fitting the calibration curve, resulting in the fitting parameters in **Table 2.4**.

$$y = \frac{A_1 - A_2}{1 + (x / x_0)^p} + A_2$$

Table 2.4. Parameters utilized for the protein assay calibration curve

	A1		A2		x_0		p		Statistics	
	Value	Error	Value	Error	Value	Error	Value	Error	Reduced Chi-Sqr	Adj. R-Square
IL-2	0	0	256	0	7659.58168	973.0838	1.12824	0.16788	91.39131	0.99224
MCP-1	0	0	256	0	65733.51686	4770.5	1.12607	0.09607	29.62623	0.99578
IL-6	0	0	256	0	16231.59942	4515.94	0.67887	0.12265	243.09932	0.95697
GMCSF	0	0	256	0	2451.99685	295.3281	1.2195	0.13013	72.59138	0.99458
MIF	0	0	256	0	7892.74068	483.8218	1.14428	0.07578	20.31714	0.99821
IFN- γ	0	0	256	0	14549.5316	2773.804	1.57222	0.26181	172.2368	0.98713
VEGF	0	0	256	0	1687.9445	225.4782	0.69008	0.05631	58.49911	0.99513
IL-1 β	0	0	256	0	2137.44388	208.9672	0.89593	0.07185	41.21361	0.99694
IL-10	0	0	256	0	3961.03661	328.4038	1.23209	0.08611	33.93572	0.99669
IL-8	0	0	256	0	1255.89317	225.9207	1.23262	0.19534	161.8703	0.98686
MMP9	0	0	256	0	70537.40022	1584.696	1.062	0.02495	2.60945	0.99961
TNF- α	0	0	256	0	4126.15703	661.2747	0.81683	0.09483	99.72583	0.99185



2.4. Conclusions

By the comparison with the commercial microarrays, I showed the superiority of NACS platform that can be inexpensively made by traditional DNA printing technologies. Also, I presented the calibration curves of DNA-encoded antibodies, which is essential for the further experiments with SCBC to analyze the protein-protein network in macrophages quantitatively and to predict the role of perturbations.

2.5. References

1. Altman, J. D. *et al.* Phenotypic analysis of antigen-specific T lymphocytes. *Science* **274**, 94–96 (1996).
2. Lamoreaux, L., Roederer, M. & Koup, R. Intracellular cytokine optimization and standard operating procedure. *Nat Protoc* **1**, 1507–1516 (2006).
3. Nomura, L., Maino, V. C. & Maecker, H. T. Standardization and optimization of multiparameter intracellular cytokine staining. *Cytometry A* **73**, 984–991 (2008).
4. Ariztia, E. V., Lee, C. J., Gogoi, R. & Fishman, D. A. The tumor microenvironment: key to early detection. *Crit Rev Clin Lab Sci* **43**, 393–425 (2006).
5. Chen, D. S. *et al.* Marked differences in human melanoma antigen-specific T cell responsiveness after vaccination using a functional microarray. *PLoS Med.* **2**, e265 (2005).
6. Soen, Y., Chen, D. S., Kraft, D. L., Davis, M. M. & Brown, P. O. Detection and characterization of cellular immune responses using peptide-MHC microarrays. *PLoS Biol* **1**, E65 (2003).
7. Stone, J. D., Demkowicz, W. E. & Stern, L. J. HLA-restricted epitope identification and detection of functional T cell responses by using MHC-peptide and costimulatory microarrays. *P Natl Acad Sci USA* **102**, 3744–3749 (2005).
8. Deviren, G., Gupta, K., Paulaitis, M. E. & Schneck, J. P. Detection of antigen-specific T cells on p/MHC microarrays. *J. Mol. Recognit.* **20**, 32–38 (2007).

9. Haab, B. B., Dunham, M. J. & Brown, P. O. Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. *Genome Biol.* **2**, RESEARCH0004 (2001).
10. Butler, J. E. *et al.* The physical and functional behavior of capture antibodies adsorbed on polystyrene. *J. Immunol. Methods* **150**, 77–90 (1992).
11. Butler, J. E. *et al.* The immunochemistry of sandwich ELISAs--VI. Greater than 90% of monoclonal and 75% of polyclonal anti-fluorescyl capture antibodies (CAbs) are denatured by passive adsorption. *Mol. Immunol.* **30**, 1165–1175 (1993).
12. MacBeath, G. & Schreiber, S. L. Printing proteins as microarrays for high-throughput function determination. *Science* **289**, 1760–1763 (2000).
13. Lesaichere, M.-L., Lue, R. Y. P., Chen, G. Y. J., Zhu, Q. & Yao, S. Q. Intein-mediated biotinylation of proteins and its application in a protein microarray. *J. Am. Chem. Soc.* **124**, 8768–8769 (2002).
14. Peluso, P. *et al.* Optimizing antibody immobilization strategies for the construction of protein microarrays. *Anal. Biochem.* **312**, 113–124 (2003).
15. Kwon, Y., Han, Z., Karatan, E., Mrksich, M. & Kay, B. K. Antibody arrays prepared by cutinase-mediated immobilization on self-assembled monolayers. *Anal. Chem.* **76**, 5713–5720 (2004).
16. Arenkov, P. *et al.* Protein microchips: use for immunoassay and enzymatic reactions. *Anal. Biochem.* **278**, 123–131 (2000).
17. Kiyonaka, S. *et al.* Semi-wet peptide/protein array using supramolecular hydrogel. *Nat Mater* **3**, 58–64 (2004).
18. 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**, 1959–1967 (2007).
19. Boozer, C. *et al.* DNA directed protein immobilization on mixed ssDNA/oligo(ethylene glycol) self-assembled monolayers for sensitive biosensors. *Anal. Chem.* **76**, 6967–6972 (2004).
20. Niemeyer, C. M. Functional devices from DNA and proteins. *Nano Today* **2**, 42–52 (2007).

21. Chandra, R. A., Douglas, E. S., Mathies, R. A., Bertozzi, C. R. & Francis, M. B. Programmable cell adhesion encoded by DNA hybridization. *Angew Chem Int Ed Engl* **45**, 896–901 (2006).
22. Fan, R. *et al.* Integrated barcode chips for rapid, multiplexed analysis of proteins in microliter quantities of blood. *Nat Biotechnol* **26**, 1373–1378 (2008).
23. Lin, W.-W. & Karin, M. A cytokine-mediated link between innate immunity, inflammation, and cancer. *J Clin Invest* **117**, 1175–1183 (2007).
24. Gnechi, M. *et al.* Paracrine action accounts for marked protection of ischemic heart by Akt-modified mesenchymal stem cells. *Nature Publishing Group* **11**, 367–368 (2005).
25. Croci, D. O. *et al.* Dynamic cross-talk between tumor and immune cells in orchestrating the immunosuppressive network at the tumor microenvironment. *Cancer Immunol. Immunother.* **56**, 1687–1700 (2007).
26. Seruga, B., Zhang, H., Bernstein, L. J. & Tannock, I. F. Cytokines and their relationship to the symptoms and outcome of cancer. *Nat Rev Cancer* **8**, 887–899 (2008).
27. Polyak, K. & Weinberg, R. A. Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. *Nat Rev Cancer* **9**, 265–273 (2009).
28. Sachs, K., Perez, O., Pe'er, D., Lauffenburger, D. A. & Nolan, G. P. CaUSAI protein-signaling networks derived from multiparameter single-cell data. *Science* **308**, 523–529 (2005).
29. Cox, J. H., Ferrari, G. & Janetzki, S. Measurement of cytokine release at the single-cell level using the ELISPOT assay. *Methods* **38**, 274–282 (2006).
30. Song, M. & Phelps, D. S. Comparison of SP-A and LPS effects on the THP-1 monocytic cell line. *Am. J. Physiol. Lung Cell Mol. Physiol.* **279**, L110–7 (2000).
31. SANO, T. & Cantor, C. R. Expression of a cloned streptavidin gene in *Escherichia coli*. *P Natl Acad Sci USA* **87**, 142–146 (1990).
32. Garboczi, D. N., Hung, D. T. & Wiley, D. C. HLA-A2-peptide complexes: refolding and crystallization of molecules expressed in *Escherichia coli* and complexed with single antigenic peptides. *P Natl Acad Sci USA* **89**, 3429–3433

- (1992).
33. Shin, Y. S. *et al.* Chemistries for patterning robust DNA microbarcodes enable multiplex assays of cytoplasm proteins from single cancer cells. *Chemphyschem* **11**, 3063–3069 (2010).
 34. Quake, S. R. & Scherer, A. From micro- to nanofabrication with soft materials. *Science* **290**, 1536–1540 (2000).
 35. Millet, L. J., Stewart, M. E., Sweedler, J. V., Nuzzo, R. G. & Gillette, M. U. Microfluidic devices for culturing primary mammalian neurons at low densities. *Lab Chip* **7**, 987–994 (2007).
 36. Lee, J. N., Park, C. & Whitesides, G. M. Solvent compatibility of poly(dimethylsiloxane)-based microfluidic devices. *Anal. Chem.* **75**, 6544–6554 (2003).
 37. Johnson, L. A. *et al.* Gene transfer of tumor-reactive TCR confers both high avidity and tumor reactivity to nonreactive peripheral blood mononuclear cells and tumor-infiltrating lymphocytes. *J Immunol* **177**, 6548–6559 (2006).

Chapter 3

Fast Metabolic Response to Drug Intervention through Analysis on a Miniaturized, Highly Integrated Molecular Imaging System

3.1. Introduction

In vivo molecular imaging assays, employing a variety of probes of specific biological processes, have been developed for positron emission tomography (PET). The most common probe in patients care and research is the glucose analog, 2-deoxy-2- ^{18}F fluoro-deoxy-D-glucose (^{18}F FDG) for imaging and measuring rates of glycolysis. In cancer patients, ^{18}F FDG assays are used for diagnosis and assessing therapeutic responses¹. Around 3,000 molecular imaging probes for PET have been reported for various metabolic and other processes associated with disease states¹.

In vitro metabolic assays using radio-labeled probes have been adapted to 96-well plates and microchip formats². We explore here the use of such assays for quantitating the kinetics of cellular responses to targeted drugs. We first introduce a microfluidic chip design (the RIMChip) that couples to a beta-particle imaging camera²⁻⁴ to form the betabox. The betabox is designed for the quantitative analysis of the metabolic response of small numbers of cells to pharmaceuticals. Most assay steps are similar to those of standard 96-well plate radioassays, but require far fewer cells, permit quantitation of signal per cell, and utilize live cells. The platform is validated on various adherent and suspension cancer cells. We characterize the influence of the monocarboxylate transporter (MCT1) protein, as well as a glycolytic inhibitor, on ^{18}F FDG uptake in isogenic liposarcoma cells. Using different liposarcoma cells, we investigate the influence of

deoxycytidine kinase (dCK)⁵ on the uptake of the deoxycytidine analog molecular imaging probe [¹⁸F]-FAC⁶.

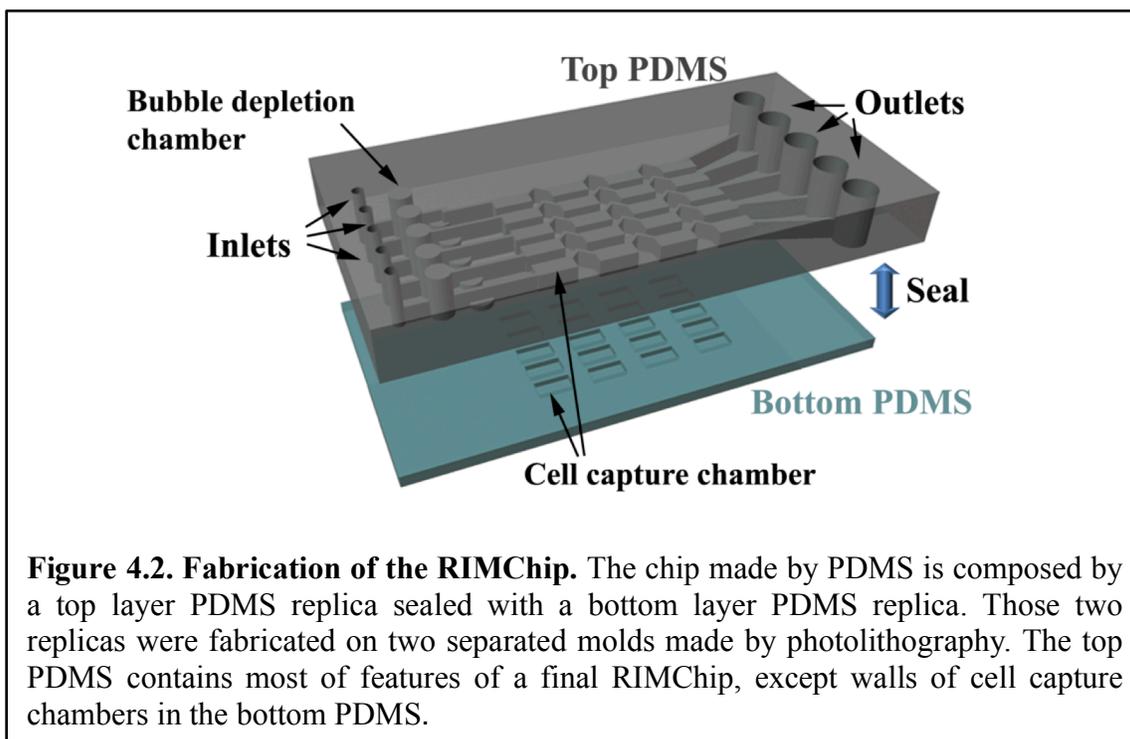
We then explore how certain cancer cell lines respond to two mechanistically distinct targeted inhibitors. We quantitate the response kinetics of liposarcoma cells to gemcitabine⁷ by correlating cell-cycle arrest with [¹⁸F]FDG uptake. We then quantitate the kinetic response of model glioblastoma multiforme (GBM) cancer cells to an epidermal growth factor receptor (EGFR) inhibitor, by correlating changes in [¹⁸F]FDG uptake with the levels of phosphoproteins associated with EGFR signaling. Glucose consumption consistently provides a rapid (~ 30 min.) indicator of positive therapeutic response, and the betabox platform provides a simple tool for quantitating those kinetics.

3.2. Experimental Methods

3.2.1. RIMChip design concepts

The RIMChip design has 5 microchannels for executing 5 distinct assay conditions. Each microchannel has 4 microchambers in which the assays are executed, and so a single assay condition is repeated four times. Several key features are designed into each microchannel to ease the execution of processes such as imaging, cell loading, and reagent flushing and delivery. First, the spatial resolution of the beta camera is 600 μm , and nearest edges of neighboring microchambers are separated by 800 μm to avoid signal overlap (**Fig. 3.1**). Second, the microchannels are designed so that cells are only captured within the microchambers. An individual cell chamber contains a 7 \times 3 waffle-structured array of 200 μm deep microwells, separated by 30 μm thick walls. This structure is

designed to minimize the shear stress exerted onto cells that are attached to the bottom of microwells when flushing through new medium or reagents. This waffle design also helps maintain the rigidity of the bottom surface. Third, this bottom surface is only 50 μm thick due to the constraint that the assay sensitivity depends upon the proximity of the cells to the camera. This thin-floored PDMS chip provides a $\sim 130\%$ increase in collected signal over a common glass microscope coverslip. It also means that cells that do get trapped in places away from the microchambers do not provide detectable signal during an assay, since those cells will have up to 1 mm of PDMS separating them and the beta camera. Fourth, the microchannel inlet is designed to avoid clogging of the channels by either bubbles, clumps of cells, or other debris, and for ease of use. The microchannel inlets accommodate a 20-200 μl pipette tip so that standard microwell procedures may be used. The bubble depletion chamber has a volume of 0.95 μl , and successfully removes any bubbles that are introduced in the micropipetting steps. The bubble depletion volume is followed by a filtering area of posts that have a tailored depth of between 25 and 35 μm . This dimension is customized for specific cell-types. In this study, we designed a 32 μm gap for liposarcoma cell lines and the U87 isogenic cell lines, and 20 μm gap for all suspended cells (which have a smaller cell size). The filter design also significantly enhances the hydrodynamic resistance of a channel and prevents backflow at the completion of solution injection. This means that cell loading uniformity is significantly increased. The outlet volume of the microchip holds all of the overflow from a given micropipetting step, and that overflow is removed using a micropipette. The microchamber area of the RIMChip area has a 1 cm \times 1 cm footprint, and which is designed to match the beta-particle camera.



3.2.2. RIMChip fabrication

The microchip is fabricated from multiple layers of the elastomer polydimethylsiloxane (PDMS) using standard soft lithography methods as previously described³. PDMS is air permeable, which is helpful for both removing bubbles through the bubble depletion chamber, and also for assisting in on-chip cell incubation. The filter layer was patterned on a 4" silicon wafer, using negative photoresist SU8 2025 (Microchem). On the same wafer, the fabrication of a layer of SU8 2100 generated the microchannel features that mate to the filters. This layer has a depth of 200 μm , which is also the microchannel depth. This ensures a sufficient nutrition supply to allow several hours of on-chip cell incubation. The cell chamber features were patterned using SU8 2025 on a separate wafer. Those two masters were used to mold the PDMS. Precure PDMS mixture at the ratio of 5:1

was poured onto the master with the microchannel and filter layers, and baked at 80 °C for 1 h to cure the PDMS (top PDMS in **Fig. 3.1.**). Meanwhile another PDMS precure mixture at the ratio of 20:1 was molded on to the second master. This PDMS replica was peeled off from this master, flipped and laid flat on a clean wafer (bottom PDMS in **Fig. 3.1.**). The top PDMS and the bottom PDMS layers were then trimmed and punched with appropriate inlet and outlet holes at 1.5 mm and 5 mm diameters, respectively, and then aligned and mated. The microchip was baked at 80 °C overnight to seal the PDMS layers. Finally, the entire microchannel was filled with PBS solution through dead-end filling by exerting 3 psi N₂ pressure in a tubing connected to the microchip

3.2.3. The beta-particle camera (Betabox)

The Betabox is a direct detection beta-particle detection camera utilizing a 13.5 x 13.5 mm² active area position sensitive avalanche photodiode (PSAPD) (Radiation Monitoring Devices). The five outputs of the PSAPD first pass through CR-110 charge sensitive preamplifiers and then shaping amplifiers with a 200 ns shaping time. The shaped sum signal then passes to a threshold comparator which produces event trigger pulses to four sample and hold (S/H) circuits (53ns acquisition time, 6μs hold time) to initiate an analog to digital conversion of the positioning channels using a simultaneous sampling data acquisition system (DAQ) (National Instruments PCI-6143).

3.2.4. Cell sample preparation, viability, and cell-cycle assays

Liposarcoma cell lines LPS1 and LPS2 with dCK- and MCT1-knockdown, respectively, were derived from patient samples. Lentiviral-based, shRNA-mediated

knockdown of MCT1 and dCK were described in the previous literature⁸, The murine leukemic lines (L1210 wt and L1210-10K)⁹ were a kind gift from Charles Dumontet (Université Claude Bernard Lyon I, Lyon, France)¹⁰. The human lymphoma line CEM was purchased from ATCC (#CCL-119) and the sub-line, CEM-dCK negative, was generated via selection with ara-C¹¹ and was a gift of Margaret Black (Washington State University). A human leukemia T cell line (Jurkat T) was purchased from ATCC. The human glioblastoma cell line U87 EGFRvIII/PTEN were prepared as described¹². The LIVE/DEAD® Viability/Cytotoxicity Kit (Invitrogen) was used to distinguish live cells from dead cells. For the cell-cycle assay, 2×10^6 cells were collected and washed with PBS. DNA content was determined through staining with 50 $\mu\text{g}/\text{mL}$ Propidium Iodide (Sigma) for L1210 cells or BrdU kit (R&D Systems) for U87 EGFRvIII/PTEN cells. Data were acquired on 4 and 5-laser LSRII cytometers (BD Biosciences) and analyzed as previously described¹³.

3.2.5. Betabox radioassay

Each RIMChip assay relies on the use of a pipettor for fluidic control. To replace a solution in microchannels, we always flow a solution from inlets and remove the excess at the outlets using a 100 μl pipettor. For adherent cells, the cells were prepared with fibronectin surface coating. Immediately before cell loading, the outlets were drained of solution. Cells at a concentration of 3×10^6 were injected into the microchip from inlets by pressing a pipettor for 5 sec to sufficiently cover all microchannels with cells. Then the outlets were refilled with cell culture medium to prevent the connected microchannels dried up owing to evaporation during incubation in a CO₂ incubator at 37 °C for 4 h. After

incubation, the microchannels were flushed 2× with flowing ~20 µl PBS using the same pipettor, each time for 30 sec, to flush out any residual glucose. About 10 µl ^{18}F -FDG or ^{18}F -FAC in PBS was loaded into the microchannels by pressing the pipettor for 15 sec. The whole microchip was then incubated for 30 min in a CO_2 incubator at 37 °C. It was then flushed 2× with flowing ~20 µl PBS using the same pipettor. The cells in the microchip were then imaged by the beta-box.

For on-chip treatment of cells, we introduced the cells onto the RIMChip 4 h prior to the radioassay. 5mM 2-deoxyglucose (2DG), 10 µM gemcitabine, or 5µM erlotinib, in RPMI 1640 or DMEM medium supplemented with 10% FBS was then added to the microchannels by pipetting, and incubated with cells for 2 hours or designated periods for kinetic study. Drug molecules were removed from the microchannels using two PBS flushing steps. Immediately following, 1 mCi activity of ^{18}F -FDG in PBS was introduced into the microchambers. One microchannel was loaded with cells, but not treated with gemcitabine. This provided a control assay. The signal from this control was converted into activity/cell, and then used for normalization of the similarly converted signal from the drug-treated cells.

The imaging acquisition time depends on the activity of radioactive probe used for radioassay, with other experimental conditions set. We used 1mCi activity of ^{18}F -FDG for tests on both adherent and suspended cells, with a fixed incubation time of 30 min. The imaging acquisition time was set for 5 min. We also explored lowering the dosage of ^{18}F -FAC to 10 µCi/ml for suspended cells, and using an acquisition time of 10 min. Even with such low activity, the radioassay on our platform is still able to differentiate leukemic lines L1210 and its dCK knockdown line.

3.2.6. Off-chip radioassay

In parallel with betabox radioassay, off-chip radioassays were performed for each cell sample. Liposarcoma cells were detached from a Petri dish and transferred to a 12 well plate, with each well containing $\sim 10^4$ cells. The poly-D-lysine coated plates were then placed in a CO₂ incubator at 37 °C for 4 h to allow for cell attachment to well plate bottom. 10 μ Ci ¹⁸F-FDG or ¹⁸F-FAC was added to each well, followed by 30 minute incubation. Subsequently each well was washed by PBS twice. The cells were then lysed, and the cell lysates were transferred to plastic vials. The radioactivity of each cell sample was measured using a well-type γ -counter (1480 Wizard 3; Perkin Elmer).

3.2.7. Phosphoprotein assay

Confluent Petri dishes containing 5×10^6 cells were prepared. Lysis buffer (Cell Signaling) with protease inhibitor (Roche) and phosphatase inhibitor (Sigma), was prepared according to the manufacturer's instructions. Following cell lysis, protein concentrations of cell lysates were quantified with a BCA kit (Pierce). A panel of phosphoproteins from the lysates were assayed using a multiplex antibody array¹⁴ and sandwich-type enzyme-linked immunoassays (ELISAs). All proteins in the panel were measured simultaneously from each sample.

3.2.8. Data processing

The radioimage was segmented into 20 rectangular regions of interest (ROIs), each of which overlaps a cell chamber and contains > 95% of the beta-particle counts from that

cell chamber. Beta-particle counts of each ROI were quantitated by a custom-written MATLAB program written. The background level was defined from the averaged beta-particle counts from 4 ROIs that covered the 4 cell capture chambers in a microchannel that was not loaded with any cells. For facilitation of cell counting, we stained cells with nuclear dye Hoechst 33342 to visualize each cell under a fluorescence microscope. A script written in ImageJ was used to automate cell counting. Radioactivity of each ROI was normalized by its cell number, and statistics for a given assay condition were calculated from the 4 repeat assays within a given microchannel. We limited comparisons of absolute beta-particle counting results across different experimental conditions to only those assays that were executed on the same RIMChip. To compare the results between RIMChips, we choose one of five microchannels in each RIMChip, and repeat the same experiment across all RIMChips. The radioactivity of other microchannels was normalized to that of the reference microchannel. This meant that we did not need to account for the changing radioactivity from the ^{18}F -labeled radiopharmaceutical imaging probe.

3.3. Results

3.3.1. Design of RIMChip and cell loading

Each RIMChip (**Fig. 3.2.**) permits 4 repeats of 5 independent assays. Each assay microchannel inlet contains a bubble depletion chamber and a debris-trapping filter (**Fig. 3.1., Fig. 3.2. A and B**). These permit the use of standard micropipetting for cell and reagent introduction. The 50 μm separation between the cell capture chamber floors and the camera yields an 11-fold increase in signal level relative to previous designs², which used a microscope cover slip as the chamber floor. Fibronectin coating of the

microchannel surfaces promoted attachment and spreading of the adherent cells (**Fig. 3.3.**), but did not influence the uniformity of cell loading or the [^{18}F]FDG uptake (**Fig. 3.3. and Fig. 3.4.**). Suspension cells were found to attach to the untreated hydrophobic PDMS surface. (**Fig. 3.5.**)

Cells were counted within each cell capture chamber (**Fig. 3.2.C**), before and after the radioassay, to permit per cell quantitation of the radioassay results. Cell loading and [^{18}F]FDG uptake exhibited ~8% variations across the different cell capture chambers associated with a single microchannel. For T cell assays (**Fig. 3.2.E**), chambers with ~ 70-110 cells exhibited a ~30-fold higher signal than control chambers with zero cells.

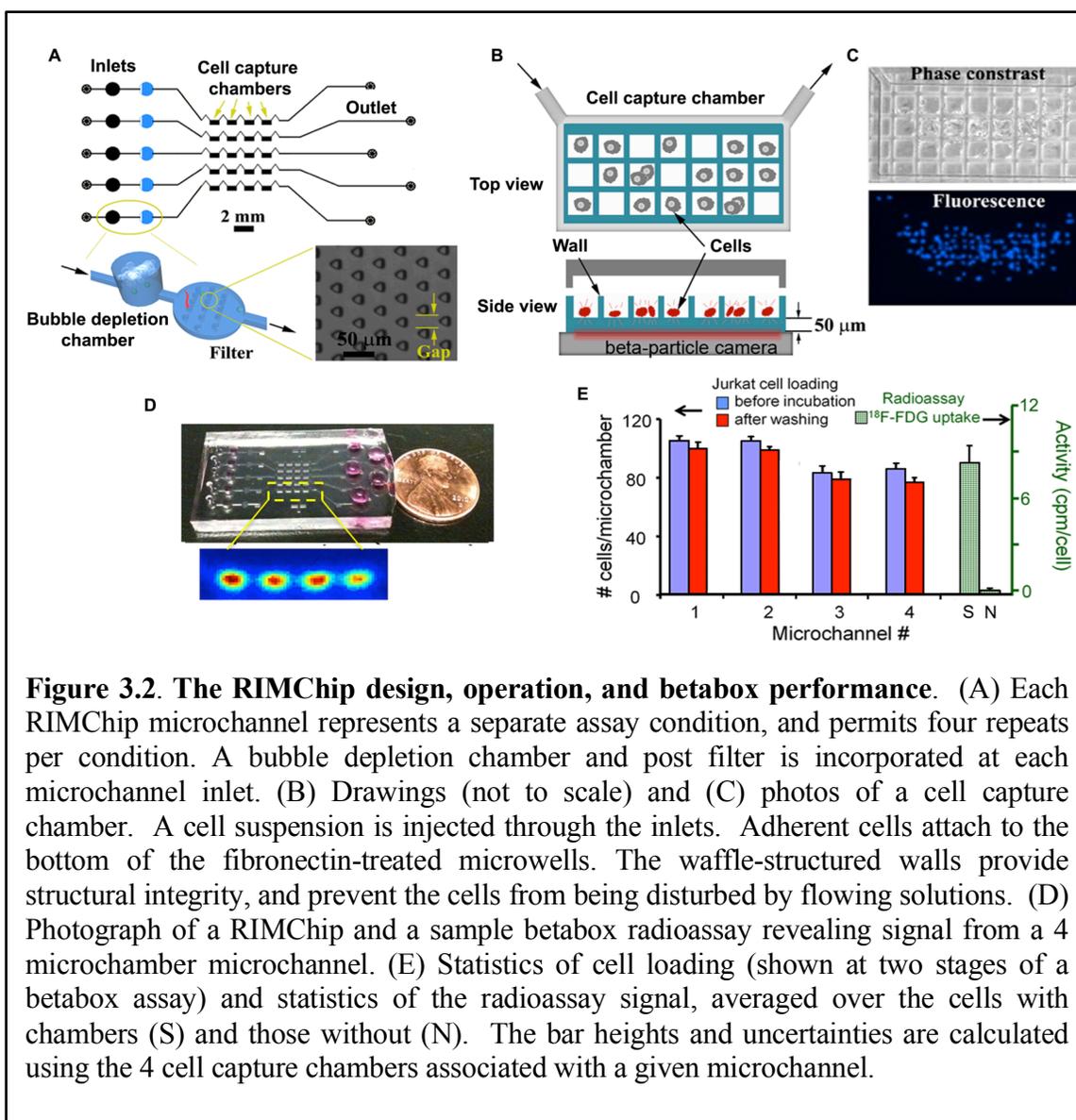
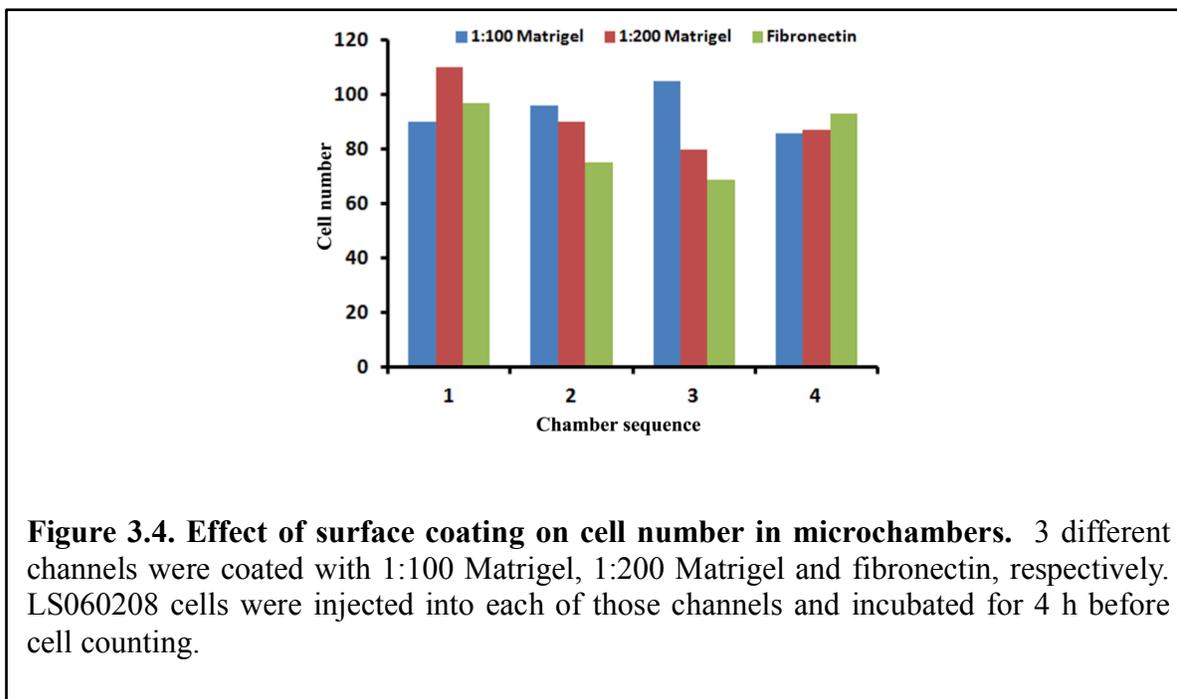
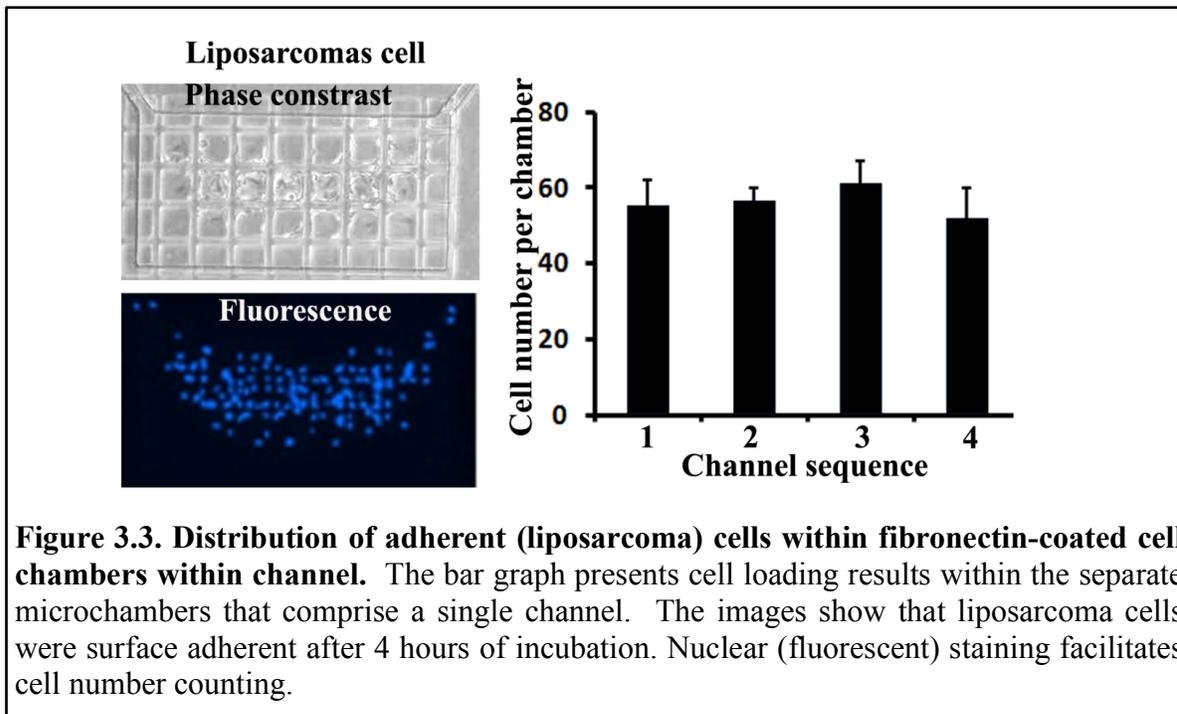


Figure 3.2. The RIMChip design, operation, and betabox performance. (A) Each RIMChip microchannel represents a separate assay condition, and permits four repeats per condition. A bubble depletion chamber and post filter is incorporated at each microchannel inlet. (B) Drawings (not to scale) and (C) photos of a cell capture chamber. A cell suspension is injected through the inlets. Adherent cells attach to the bottom of the fibronectin-treated microwells. The waffle-structured walls provide structural integrity, and prevent the cells from being disturbed by flowing solutions. (D) Photograph of a RIMChip and a sample betabox radioassay revealing signal from a 4 microchamber microchannel. (E) Statistics of cell loading (shown at two stages of a betabox assay) and statistics of the radioassay signal, averaged over the cells with chambers (S) and those without (N). The bar heights and uncertainties are calculated using the 4 cell capture chambers associated with a given microchannel.



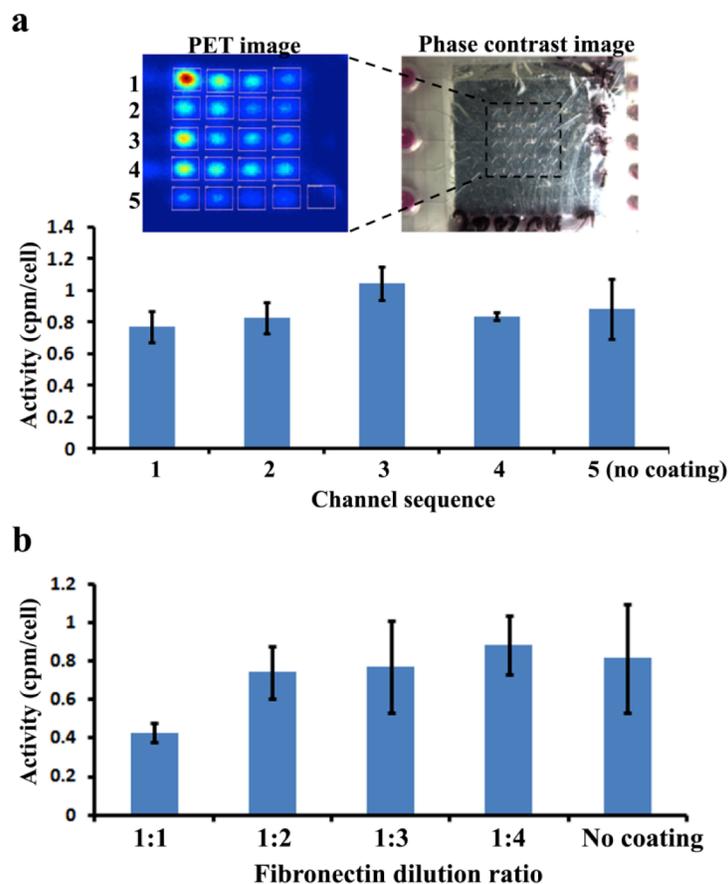
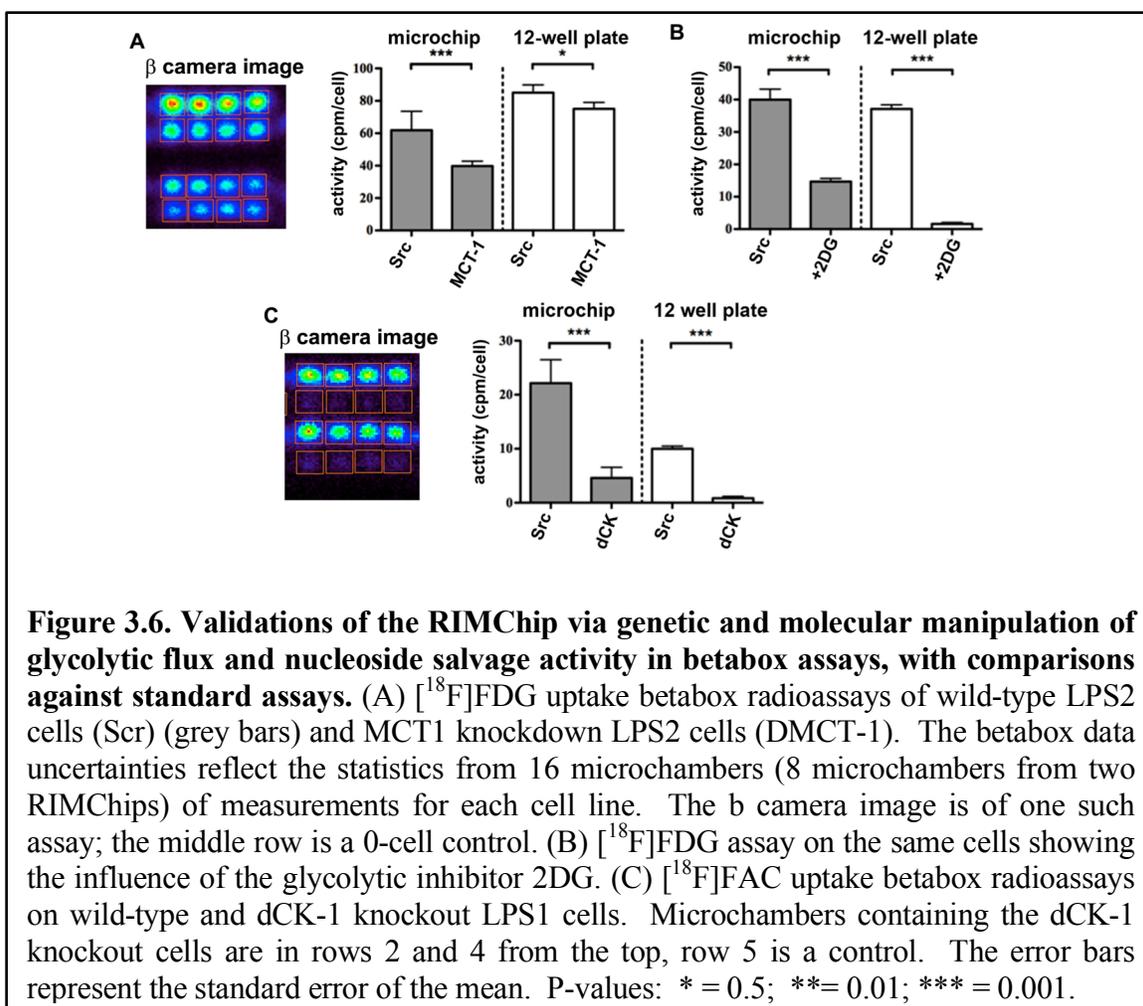


Figure 3.5. Effect of fibronectin coating on ^{18}F -FDG uptake by liposarcoma cells. (a) Channel 1, 2, 3, 4 have been coated with fibronectin for 4 days, 3 day, 2 days, and 1 day, respectively. Channel 5 was not coated with fibronectin. Then cells were loaded to the chambers of each channel and fed with ^{18}F -FDG following the same procedure as described elsewhere. No significant differences from channel 1 to channel 5 are observed. In channel 5, half of cells were lost after the completion of whole procedure. Cells were not spreading out on an uncoated surface. (b) Effect of fibronectin concentration for coating on uptake of ^{18}F -FDG. The initial concentration is 1 mg/ml which is diluted 2 times, 3 times and 4 times, respectively. Cells do not demonstrate significant difference with respect to ^{18}F -FDG uptake except in the cell chambers coated with the highest concentration of fibronectin.

3.3.2. Validation of RIMChip

We executed proof-of-principle betabox radioassays in which the glycolytic or nucleoside salvage pathways were genetically or molecularly manipulated in isogenic adherent liposarcoma cell lines (**Fig. 3.6.**). We altered monocarboxylate transporter 1 (MCT-1) levels via stable expression of shRNA, and examined the resulting changes in [^{18}F]FDG uptake (**Fig. 3.6.A**). In these cells, MCT-1 enhances glycolytic flux, so knockdown of MCT1 should result in reduction of [^{18}F]FDG uptake. The RIMChip assays detected a 35% relative decrease in [^{18}F]FDG uptake between Scr and DMCT-1 cells. The conventional assays revealed a 12.5% reduction (**Fig. 3.6.A**). Introduction of the glycolytic inhibitor 2-deoxy-glucose (2DG) resulted in a stronger (3-10 fold) repression of [^{18}F]FDG uptake, as recorded in the RIMChip assay and bulk assays, respectively (**Fig. 3.6.B**).

Certain liposarcomas exhibit nucleoside salvage activity, which can be imaged using the nucleoside analog 1-(2'-deoxy-2' [^{18}F]fluoroarabinofuranosyl) cytosine ([^{18}F]FAC)⁸. [^{18}F]FAC is a substrate for enzyme dCK, and so dCK knockdown should exhibit decreased [^{18}F]FAC uptake. The third betabox validation study supported this hypothesis (**Fig. 3.6.C**).



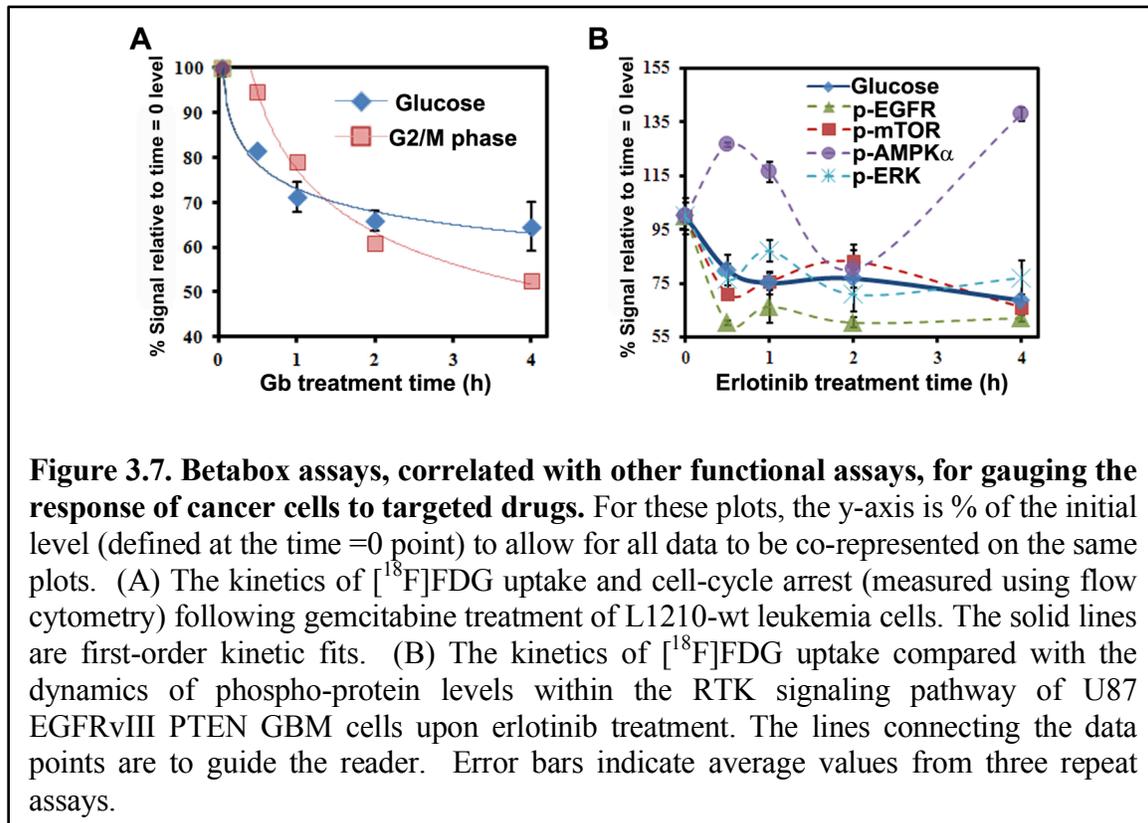
3.3.3. Kinetic study of drug response with RIMChip

The betabox was also used for interrogating the kinetics of cellular responses to targeted therapies. All time points in a given kinetics run are imaged simultaneously, for easy comparisons. The influence of gemcitabine on L1210-wt leukemia cells was first studied. Gemcitabine is an anti-cancer pro-drug nucleoside analogue that will impede the cell-cycle in dCK positive tumors⁹. Since glucose metabolism can help fuel the cell-cycle, we reasoned that gemcitabine treatment could influence both $[^{18}\text{F}]\text{FDG}$ uptake and the cell-cycle. Indeed $[^{18}\text{F}]\text{FDG}$ uptake decreases upon cell exposure to gemcitabine, with first-

order kinetics and a half-life ($\tau_{1/2}$) of ~30 minutes cell-cycle arrest exhibits similar kinetics ($\tau_{1/2}$ ~50 minutes), and lags about 25 minutes behind changes in [^{18}F]FDG uptake. (**Fig. 3.7.A**).

A second class of targeted drugs is aimed at blocking growth factor signaling of receptor tyrosine kinases (RTKs). For example, the ATP-competitive EGFR tyrosine kinase inhibitor erlotinib blocks wild-type and mutant EGFR (EGFRvIII) signaling, inhibiting the phosphorylation of kinases that represent downstream effectors of EGFR^{15,16} and thus inhibiting the growth of PTEN-expressing glioblastomas¹². An activated growth factor signaling pathway implies energy flux through that pathway, so we reasoned that inhibiting that pathway would likely reduce cellular glucose consumption. Thus, we treated EGFRvIII and PTEN-expressing model glioblastoma multiforme (GBM) cell lines with erlotinib, and measured [^{18}F]FDG uptake kinetics, plus the levels of a panel of phosphoproteins that are downstream effectors of EGFR (**Fig. 3.7.B** and **Fig. 3.8**). The full panel of assayed proteins is provided as **Table 3.1**. Again, [^{18}F]FDG uptake drops sharply within 30 min, but the kinetics yield behavior reminiscent of a damped oscillator. Interestingly, this oscillatory behavior is reflected in the changing levels of the assayed phosphoproteins. The protein phosphorylated-AMP activated protein kinase (p-Ampk α) functions as an energy regulator within the cell¹⁷. Its level appears to initially oscillate out of phase with [^{18}F]FDG uptake, implying a compensatory mechanism for loss of glucose consumption. The levels of p-EGFR, p-Erk and p-mTOR oscillate mostly in phase with each other, and are partially synchronized with changes in [^{18}F]FDG uptake. The amplitude of the changing levels of a given phosphoprotein may reflect its proximity to EGFR in the RTK signaling pathway. For example, p-EGFR itself exhibits the largest

amplitude response, with p-mTOR exhibiting the weakest response. Over the course of the 4-hour drug treatment window, the cell-cycle was relatively unaffected.



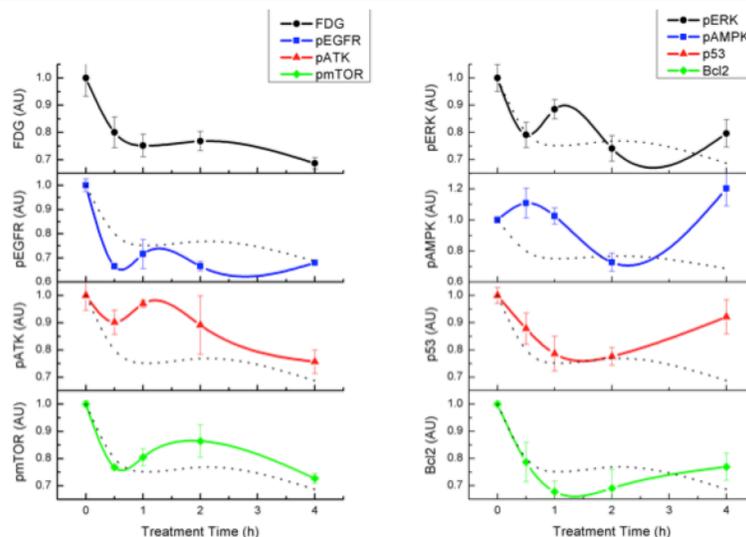


Figure 3.8. The kinetics of glucose consumption rate and protein level upon Erlotinib treatment. Each protein is measured and converted to the percentage of the measured level at time 0. Glucose consumption rate has been calculated from ^{18}F -FDG uptake rate. Error bars indicate fluctuations of three repeats.

Table 3.1. List of antibodies used for GBM cell proteomic assay

DNA label	Antibody (ventor: clone)	Source
D	mouse anti-hu phospho-EGFR	R&D Y1068
	biotin-labeled goat anti-hu EGFR	R&D BAF231
E	anti-hu phospho-Ampka kit	R&D DYC 3528
F	anti-hu p53 kit	R&D DYC 1746
G	anti-hu Bcl2 kit	R&D DYC 827B
I	anti-hu phospho-mTOR kit	R&D DYC 1665
K	anti-hu phoshpo-ERK kit	R&D DYC 1018B
L	anti-hu phospho-Akt1 kit	R&D DYC 2289

3.4. Discussions

The comparisons of **Fig. 3.6.** between the betabox assays and the 12-well plate (bulk) assays reveal qualitative, but not fully quantitative agreement between the two approaches. A concern might be that since the RIMChip assays a relatively small number of cells, the corresponding statistical spread of results would be significantly larger than for the bulk assays. However, this spread, which should scale as the square root of the numbers of cells, is not large. For the comparisons of **Fig 3.6.**, 4 microchannels \times 4 microchambers per channel \times 100 cells per microchamber yields a 2.5% error, relative to an expected $\sim 0.5\%$ error for a ($\sim 50,000$ cells) bulk assay. Cell counting errors for the RIMChip assays may add a few additional percent, as will the experimental errors in manipulating the cells for the bulk assay. Small variations in the 50 μm thick PDMS membrane separating the cells from the camera can also contribute a few percent error. However, given that the two techniques have independent sources of error, and that the RIMChip assays for ~ 50 -fold fewer cell numbers, the agreement between the two techniques is good. There is a flow cytometry-based assay that utilizes a fluorophore-labeled [^{18}F]FDG analogue, but recent literature¹⁸ has called into question the validity of that assay, and so we did not compare against it here.

The kinetic responses recorded in **Fig. 3.7.** reveal that the [^{18}F]FDG RIMChip assay provides a rapid (< 1 hour) tool for detecting the response of small cell numbers to a therapy. However, the assays also reveal different response kinetics for the different cell lines and drugs. An exact resolution of such responses can be accomplished by either capturing the dynamic trajectories of individual cells¹⁹, or the fluctuations of the functional proteins at the single-cell level²⁰. Obviously, the RIMChip assay does not resolve such

trajectories of fluctuations, but the cited literature can provide some insight into the observed responses. For example, the first-order relaxation kinetics observed for the response of the L1210 leukemia cells to gemcitabine exposure is consistent with a transition between two distinct, steady state descriptions of those cells. An alternative explanation, which we can rule out by our observations, is that gemcitabine exposure killed a fraction of the cells. No dead cells were detected during the course of the experiment. The oscillatory dynamics observed for the erlotinib treated GBM cells imply that there are competing networks that influence the kinetics. The overall result may be the same-i.e., the cells are switching between two states, but that is not as clear from our data. Again, however, no dead cells were detected during the course of the experiments.

3.5. Conclusions

The betabox platform, comprised of a microfluidic chip (the RIMChip) mated to a beta-particle imaging camera, enables robust, user-friendly execution of sensitive and quantitative cell-based radioassays. Each radioassay requires ~100 cells. Betabox radioassays provide a useful and rapid screening platform for investigating the response of various cell lines to mechanistically distinct, targeted drugs. The betabox platform provides a rapid screening tool for a variety of drug/cell line combinations, as well as a powerful tool for mechanistic investigations.

3.6. References

1. Czernin, J. & Phelps, M. E. Positron emission tomography scanning: current and future applications. *Annu. Rev. Med.* **53**, 89–112 (2002).
2. Vu, N. T. *et al.* A Camera Integrated with a Microfluidic Chip for Radioassays Based on

- Real-Time Imaging of Glycolysis in Small Cell Populations. *J Nucl Med* **52**, 815–821 (2011).
3. Dooraghi, A. A. *et al.* Betabox: a beta-particle imaging system based on a position sensitive avalanche photodiode. *Phys Med Biol* **58**, 3739–3753 (2013).
 4. Fang, C. *et al.* Integrated Microfluidic and Imaging Platform for a Kinase Activity Radioassay to Analyze Minute Patient Cancer Samples. *Cancer Res* **70**, 8299–8308 (2010).
 5. Usova, E., Maltseva, T., Földesi, A., Chattopadhyaya, J. & Eriksson, S. Human deoxycytidine kinase as a deoxyribonucleoside phosphorylase. *Journal of Molecular Biology* **344**, 1347–1358 (2004).
 6. Radu, C. G. *et al.* Molecular imaging of lymphoid organs and immune activation by positron emission tomography with a new [18F]-labeled 2'-deoxycytidine analog. *Nature Medicine* **14**, 783–788 (2008).
 7. Shipley, L. A. *et al.* Metabolism and disposition of gemcitabine, and oncolytic deoxycytidine analog, in mice, rats, and dogs. *Drug Metab. Dispos.* **20**, 849–855 (1992).
 8. Braas, D. *et al.* Metabolomics strategy reveals subpopulation of liposarcomas sensitive to gemcitabine treatment. *Cancer Discov* **2**, 1109–1117 (2012).
 9. Laing, R. E. *et al.* Noninvasive prediction of tumor responses to gemcitabine using positron emission tomography. *P Natl Acad Sci USA* **106**, 2847–2852 (2009).
 10. Jordheim, L. P. *et al.* Characterization of a gemcitabine-resistant murine leukemic cell line: reversion of in vitro resistance by a mononucleotide prodrug. *Clin Cancer Res* **10**, 5614–5621 (2004).
 11. Owens, J. K., Shewach, D. S., Ullman, B. & Mitchell, B. S. Resistance to 1-beta-D-arabinofuranosylcytosine in human T-lymphoblasts mediated by mutations within the deoxycytidine kinase gene. *Cancer Res* **52**, 2389–2393 (1992).
 12. Mellinghoff, I. K. *et al.* Molecular determinants of the response of glioblastomas to EGFR kinase inhibitors. *N Engl J Med* **353**, 2012–2024 (2005).
 13. Fenton, T. R. *et al.* Resistance to EGF receptor inhibitors in glioblastoma mediated by phosphorylation of the PTEN tumor suppressor at tyrosine 240. *P Natl Acad Sci USA* **109**, 14164–14169 (2012).
 14. 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**, 1959–1967 (2007).
 15. Huang, P. H. *et al.* Quantitative analysis of EGFRvIII cellular signaling networks reveals a combinatorial therapeutic strategy for glioblastoma. *P Natl Acad Sci USA* **104**, 12867–

- 12872 (2007).
16. Shi, Q. *et al.* Single-cell proteomic chip for profiling intracellular signaling pathways in single tumor cells. *P Natl Acad Sci USA* **109**, 419–424 (2012).
 17. Inoki, K., Kim, J. & Guan, K.-L. AMPK and mTOR in Cellular Energy Homeostasis and Drug Targets. *Annu. Rev. Pharmacol. Toxicol.* **52**, 381–400 (2012).
 18. Tseng, J.-C., Wang, Y., Banerjee, P. & Kung, A. L. Incongruity of imaging using fluorescent 2-DG conjugates compared to 18F-FDG in preclinical cancer models. *Mol Imaging Biol* **14**, 553–560 (2012).
 19. Gascoigne, K. E. & Taylor, S. S. Cancer cells display profound intra- and interline variation following prolonged exposure to antimitotic drugs. *Cancer Cell* **14**, 111–122 (2008).
 20. Wei, W. *et al.* Hypoxia induces a phase transition within a kinase signaling network in cancer cells. *P Natl Acad Sci USA* **110**, E1352–60 (2013).

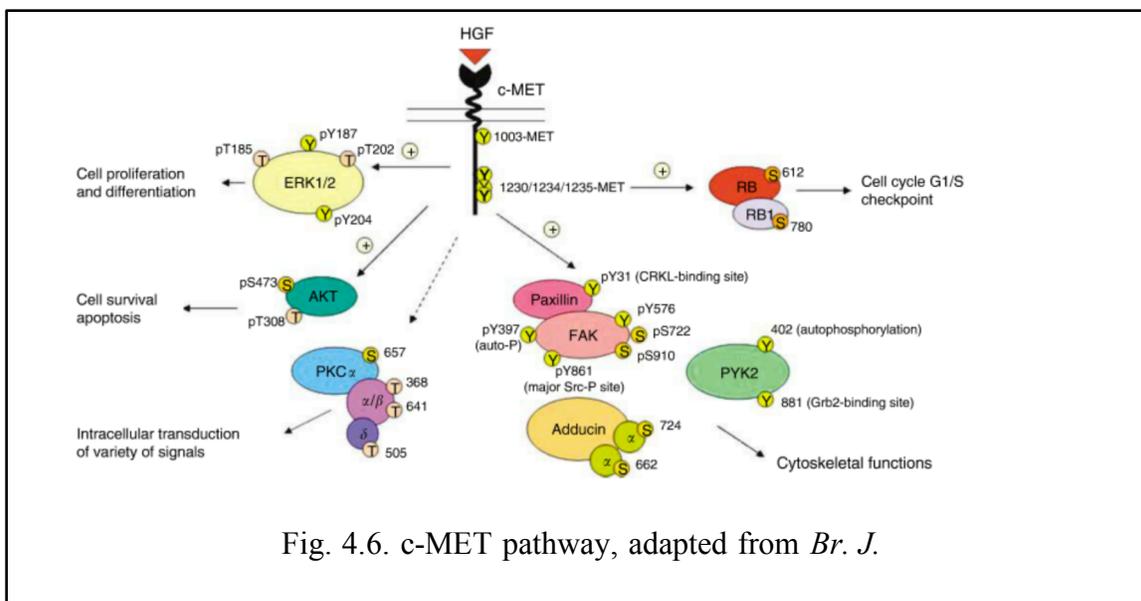
Chapter 4

Development and characterization of the capture agents library targeting c-MET for in vivo cancer imaging

4.1. Introduction

As I mentioned in the previous chapter, positron emission tomography (PET) enables imaging of biological process in noninvasive and quantitative way¹, it is requested to develop new radio probes visualizing specific biomarkers, especially cancer markers to diagnose and stage the disease and guide therapeutic decisions². For example, [4-¹⁸F]fluorobenzaldehyde-conjugated aminoxy-protein scaffolds are used to detect tumoral Her2 expression³, and [¹⁸F]fluoroestradiol is used to stratify breast cancer patients by measuring estrogen receptor expression^{4,5}. However, radio probes able to image-specific surface markers for cancer cells are still lacking.

Hepatocyte growth factor receptor (HGFR/c-MET) is a receptor tyrosine kinase whose expression has been shown to be associated with tumor invasion and metastasis. Since its discovery in mid-1980s⁶, many reports elucidate that c-MET can be a cancer marker as well as a target of cancer therapy due to its high expression level in various tumors, such as lung, kidney, colon, gastric, thyroid, pancreas and prostate cancers⁷⁻⁹. As shown in Fig.4.1.¹⁰, numerous oncogenic pathways are engaged downstream of c-MET, which may lead to multiple strategies for therapeutic inhibition. As recent clinical studies reported c-MET antibody improved the overall survival when it was treated to well stratified cancer patients, targeting c-MET shows promising prospect¹¹.

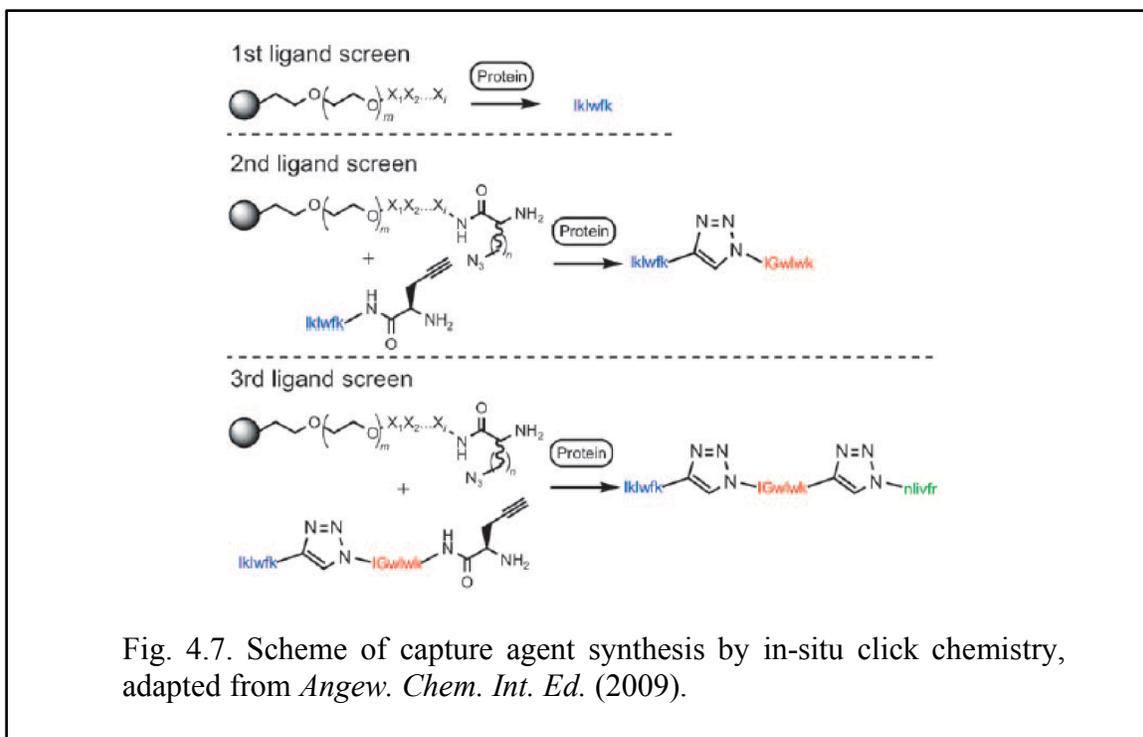


In this chapter summarizing the current work, I present a novel library of capture agents synthesized from artificial amino acids to recognize extracellular domain of c-MET for avoiding disadvantages of antibodies, such as high cost, batch to batch variation and instability for long-term storage^{12,13}. Researchers reported that these protein-catalyzed capture agents (PCC) could be synthesized for recognizing their specific target biomolecules, bCAII and Akt1, with high affinity compared to the commercialized antibody counterpart, and with solid stability for long-term storage at room temperature. Since it is also possible to synthesize the capture agents in bulk size by using standard solid-phase method, their production costs also have room to lower upon the demand^{12,13}.

4.2. Experimental Methods

4.2.1. Synthesis of PCC for c-MET

Multiligand protein-catalyzed capture agents (PCC) library can be generated by using in situ click chemistry, as shown in **Fig. 4.2**. Detailed synthetic method is explained in the previous paper^{12,13}. In short, capture agents were generated by iteratively screening one-bead-one-compound (OBOC) peptide library on TentaGel beads against recombinant human c-MET extracellular domain (R&D Systems). The target protein was exploited to promote click reactions between individual azide- and alkyne-functionalized library elements, first to produce biligands and second to produce triligand candidates. At each screening stage, beads were selected both for binding to the target and evolution of a triazole-linked product. To remove beads that are prone to non-selective binding, an anti-screen was performed against human serum and detected via a labeled anti-human serum antibody. Only the peptides passing through both of positive and negative selections were employed as triligand c-MET capture agents.



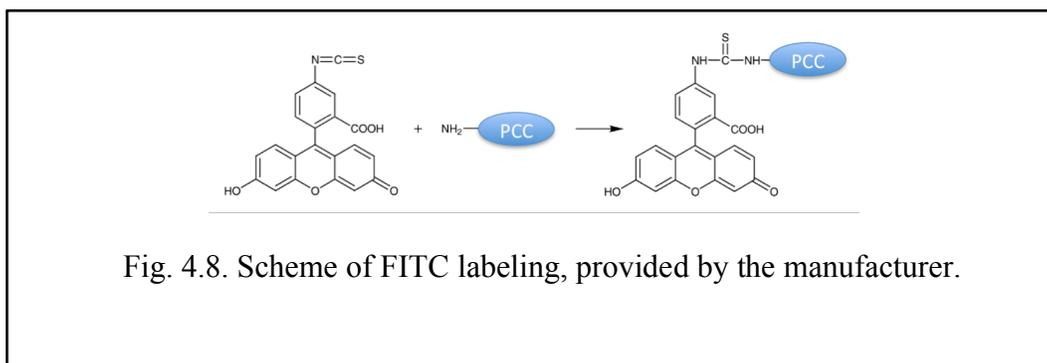
Once several candidates were found, these triligand c-MET capture agents were prepared in bulk quantities by using solid-phase synthesis of the individual ligands followed by copper(I)-catalyzed ligation of azide and terminal alkyne¹⁴. Final products were purified by HPLC and characterized by mass spectrometry.

4.2.2. Functionalization of capture agents.

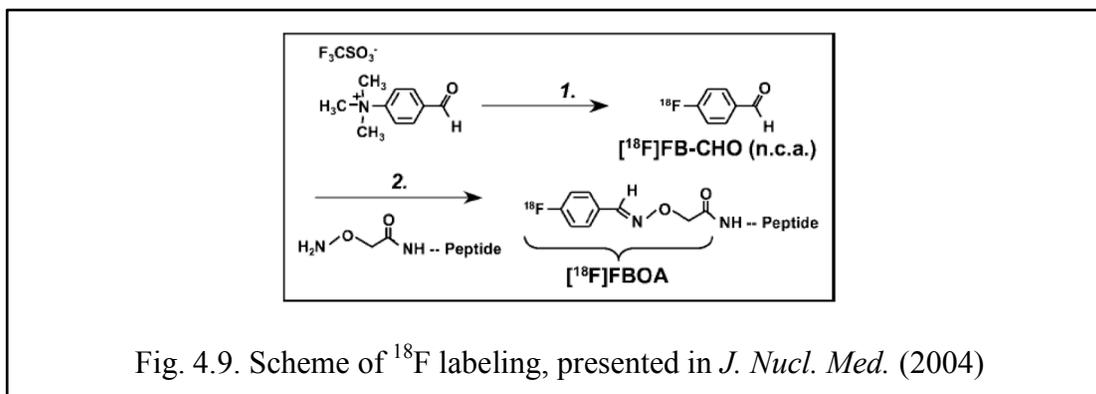
To work as in vivo imaging probes, the capture agents synthesized as above were labeled with a radioactive atom. However, it is not practical to characterize the radio-labeled capture agents at the beginning, because their short life time limits available time for experiments and their radioactivity demands extra safety and special equipment for handling and visualizing. Hence, the characterization of capture agents was achieved with dye-labeled materials, and the fluorescent signals from the capture agents were quantified

first. Only candidates showing superior performance to the commercialized c-MET antibody was selected for the next step, the labeling with a radioactive atom.

As a dye molecule, fluorescein isothiocyanate (FITC) was used for its ease of functionalization. The N-terminus of the capture agent was labeled with FITC (AnaSpec) as a final step in its solid-phase synthesis. The reaction proceeded as follows **Fig. 4.3**.



For the radio-labeling of capture agents, 4-¹⁸F]fluorobenzaldehyde was synthesized first, ([¹⁸F]FB-CHO), then it was conjugated with the N-terminus of the capture agent as shown in **Fig. 4.4**, as described in the literature¹⁵. After the synthesis, the product was rapidly purified by HPLC and its radioactivity was measured with a gamma counter.



4.2.3. Preparation of cells

In order to verify the binding capacity of capture agents, human prostatic carcinoma cell lines, PC3, DU145, LNCaP and 22Rv1 cells, were purchased from ATCC as described in the literature¹⁶. All cell lines were cultured in RPMI medium 1640, supplemented with 10% FBS and 1% pen/strep solution, in a 37 °C incubator.

Also, HGF, the ligand of c-MET, was added to cells in c-MET inhibition experiment to check that synthetic capture agents actually bind to c-MET to form a competitive relationship with the original ligand, HGF. In this case, just before the harvest of cells, cells were treated with recombinant HGF (Sigma) at the concentration of 25ng/ml for 15 min at the cell incubator as suggested in a previous work¹⁷.

4.2.4. Confocal experiment for 3D imaging of live cells incubated with capture agents

Synthetic capture agents must bind to the extracellular domain of c-MET for working as imaging probes. In order to confirm the binding site, c-MET expressing cells were incubated with FITC-conjugated capture agents to be analyzed with a confocal microscope. As a control, cells were also incubated with a commercial FITC labeled antibody for the visualization, and the 3D images from both conditions were compared with ImageJ (NIH).

The detailed procedure will be as follows: 100,000 cells were seeded into each well of microwell slides with 2ml of cell culture media 24 hour before the imaging experiment. On the next day, cell culture media was aspirated, and cells were stained with 100µl of cell culture media containing different concentration of dye-conjugated capture agent or

antibody for 1 hour at cell incubator to find an optimal condition. After washing with cell culture media twice, cells were on top of LSM510 microscope (Zeiss) for imaging.

4.2.5. Flow cytometry for the quantification of binding affinity

Flow cytometry is a powerful tool to compare the interactions between ligands and receptors on the surface of cells as well as to find an optimal concentration of materials for staining¹⁸. To quantify the binding capacity of capture agents to c-MET on the surface of cells, the fluorescence intensity of cells incubated with FITC-labeled capture agents were analyzed with a flow cytometer by following procedures:

Cells were harvested and aliquoted into 1.7ml falcon tubes at the volume of 100 μ l having 10million/ml concentration in 1% BSA-PBS. In order to make calibration curves, each aliquot was treated with serial dilutions of FITC-conjugated capture agents and antibody, as a control, for 1hr on ice. After the treatment, by using a centrifuge, cells were washed and re-suspended into 0.5ml of 1% BSA-PBS. Data was obtained and analyzed with FACSCalibur (BD bioscience) and Flow Jo (Tree Star), respectively.

4.2.6. In vitro radioassay with RIMChip

In order to do an in vitro radioassay, RIMChip was fabricated with PDMS material as explained at the previous Chapter 3. Cells were prepared at 3×10^6 cells per ml and injected into the RIMChip. Cell numbers were counted on an optical microscope. For the HGF inhibition studies, 25ng/ml of HGF in RPMI 1640, supplemented with 10% FBS, was added to the cells for 15 min. 100 μ Ci of radio-labeled capture agents were pipetted into the RIMChip microchannels, and the RIMChip were then incubated for 1 hour at cell culture

incubator and flushed with PBS twice. Finally, a betabox having beta camera was assembled for the measurement of radioactivity.

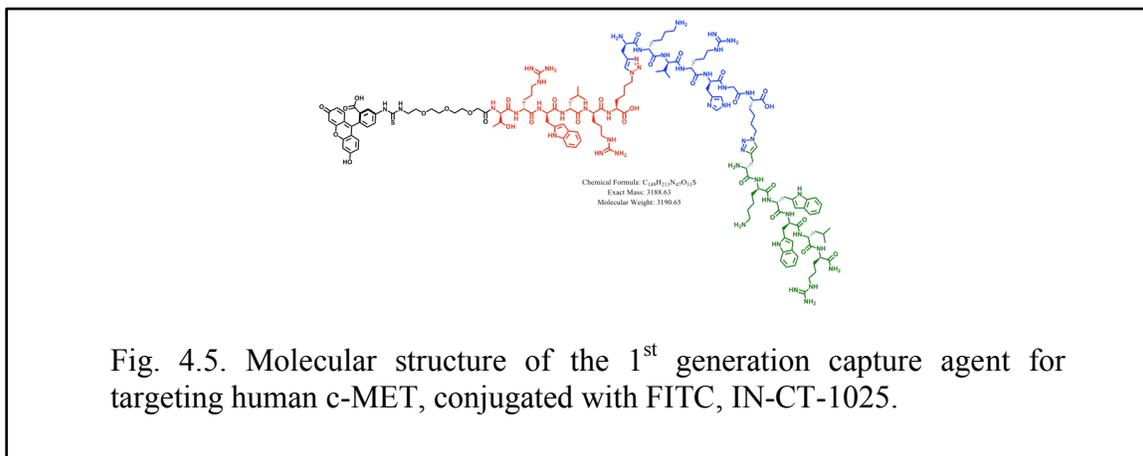
4.2.7. In vivo treatment and imaging with microPET/CT

To demonstrate the PCC work as imaging probes, in vivo experiment with mice is required. All animal studies will be done according to the guideline of the office of laboratory animal resources at Caltech. SCID mice will be injected with 1 million PC3 cells, and tumors will be allowed to grow for one week. Mice will be killed after imaging on day 14 for biodistribution studies. For imaging, mice will be injected with 100 μCi of radio-labeled capture agents. Data will be obtained in a dynamic scan for 1.5 hours after the probe administration by using Siemens Preclinical Solutions microPET focus 220 and a MicroCAT II CT instrument as described previously².

4.3. Results and Discussion

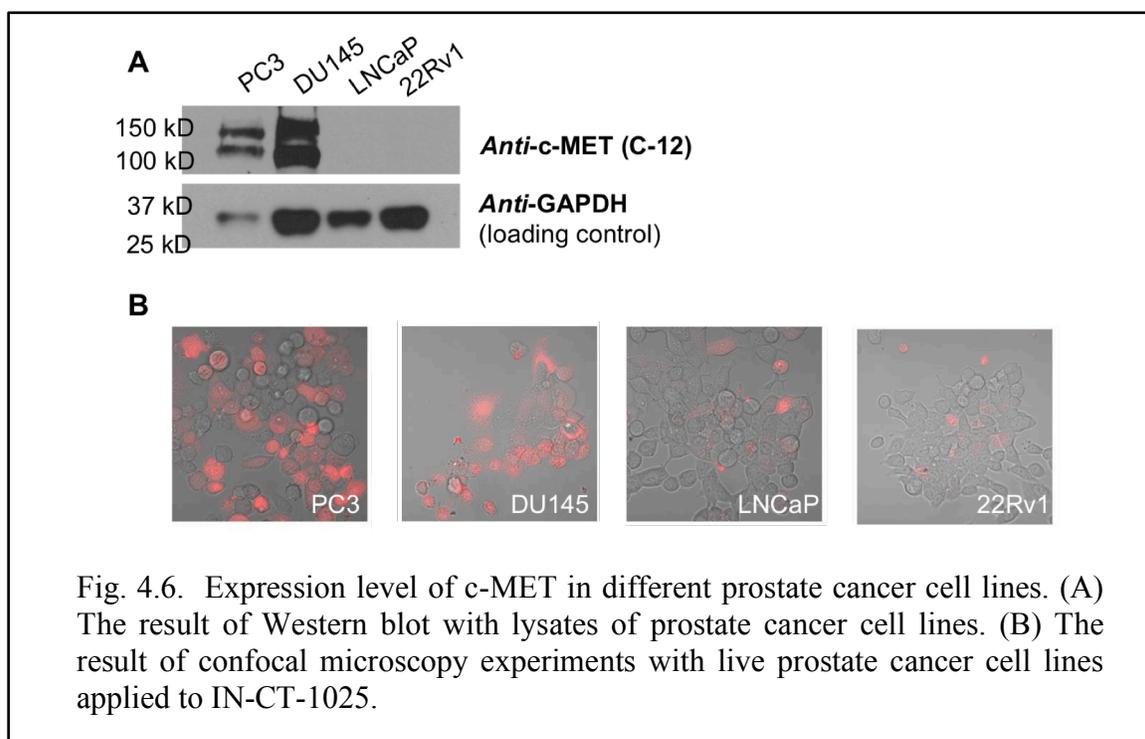
4.3.1. Synthesis of capture agents for c-MET

IN-CT-1025, the triligand capture agent against c-MET, was discovered and synthesized by Indi Molecular in Culver City, CA (data unpublished, **Fig. 4.5**).

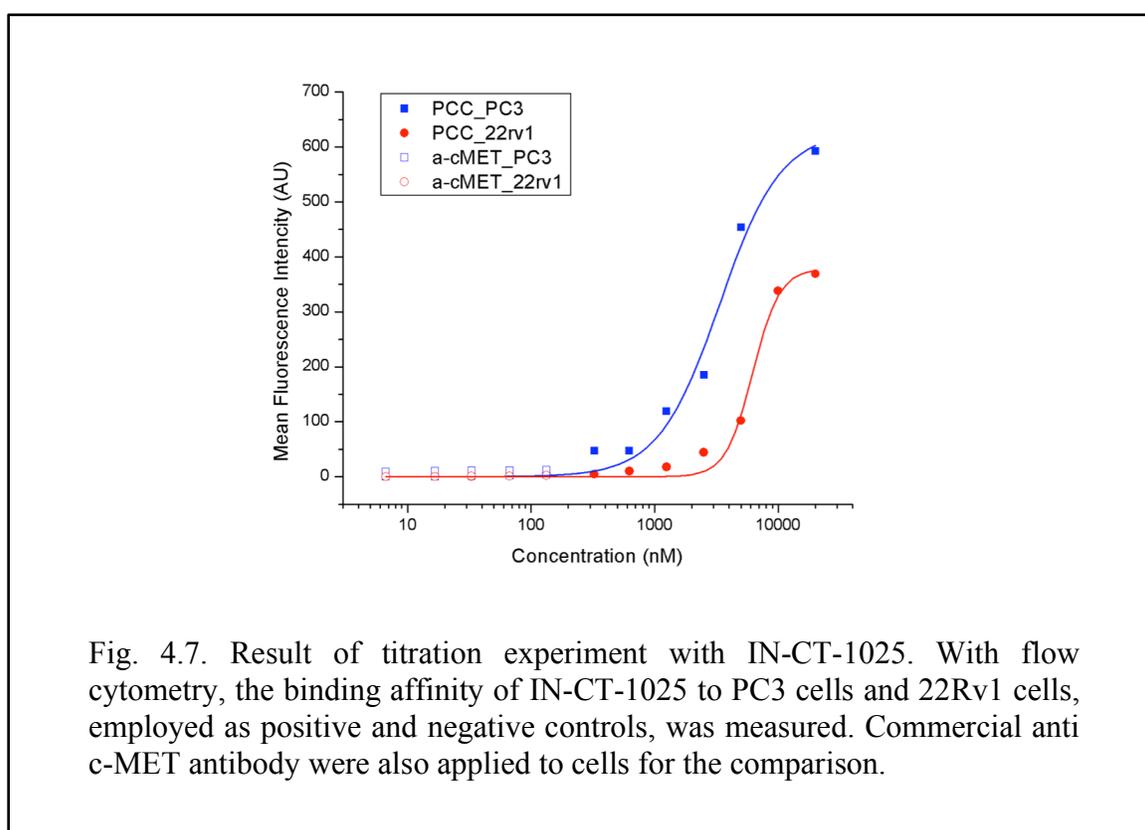


4.3.2. Cell-based fluorescence imaging assays

In order to find the best positive and negative controls for the cell-based assays, the expression level of c-MET in four prostate cancer cell lines, PC3, DU145, LNCaP and 22Rv1, was first examined by Western blot with lysates of those cell lines as shown in **Fig. 4.6.A**, which is accordance with the previous literature¹⁷. To check the affinity of IN-CT-1025 to c-MET on the surface of prostate cancer cells, with a confocal microscope, we did a proof-of-concept experiment by measuring fluorescent signals from four kinds of living prostate cancer cell lines applied with IN-CT-1025. As presented in **Fig. 4.6.B**, the fluorescent signals were comparable with the result from Western blot, which indicates that IN-CT-1025 is able to recognize the target protein, c-MET, at the extracellular domain of cell membrane.



After visualizing the binding of IN-CT-1025 to cells, its affinity to prostate cancer cells was quantitatively measured with a flow cytometer. By titrating the concentration of IN-CT-1025 with 1 million PC3 and 22Rv1 cells as positive and negative controls, the mean fluorescent signals from each sample were obtained to make a plot as shown in **Fig. 4.7**. A commercial c-MET antibody (eBioscience) was also applied to cells for the comparison, but we could not explore the region of high concentration of antibody (more than 375nM) due to its high cost.



The data was fitted with Hill function which gives the dissociation constant, K_d , of IN-CT-1025 with PC3 cells to be $3.36\mu\text{M}$ while K_d of IN-CT-1025 with 22rv1 cells is $6.23\mu\text{M}$. Interestingly, compared to the dynamic range of the commercial antibody, 30 (11.84 to 0.39), IN-CT-1025 showed much wider dynamic range of about 2,800 (592.7 to

0.21), which implies that IN-CT-1025 can work as a better imaging probe to distinguish c-MET positive and negative cells than antibodies.

4.4. Conclusions and Future Directions

As written above, c-MET is highly overexpressed in cancer patients. Having a specific imaging probe to this cancer marker, diagnosis and/or staging of cancer will be more accurate and lead to more rapid therapeutic decision making. With the first-generation protein-catalyzed capture agent, IN-CT-1025, we checked the possibility of PCC as an in vivo imaging probe by measuring its binding affinity to prostate cancer cells through cell-based fluorescent imaging assays. Compared to a commercialized c-MET antibody, IN-CT-1025 is shown to be a better agent to discriminate c-MET positive and negative cells due to its wide dynamic range. With this PCC, we will examine its capacity to work as an in vivo imaging probe by using an in vitro radio assay with RIMChip and Betabox as described in the previous Chapter 3, which will be then followed by the in vivo experiments with mice imaging on micro PET/CT.

4.5. References

1. Phelps, M. E. Positron emission tomography provides molecular imaging of biological processes. *P Natl Acad Sci USA* **97**, 9226–9233 (2000).
2. Laing, R. E. *et al.* Noninvasive prediction of tumor responses to gemcitabine using positron emission tomography. *P Natl Acad Sci USA* **106**, 2847–2852 (2009).
3. Cheng, P. *et al.* Inhibition of dendritic cell differentiation and accumulation of myeloid-derived suppressor cells in cancer is regulated by S100A9 protein. *Journal of Experimental Medicine* **205**, 2235–2249 (2008).
4. Linden, H. M. *et al.* Quantitative fluoroestradiol positron emission tomography

- imaging predicts response to endocrine treatment in breast cancer. *J. Clin. Oncol.* **24**, 2793–2799 (2006).
5. Peterson, L. M. *et al.* Quantitative imaging of estrogen receptor expression in breast cancer with PET and 18F-fluoroestradiol. *J Nucl Med* **49**, 367–374 (2008).
 6. Cooper, C. S. *et al.* Molecular cloning of a new transforming gene from a chemically transformed human cell line. *Nature* **311**, 29–33 (1984).
 7. Ma, P. C. *et al.* Expression and mutational analysis of MET in human solid cancers. *Genes Chromosomes Cancer* **47**, 1025–1037 (2008).
 8. Liu, N., Furukawa, T., Kobari, M. & Tsao, M. S. Comparative phenotypic studies of duct epithelial cell lines derived from normal human pancreas and pancreatic carcinoma. *Am J Pathol* **153**, 263–269 (1998).
 9. Jacobsen, F. *et al.* High c-MET expression is frequent but not associated with early PSA recurrence in prostate cancer. *Exp Ther Med* **5**, 102–106 (2013).
 10. Ma, P. C., Tretiakova, M. S. & Nallasura, V. Downstream signalling and specific inhibition of c-MET/HGF pathway in small cell lung cancer: implications for tumour invasion. *British journal of Cancer* (2007).
 11. Gherardi, E., Birchmeier, W., Birchmeier, C. & Vande Woude, G. Targeting MET in cancer: rationale and progress. *Nat Rev Cancer* **12**, 89–103 (2012).
 12. Agnew, H. D. *et al.* Iterative in situ click chemistry creates antibody-like protein-capture agents. *Angew Chem Int Ed Engl* **48**, 4944–4948 (2009).
 13. Millward, S. W. *et al.* Iterative in Situ Click Chemistry Assembles a Branched Capture Agent and Allosteric Inhibitor for Akt1. *J. Am. Chem. Soc.* **133**, 18280–18288 (2011).
 14. Rostovtsev, V. V., Green, L. G., Fokin, V. V. & Sharpless, K. B. A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective ‘ligation’ of azides and terminal alkynes. *Angew Chem Int Ed Engl* **41**, 2596–2599 (2002).
 15. Poethko, T., Schottelius, M. & Thumshirn, G. Two-step methodology for high-yield routine radiohalogenation of peptides: 18F-labeled RGD and octreotide analogs. *J Nucl Med* (2004).
 16. Aggarwal, S., Singh, P., Topaloglu, O., Isaacs, J. T. & Denmeade, S. R. A dimeric peptide that binds selectively to prostate-specific membrane antigen and inhibits its

- enzymatic activity. *Cancer Res* **66**, 9171–9177 (2006).
17. Van Leenders, G. J. L. H. *et al.* Activation of c-MET induces a stem-like phenotype in human prostate cancer. *PLoS ONE* **6**, e26753 (2011).
 18. Sklar, L. A. & Finney, D. A. Analysis of ligand-receptor interactions with the fluorescence activated cell sorter. *Cytometry* **3**, 161–165 (1982).

Chapter 5

Imaging of glycolytic metabolism in primary glioblastoma cells with RIMChip

5.1. Introduction

Glioblastoma(GBM) is one of the most common brain tumors¹. It is composed of heterogeneous subpopulations of phenotypically distinctive tumor that express different level of oncogenes such as epidermal growth factor receptor variant III (EGFRvIII)². GBM is also remarkably resistant to targeted therapies, including EGFR inhibitors despite frequent expression of EGFRvIII. Previous works have presented a number of important mechanisms underlying resistance^{3,4}, but the contribution of cellular heterogeneity to resistance is not well understood.

In spite of cellular heterogeneity, some solid tumors exhibit hallmarks of homogeneous systems. For example, Lander's group has reported that luminal, basal, and stem-cell phenotypes purified from a breast tumor can, over time, regenerate the characteristic cellular heterogeneity (the phenotypic equilibrium) of the original tumor⁵. This behavior, in turn, has consequences for therapeutic resistance⁶.

In this chapter summarizing the current work, I will introduce the preliminary results showing the heterogeneous glucose metabolism in cancer cells and the possibility of RIMChip as a screening tool for drug treatments. We begin with a patient-derived, primary glioblastoma cell line (GBM39)⁷, which has heterogeneous EGFRvIII expression. By examining the glycolytic activity of cells according to their expression level of EGFRvIII,

we expect to see the relationship between glucose metabolism and EGFR signaling pathway.

Also, we will explore the possibility of RIMChip platform as a drug screening tool by applying EGFR+ subpopulations of GBM39 cells undergoing different types of drugs.

5.2. Experimental Methods

5.2.1. Preparation of cells

Primary glioblastoma cells, GBM39 cells, were cultured to form neurospheres in DMEM F-12 (Invitrogen) supplemented with 2% of B27 (GIBCO), 1% of Glutamax (GIBCO), 1% of PSQ (GIBCO), 20 ng/mL of EGF (Sigma), 20 ng/mL of FGF (Sigma) and 1 μ g/mL of Heparin (Sigma). Every 3 days, growth factors were added to the cell culture media, and the growth media was changed every 2 weeks. In the case of ERZ cells, GBM39 cells having the resistance to erlotinib, neurospheres were cultured in the growth media having 5 μ M of erlotinib (Chemietek) to maintain the resistance.

In order to make a single-cell suspension, neurospheres were collected by the centrifugation at 400 g for 4 min. These neurospheres were then re-suspended in 3mL of Triple (Invitrogen) for 5 min at cell incubator, followed by the addition of the growth media to neutralize Triple. This solution was spined down again, and the single-cell suspension was obtained after the discard of supernatant followed by the resuspension of cell pellet into the growth media.

5.2.2. Cell sorting

In order to study the glycolytic metabolism along with the expression level of EGFRvIII in glioblastoma cells, single-cell suspension of GBM39 cells were prepared. Cells were washed with cold PBS followed by the resuspension to 20 million/mL in FACS buffer, PBS supplemented with 2% of FBS. After the removal of 20uL cell solution for the negative control, EGFRvIII antibody was added to the single-cell suspension to be reacted with cells for 20 min on ice. Cells were washed twice with ice-cold PBS, and re-suspended in FACS buffer containing secondary Anti-Mouse 488 (Jackson Immunolabs) for tagging with chromophore. After the incubation for 20 min on ice, cells were washed with ice-cold PBS twice, followed by the resuspension into Sorting Buffer, PBS supplemented with 0.5% FBS, at the concentration of 20 million/mL. According to their fluorescent signals, cells were sorted by a flow cytometer.

In order to measure the effect of drug treatment to metabolism of the EGFR+ subpopulation of GBM39 cells, the single-cell suspension was applied to the magnetic beads conjugated with EGFR antibody. After the sorting, cells were released and placed in a laminin coated cell culture petri for overnight to eliminate dead cells floating among growth media instead of binding to the laminin coated surface.

5.2.3. In vitro radio assay with RIMChip

The detailed procedure to fabricate RIMChip is introduced in Chapter 3. Cells at a concentration of 3×10^6 were injected into the microchip from inlets by pressing a pipettor for 5 sec to sufficiently cover all microchannels with cells followed by the incubation in a CO₂ incubator at 37 °C for 4 h. After incubation, the microchannels were flushed 2× with

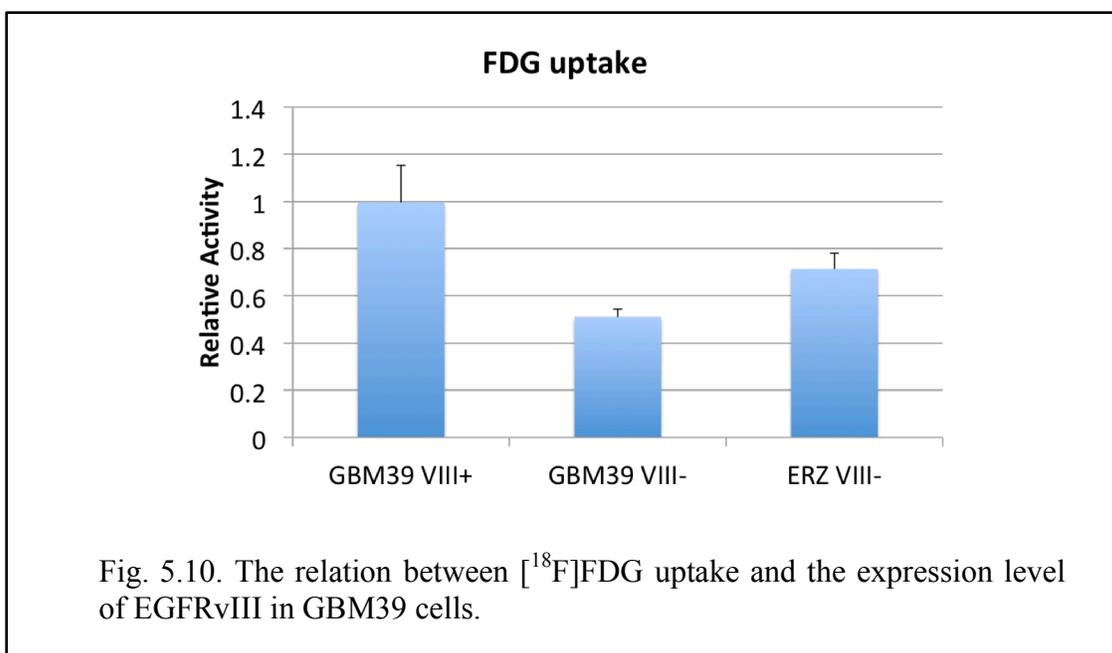
flowing ~ 20 μ l PBS using the same pipettor, each time for 30 sec, to flush out any residual glucose. About 10 μ l ^{18}F -FDG in PBS was loaded into the microchannels by pressing the pipettor for 15 sec. The whole microchip was then incubated for 30 min in a CO₂ incubator at 37 °C. It was the flushed 2 \times with flowing ~ 20 μ l PBS using the same pipettor. The cells in the microchip were then for imaging by the beta-box.

For the imaging of glycolytic metabolism in cells under drug treatment, we introduced CC214, U012610, or the combination of CC214 and 5 μ M erlotinib, to EGFR+ GBM39 cells 24 hour prior to the experiment. The remaining procedure was the same as above.

5.3. Results and Discussion

5.3.1. Imaging of glycolytic metabolism in glioblastoma cells according to their expression level of EGFRvIII

In order to study the glycolytic metabolism in the subpopulation of glioblastoma cells, we first did RIMChip experiment to image [^{18}F]FDG uptake in EGFRvIII+ and EGFRvIII- GBM39 cells based on the hypothesis that EGFRvIII+ cells, compared with EGFRvIII- cells, would show higher uptake of [^{18}F]FDG due to their uncontrolled active signaling. ERZ cells, GBM39 cells evolved to have erlotinib resistance during the treatment of the drug, were also studied to investigate the effect of erlotinib resistance as shown in **Fig. 5.1**.



As I expected, EGFRvIII+ subpopulation showed two times higher uptake of [^{18}F]FDG than EGFRvIII- counterpart. Interestingly, ERZ cells which rarely express EGFRvIII- on their cellular membrane consumed 20% more [^{18}F]FDG than EGFRvIII- subpopulation, which might be related with the evolution of bypass signaling pathway leading to the resistance to the inhibition of EGFR caused by erlotinib treatment.

5.3.2. Imaging of glycolytic metabolism in glioblastoma cells under drug treatment

As mentioned in Chapter 3, RIMChip can work as a rapid screening tool for drug treatments. Here we confirm the capability of RIMChip again with EGFR+ subpopulation of GBM cells undergoing different types of treatment: mTOR inhibition by CC214⁸, MEK inhibition by U0126⁹ and the dual inhibition of mTOR and MEK by the combinatorial treatment with CC214 and U0126.

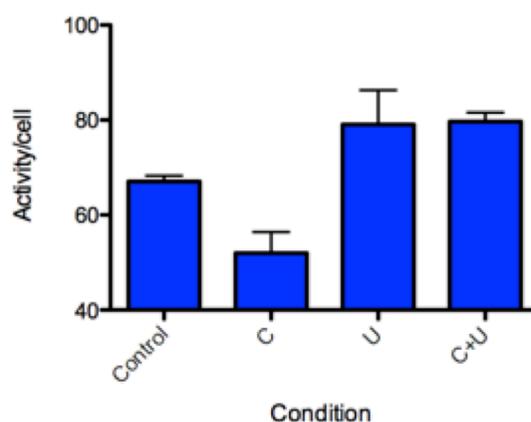


Fig. 5.11. The relation between [^{18}F]FDG uptake and the drug treatment to EGFR+ subpopulations of GBM39 cells C represents CC214, U represents U0126 and C+U represents the combinatorial treatment with CC214 and U0126.

As shown in **Fig. 5.2.**, [^{18}F]FDG uptake is decreased with the inhibition of mTOR by CC214, which can be explained by the diminished activation of glucose transporter at the downstream of mTOR signaling pathway¹⁰. The increased [^{18}F]FDG uptake was observed with the treatment of U0126, MEK inhibitor, which may be related with the interrupted activity of mitochondria as described in literature¹¹.

5.4. Conclusions and Future Directions

By using RIMChip platform, we observed the heterogeneity of energy metabolism in tumor cells with the imaging of glycolytic metabolism in the subpopulations of primary glioblastoma cells sorted by the expression level of oncoprotein, EGFRvIII. Also, we confirm the capability of RIMChip as a drug screening tool with the preliminary results of primary glioblastoma cells under treatment of different kinds of drugs. Combined with

further study of drug kinetics, with this platform, it would be possible to screen drugs rapidly with low requirement of biopsy samples.

5.5. References

1. Furnari, F. B. *et al.* Malignant astrocytic glioma: genetics, biology, and paths to treatment. *Gene Dev* **21**, 2683–2710 (2007).
2. Inda, M. D. M. *et al.* Tumor heterogeneity is an active process maintained by a mutant EGFR-induced cytokine circuit in glioblastoma. *Gene Dev* **24**, 1731–1745 (2010).
3. Vivanco, I. *et al.* The phosphatase and tensin homolog regulates epidermal growth factor receptor (EGFR) inhibitor response by targeting EGFR for degradation. in **107**, 6459–6464 (2010).
4. Stommel, J. M. *et al.* Coactivation of receptor tyrosine kinases affects the response of tumor cells to targeted therapies. *Science* **318**, 287–290 (2007).
5. Gupta, P. B., Fillmore, C. M., Jiang, G., Shapira, S. D. & Tao, K. Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* (2011).
6. Polyak, K. Heterogeneity in breast cancer. *The Journal of clinical investigation* (2011).
7. Giannini, C., Sarkaria, J. N. & Saito, A. Patient tumor EGFR and PDGFRA gene amplifications retained in an invasive intracranial xenograft model of glioblastoma multiforme. *Oncology-Basel* (2005).
8. 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**, 1588–1591 (2013).
9. Zohrabian, V. M., Forzani, B., Chau, Z., Murali, R. & Jhanwar-Uniyal, M. Rho/ROCK and MAPK signaling pathways are involved in glioblastoma cell migration and proliferation. *Anticancer Res* **29**, 119–123 (2009).
10. Buller, C. L. *et al.* A GSK-3/TSC2/mTOR pathway regulates glucose uptake and GLUT1 glucose transporter expression. *Am. J. Physiol., Cell Physiol.* **295**, C836–43

(2008).

11. Freeman, M. R., Kim, J., Lisanti, M. P. & Di Vizio, D. A metabolic perturbation by U0126 identifies a role for glutamine in resveratrol-induced cell death. *Cancer Biol. Ther.* **12**, 966–977 (2011).

Appendix A

Modular nucleic acid assembled p/MHC microarrays for multiplexed sorting of antigen-specific T cells

A.1. Introduction

T cells, which play a central role in an adaptive immune response, recognize different antigens via T cell receptor (TCR) generated by gene rearrangement and secondary modification. When a naive T cell triggered through the interaction between TCR and its matching antigenic peptides bound to Major Histocompatibility Complex (p/MHC) on the surface of antigen presenting cells (APC), it begins to proliferate rapidly and form a large population of matured effector cells to defend the host. Even though the pathogen is cleared, some portion of activated cells survives and transforms into memory cell able to give the host a quick immune response toward the previously faced pathogen. As the result, each individual has his/her own T cell repertoire showing immune response toward self and foreign antigens. To detect and characterize these T cell populations, consequently, has fundamental and clinical importance.

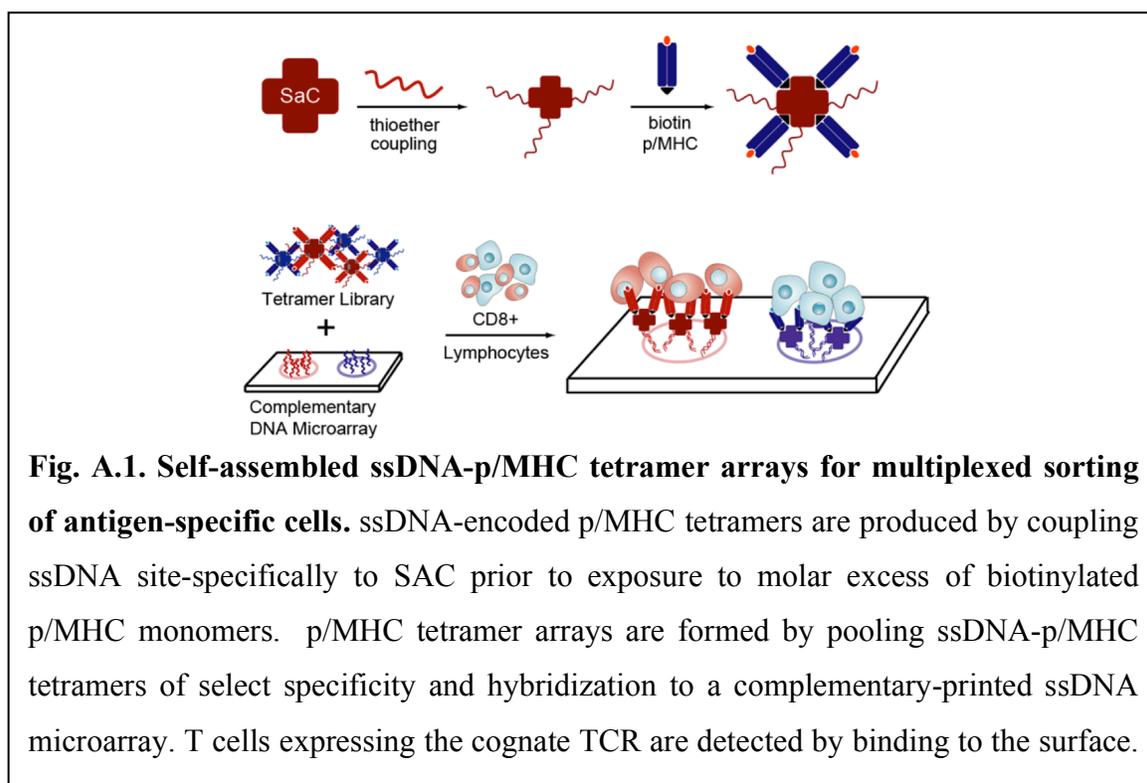
In order to detect antigen-specific T cells, soluble p/MHC tetramer has been developed. Four enzymatically biotinylated p/MHC monomers are coupled with a streptavidin molecule primed by chromophore to form a p/MHC tetramer, and the tetramer enables its counter part T cell to be detected via flow cytometry¹. Because the p/MHC tetramer uses chromophore for its detection, however, this method has inevitable restrictions. At first, each antigen-specificity requires unique optical dye

molecule having little spectral overlap with others. Secondly, the existence of detection limit in flow cytometry, serial flow cytometry detection for multiplexing study is restricted by the sample size. Due to the existence of diverse optical characters in each cell population, moreover, the matter of compensation always follows. Although researchers used polychromatic flow cytometry utilizing quantum dots to resolve those problems^{2,3}, high cost, long sample preparation time, and complex color compensation still hinder further study.

To avoid such problems, several researchers have reported antigen-specific cell sorting by microarrays, which contains directly printed p/MHC tetramers on a supporting substrates⁴⁻⁷. When a population of cells is applied on the array, only target antigen-specific T cells bind to the region having p/MHC tetramers. After that, the captured cells are observed by a microscopy. Because this method uses the location of p/MHC tetramers to separate a target antigen-specific T cell from the variety of other cells instead of chromophore's emission spectrum, multiplexing is only restricted the number of reagent able to fit on a substrate.

In order to increase the capture efficiency, orientation of p/MHC tetramer should be preserved in a way to react approaching cells and the tetramer must not be denatured through the microarray fabrication. Analogous protein arrays, made via antibody adsorption to unmodified and derivatized surfaces, have shown that the adsorbed antibodies could fully react because of surface-induced effects such as protein denaturation and orientational change to inactive. As the result, amount of functional

antibody is decreased, immobilization occurs heterogeneously, and required concentrations of antibody is increased to compensate for the protein loss⁸⁻¹⁰. To avoid these problems, several mild chemistries have been studied for protein immobilization¹¹⁻¹⁷, but often the surface that meets the demands of application requires a high level of technical expertise and/or is limited in accessibility. Therefore, the development of a new technology to fabricate p/MHC array in an easy way is important.



Here I introduce a new p/MHC arrays conjugated with nucleic acid for multiplexed antigen-specific lymphocytes sorting. Instead of direct printing, each specific p/MHC tetramer is conjugated to unique sequence of ssDNA designed to be orthogonal to other sequences. On the glass surface where the complementary DNA sequences are printed, the ssDNA-p/MHC tetramer conjugates are self-assembled by DNA hybridization, then used as a p/MHC array to sort mixed population of antigen-specific T

cells (**Fig. A.1**). This method is called “Nucleic Acid Cell Sorting (NACS).” Because NACS employs DNA as a linkage molecule, fabrication of p/MHC arrays is simple and highly modular¹⁸⁻²¹. Traditional DNA microarray technologies are also available to make cDNA glass substrates. In order to produce ssDNA-p/MHC tetramer conjugates having maximum capacity to immobilize cells, cysteine-engineered SA (SAC) is used as a scaffold. NACS arrays have better performance than traditional p/MHC arrays made by direct printing, and the specificity, multiplexing and sensitivity of NACS are studied. Selective detachment after cell sorting by NACS is also illustrated, and its application to primary human T cells is introduced.

A.2. Experimental Methods

A.2.1. Microarray fabrication

All DNAs were acquired from IDT. By the facility at the Institute for Systems Biology (ISB, Seattle), DNA strands were printed on amine coated glass slides (GAPSII, Corning) in the way to form 12x12 arrays having alternative rows of A, B, and C, or A_{EcoRI} and B_{BamHI} spots with a SMPXB15 pin (Arrayit). Sequence of each strand and its counterparts is written at the following **Table A.1**.

Table 2.1. Orthogonal DNA sequences for spatial encoding of p/MHC tetramers

Name	Sequence*
A	5' - AAA AAA AAA AAA AAT CCT GGA GCT AAG TCC GTA AAA AAA AAA AAT CCT GGA GCT AAG TCC GTA AAA AAA AAA AAA A
A'	5' - NH ₂ - AAA AAA AAA ATA CGG ACT TAG CTC CAG GAT
B	5' - AAA AAA AAA AAA AGC CTC ATT GAA TCA TGC CTA AAA AAA AAA AGC CTC ATT GAA TCA TGC CTA AAA AAA AAA AAA A
B'	5' - NH ₂ - AAA AAA AAA ATA GGC ATG ATT CAA TGA GGC
C	5' - AAA AAA AAA AAA AGC ACT CGT CTA CTA TCG CTA AAA AAA AAA AGC ACT CGT CTA CTA TCG CTA AAA AAA AAA AAA A
C'	5' - NH ₂ - AAA AAA AAA ATA GCG ATA GTA GAC GAG TGC
A _{EcoRI}	5' - AAA AAA AAA AAA GAG CTA AGT CCG TAG AAT TCA AAA AAA AAA GAG CTA AGT CCG TAG AAT TCA AAA AAA AAA AAA
A _{EcoRI} '	5' - NH ₂ - AAA AAA AAA AGA ATT CTA CGG ACT TAG CTC CAG GAT
B _{BamHI}	5' - AAA AAA AAA AAA TTG AAT CAT GCC TAG GAT CCA AAA AAA AAA TTG AAT CAT GCC TAG GAT CCA AAA AAA AAA AAA
B _{BamHI} '	5' - NH ₂ - AAA AAA AAA AGG ATC CTA GGC ATG ATT CAA TGA GGC

* All sequences to be conjugated to SAC (A', B', C', A_{EcoRI}', and B_{BamHI}') were designed with a polyA linker followed by a 20mer hybridization region. The 5' amine is required for the attachment of the hetero-bifunctional maleimide derivative MHPH. Sequences printed on glass substrates (A, B, C, A_{EcoRI}, and B_{BamHI}) were designed with two hybridization regions separated by polyAs. This was designed to facilitate electrostatic adsorption to amine glass substrates.

A.2.2. Synthesis of DNA-SAC conjugates

In order to express SAC, pET-3a plasmid, as a kind gift from Takeshi Sano (Harvard Medical School), was used. SAC was expressed according to the previously published protocol²², and buffer exchanged to PBS with 5mM Tris(2-carboxyethyl) phosphine hydrochloride (Solulink), using zeba desalting columns (Pierce). MHPH (3-N-

Maleimido-6-hydraziniumpyridine hydrochloride, Solulink) was added to SAC at a molar excess of 300:1. Meanwhile, SFB (Succinimidyl 4-formylbenzamide, Solulink) in DMF was added to 5' aminated DNA oligos at a molar excess of 40:1. In order to finish reactions, these mixtures were incubated at room temperature (RT) for 3-4 hours. Then, both mixture were buffer exchanged to pH 6.0 buffer solution (50mM sodium citrate, 150mM NaCl), and SFB-conjugated DNA oligos were added to MHPH-conjugated SAC at a molar excess of 20:1, and incubated overnight at RT. With Phamacia Superdex 200 gel filtration column at 0.5ml/min isocratic flow of PBS, SAC-DNA oligo conjugates was purified, and concentrated by 10K mwco concentration filters (Millipore).

A.2.3. Preparation of T cells

The cDNA from alpha and beta chains of TCR specific for tyrosinase₃₆₈₋₃₇₆ was obtained from Michael I. Nishimura (Medical University of South Carolina, Charleston, SC). These alpha and beta chain were cloned into a lentiviral vector having transgenes linked by a 2A self-cleaving sequence as written²³. Supernatant from the lentiviral vector was concentrated and applied to infect Jurkat cells to generate Jurkat^{a-Tyro} cells. In order to make Jurkat^{a-MART-1} cells, MSGV1-F5Aft2AB retroviral vector expressing the F5 MART-1 TCR, obtained from Steven A. Rosenberg and Richard Morgan (Surgery Branch, National Cancer Institute Bethesda, MD), was used to infect Jurkat cells. For generating primary human T lymphocytes expressing the F5 MART-1 TCR, the same vector was applied to PBMCs from leukapheresis. Prior to the infection, PBMCs were activated with 50ng/ml of OKT3 (muromonab anti-human CD3 antibody, Ortho-Biotech, Bridgewater, NJ) and 300 U/ml of IL-2 (adesleukin, Novartis, Emeryville, CA) for 48

hours. Then, on the retronectin-coated wells (Takara Bio Inc., Japan) containing MSGV1-F5Aft2AB retroviral vector, activated PBMC in RPMI with 5% human AB serum supplemented by 300 IU of IL-2 were added and incubated at 37°C for overnight. On the next day, a second set of pre-coated retronectin retroviral vector tissue culture plate was used to transfer the activated PBMC and incubated at 37°C for overnight. Subsequently, with the same media as above, cells were washed and re-suspended. In RPMI supplemented with 10% human AB serum and 1% penicillin, streptomycin, and amphotericin (Omega Scientific), frozen leukapheresis fractions from patients NRA11 and NRA 13 (UCLA IRB#03-12-023) were thawed and incubated overnight. Then using an AutoMACS machine according to the manufacturer's instructions, CD8⁺ enrichment (anti-CD8 microbeads, Miltenyi Biotech) was done. Following separation, the cells were cultured at in RPMI-humanAB media containing 30 U IL2/mL.

A.2.4. Sorting cells

The HLA-A*0201 restricted MHC class I monomers loaded with tyrosinase369-377 (YMDGTMSQV) and MART-126-35 (ELAGIGILTV) were made in house with previous published protocols²⁴. A2.1-restricted EBV BMLF1259-267 (GLCTLVAML), CMV pp65495-503 (NLVPMVATV), murine H-2Kb/-OVA257-264 (SIINFEKL), and murine H-2Db/-gp10025-33 (KVPRNQDWL) as well as all fluorescent HLA-A*0201 tetramers were purchased from Beckman Coulter, and lipophilic cell membrane staining dyes DiO, DiD, and DiL were obtained from Invitrogen.

To prevent non-specific cell binding, 1mg/ml PEG-NHS ester (Sunbio) in PBS was applied on microarray slides for 2 hours at RT. p/MHC monomers were combined in a 4:1 molar excess with ssDNA-SAC at 37°C for 20 min to form ssDNA-p/MHC tetramers. With the tetramers in 200µl media, DNA microarrays was hybridized for 1 hour at 37°C, and washed with 3% FBS in PBS. T cells (10^6 cells/ml) were incubated on the array for 30 min at 37°C, and washed with the same media.

In the comparative study, SuperEpoxy(covalent) and SuperProtein(hydrophobic) slides were obtained from Arrayit, GAPSII(electrostatic) slide was purchased from Corning, and Hydrogel(hydrophilic) slide was acquired from Xantec. According to manufacturer's protocol, fluorescent MART-1 tetramers were immobilized on each slide, and the result was analyzed with ImageJ and Origin.

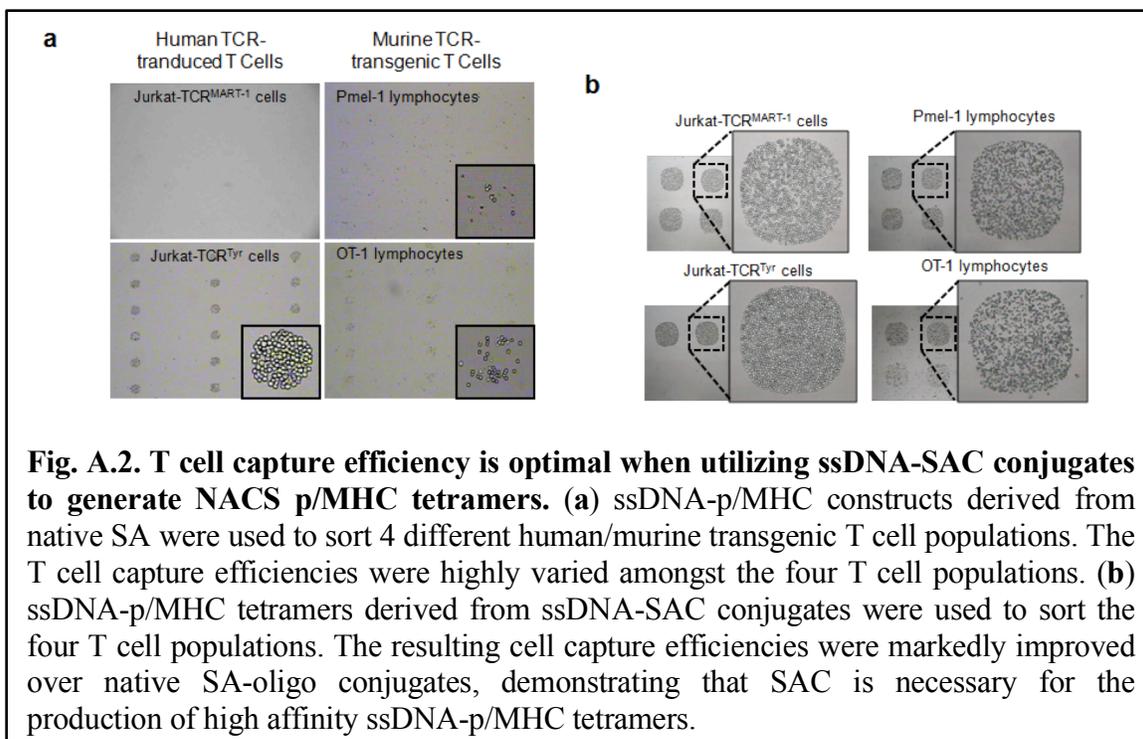
Staining p/MHC tetramer after T cell capture was done with adding fluorescent cDNA (Cy5-A' and Cy3-B). For selective T cell release, EcoRI, BamHI or DNase in RPMI media was applied onto the immobilized cells for 1-2 hours at 37°C. Two restrict enzymes were purchased from NEbiolabs, and DNase was obtained from Sigma. All cell capture images were obtained by bright filed (Nikon Eclipse TE2000) and/or confocal microscopy (Nikon E800).

A.3. Results and Discussions

A.3.1. Design of ssDNA-p/MHC tetramers

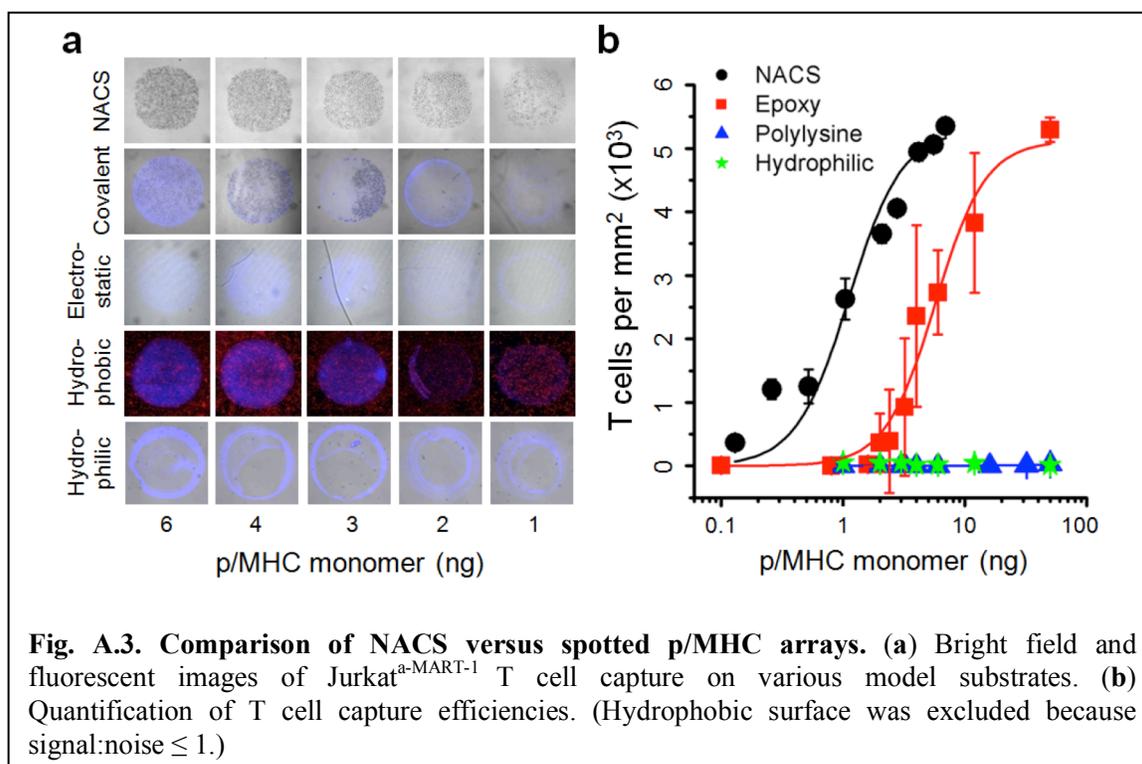
Mostly used reagent to assemble p/MHC monomer into tetramers is SA-phycoobiliprotein (PE or APC) conjugates, which is linked via chemical cross-linking. However, we could not use them, because their functional groups for attaching ssDNA are modified during the conjugation. Unmodified SA was not the best choice because its functional groups close to the biotin binding sites prohibit the access of biotin as published^{25,26}. We solve the problem with employing a mutant SA, having additional cysteine residue at the carboxy-terminus, which was first introduced by Ramachandiran and co-workers²⁵. Because native SA has no cysteine residue, the reacting site can be restricted at the end of carboxy-terminus²⁷, far from biotin binding pocket, when cysteine-maleimide conjugation was applied.

In order to compare the ability of biotin binding between SA-ssDNA conjugate and SAC-ssDNA conjugate, 2-(4'-Hydroxyazobenzene) benzoic acid (HABA)²⁸ was used. Because it is a molecular mimic of biotin with distinct optical density coefficients dependent on whether biotin is bound to SA or not, the biotin binding capacity is verified through absorption spectra. The result of HABA absorption spectra showed the binding capacity of SAC-ssDNA conjugate is 3.7 while that of native SA-ssDNA is 2.9. When these conjugates were applied to capture 4 different monoclonal T cells, the difference of binding capacity was signified. NACS p/MHC tetramers assembled by SAC-ssDNA conjugates immobilized T cells with a high efficiency, while SA-ssDNA conjugates coupled with p/MHC showed little cells capture (**Fig. A.2.**). Therefore, all NACS tetramers as follow were assembled by SAC-ssDNA.



A.3.2. Comparison between NACS and conventional protein microarray

For comparison study between NACS and conventional microarrays based on direct protein spotting method, we chose various substrates to cover different spectrum of commonly used strategies (covalent, electrostatic, hydrophobic, and hydrophilic adsorption) for protein immobilization. According to manufacturer's instruction, serial dilutions of MART-1 SA-PE tetramer (HLA-A2.1 MHC molecules loading melanoma epitope MART-1₂₆₋₃₅) were directly printed. Onto these microarray, Jurkat^{a-MART-1} T cells (Jurkat cells transduced with the F5 MART-1 TCR²⁹ specific for peptide epitope MART-1₂₆₋₃₅) were applied. Collected images of immobilized cells on the various substrates (**Fig. A.3.a**) and quantified analysis (**Fig. A.3.b**) were obtained.



Compared to conventional microarrays with identical concentration of p/MHC tetramers, NACS was superior. Electrostatic and hydrophilic immobilization have little/no captured T cells, and hydrophobic immobilization has huge noise. Even though covalent immobilization succeeded to capture T cells, it had intra-spot and inter-spot heterogeneity while NACS gave homogeneous result at the same concentration of tetramers. Furthermore, as shown in the quantified analysis, only with a fifth of materials required in covalent immobilization, NACS could capture equivalent T cell. (p/MHC monomer at half max $\equiv K_{1/2} = 1.1$ ng for NACS and 5.7 ng for covalent immobilization).

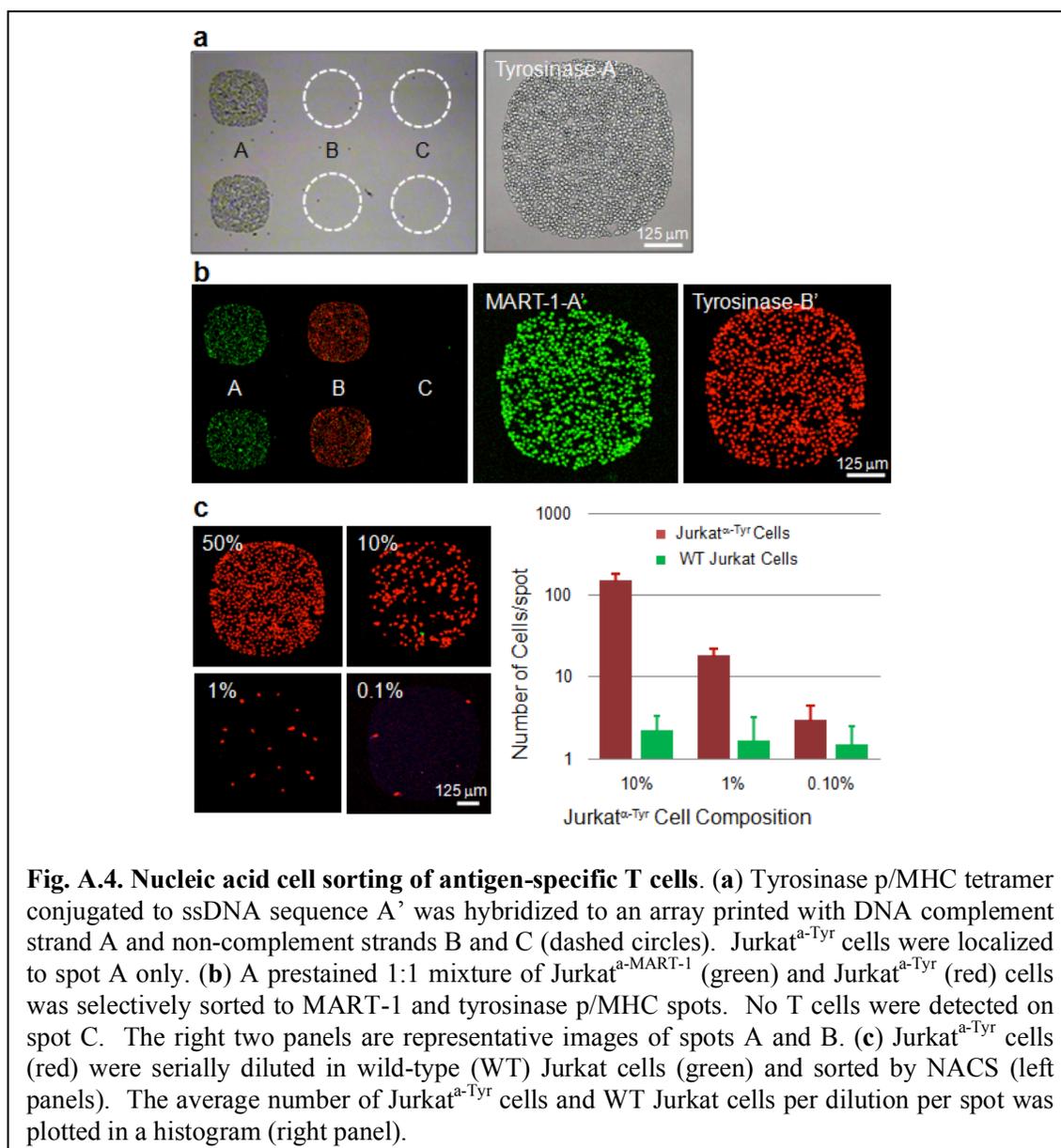
Two factors can be reasons of superior capturing efficiency in NACS. At first, linked by DNA-cDNA hybridization, p/MHC tetramers of NACS have great orientational freedom while those of conventional microarrays hardly have other options except just absorbed on the surface. This may increase the density of functional p/MHC tetramers

and result in reduced $K_{1/2}$. Secondly, NACS does not disrupt hydration state of arrayed proteins, which can be modified during production and storage process of conventional protein microarray to decrease reproducibility^{11,16,21}. Because p/MHC tetramer arrays are self-assembled in solution immediately prior to an experiment and only stable DNA oligos are printed on the surface when the chip is stored, NACS may show intra-spot and inter-spot homogeneity and guarantee the array reproducibility.

A.3.3. Specificity of NACS and its detection limit

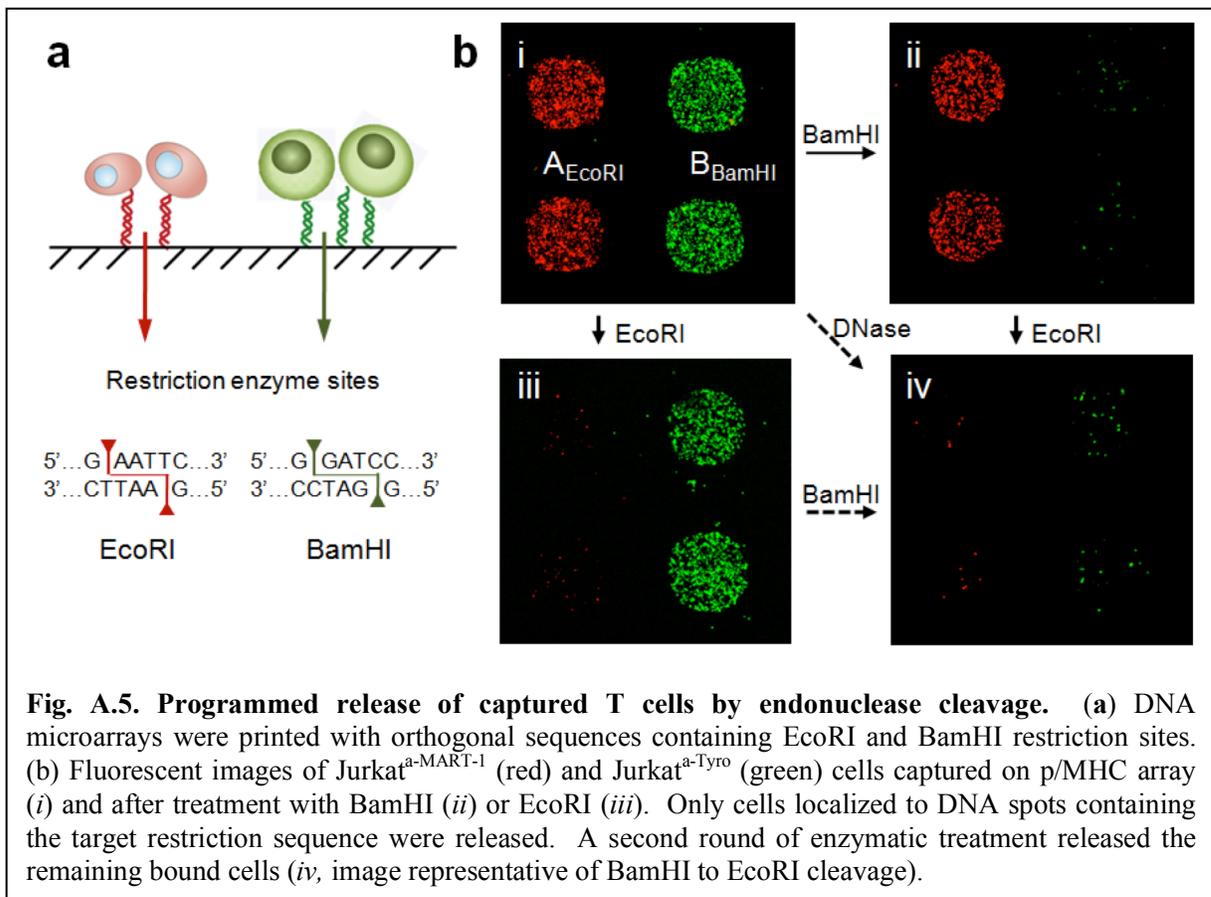
In order to study the specificity of NACS, SAC-ssDNA p/MHC tetramer (human HLA-A*0201 MHC molecules loaded with melanoma antigen peptide epitope tyrosinase368-376 with pendant DNA sequence A') was applied to a DNA microarray made with the complementary strand A and two other orthogonal strand B and C. Then, Jurkat^{α-Tyr} cells (human T leukemia cell line transduced with a TCR specific for tyrosinase368-376)³⁰ were applied to the array. T cells were immobilized only on A spots hybridized appropriate p/MHC, not B or C spots (**Fig. A.4.a**), and the mean binding number of Jurkat^{α-Tyr} cells calculated from three spots was $\sim 1486 \pm 62$. When a 1:1 mixture of Jurkat^{α-MART-1} and Jurkat^{α-Tyr} cells pre-stained with lipophilic dyes (green and red respectively) was applied to the array, those cells were captured into alternating columns and few cells were localized to spot C. The average number of captured cell was a factor of two less than that of homogeneous sorting. (661 ± 19 T cells/spot) (**Fig. A.4.b**). In order to find the detection limit of NACS, populations of Jurkat^{α-Tyr} were spiked at 10%, 1%, and 0.1% into wild Jurkat T cells and sorted. The number of captured cells per spot was analyzed, which shows linear correlation with the fractional

composition of Jurkat^{a-Tyr} cells in the mixture (**Fig A.4.c**). The limit of detection is around 1 in 1000 cells, and non-specific binding between wild Jurkat T cells and the array was always constant.



A.3.4. Selective release of immobilized T cells with restriction endonucleases

After immobilization of cells, immunohistochemistry(IHC), fluorescent in situ hybridization(FISH), and cytokine secretion assays were traditionally performed^{4,6}. In order to study functional status of captured cells through mRNA or TCR, however, selective release of cell is required for eliminating the interference of back ground noise. Contrast to the conventional protein arrays made by direct printing, NACS can employ DNA sequence having unique cleavage site by restriction enzyme to selectively release captured cells (**Fig. A.5.a**). Before the specific release by restriction enzyme, Jurkat^{a-}MART-1 and Jurkat^{a-Tyr} cells prestained by lipophilic dyes (red and green, respectively) were immobilized on a NACS array employed by DNA stands A_{EcoRI} and B_{BamHI} (**Fig. A.5.bi**). Then BamHI was applied on the array for 1 hour to cleave B_{BamHI} strands. Meanwhile, on a separate but identical array, EcoRI was applied to cleave A_{EcoRI} strands at the same condition. In each case, selective release was observed (**Fig. A.5.bii and biii**). When a complementary restrict endonuclease was applied on those arrays, remaining cells were released (**Fig. A.5.biv**). Of course, non-selective release was achieved in a single step with an addition of DNase (data not shown).



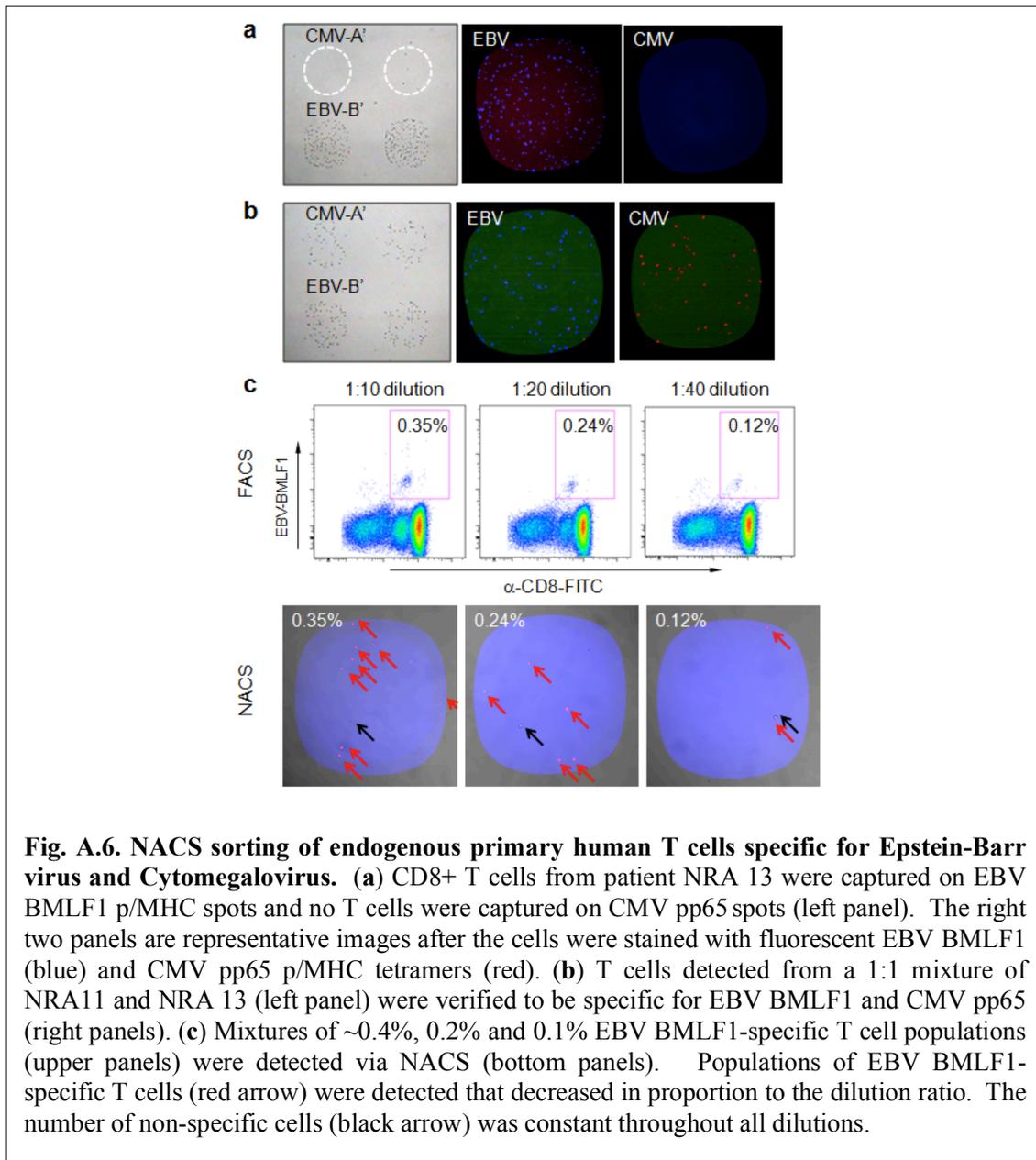
A.3.4. Sorting of TCR-engineered and endogenous primary human T cells by NACS

Recently, generating a huge number of tumor antigen-specific T cells and developing a therapy using those cells in patients with melanoma and other cancers have been reported^{31,32}. In the study, T cells of patients was collected and engineered with a TCR against a target cancer antigen to get rid of tumor cells. In order to verify its clinical relevance, detecting TCR-engineered human lymphocytes is important, and it is achieved by NACS in this study. Via leukapheresis, human peripheral blood mononuclear cells (PMBCs) containing CD8⁺ cells of patients were extracted, and those cells were expanded and transduced with a retrovirus vector containing the F5 MART-1 TCR. Then, the cells were applied on a NACS array containing MART-1 and Cytomegalovirus (CMV)

pp65₄₉₅₋₅₀₃/HLA-A2.1 p/MHC tetramers. Only on the MART-1 spots, transduced T cells were captured. On the array, the antigen-specific sorting was validated again via antigen-specific staining with fluorescent MART-1 and CMV p/MHC tetramers (red and blue, respectively) (**Fig. A.6.a**)

Because there are huge number of T cells expressing different monoclonal or polyclonal TCR, detection of specific primary human T cell isolated from peripheral blood is much more difficult than cultured cell lines. Also, these T cells have endogenous expression level of TCR. In order to show that the specificity, multiplexing, and sensitivity of NACS can be equally applied to endogenous primary human T cells as like to cultured cell lines, frozen leukapheresis samples from patient NRA13 was studied. These cells were CD8⁺ enriched and applied to a CMV and Epstein-Barr virus (EBV BMLR1₂₅₉₋₂₆₇/HLA-A2.1) p/MHC array. Only on the EBV spots, T cells were immobilized, and the result of flow cytometry showed ~5% of NRA13 CD8⁺ cells were EBV-specific, but ~0% of them were CMV-specific.

Multiplexed detection was also available. When a 1:1 mixture of EBV-specific and CMV-specific CD8⁺ T cells, produced by combining NRA 13 lymphocytes with CMV-specific T cells from NRA11, was applied on a NACS array and stained with fluorescent p/MHC tetramer, immobilized cells showed appropriate color of fluorescence (**Fig. A.6.b**). In order to verify the detection limit, serial dilutions of EBV-specific T cell mixtures (~0.4%, ~0.2% and ~0.1% by FACS) were applied on NACS array, and the result demonstrated the resolution can be lowered at 0.1% (**Fig. A.6.c**).



A.4. Conclusions

We showed the NACS platform is efficient, facile, and modular on-chip strategy for the immobilization of antigen-specific T cells. Because the DNA-printed glass substrate can be easily made by traditional DNA printing technologies, NACS is also inexpensive. Streptavidin-cysteine conjugated with single stranded-DNA is rationally designed to bind any family of proteins or small molecules labeled with biotin, therefore, the application of NACS can be extended to other capture agents, such as biotinylated antibodies.

A.5. References

1. Altman, J. D. *et al.* Phenotypic analysis of antigen-specific T lymphocytes. *Science* **274**, 94–96 (1996).
2. McLaughlin, B. E. *et al.* Nine-color flow cytometry for accurate measurement of T cell subsets and cytokine responses. Part II: Panel performance across different instrument platforms. *Cytometry A* **73**, 411–420 (2008).
3. Chattopadhyay, P. K. *et al.* Quantum dot semiconductor nanocrystals for immunophenotyping by polychromatic flow cytometry. *Nature Publishing Group* **12**, 972–977 (2006).
4. Chen, D. S. *et al.* Marked differences in human melanoma antigen-specific T cell responsiveness after vaccination using a functional microarray. *PLoS Med.* **2**, e265 (2005).
5. Soen, Y., Chen, D. S., Kraft, D. L., Davis, M. M. & Brown, P. O. Detection and characterization of cellular immune responses using peptide-MHC microarrays. *PLoS Biol* **1**, E65 (2003).
6. Stone, J. D., Demkowicz, W. E. & Stern, L. J. HLA-restricted epitope identification and detection of functional T cell responses by using MHC-peptide and costimulatory microarrays. *P Natl Acad Sci USA* **102**, 3744–3749 (2005).
7. Deviren, G., Gupta, K., Paulaitis, M. E. & Schneck, J. P. Detection of antigen-

- specific T cells on p/MHC microarrays. *J. Mol. Recognit.* **20**, 32–38 (2007).
8. Haab, B. B., Dunham, M. J. & Brown, P. O. Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. *Genome Biol.* **2**, RESEARCH0004 (2001).
 9. Butler, J. E. *et al.* The physical and functional behavior of capture antibodies adsorbed on polystyrene. *J. Immunol. Methods* **150**, 77–90 (1992).
 10. Butler, J. E. *et al.* The immunochemistry of sandwich ELISAs--VI. Greater than 90% of monoclonal and 75% of polyclonal anti-fluorescyl capture antibodies (CAbs) are denatured by passive adsorption. *Mol. Immunol.* **30**, 1165–1175 (1993).
 11. MacBeath, G. & Schreiber, S. L. Printing proteins as microarrays for high-throughput function determination. *Science* **289**, 1760–1763 (2000).
 12. Lesaichere, M.-L., Lue, R. Y. P., Chen, G. Y. J., Zhu, Q. & Yao, S. Q. Intein-mediated biotinylation of proteins and its application in a protein microarray. *J. Am. Chem. Soc.* **124**, 8768–8769 (2002).
 13. Peluso, P. *et al.* Optimizing antibody immobilization strategies for the construction of protein microarrays. *Anal. Biochem.* **312**, 113–124 (2003).
 14. Kwon, Y., Han, Z., Karatan, E., Mrksich, M. & Kay, B. K. Antibody arrays prepared by cutinase-mediated immobilization on self-assembled monolayers. *Anal. Chem.* **76**, 5713–5720 (2004).
 15. Arenkov, P. *et al.* Protein microchips: use for immunoassay and enzymatic reactions. *Anal. Biochem.* **278**, 123–131 (2000).
 16. Kiyonaka, S. *et al.* Semi-wet peptide/protein array using supramolecular hydrogel. *Nat Mater* **3**, 58–64 (2004).
 17. 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**, 1959–1967 (2007).
 18. Boozer, C. *et al.* DNA directed protein immobilization on mixed ssDNA/oligo(ethylene glycol) self-assembled monolayers for sensitive biosensors. *Anal. Chem.* **76**, 6967–6972 (2004).
 19. Niemeyer, C. M. Functional devices from DNA and proteins. *Nano Today* **2**, 42–

- 52 (2007).
20. Chandra, R. A., Douglas, E. S., Mathies, R. A., Bertozzi, C. R. & Francis, M. B. Programmable cell adhesion encoded by DNA hybridization. *Angew Chem Int Ed Engl* **45**, 896–901 (2006).
 21. Fan, R. *et al.* Integrated barcode chips for rapid, multiplexed analysis of proteins in microliter quantities of blood. *Nat Biotechnol* **26**, 1373–1378 (2008).
 22. SANO, T. & Cantor, C. R. Expression of a cloned streptavidin gene in *Escherichia coli*. *P Natl Acad Sci USA* **87**, 142–146 (1990).
 23. Szymczak, A. L. *et al.* Correction of multi-gene deficiency in vivo using a single ‘self-cleaving’ 2A peptide-based retroviral vector. *Nat Biotechnol* **22**, 589–594 (2004).
 24. Garboczi, D. N., Hung, D. T. & Wiley, D. C. HLA-A2-peptide complexes: refolding and crystallization of molecules expressed in *Escherichia coli* and complexed with single antigenic peptides. *P Natl Acad Sci USA* **89**, 3429–3433 (1992).
 25. Ramachandiran, V. *et al.* A robust method for production of MHC tetramers with small molecule fluorophores. *J. Immunol. Methods* **319**, 13–20 (2007).
 26. Cameron, T. O., Cochran, J. R., Yassine-Diab, B., Sékaly, R. P. & Stern, L. J. Cutting edge: detection of antigen-specific CD4⁺ T cells by HLA-DR1 oligomers is dependent on the T cell activation state. *J Immunol* **166**, 741–745 (2001).
 27. Reznik, G. O., Vajda, S., Cantor, C. R. & SANO, T. A streptavidin mutant useful for directed immobilization on solid surfaces. *Bioconjug. Chem.* **12**, 1000–1004 (2001).
 28. Green, N. M. in *Methods in Enzymology* **18**, 418–424 (Elsevier, 1970).
 29. Johnson, L. A. *et al.* Gene transfer of tumor-reactive TCR confers both high avidity and tumor reactivity to nonreactive peripheral blood mononuclear cells and tumor-infiltrating lymphocytes. *J Immunol* **177**, 6548–6559 (2006).
 30. Nishimura, M. I. *et al.* MHC class I-restricted recognition of a melanoma antigen by a human CD4⁺ tumor infiltrating lymphocyte. *Cancer Res* **59**, 6230–6238 (1999).
 31. Morgan, R. A. *et al.* Cancer regression in patients after transfer of genetically

- engineered lymphocytes. *Science* **314**, 126–129 (2006).
32. Schumacher, T. N. M. T-cell-receptor gene therapy. *Nat Rev Immunol* **2**, 512–519 (2002).

Appendix B

Protein-Signaling Networks from Single-cell Fluctuations and Information Theory Profiling

B.1. Introduction

Protein-signaling pathways play important roles in tissue processes ranging from tumorigenesis to wound healing¹⁻⁵. Elucidation of these signaling pathways is challenging, in large part, because of the heterogeneous nature of tissues⁶. Such heterogeneity makes it difficult to separate cell-autonomous alterations in function from alterations that are triggered via paracrine signaling, and it can mask the cellular origins of paracrine signaling. Intracellular signaling pathways can be resolved via multiplex protein measurements at the single-cell level⁷. For secreted protein signaling, there are additional experimental challenges. Intracellular staining flow cytometry (ICS-FC) requires the use of protein transport inhibitors which can influence the measurements⁸. In addition, the largest number of cytokines simultaneously assayed in single-cells by ICS-FC is only 5⁹. Finally, certain biological perturbations, such as the influence of one cell on another, are difficult to decipher using ICS-FC. Other methods, such as multiplex fluorospot assays¹⁰, have even more significant limitations.

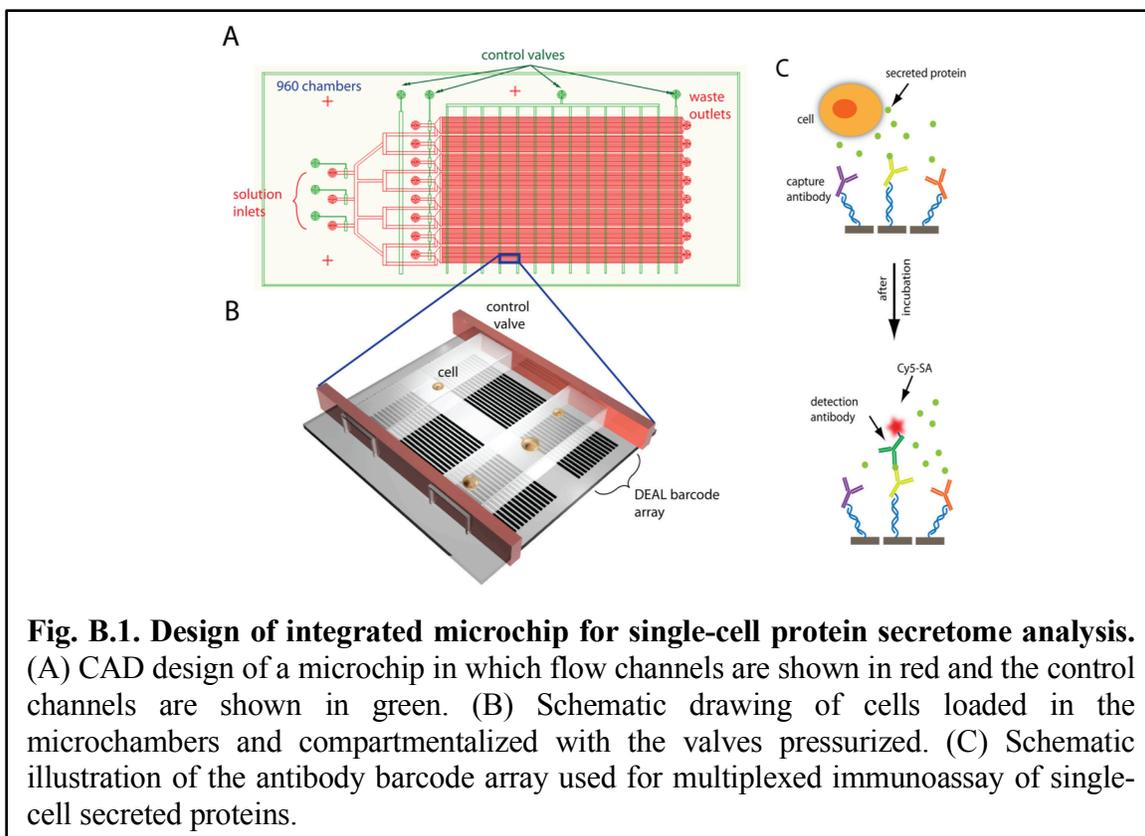
We describe here an experimental/theoretical approach designed to unravel the coordinated relationships between secreted proteins, and to understand how molecular and cellular perturbations can influence those relationships. Our starting points are single, lipopolysaccharide (LPS)-stimulated, human macrophage cells¹¹. LPS stimulation activates the Toll-like Receptor-4 (TLR-4), and emulates the innate immune response to Gram-negative bacteria. We characterize the secretome, at the single-cell level, through the use of a microchip platform in which single, stimulated macrophage cells are isolated into 3 nanoliter (nl) volume microchambers, with ~1000 microchambers per chip. Each microchamber permits duplicate assays for each of a dozen proteins that are secreted over the course of a several-hour incubation period following cell stimulation. The barcode assays are developed using detection antibodies and fluorescent labels, and then converted into numbers of molecules detected. We demonstrate that the observed spread in protein

levels is dominated by the cellular behaviors (the biological fluctuations), rather than the experimental error. These fluctuations are utilized to compute a covariance matrix linking the different proteins. This matrix is analyzed at both coarse and fine levels to extract the protein-protein interactions. We demonstrate that our system has the stability properties requisite for the application of a quantitative version of a Le Chatelier-like principle, which permits a description of the response of the system to a perturbation. This is a prediction in the strict thermodynamic sense. The fluctuations, as assessed from the multiplexed protein assays from unperturbed single-cells, are used to predict the results when the cells are perturbed by the presence of other cells, or through molecular (antibody) perturbations.

B.2. EXPERIMENTAL METHODS

B.2.1. Microchip fabrication.

The SCBCs were assembled from a DNA barcode microarray glass slide and a PDMS slab containing a microfluidic circuit^{12,13}. The DNA barcode array was created with microchannel-guided flow patterning technique¹³. Each barcode was comprised of thirteen stripes of uniquely designed ssDNA molecules. PDMS microfluidic chip was fabricated using a two-layer soft lithography approach¹⁴. The control layer was molded from a SU8 2010 negative photoresist (~20 μm in thickness) silicon master using a mixture of GE RTV 615 PDMS prepolymer part A and part B (5:1). The flow layer was fabricated by spin-casting the pre-polymer of GE RTV 615 PDMS part A and part B (20:1) onto a SPR 220 positive photoresist master at ~2000 rpm for 1minute. The SPR 220 mold was ~18 mm in height after rounding via thermal treatment. The control layer PDMS chip was then carefully aligned and placed onto the flow layer, which was still situated on its silicon master mold, and an additional 60 min thermal treatment at 80 °C was performed to enable bonding. Afterward, this two-layer PDMS chip was cut off and access holes drilled. In order to improve the biocompatibility of PDMS, we performed a solvent extraction step, which removes uncrosslinked oligomers, solvent and residues of the curing agent through serial extractions/washes of PDMS with several solvents^{15,16}. We noticed that this step significantly improves the biocompatibility and the reproducible protein detection. Finally, the microfluidic-containing PDMS slab was thermally bonded onto the barcode-patterned glass slide to give a fully assembled microchip.



B.2.2. Preparation of barcode arrays

The barcode array initially consists of 13 uniquely designed DNA strands labeled in order as A through M. Prior to loading cells, a cocktail containing all capture antibodies conjugated to different complementary DNA strands (A'-L') is flowed through the chambers, thus transforming, via DNA-hybridization, the DNA barcode into an antibody array. These dozen proteins that comprised the panel used here were encoded by the DNA strands A through L, respectively. Calibration and cross-reactivity curves for each protein assay are in **Fig. B.2.**, The DNA oligomer sequences and the antibody pairs used are listed in **Table B.1.** and **Table B.2.**

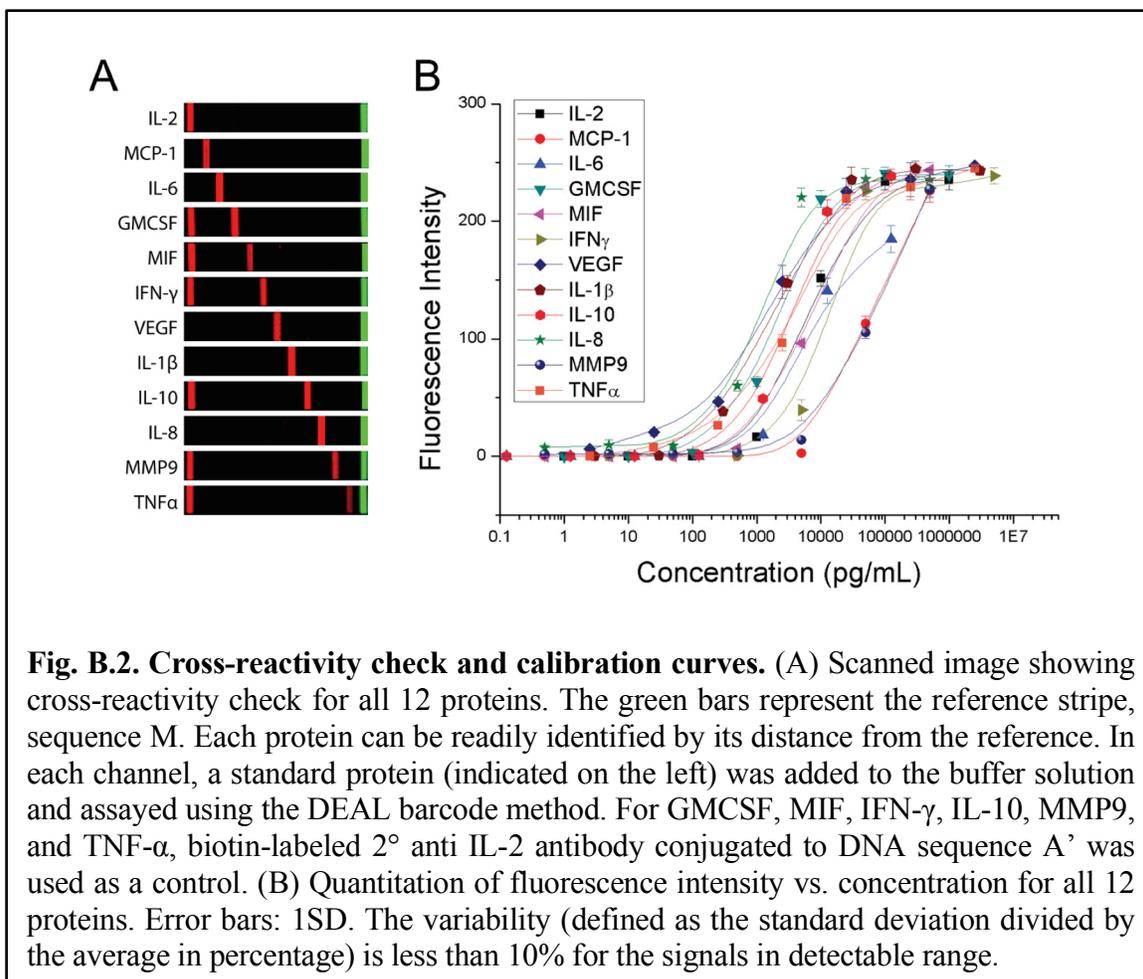


Table B.1. Sequences and terminal functionalization of oligonucleotides*.

Name	Sequence
A	5'- AAA AAA AAA AAA AGT CCT CGC TTC GTC TAT GAG-3'
A'	5' NH3-AAA AAA AAA ACT CAT AGA CGA AGC GAG GAC-3'
B	5'-AAA AAA AAA AAA AGC CTC ATT GAA TCA TGC CTA -3'
B'	5' NH3-AAA AAA AAA ATA GGC ATG ATT CAA TGA GGC -3'
C	5'- AAA AAA AAA AAA AGC ACT CGT CTA CTA TCG CTA -3'
C'	5' NH3-AAA AAA AAA ATA GCG ATA GTA GAC GAG TGC -3'
D	5'-AAA AAA AAA AAA AAT GGT CGA GAT GTC AGA GTA -3'
D'	5' NH3-AAA AAA AAA ATA CTC TGA CAT CTC GAC CAT -3'
E	5'-AAA AAA AAA AAA AAT GTG AAG TGG CAG TAT CTA -3'
E'	5' NH3-AAA AAA AAA ATA GAT ACT GCC ACT TCA CAT -3'
F	5'-AAA AAA AAA AAA AAT CAG GTA AGG TTC ACG GTA -3'
F'	5' NH3-AAA AAA AAA ATA CCG TGA ACC TTA CCT GAT -3'
G	5'-AAA AAA AAA AGA GTA GCC TTC CCG AGC ATT-3'
G'	5' NH3-AAA AAA AAA AAA TGC TCG GGA AGG CTA CTC-3'
H	5'-AAA AAA AAA AAT TGA CCA AAC TGC GGT GCG-3'
H'	5' NH3-AAA AAA AAA ACG CAC CGC AGT TTG GTC AAT-3'
I	5'-AAA AAA AAA ATG CCC TAT TGT TGC GTC GGA-3'
I'	5' NH3-AAA AAA AAA ATC CGA CGC AAC AAT AGG GCA-3'
J	5'-AAA AAA AAA ATC TTC TAG TTG TCG AGC AGG-3'
J'	5' NH3-AAA AAA AAA ACC TGC TCG ACA ACT AGA AGA-3'
K	5'-AAA AAA AAA ATA ATC TAA TTC TGG TCG CGG-3'
K'	5' NH3-AAA AAA AAA ACC GCG ACC AGA ATT AGA TTA-3'
L	5'-AAA AAA AAA AGT GAT TAA GTC TGC TTC GGC-3'
L'	5' NH3-AAA AAA AAA AGC CGA AGC AGA CTT AAT CAC-3'
M	5'-AAA AAA AAA AGT CGA GGA TTC TGA ACC TGT-3'
M'	5' Cy3-AAA AAA AAA AAC AGG TTC AGA ATC CTC GAC-3'

* All oligonucleotides were synthesized by Integrated DNA Technology (IDT) and purified via high-performance liquid chromatography (HPLC).

Table B.2. Summary of antibodies used for macrophage experiments

DNA label	primary antibody (vendor)	secondary antibody (vendor)
A'	mouse anti-hu IL-2 (BD Biosciences)	biotin-labeled mouse anti-hu IL-2 (BD Biosciences)
B'	mouse anti-hu MCP-1 (eBioscience)	biotin-labeled armenian hamster anti-hu MCP-1 (eBioscience)
C'	rat anti-hu IL-6 (eBioscience)	biotin-labeled rat anti-hu IL-6 (eBioscience)
D'	rat anti-hu GMCSF (Biolegend)	biotin-labeled rat anti-hu GMCSF (Biolegend)
E'	goat anti-hu MIF(R&D systems)	biotin-labeled goat anti-hu MIF(R&D systems)
F'	mouse anti-hu IFN- γ (eBioscience)	biotin-labeled mouse anti-hu IFN- γ (eBioscience)
G'	mouse anti-hu VEGF (R&D systems)	biotin-labeled goat anti-hu VEGF (R&D systems)
H'	mouse anti-hu IL-1 β (eBioscience)	biotin-labeled mouse anti-hu IL-1 β (eBioscience)
I'	rat anti-hu IL-10 (eBioscience)	biotin-labeled rat anti-hu IL-10 (eBioscience)
J'	mouse anti-hu IL-8 (R&D systems)	biotin-labeled mouse anti-hu IL-8 (R&D systems)
K'	mouse anti-hu MMP9 (R&D systems)	biotin-labeled goat anti-hu MMP9 (R&D systems)
L'	mouse anti-hu TNF- α (eBioscience)	biotin-labeled mouse anti-hu TNF- α (eBioscience)

B.2.3. Culture and stimulation of THP-1 cells.

We cultured human monocyte THP-1 cells (clone TIB 202) in RPMI-1640 (ATCC) medium supplemented with 10% fetal bovine serum and 10 μ M 2-mercaptoethanol. Cells grown close to the maximum density (0.8×10^6 cells/mL) were chosen for the experiment. Cells were first treated with 100 ng/mL phorbol 12-myristate 13-acetate (PMA) for 12 hours during which a characteristic morphological change was noticed as an indication of the induction to the macrophages. Cells were washed with fresh media and re-suspended in media with PMA (100 ng/mL) and lipopolysaccharide (LPS, 200 ng/mL) at 0.5×10^6 cells/mL for the further differentiation and the TLR-4 activation.

B.2.4. On-chip secretion profiling

Prior to loading cells on chip, the DNA barcode array was transformed into an antibody microarray through the following steps. First, 1% bovine serum albumin (BSA) in phosphate buffered saline (PBS) was flowed and dead-end filled into the chip to block non-specific binding. Second, a 200 ml cocktail containing all 12 DNA-antibody conjugates at 1.25 μ g/mL in 1% BSA/PBS buffer was flowed through all microfluidic channels for a period of 1 h. Then, 100 ml of fresh buffer was flowed into the device to replace DNA conjugated primary antibody solutions. The chip is then ready for use. Cells stimulated

with PMA/LPS were loaded into the SCBC chip within 10 min in order to minimize pre-loading secretion. Then, the pneumatic valves were pressed down by applying 15-20 psi constant pressure to divide 80 microfluidic channels into 960 isolated microchambers. Next, the cells in every microchamber were imaged under a Nikon LV100 microscope and their numbers were counted. Afterwards the chip was placed in a cell incubator ($\sim 37^{\circ}\text{C}$ and 5% CO_2) for 24 hours to perform on chip secretion. The chip was removed from the incubator and a 200 μl cocktail containing all detection antibodies (each at 0.5 $\mu\text{g}/\text{mL}$ concentration) tagged with biotin flowed through the microchannels by releasing the valves. Then, 200 μl of the fluorescent probe solution (1 $\mu\text{g}/\text{ml}$ Cy5-labeled streptavidin and 25 nM Cy3-labeled M' ssDNA) was flowed through to complete the immunosandwich assay. Finally, the PDMS slab was peeled off and the microarray slide was rinsed with 1 \times PBS, 0.5 \times PBS and DI water twice, sequentially, and spin-dried.

B.2.5. Bulk secretion profiling

Bulk measurements on the same panel of secreted proteins as were assessed within the SCBC microchambers were also carried out for the THP-1 cells with no stimulation, PMA stimulation, and PMA+LPS stimulation. Cells were cultured at 0.3×10^6 cells/mL, a comparable density to a single-cell in a chamber. The media were collected after 24 hours and the secreted proteins were detected as described below. For the PMA+LPS stimulation condition, the media were collected at multiple time points (2, 4, 6, 8, and 10 hours) for the time-dependent analysis as well. For the bulk test, SCBC chip was utilized without using valves for the microchannel to microchamber conversion. The same conditions as for the on-chip secretion profiling were applied except for the cell incubation step. Instead, the collected media was introduced to the channel sets and incubated for 3 hours in the incubator.

B.2.6. Quantification and statistics.

All the barcode array slides used for quantification were scanned using an Axon GenePix 4400a two-color laser microarray scanner at the same instrumental settings—50% and 15%

for the laser power of 635 nm and 532 nm, respectively. Optical gains are 500 and 450 for 635 nm and 532 nm fluorescence signals, respectively. The brightness and contrast were set at 90 and 93. The averaged fluorescence intensities for all barcodes in each chamber were obtained and matched to the cell number by custom-developed MATLAB (the mathworks, Natick, MA) codes. Heat maps were generated using cluster 3.0 and java *treeview* (<http://rana.lbl.gov/eisensoftware.htm>).

B.2.7. Data Analysis: Conversion to the number of molecules

The collected raw data is based on the fluorescence. In order to convert the fluorescence to the number of protein molecules, we used the calibration curves (**Fig. 3.2.**). We used the four parameter logistic model which is commonly used for fitting ELISA calibration curve. The fitting parameters can be found from the **Table 3.3.**

$$y = \frac{A_1 - A_2}{1 + (x / x_0)^p} + A_2$$

Table 3.3. Parameters utilized for the protein assay calibration curve

	A1		A2		x0		p		Statistics	
	Value	Error	Value	Error	Value	Error	Value	Error	Reduced Chi-Sqr	Adj. R-Square
IL-2	0	0	256	0	7659.58168	973.0838	1.12824	0.16788	91.39131	0.99224
MCP-1	0	0	256	0	65733.51686	4770.5	1.12607	0.09607	29.62623	0.99578
IL-6	0	0	256	0	16231.59942	4515.94	0.67887	0.12265	243.09932	0.95697
GMCSF	0	0	256	0	2451.99685	295.3281	1.2195	0.13013	72.59138	0.99458
MIF	0	0	256	0	7892.74068	483.8218	1.14428	0.07578	20.31714	0.99821
IFN- γ	0	0	256	0	14549.5316	2773.804	1.57222	0.26181	172.2368	0.98713
VEGF	0	0	256	0	1687.9445	225.4782	0.69008	0.05631	58.49911	0.99513
IL-1 β	0	0	256	0	2137.44388	208.9672	0.89593	0.07185	41.21361	0.99694
IL-10	0	0	256	0	3961.03661	328.4038	1.23209	0.08611	33.93572	0.99669
IL-8	0	0	256	0	1255.89317	225.9207	1.23262	0.19534	161.8703	0.98686
MMP9	0	0	256	0	70537.40022	1584.696	1.062	0.02495	2.60945	0.99961
TNF- α	0	0	256	0	4126.15703	661.2747	0.81683	0.09483	99.72583	0.99185

B.2.8. Signal-to-noise calculations

Since the signal range highly depends on the activities of the antibodies as well as the cell biology, it is required to decide if the signal is real and reliable. Certain assayed

proteins were identified as positively detected from single-cells based upon signal-to-noise ratio (S/N), which was measured as follows: For each microchamber, the averaged fluorescence from the two barcode stripes used to capture and detect a given protein and the averaged fluorescence from the barcode stripes designed to capture and detect IL-2 were obtained. The ratio of the averaged values over all single-cell experiments (specific protein to IL-2) yields a S/N value. An S/N of 4 was utilized as a minimum for positive detection. Eight secreted proteins were thus identified from the single-cell measurements. Those proteins were (with S/N included in the parenthesis after the protein name): MCP-1 (4.65), MIF (1381.13), IFN- γ (4.33), VEGF (77.32), IL-1 β (94.70), IL-8 (2622.40), MMP9 (119.50), and TNF- α (410.74).

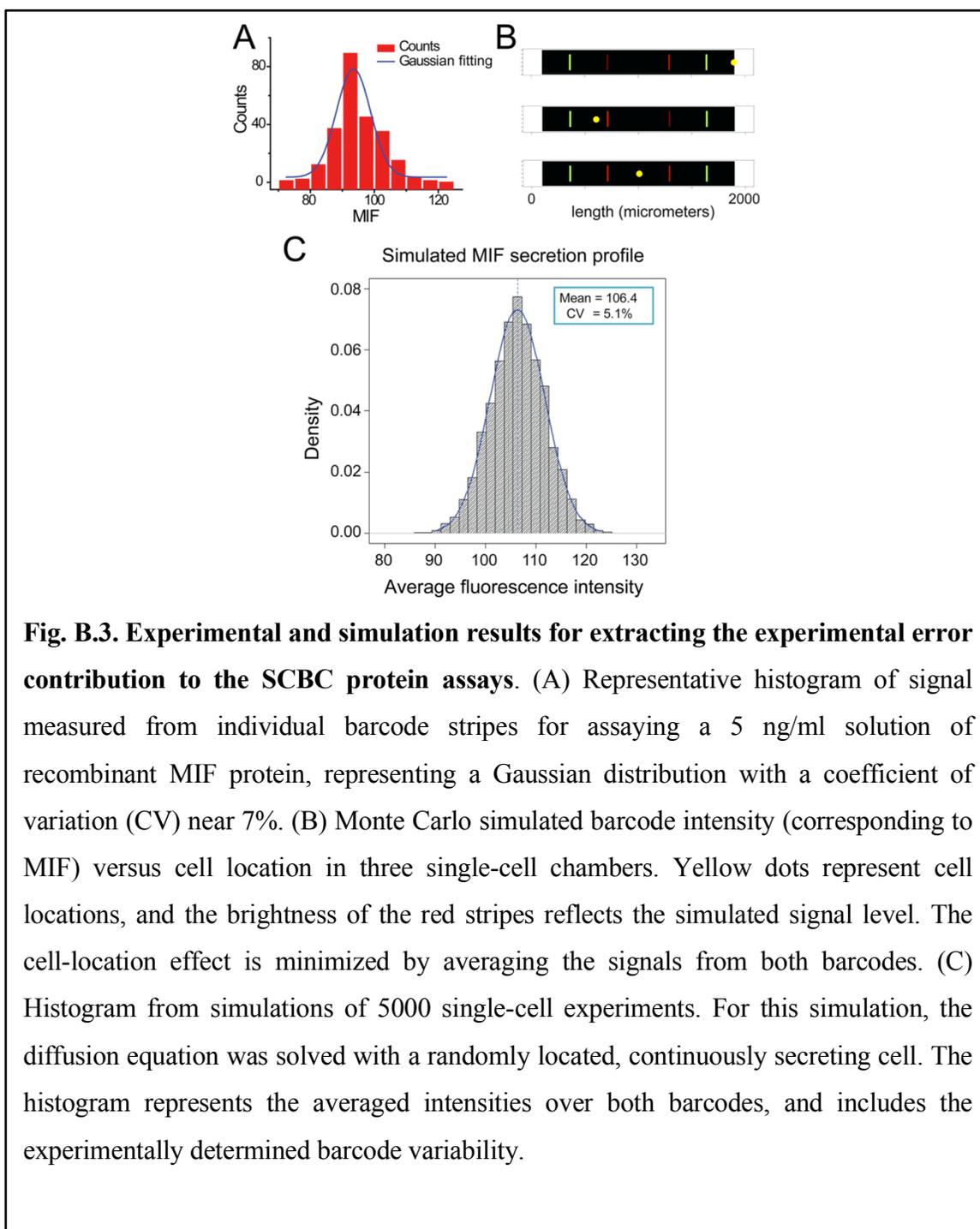
B.2.9 Analysis of experimental and biological variation from SCBC-based single-cell measurement

One of the major characteristics of SCBC analysis is the heterogeneous cellular behavior at single-cell level. The experimental variation of the SCBC platform which reflects the system error as well as the biological variation due to the cellular heterogeneity is contributing to the fluctuation of the total signal. Thus, we need to check if the heterogeneous signal responses are from the cells or the device itself.

The experimental error mainly includes the variation from non-uniform DNA barcode patterns 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 the 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 reflects the detection profile of the DNA barcode.

Fig. B.3.A shows a representative histogram of signal derived from recombinant MIF protein at 5 ng/ml. The histogram shows a nice Gaussian distribution with a coefficient of variation (CV) around 7%. In the calibration experiment, basically the intensities of all the recombinant proteins at detectable concentrations follow a Gaussian distribution with CVs typically lower than 10%.

The cell location is another important factor for the system error. Even though the chamber size is small, it is still big for a single-cell. So the protein signal is dependent on diffusion and that is why the cell location can be a source of the variation. 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. However, the barcode close to the cell will undergo a higher local protein concentration than its counterpart and the different intensities of two sets of barcodes are amplified during the long incubation time. The diffusion process will lead the system close to the equilibrium but the cell that keeps secreting proteins with different kinetics makes it difficult for the chamber to reach its full equilibrium. In that sense, the randomly located cells can add an extra uncertainty to the SCBC system.



Because it is difficult to isolate the system error (especially for the cell-location effect) from the heterogeneous cell response experimentally, we performed a Monte Carlo simulation by R (R Foundation for Statistical Computing, version 2.10.1). First of all, we investigate the case of MIF as a representative case. We assumed one chamber has

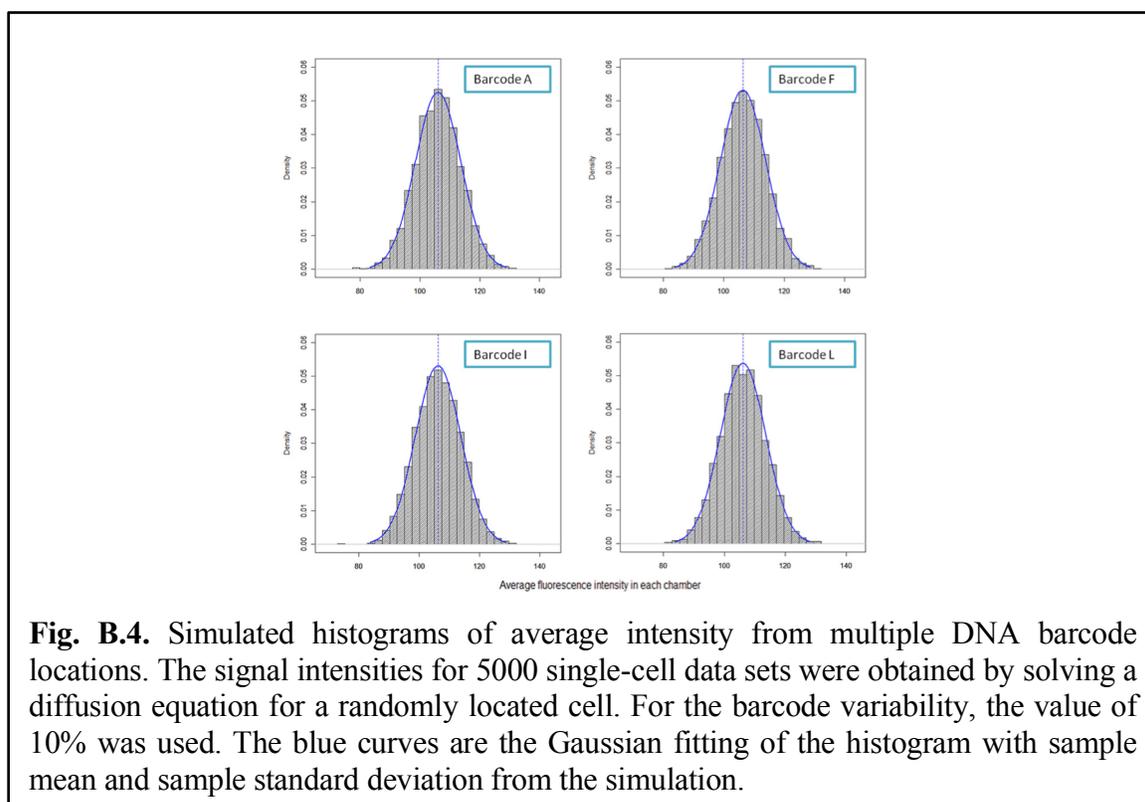
two sets of 13 barcodes such that all of them have MIF antibodies. By randomly positioning a cell with a fixed protein secretion rate and getting the protein concentration at specific barcode positions, we can find out what is the variation that purely depends on the cell location and barcode non-uniformity. The total amount of secreted MIF during 24 hours was estimated based on our experimental result. The secretion rate was 4.84 pg/mL per min from the SCBC (used for the simulation) and 11 pg/mL per min from the bulk condition. The corresponding secretion rate of a single-cell, back-calculated based on the chamber and cell size ($10\mu\text{m}^3$), was 0.065 nM/min. Values of parameters used in simulation can be found from **Table B.4**. 5000 data sets for the protein concentration distributions from randomly located single-cell were generated by solving a diffusion equation with a custom made MATLAB code and the results were analyzed with R. The parameters used in the simulation are exactly the same as our experimental environment. The chamber is 2000 μm in length and 100 μm in width with two sets of DNA barcodes M-A and A-M from left to right. Each barcode is 20 μm in width with 50 μm in pitch (30 μm gap between barcodes). The detection variation of the MIF protein due to the DNA uniformity obtained from the histogram of the calibration data set was incorporated to the analysis. **Fig. B.3.C** shows the histogram of the average fluorescence intensity from DNA sequence E (corresponding to MIF in the actual experiment) for 5000 single-cell cases. For the barcode variability, the actual value of 7.3% was used. The final system error was 5.1% which is a lot smaller than the assay error from the experimental data sets, 55.2 %.

Table B.4. Values of parameters used in simulation

Chamber size	2000 μm × 100 μm × 18 μm
Cell diameter	10 μm
Diffusion Coefficient(11)	10^{-6} cm^2/sec
Protein secretion rate (MIF)	0.065 nM/min
Molecular weight	12500 Da

In order to think of the worst case, we used the barcode variability of 10% for the rest of the analysis. If the cell-location effect is significant, we are supposed to see different

errors on different barcode positions. **Fig. B.4.** illustrate the histograms of average intensities from multiple barcode locations. The blue curves are line profiles of Gaussian distribution fitted with the mean and the standard deviation obtained from the corresponding simulation. The nice fitting between the Gaussian curves and the histogram indicates that the average intensity per chamber follows a Gaussian distribution with a predictable mean and CV. The CVs from this simulation represent the distribution of our measurements for single-cell chambers without considering the cellular heterogeneity, i.e. the system error. The experimental CVs for different barcode locations based on the system error were quite similar to one another ($\sim 7\%$).



We can define CV_{system} as the system error estimated by the simulation. We can also calculate the assay error from our experimental data set such that CV_{assay} refers to the total CV of our experimental data. Consequently, the biological variation for single-cell experiment can be quantitatively estimated by the formula below:

$$CV_{\text{assay}} = (CV_{\text{system}}^2 + CV_{\text{biological}}^2)^{1/2}$$

An estimation of biological variations of proteins for different barcode locations are shown in **Table B.5**. It can be noticed that the biological variation is dominant in the total error of the assay. This analysis verifies that the signal fluctuation that we can see from the single-cell experiment is a good representation for the single-cell heterogeneity rather than the systemic error from our platform.

Table B.5. The coefficients of variation for each of the assayed proteins from single-cell experiments. The experimental CVs are estimated from the Monte Carlo simulations. The biological CVs, which clearly dominate the experiment, are calculated from $CV_{\text{assay}} = (CV_{\text{system}}^2 + CV_{\text{biological}}^2)^{1/2}$.

Barcode/Protein	Experimental CV (%)	Assay CV (%)	Biological CV (%)
B / MCP-1	7.12	380.4	380.3
E / MIF	7.05	55.2	54.7
F / IFN- γ	7.04	131.5	131.3
G / VEGF	7.03	149.7	149.5
H / IL-1 β	7.02	300.6	300.5
J / IL-8	7.00	14.4	12.6
K / MMP9	6.98	192.6	192.5
L / TNF- α	6.97	132.9	132.7

B.2.10 Signal-to-noise calculations and experimental error

An Axon GenePix 4400A scanner coupled with a custom algorithm was used to quantify the fluorescence intensities of each protein from each microchamber (Fig 1B). Certain proteins were positively detected based upon signal-to-noise (S/N) > 4. S/N was calculated as follows. Each protein was measured twice per microchamber. The averaged fluorescence values from the two barcode stripes for all proteins were used as signals from each chamber. The ratio of the averaged signal over all single-cell experiments for a specific protein to IL-2 yields a S/N. The following eight proteins were detected (S/N is indicated after the protein name): MCP-1 (4.7), MIF (1380), IFN-g (4.3), VEGF (77), IL-1b (95), IL-8 (2620), MMP9 (120), and TNF-a (411).

Macrophages are highly responsive to their environment, and so experimental conditions can influence macrophage behavior. Thus, we sought confirmation that our protocols could lead to reproducible results. We executed identical sets of experiments on different SCBCs, and showed that the distributions of the unambiguously detected proteins

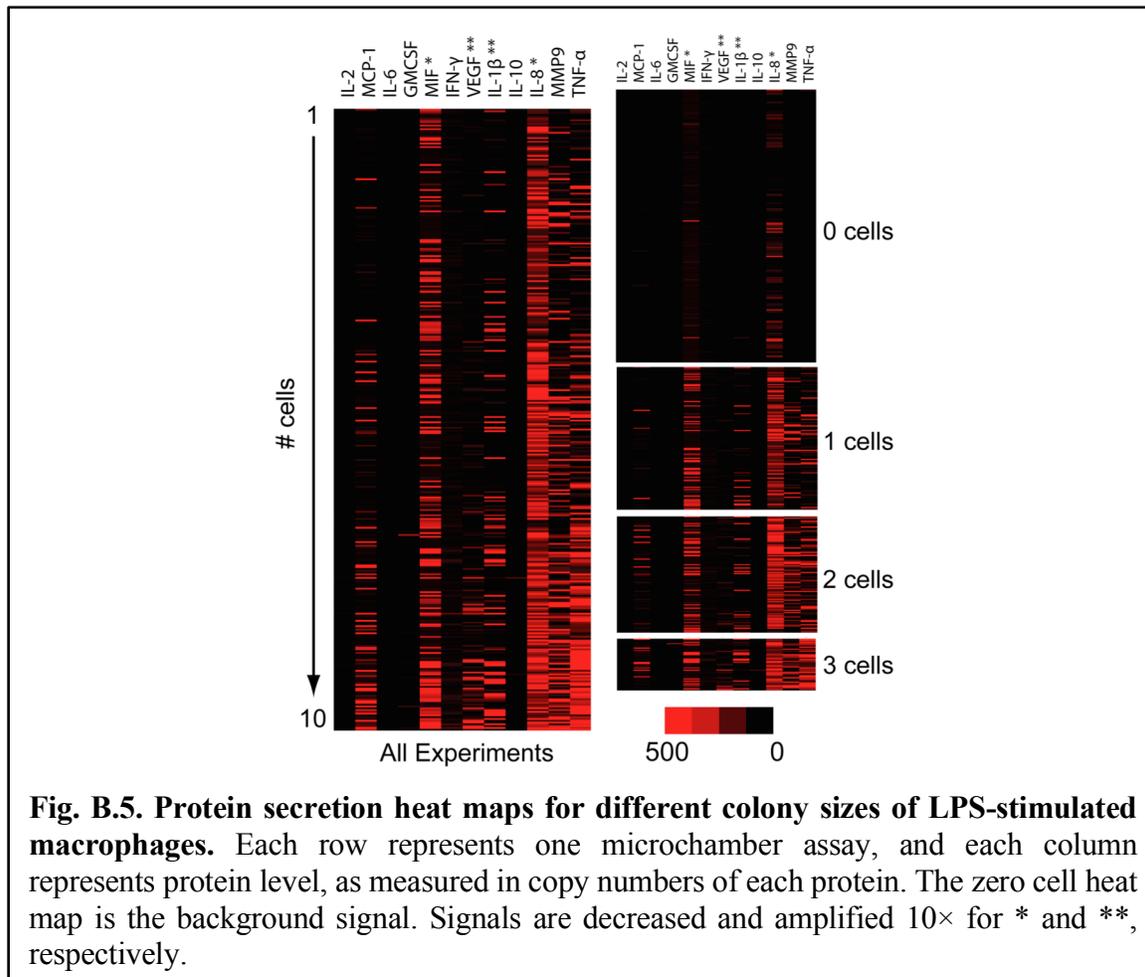
were effectively identical (p-value > 0.25). The results presented here do depend on the amount of PMA or LPS used and, to a lesser extent, the passage number of THP-1 cells. In addition, a solvent extraction of the PDMS improves the SCBC biocompatibility and the assay reproducibility¹⁵.

Levels of proteins secreted from single-cells can exhibit a variability that reflects the stochastic nature of biology¹⁷ and, in fact, represents the biological fluctuations. The SCBC experimental error must be compared against the measured variations for extracting the true macrophage fluctuations. One contribution to the experimental error arises from the variability of the flow-patterned antibody barcodes. We characterized that variability via protein assays executed within a complex biological environment (serum), and within the microchambers of an SCBC, but using cocktails spiked with known quantities of standard proteins. In both cases, we found a variability of < 10%¹⁸ and **Fig. B.2.**), depending upon the protein. Averaging the two identical protein assays per microchamber lowers the variability within a microchamber by a factor of $2^{1/2}$. A second experimental error arises from the competition between protein capture by surface-bound antibody, and protein diffusion. When a cell is proximal to a barcode, that barcode may exhibit a higher signal intensity than the more distant barcode. A Monte Carlo calculation allowed for an estimation of the total system error by simulating the location-dependent experimental variation. Using MIF as a representative protein for the simulation (it has a barcode variability of 7.3%; **Fig. B.3.A**) the experimental error of the system is estimated to be 5.1% (**Fig. B.3.B, C**). For the worst case of a 10% barcode variability, the total experimental error is estimated to be ~7% (**Table B.5.** and **Fig. B.4.**). Based upon these results, we can calculate the biological coefficient of variation ($CV_{\text{biological}}$) from $CV_{\text{assay}} = (CV_{\text{system}}^2 + CV_{\text{biological}}^2)^{1/2}$, where CV_{assay} is the measured spread in secretion levels for a given protein across all measurements for a given number of cells. For IL-8, the biological CV was only ~2-fold larger than the experimental CV, but for the other 7 detected proteins, the biological CV was 7-50× larger than the experimental CV (**Table B.5.**). Thus, the fluctuation extracted from our single-cell experiments reflects the cellular behaviors.

The individual protein assays were evaluated for cross-reactivity and calibrated using standard proteins (**Fig. B.2.**). Calibration curves were fitted by a four parameter logistic

model¹⁹. The SCBC assay sensitivities are comparable to commercial ELISAs (e.g., a few measured limits-of-detection are MIF ~100 pg/ml, IL-8 ~50 pg/ml, IL-1b~20 pg/ml, and VEGF ~2.5 pg/ml), with each exhibiting a $\sim 10^3$ linear detection range. The SCBC barcode assay results can be translated into numbers of detected molecules using the molecular weight of the standard proteins and the microchamber volume (**Fig. B.2.**). This quantitative representation of the data is used for the calculations described below. However, the standard proteins may differ from the proteins secreted by the macrophages (for example, glycosylation patterns may vary). Such variations can translate into differences in molecular weight, as well as differences in assay sensitivity.

The experimental results, presented as the number of cells per experiment, are shown in the heat maps of **Fig. B.5.**



B.3. THEORETICAL METHODS

B.3.1. The fluctuations in the secretome

The calibrated experimental data can be organized into digital tables of twelve columns, each representing a different protein, with different tables representing different numbers of cells in the microchamber. For a given table, each row represents the copy numbers of the twelve proteins for a single-cell, or small cell colony. For a given table, if the number of measurements is large enough, we can bin the data for each individual protein into a histogram with each bin representing a defined range of measured levels (**Fig. B.6**). With even more measurements one could generate joint distributions between two proteins, etc. However, we first confine our attention to the individual protein histograms because they provide a natural meeting place for experiment and theory. The theoretical prediction is made by seeking that distribution of copy numbers that is of maximal entropy, meaning that the distribution is as uniform as possible subject to a given mean number of copies²⁰⁻²³. As described in detail in the Supplement, We use the distribution of maximal physical entropy. This means that at the very global maximum of the entropy, the probabilities of the different proteins are not equal. Rather, as in any multi-component system at thermal equilibrium, each protein will be present in proportion to its partition function²⁴ where the partition function is the effective thermodynamic weight of a species at thermal equilibrium. We show below that in our system there is a network structure that imposes (at least) two overriding constraints that preclude the system from being in thermal equilibrium.

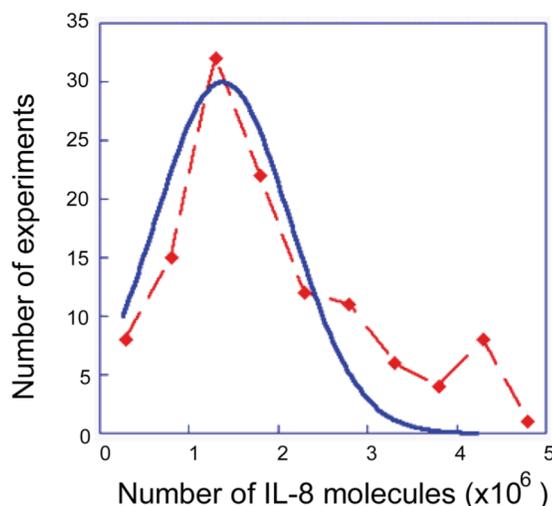


Fig. B.6. Fluctuations in the numbers of secreted IL-8 proteins, for all single-cell experiments. The fit to the theoretical distribution is shown as the continuous curve. Even for one cell there can be deviations from the bell-shaped theoretical functional form in the higher tail of the histogram due to autocrine signaling.

B.3.2. Theoretical approach

The essence of our approach is to regard the system, a single-cell (or a small colony), as not being in an equilibrium state because it is under the action of constraints. When the constraints are present the system is in that state of equilibrium that is possible under the constraints. This allows us to derive a quantitative version of the principle of Le Chatelier. Thereby we can quantitatively predict the response of the system to a (small) perturbation. Early on mathematical biologists expressed caution about the application of the Le Chatelier's principle to biological systems²⁵. It is possible to directly use the measured experimental results to validate our point of view. The qualitative reasoning is straightforward and so we give it here. It is valid to apply the principle of Le Chatelier when the system is in a stable equilibrium. When is the system in a stable equilibrium? – when, under a small perturbation, it returns to its equilibrium state. Here we simply state that if the observed fluctuations in protein copy number are about a stable state then we can apply the principle of Le Chatelier. The stability of the state is decided by the experimental measurements. Both the notion of stability and the response to perturbations, as quantified in the principle of Le Chatelier, require that the departure from equilibrium be small. Neither textbook equilibrium thermodynamics applied to a macroscopic system nor the

extended theory used here to describe one or a few cells implies that under a ‘large’ perturbation it should be possible to displace a cell to a new stable state that is distinct from its unperturbed state. For a single-cell or small cell colony, the experiments reveal that cell-cell perturbations are indeed small. For larger cell colonies the statistics are not secure enough to make a clear-cut statement. We have, however, numerical indications that the unperturbed state of the single-cell is possibly unstable in the presence of many other cells.

B.3.3. Theory of fluctuations

We begin by considering a compartment containing a single-cell secreting different proteins. For different compartments the experiment shows a possibly different number of secreted proteins of a given type. We denote the experimentally measured copy number of protein i in a given microchamber by N_i . We impose the constraints that the distribution for each protein is characterized by the mean number of its molecules. Then the distribution, $P(N_i)$ of copy number fluctuations of a protein i that is of maximal physical entropy (= the distribution at physical equilibrium subject to constraints), is derived in Supplement, Eq. S2. It is a bell-shaped function of N_i with a single maximum.

In seeking the maximum of the entropy we require that the energy is conserved. This constraint is imposed by the method discussed in Supplement. This method introduces parameters into the distribution. b is determined by the constraint of conservation of 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 analogs of the chemical potentials as introduced in the thermodynamics of systems of more than one component. Here, however, we are dealing with many replicas of a single-cell isolated within a microchamber. Even though we deal with just a single-cell, the μ_i ’s will be shown in Eq. 1 below to also play the role of potentials. This means, for example, that the mean copy number \bar{N}_i of protein i increases when its potential μ_i is increased. The mean number, $\bar{N}_i = \sum_i N_i P(N_i)$, is the average computed over the distribution. In operational terms this is an average computed over the different microchamber assays of protein i . We take it that the copy number distribution is normalized meaning that $\sum_i P(N_i) = 1$.

We next discuss the effect of perturbations on the distribution for a single-cell in the compartment. The regime of small perturbations is one in which the distribution, although perhaps distorted from a simple bell-shaped curve, still exhibits only a single maximum. The signature of large perturbations is that secondary maxima appear. When these become dominant a new state of the cell is prevailing.

To theoretically characterize the response of the cellular secretion to a perturbation we compute first the change in the distribution for the special case in which a perturbation changes the potential of protein i from μ_i to $\mu_i + \delta\mu_i$, where $\delta\mu_i$ is a small increment. We show (Eq. A2 in 3.6.3) that, to first-order in the change of the potential, the distribution changes by $\delta P(N_i) = \beta(\bar{N}_i - N_i)P(N_i)\delta\mu_i$. The result for δP has two immediate implications. One is that a perturbation will distort the shape of the distribution of the copy numbers of a given protein. Specifically, the change is proportional to the unperturbed distribution but its magnitude is weighted by the factor $(\bar{N}_i - N_i)$ so as to favor higher values of protein numbers. Thus, it is the high-end tail of the distribution that is most strongly influenced by the perturbation (see **Fig. B.6.**, for example).

The other immediate implication of the change in the distribution is that the mean values will change. Specifically the updated mean value of the copy number of protein i when we change from μ_i to $\mu_i + \delta\mu_i$ is $\bar{N}_i + \delta\bar{N}_i = \sum_i N_i [P(N_i) + \delta P(N_i)]$. A technical point is that because the distribution needs to be normalized we must have $\sum_i \delta P(N_i) = 0$. Using the result above that the change $\delta P(N_i)$ in the distribution is proportional to the unperturbed distribution and the normalization we arrive at the explicit result for the change in the mean copy number under a small disturbance.

$$\delta\bar{N}_i = \sum_i N_i \delta P(N_i) = \beta\delta\mu_i \sum_i N_i (\bar{N}_i - N_i) P(N_i) = \beta\delta\mu_i \overline{(\bar{N}_i - N_i)^2} \quad (1)$$

This equality states that 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 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 the variance of the distribution of fluctuations. Therefore, the lesser are the fluctuations (i.e., the narrower is the histogram), the more resilient to change is the distribution. As an example, IL-8 (**Fig. B.6.**) will be shown to be a very strongly coupled protein. IL-8 also has a particularly large variance as compared to the other proteins. Therefore there is some perturbation via autocrine signaling as seen in the hump in the higher tail of the histogram.

B.3.4. A quantitative Le Chatelier equation

With good measurement statistics one can examine the histogram for a joint distribution of two proteins and verify that pairs of proteins are correlated. Therefore the mean value (and other averages) of a protein i will change when protein j is perturbed. In the linear regime the result (see B.6.4.) is

$$\delta \bar{N}_i = \beta \sum_j \left[\overline{(\bar{N}_i - N_i)(\bar{N}_j - N_j)} \right] \delta \mu_j \quad (2)$$

where the covariance is computed over the unperturbed distribution. Eq. 2 is valid in the linear regime of small perturbations, and indicates that the contributions of different perturbations add up. The covariance matrix Σ , whose elements are $\Sigma_{ij} = \overline{(N_j - \bar{N}_j)(N_i - \bar{N}_i)}$, is what is called in matrix algebra a positive matrix²⁶. The implications of positivity are explored in B.6.5.

We prove in the B.6.4 that Eq. 2 is a quantitative statement of the principle of Le Chatelier in the meaning that a response to a perturbation changes the system in the direction of restoring a stable equilibrium. This is the analog of the observation that when we add energy (i.e., heat the system) the temperature goes up (rather than down). By equilibrium we mean a state of maximal entropy subject to the current value of all the constraints operating on the system. A system can therefore be maintained at equilibrium by imposing constraints such as keeping a gas under higher pressure at a fraction of the available volume of a cylinder. When these constraints are changed the system can move to a new equilibrium.

The covariance matrix is used in statistics as input in such methods of data analysis as principal component analysis^{27,28}. We emphasize that for us the covariance matrix is derived by physical considerations leading to Eq. 2. We can thereby state that Σ_{ij} is quantitatively the change in the number of copies of protein i when protein j is perturbed. Note that while the covariance is a positive matrix, individual off-diagonal elements can be negative signifying inhibition. The covariance matrix in digital form is provided in **Table B.6**.

Table B.6. Digital representation of the covariance matrix for 1 cell measurements

COV	IL-2	MCP-1	IL-6	GMCSF	MIF	IFN- γ	VEGF	IL-1 β	IL-10	IL-8	MMP9	TNF- α
IL-2	1.18E+06	-27830	7966.7	1.34E+05	2.30E+08	4.10E+06	85756	1.27E+06	1.60E+05	-1.06E+08	1.50E+07	2.31E+07
MCP-1	-27830	6.34E+09	-1.83E+05	-1.14E+06	-9.54E+09	-9.51E+07	6.10E+05	-3.31E+07	-2.52E+06	-8.51E+09	4.21E+09	-7.34E+08
IL-6	7966.7	-1.83E+05	9050.5	11507	1.56E+07	5.89E+05	734.25	4.20E+05	24714	-81623	-7.38E+05	1.04E+06
GMCSF	1.34E+05	-1.14E+06	11507	3.39E+05	3.75E+08	3.97E+06	53462	5.71E+06	2.07E+05	1.05E+07	-1.66E+07	1.75E+06
MIF	2.30E+08	-9.54E+09	1.56E+07	3.75E+08	3.12E+12	1.48E+10	-4.32E+08	1.19E+10	3.35E+08	-4.33E+11	-7.53E+10	-5.22E+10
IFN- γ	4.10E+06	-9.51E+07	5.89E+05	3.97E+06	1.48E+10	3.09E+08	-2.28E+05	1.40E+08	5.83E+06	-1.70E+09	-5.01E+08	-1.00E+08
VEGF	85756	6.10E+05	734.25	53462	-4.32E+08	-2.28E+05	2.48E+06	-2.65E+06	1.04E+05	9.34E+07	2.92E+07	2.59E+07
IL-1 β	1.27E+06	-3.31E+07	4.20E+05	5.71E+06	1.19E+10	1.40E+08	-2.65E+06	4.78E+08	4.46E+06	1.86E+09	-4.08E+08	3.16E+08
IL-10	1.60E+05	-2.52E+06	24714	2.07E+05	3.35E+08	5.83E+06	1.04E+05	4.46E+06	7.39E+05	2.20E+08	-5.64E+06	2.92E+07
IL-8	-1.06E+08	-8.51E+09	-81623	1.05E+07	-4.33E+11	-1.70E+09	9.34E+07	1.86E+09	2.20E+08	2.73E+12	7.05E+09	3.56E+10
MMP9	1.50E+07	4.21E+09	-7.38E+05	-1.66E+07	-7.53E+10	-5.01E+08	2.92E+07	-4.08E+08	-5.64E+06	7.05E+09	3.70E+10	5.25E+09
TNF- α	2.31E+07	-7.34E+08	1.04E+06	1.75E+06	-5.22E+10	-1.00E+08	2.59E+07	3.16E+08	2.92E+07	3.56E+10	5.25E+09	5.26E+10

To summarize, the result for the distribution of protein copy numbers for the strongly interacting protein IL-8 (**Fig. B.6.**) has just one maximum. The noticeable deviations in the tail of the distribution are likely due to autocrine signaling, because the correlation of IL-8 with itself is only comparable in magnitude to the correlation of MIF with itself. Those two correlations are larger than any other variance or covariance. As discussed below, IL-8 is also strongly correlated with other proteins. For $n \geq 3$ cells in the microchamber, there is numerical evidence for a second maximum in the distribution of IL-8 fluctuations. For other proteins, six or more cells per chamber are required before a second maximum is resolved.

We can draw two conclusions from the fit of **Fig. B.6.**, between observed fluctuations and the theoretical result. First, the experimental distribution has but one maximum, and so the state is stable. Second, the theory accounts for the shape of the experimental distribution. This implies that we have correctly identified the important constraints on the system. Therefore we have Eq. 1 for the change of the distribution and hence Eq. 2 as the quantitative statement of the Le Chatelier's theorem. If there are additional constraints one can still derive a quantitative Le Chatelier's theorem but there will be additional terms beyond those shown explicitly in Eq. 2. We reiterate that Eq. 2 is the covariance computed from the experiments for an unperturbed cell. In our work below we use Eq. 2 to predict the effect of perturbation.

B.4. Results and Discussion

B.4.1. Computing the covariance matrix

The single-cell data (the heat map of **Fig. B.5.**) can be regarded as a rectangular matrix \mathbf{X} where each row is a separate measurement and each column contains the copy number of a particular protein. For our convenience we mean center each column. When the number of measurements (= number of rows of \mathbf{X}) is not small (and is \geq than the number of columns) the covariance matrix can be immediately computed as $\Sigma_{ij} = \sum_{k=1}^K X_{ki} X_{kj} / K$ where k runs over all measurements, $k=1,2,\dots,K$. By construction of the matrix \mathbf{X} , the matrix element X_{ki} is the number recorded in the k 'th measurement for protein i minus the mean number \bar{N}_i for that protein. We divide $\mathbf{X}^T \mathbf{X}$ by the number, K , of measurements so that the covariance is the mean value. The covariance is a product of the measured numbers, so the coefficient of variation of the covariance is, for small variations, twice the coefficient of variation of the measurements. An upper estimate, see **Table B.5.** and **Fig. B.6.**, is 14% when the covariance is computed from the fluorescence intensities. The conversion from the fluorescence intensity to the number of molecules does not change the coefficient of variation when we are in the linear regime of the calibration curve, see **Fig. B.2.** However at very low or high intensities the calibration curve is non-linear, so that small changes in fluorescence intensity are amplified to larger differences in the number of molecules, and thus large values of the variance. Out of $K = 129$ single-cell

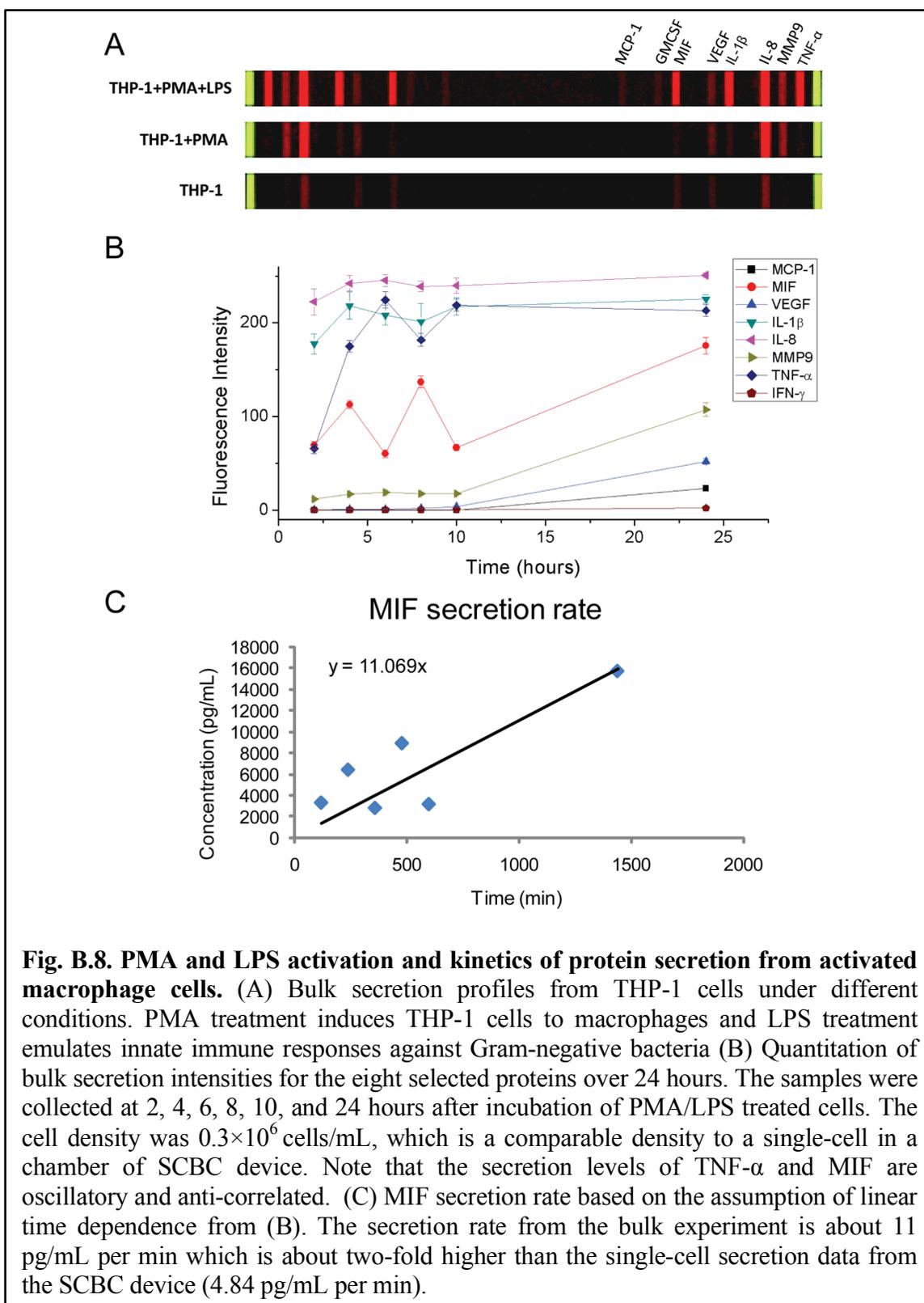
experiments, we therefore eliminated four outliers. These corresponded to one instance each for which the fluorescence levels of TNF- α , IL-1 β , MIF or IL-6 were very high. We thus used $K = 125$ values to compute the covariance matrix. The elimination of these four outliers brings the error of reading the number of molecules to be more comparable to the error in reading the fluorescence intensity.

B.4.2. The network

We analyze the covariance matrix in two stages. The first stage yields a quick (but correct and reliable) ‘global’ summary of the network, meaning which protein is coupled with which other proteins. There is finer structure, discussed below, that is not resolved in this first stage. To obtain the global network we begin by noting that the covariance matrix is symmetrical so that protein i is correlated with protein j just as much as protein j is correlated with protein i , $\Sigma_{ij} = \Sigma_{ji}$. This means that although both positive and inhibitory couplings can be extracted from the network, the direction of those coupling (i.e. protein i inhibits protein j , rather than vice-versa) is not resolved. The results for the overall network are shown in **Fig. B.7**. Panel A is the raw data for plotting the network and panel B is the network itself. The protein most strongly coupled to all others is MIF, and it is primarily anti-correlated with the other proteins. Next in strength of coupling is IL-8. Note that the symmetry between any two proteins is limited; proteins 1 and 2 may be coupled to each other, but protein 1 may be coupled to protein 3, while proteins 2 and 3 are uncorrelated. Mathematically this is possible because the total coupling strength of protein i , sum of Σ_{ij} over all j , can be quite different from the total coupling strength of protein j that is given as the sum of Σ_{ji} over all possible proteins i .

inhibition is indicated, as usual, by a bar at the end of the connector. The dashed line correlations of MIF with IFN-g is of magnitude 2×10^{10} , and so may be corrupted by noise. The dashed line correlations between MIF and both MCP-1 and IL-1b are even weaker (about 10^{10}). The more refined analysis presented in Fig. 6 shows, however, that these two correlations are likely real and above the noise level.

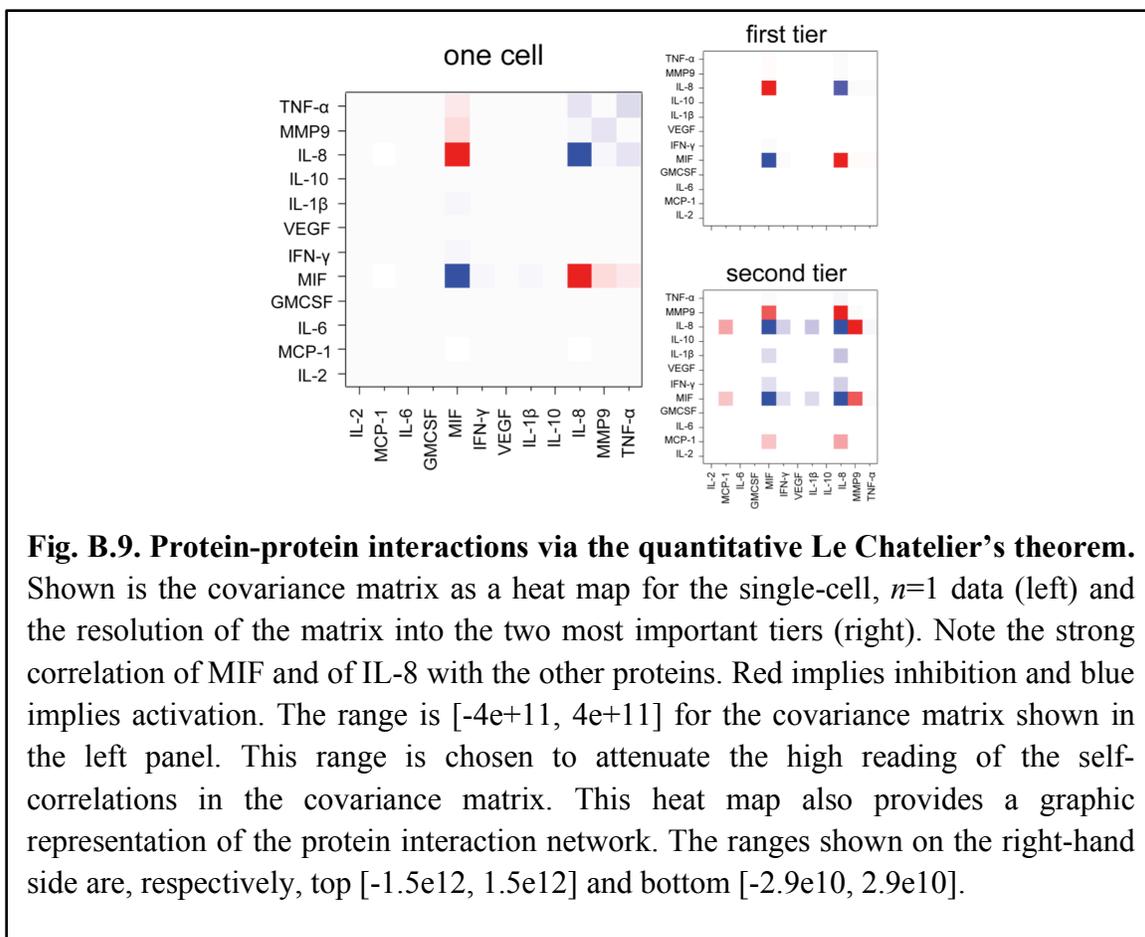
Macrophages are an important source of IL-8 and MIF²⁹⁻³¹, and IL-8 is secreted by the macrophages without LPS stimulation, while MIF is secreted upon LPS stimulation (**Fig. B.8.A**). Our derived network model indicates the MIF is inhibited by IL-8, and MIF, in turn, inhibits 3 other proteins, including TNF-a, while it promotes the production of IL-1b. These predictions are consistent with the time-dependent measurements of secreted proteins (**Fig B.8.B**). From those measurements, we find that the levels of three proteins (MIF, TNF-a, and IL-1b) that are secreted upon LPS stimulation, exhibit fluctuations over time. The MIF and TNF-a temporal fluctuations are anti-correlated, consistent with the network hypothesis. A detailed elucidation of the underlying mechanism for these dynamics will require additional experiments. However, it is encouraging that a network hypothesis derived from single-time-point, single-cell data does provide consistent insight into the dynamical responses of the macrophages to stimulation.



B.4.3. The composite networks

In the second stage in our analysis of the covariance matrix we aim to show a more resolved structure and thereby note features that are glossed over in the global network of **Fig. B.7.B**. We will show that there are several independent networks operating together to globally represent **Fig. B.7.B**. The detailed analysis also provides a more robust error estimate. To resolve independent inherent structures within the covariance matrix we consider what is known in matrix algebra as the spectral representation (See B.6.6. and B.6.7 for more details). Technically this is a ranking of the eigenvectors as also carried out in principal component analysis. We suggest, however, that for our system specifically this ranking allows an examination of tiers in the cell-cell signaling. The tiers are independent, meaning that they govern independent fluctuations. The proteins that are members of a given tier respond collectively to a perturbation.

The spectral theorem²⁶ allows us to rank the contributions according to the decreasing magnitude of the eigenvalues. At the bottom are the smallest eigenvalues and these are attributed to experimental noise rather than to real biological information. For the single-cell in the compartment we find, as expected for the linear regime, that the dominant eigenvectors are each localized around a particular protein. As shown in **Fig. B.9.**, the two largest are localized on MIF and IL-8. The leading eigenvalue = tier 1, is only about 30% bigger than the second one, $m = 2$. The third eigenvalues (not shown) is smaller by almost two orders of magnitude. **Fig. B.10.** is a plot on a logarithmic scale of all non-zero eigenvalues. There are only two eigenvectors that, judging by the value of their corresponding eigenvalues, are definitely above the noise.



In drawing **Fig. B.8.B** we could not state definitely that the correlations of MIF with IFN-g, MCP-1 and IL-1b, are above the noise level. The more refined spectral analysis shows that all these correlations are clearly evident in the second tier (**Fig. B.9.**) and so are secure. The **Fig. B.9.** results are the fluctuations measured for one cell experiments. See **Fig. B.11.** for similar results but for $n=3$ cells per microchamber.

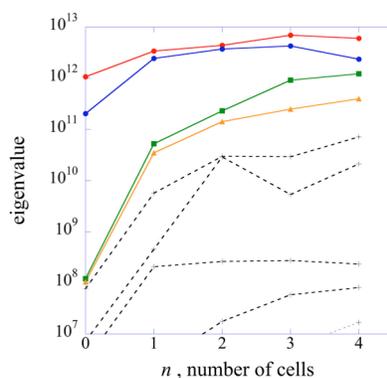


Fig. B.10. The dependence of the dominant eigenvalues of the covariance matrix on the number of cells in the sample. The result for $n = 0$, the background, is included to show the influence of the noise. The dashed lines, the fifth and higher eigenvalues are more corrupted by noise.

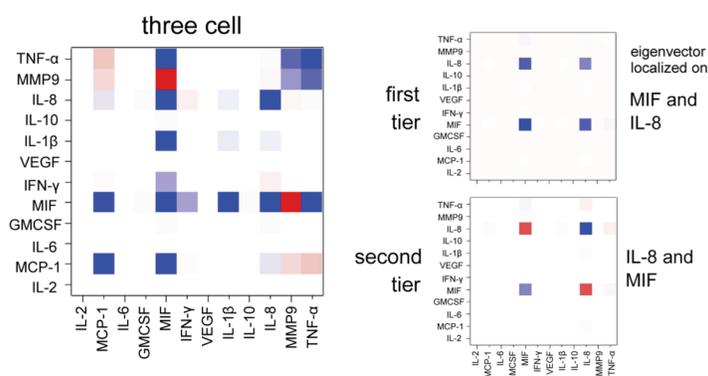


Fig. B.11. Heat map of the covariance matrix (left) and of the contributions to the first two tiers of the network (right) for measurements on chambers containing 3 cells. Similar to the single-cell case (Fig. B.9.), the entries in the tiers are scaled by the size of the eigenvalues. See the spectral representation of the covariance matrix, Eq. S11. The plot at left is the covariance matrix computed from the observed fluctuations in the 3-cell data. The color code is $-8e+10$ (red) to 0 (white) to $+8e+10$ (blue). The range is fixed so as to attenuate the effect of the self-terms in the covariance matrix. For tier 1 and tier 2, the ranges are $[-4.3e-12, 4.3e+12]$ and $[-7e+10, 7e+10]$, respectively. Note that when the numbers of cells per chamber increases, anti-correlations can get washed out.

B.4.4. The number-based network

The network presented in Fig. B.8. and Fig. B.9. is based upon experimental measurements in which raw fluorescence intensities are converted into numbers of

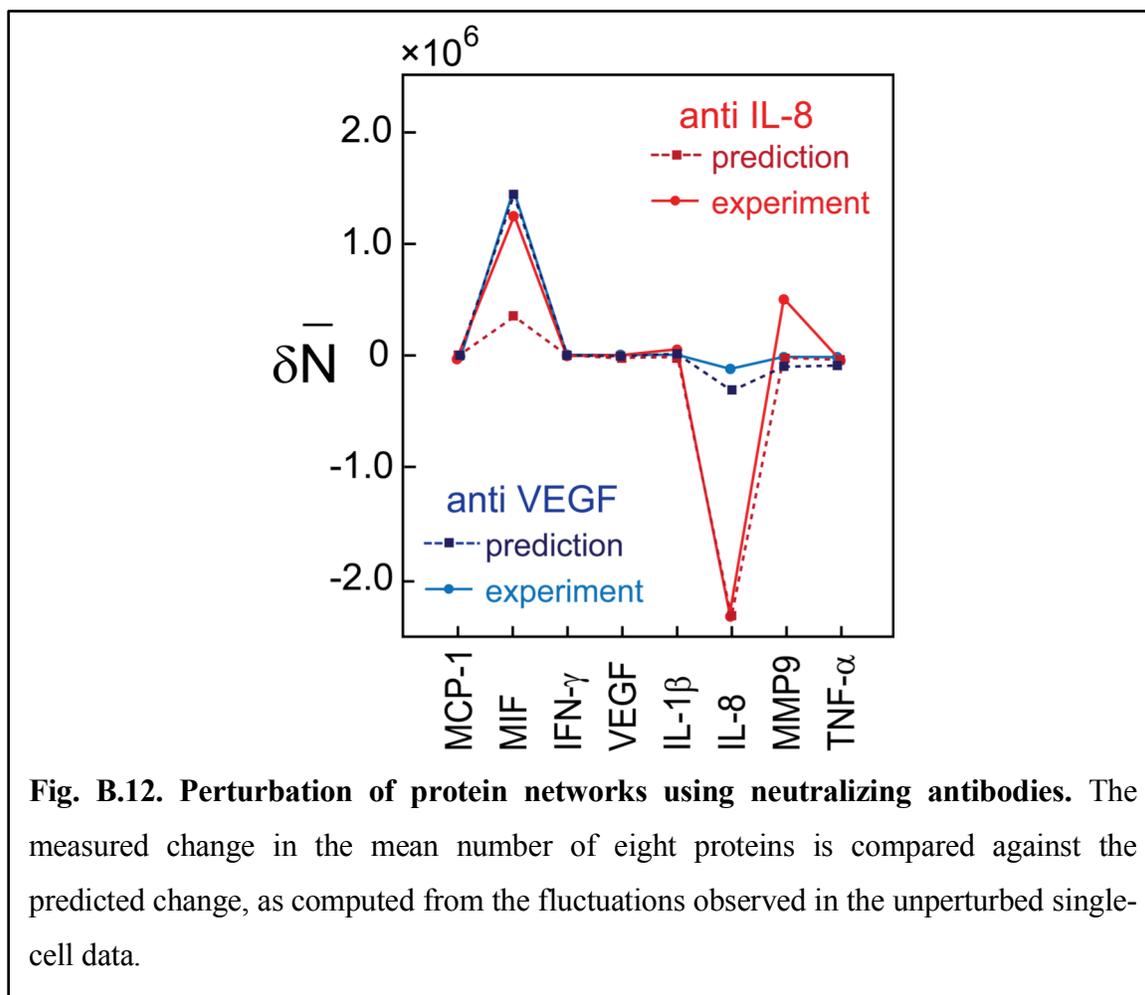
molecules. We do this conversion because it is the numbers of molecules that are secreted by the cells, or to which the cells respond, that ultimately reflects the true biology. However, this conversion seemingly introduces an additional source of noise, especially when the measured fluorescence intensity is away from the linear regime of the calibration curves. However, this conversion yields an accurate reflection of the true measurements, and the accruing benefit is worthwhile. Specifically, the number of secreted proteins is independent of the very complicated experimental response function that depends upon the fluorescence detection methods, the various capture and detection antibodies used, and the fluorescence vs. concentration profiles that characterize calibration assays. We are thus able to apply the fundamental theory to quantitative molecular measurements, and so the resultant network is a more secure representation of the true cell biology, even if the accompanying experimental uncertainties are large relative to what would be estimated from pure fluorescence measurements.

B.4.5. Antibody perturbations

We performed an inter-cellular signaling perturbation study by adding neutralizing antibodies to eliminate specific secreted cytokines. For these experiments, 4 groups of microchambers within each SCBC chip were operated independently. Three neutralizing antibodies (anti-VEGF, anti-IL-8, and anti-TNF- α) were added to the cells, with one antibody per microchamber group. A control experiment was performed without any neutralizing antibody. As shown in **Fig. B.12.**, the removal of IL-8 markedly increased the production MIF, slightly increased IL-1 β and slightly decreased TNF- α . The results are in agreement with the network hypothesis, **Fig. B.7.B.**

Using the theorem of Le Chatelier we quantitatively predict the effect of the antibody perturbations using Eq. 2. Here, the input for the prediction is the covariance matrix for the unperturbed cells. To compute the predicted mean number of protein i after an antibody for protein j is applied we need to know the change in chemical potential of protein j . We take it that an antibody for a protein lowers its chemical potential. We determine the magnitude of that reduction by requiring that the decrease in the copy number of the directly perturbed protein is reproduced. Additional details are provided in 3.6.9. The quality of the prediction in the perturbation experiments of IL-8 and VEGF is excellent, as shown in **Fig. B.12.** The

prediction of the results for the perturbation by anti-TNF- α is not in accord, likely because the change in the mean copy number of the proteins is smaller by about an order of magnitude, and so is close to the noise level.



B.5. Conclusions

The multiplexed measurements of secreted proteins by single-cells and defined, few cell colonies provide a unique opportunity to capture the fluctuations of individual cells. An information theoretic, maximal entropy analysis can be applied to reproduce the observed fluctuations in the levels of the different assayed proteins. The theoretical analysis can also account for why for some proteins exhibit broad fluctuations, while others exhibit narrow fluctuations. The experimental approach permits observations of the covariance in the fluctuations of different proteins, and how those fluctuations evolve as a single-cell is

perturbed by the presence of 1,2,3, etc., other cells. Again, with the information theory, these covariances can be analyzed to extract hypotheses about the network of interacting proteins. Measuring the role of antibodies for specific proteins provides a test of that network hypothesis, and demonstrates that the theory is able to quantitatively *predict* the results of the molecular perturbation experiments using only data obtained for the unperturbed cells. This demonstration of the Le Chatelier's principle, appears to be general, and we are currently exploring how it can be applied towards understanding the role of other perturbations (such as hypoxia, genetic modifications, etc.). The long-term goal is to extend this approach toward understanding the various protein-signaling networks that operate within complex microenvironments, such as tumors.

B.6. Supplement: Details in Theoretical Methods

B.6.1. Introduction to theoretical supplementary methods

We show how to characterize protein-protein interactions. Specifically we show (i) that the different tiers of a signaling network can be quantitatively determined from the measured fluctuations in the concentrations of signaling proteins and (ii) that the measured fluctuations in the concentrations of signaling proteins for the unperturbed cell can be used to predict the effect of introducing perturbations such as neutralizing antibodies. The approach is developed from an information theoretic perspective and it is related to the specification of the direction of change when a system responds to a perturbation, known as the principle of Le Chatelier. The corresponding result here is that we predict the sequence of tiers in the network, see **Fig. B.7.** of the article. In addition we specify which signaling proteins are at a given tier of the network and their mutual influence including inhibition, see **Fig. B.9.** of the article. Experimental measurements of the fluctuation of concentrations in samples with nanoliter volume containing n cells, $n = 0,1,2,\dots$, see Fig. S8 below, are used to validate the signaling protein network. Finally we use the protein-protein interaction as determined for the unperturbed cell to quantitatively predict, **Fig. B.12.** of the article, the effect of perturbations.

The approach we propose provides an analog and an extension of the statement that heat is transferred from a warmer to a colder body. We can understand this statement as a

statement about the direction of a process between two equilibrium states, meaning that it is a static principle. We can also think of it as a statement about the dynamics, meaning that it specifies the rate of change. We will here develop the formalism for the static interpretation. The explicit introduction of time is possible and we have the required formalism at hand but it requires a more elaborate theoretical foundation and so will be given elsewhere.

B.6.2. The ensemble: basis for making predictions

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 secreted 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 show³² 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 experiment described in the main text.

The key point is that even if the fluctuations are not small it is possible to make predictions. We discuss three types of predictions in the paper, with more details given in this section of Supplement. We predict the distribution of fluctuations, we predict the tiers in the network and, in particular and as shown in **Fig. B.12.**, we predict the response of a system to a perturbation. For these first and last predictions, we compare directly with

experimental results. We emphasize that the prediction is made strictly independently of the experiment to which it is compared.

The probability of a system in a particular composition can be shown to be given by

$$P(N_1, N_2, \dots) = \exp\left\{\beta(\sum_i \mu_i N_i - E)\right\} / \Xi \quad (\text{S1})$$

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. S1) 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^{24,33}. 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 $\mu_i, \alpha_i = \beta \mu_i$. 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. S1), 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, say, one protein is

$$P(N) \propto \left(Q^N / N! \right) \exp(-\beta\mu N) \quad (\text{S2})$$

Here Q is the factor that arises by summing over all the internal states of the protein that are occupied at the temperature T . This result is used in the main text to fit the observed distribution for a single protein (**Fig. B.6.**).

B.6.3. 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 small perturbations³³. 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 m_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. S1. In the algebraic developments in Eq. S4 below we make use of the definition of the mean concentration

$$\bar{N}_j = \sum N_j P(N_1, N_2, \dots) \quad (\text{S3})$$

The summation in Eq. S3 is over all the possible compositions, each weighted by its probability $P(N_1, N_2, \dots)$ computed as the distribution of maximal entropy. The same meaning for the summation is used also in Eq. S4 below. We denote this averaging by an over bar. From Eq. S1, the variation of the distribution that occurs when a particular chemical potential is changed by a small amount is $\delta P(N_1, N_2, \dots) = \beta \delta \mu_i N_i P(N_1, N_2, \dots)$. 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:

$$\begin{aligned}
\delta \bar{N}_j &= \sum N_j \delta P(N_1, N_2, \dots) \\
&= \sum (N_j - \bar{N}_j) \delta P(N_1, N_2, \dots) \\
&= \beta \delta \mu_i \sum (N_j - \bar{N}_j) N_i P(N_1, N_2, \dots) \\
&= \beta \delta \mu_i \sum (N_j - \bar{N}_j) (N_i - \bar{N}_i) P(N_1, N_2, \dots) \\
&= \beta \delta \mu_i \overline{(N_j - \bar{N}_j)(N_i - \bar{N}_i)}
\end{aligned} \tag{S4}$$

Note that the conservation of normalization implies that the average change in the probability must be zero, $0 = \sum \delta P(N_1, N_2, \dots)$ and we have used this result in the derivation above. In the last line in Eq. S4 we have avoided writing the summation over all compositions by the use of the over bar to designate an average over the probability $P(N_1, N_2, \dots)$, which is the notation introduced in Eq. S3.

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:

$$\delta P(N_1, N_2, \dots) = \beta \sum_i N_i P(N_1, N_2, \dots) \delta \mu_i \tag{S5}$$

The summation in Eq. S5 is an ordinary sum over the finite number S of signaling proteins, $i = 1, 2, \dots, S$. 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

$$\delta \bar{N}_j = \beta \sum_i \overline{(N_j - \bar{N}_j)(N_i - \bar{N}_i)} \delta \mu_i \tag{S6}$$

B.6.4. The principle of Le Chatelier

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. S3. When the chemical potential of protein i is changed, for an ensemble at maximal entropy the mean value of protein j changes by

$$\frac{\partial \bar{N}_j}{\partial \mu_i} = \sum N_j \frac{\partial P(N_1, N_2, \dots)}{\partial \mu_i} \quad (\text{S7})$$

where, as emphasized in Eq. S3, the distribution $P(N_1, N_2, \dots)$ is not arbitrary but is the one of maximal entropy as exhibited in Eq. S1. Eq. S4 is recovered when the derivative in Eq. S7 is evaluated. The reader may feel that this is a triviality but it is not without meaning. What we have proven is that computing a small change in the distribution $P(N_1, N_2, \dots)$ 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 $P(N_1, N_2, \dots)$ at the point where the value of the chemical potential is μ_i . Then the change in the distribution is $(\partial P(N_1, N_2, \dots) / \partial \mu_i) \delta\mu_i$. 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. S1 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. S6. 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 secretion of proteins by a single-cell and, as shown in the main text and particularly in the histogram in **Fig. 3.6.**, 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.

B.6.5. The equation for the direction of change

The (symmetric) square matrix $\overline{(N_j - \bar{N}_j)(N_i - \bar{N}_i)}$ is the covariance matrix of the (equilibrium) fluctuations in the (equilibrium) concentrations, the \bar{N}_j 's. It is an equilibrium average because, as explicitly shown in Eq. S4, it is an expectation over the equilibrium distribution as given in Eq. S3. 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.

A covariance matrix can be shown to be a non-negative matrix, also called semipositive definite, meaning that its eigenvalues are zero or positive. If the concentrations of the signaling proteins can in principle be varied independently, which is definitely not necessarily the case, then the covariance matrix $\overline{(N_j - \bar{N}_j)(N_i - \bar{N}_i)}$ is a positive matrix with positive eigenvalues. We will discuss below why it will often be the case that for reasons of both principle and practice (e.g., experimental noise) there will be eigenvalues that are effectively zero. In that case, technically, the covariance matrix is positive semidefinite³⁴.

Eq. S6 specifies how the concentration of the j 'th signaling molecule varies when the i 'th chemical potential is changed. In general the correlation coefficient $\overline{(N_j - \bar{N}_j)(N_i - \bar{N}_i)}$ between the signaling molecules i and j can be either positive or negative. Therefore, in general the change $\delta \bar{N}_j / \delta \mu_i$ is not necessarily of the same direction for all proteins j . This obvious result will be important for us below. Using the observation that the covariance matrix is semipositive definite, it is however possible to determine the direction of change by first diagonalizing the covariance matrix. This means that we can determine S distinct linear combinations of signaling molecules, where (a) each such set of molecules changes in a given direction and (b) we can order the different sets in terms of the extent of their response such that the first set is the most changing, the second set changes to a lesser extent, etc. In the time-dependent formalism, not presented here, we can outright say that the first set is the fastest changing and therefore it is the first to change. Then there follow changes in the second set, etc. It is clearly our intention to identify each set of signaling molecules as the set of molecules in a given tier in the network.

B.6.6. Tiers of the network are eigenvectors of the correlation matrix

Our next purpose is to define the tiers of the network. The set of proteins that participate in the m 'th tier is determined as follows. Let \mathbf{S}_m designate the m 'th eigenvector of the covariance matrix where the eigenvectors are listed in order of decreasing magnitude of the corresponding eigenvalue. The largest eigenvalue is $m=1$. Each eigenvector \mathbf{S}_m is a (column) vector of S components and it is determined by the matrix equation

$$\mathbf{\Sigma} \mathbf{S}_m = \sigma_m^2 \mathbf{S}_m \quad , \quad m=1,2,\dots \quad (\text{S8})$$

where $\mathbf{\Sigma}$ is the S by S symmetric covariance matrix whose elements are $\Sigma_{ij} = \overline{(N_j - \bar{N}_j)(N_i - \bar{N}_i)}$ and we explicitly indicated that the eigenvalues are positive or zero but not negative (which defines a positive semidefinite matrix). The eigenvectors of the symmetric covariance matrix are orthogonal to one another and can be chosen to be normalized

$$\mathbf{S}_{m'}^T \cdot \mathbf{S}_m = \begin{cases} 0, & m' \neq m \\ 1, & m' = m \end{cases} \quad (\text{S9})$$

Here the superscript T designates the transpose so that $\mathbf{S}_{m'}^T$ is a row vector and Eq. S9 is the scalar product.

For each value of the number of cells, n , in the compartment the eigenvalues are arranged in the order of decreasing magnitude the largest eigenvalue being labeled as $m=1$ and the smallest as $m=12$. See **Fig. B.10.** for the dependence of the largest eigenvalues vs. cell number.

B.6.7. The spectral representation of the covariance matrix

Fig. B.9. shows the covariance matrix computed for experiments with one cell in the compartment. Also shown in **Fig. B.9.** is the resolution of the covariance matrix into tiers defined as follows. From each eigenvector \mathbf{S}_m we can define an S by S symmetric matrix \mathbf{P}_m as follows

$$\mathbf{P}_m = \mathbf{S}_m \cdot \mathbf{S}_m^T \quad (\text{S10})$$

The spectral theorem (10) is the result that the covariance matrix $\mathbf{\Sigma}$ can be resolved into tiers as

$$\mathbf{\Sigma} = \sum_m \sigma_m^2 \mathbf{P}_m \quad (\text{S11})$$

The eigenvalues σ_m^2 are arranged in a decreasing order so that each subsequent tier makes a smaller contribution. The very dominant contribution is from $m = 1$. The leading eigenvalue = tier 1, is only about 30% bigger than the second one, $m = 2$. The third eigenvalue is smaller by almost two orders of magnitude. **Fig. B.10.** is a plot on a logarithmic scale of all non-zero eigenvalues. There are only two eigenvectors that, judging by the value of their corresponding eigenvalues are definitely above the noise. The dominant ($m=1$) and the $m = 2$ eigenvectors for 1 cell measurements are shown in **Fig. B.9.** and for three cells in **Fig. B.11.**

B.6.8. The role of the number of cells in the sample

It was possible to make repeated measurements of the protein concentrations for different values of the number of cells in the sample. In this section we argue that the direction of increasing n can be semi-quantitatively regarded as a direction of increasing time. Therefore by examining how the eigenvectors of the covariance matrix change with n we have an independent determination of the direction of the *dynamic* response of the system.

Fig. B.10. shows are the largest eigenvalues for $n = 0, 1, 2, 3$ and 4 cells. To interpret **Fig. B.10.** within the point of view as used in this paper we argue as follows. A single-cell secretes a number of different signaling proteins and therefore even the data measured for a single-cell can show the role of protein-protein interactions. If two cells are in the sample these interactions increase in importance. If we think of n as a measure of concentrations of proteins then $\bar{N} \propto n$ but to compute the covariance we need to divide by the number of protein molecules. So for both paracrine and endocrine signaling we expect

the covariance to increase with n . When n becomes high there may be three or more cells interacting and the simple considerations break down.

B.6.9. Antibody perturbations

Fig. B.12. shows a quantitative comparison of the measured results as compared to the purely theoretical prediction when neutralizing antibodies for specific proteins are added. We emphasize that it is a prediction because the results shown are based on using Eq. S4 that we repeat here:

$$\delta \bar{N}_j = \beta \overline{(N_j - \bar{N}_j)(N_i - \bar{N}_i)} \delta \mu_i = \beta \Sigma_{ji} \delta \mu_i$$

The addition of a neutralizing antibody for protein i means that $\delta \mu_i$ is negative. We emphasize that the experimental results shown in **Fig. B.12.** are for single-cells in the compartment. This means, see **Fig. B.10.** that the largest eigenvalue, $\sigma_{m=1}^2$, of the covariance matrix is large indeed. Then, from Eq. S11, the contribution from the first tier dominates. It is the two proteins in this tier that are shown in the panel. There are bigger discrepancies between theory and experiment for tiers 2 or 3 for which the experimental signal is weak.

B.7. References

1. Lin, W.-W. & Karin, M. A cytokine-mediated link between innate immunity, inflammation, and cancer. *J Clin Invest* **117**, 1175–1183 (2007).
2. Gnechi, M. *et al.* Paracrine action accounts for marked protection of ischemic heart by Akt-modified mesenchymal stem cells. *Nature Publishing Group* **11**, 367–368 (2005).
3. Croci, D. O. *et al.* Dynamic cross-talk between tumor and immune cells in orchestrating the immunosuppressive network at the tumor microenvironment. *Cancer Immunol. Immunother.* **56**, 1687–1700 (2007).
4. Seruga, B., Zhang, H., Bernstein, L. J. & Tannock, I. F. Cytokines and their relationship to the symptoms and outcome of cancer. *Nat Rev Cancer* **8**, 887–899

- (2008).
5. Polyak, K. & Weinberg, R. A. Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. *Nat Rev Cancer* **9**, 265–273 (2009).
 6. Ariztia, E. V., Lee, C. J., Gogoi, R. & Fishman, D. A. The tumor microenvironment: key to early detection. *Crit Rev Clin Lab Sci* **43**, 393–425 (2006).
 7. Sachs, K., Perez, O., Pe'er, D., Lauffenburger, D. A. & Nolan, G. P. CaUSA1 protein-signaling networks derived from multiparameter single-cell data. *Science* **308**, 523–529 (2005).
 8. Nomura, L., Maino, V. C. & Maecker, H. T. Standardization and optimization of multiparameter intracellular cytokine staining. *Cytometry A* **73**, 984–991 (2008).
 9. Lamoreaux, L., Roederer, M. & Koup, R. Intracellular cytokine optimization and standard operating procedure. *Nat Protoc* **1**, 1507–1516 (2006).
 10. Cox, J. H., Ferrari, G. & Janetzki, S. Measurement of cytokine release at the single-cell level using the ELISPOT assay. *Methods* **38**, 274–282 (2006).
 11. Song, M. & Phelps, D. S. Comparison of SP-A and LPS effects on the THP-1 monocytic cell line. *Am. J. Physiol. Lung Cell Mol. Physiol.* **279**, L110–7 (2000).
 12. Fan, R. *et al.* Integrated barcode chips for rapid, multiplexed analysis of proteins in microliter quantities of blood. *Nat Biotechnol* **26**, 1373–1378 (2008).
 13. Shin, Y. S. *et al.* Chemistries for patterning robust DNA microbarcodes enable multiplex assays of cytoplasm proteins from single cancer cells. *Chemphyschem* **11**, 3063–3069 (2010).
 14. Quake, S. R. & Scherer, A. From micro- to nanofabrication with soft materials. *Science* **290**, 1536–1540 (2000).
 15. Millet, L. J., Stewart, M. E., Sweedler, J. V., Nuzzo, R. G. & Gillette, M. U. Microfluidic devices for culturing primary mammalian neurons at low densities. *Lab Chip* **7**, 987–994 (2007).
 16. Lee, J. N., Park, C. & Whitesides, G. M. Solvent compatibility of poly(dimethylsiloxane)-based microfluidic devices. *Anal. Chem.* **75**, 6544–6554 (2003).

17. Altschuler, S. J. & Wu, L. F. Cellular heterogeneity: do differences make a difference? *Cell* **141**, 559–563 (2010).
18. Wang, J. *et al.* A self-powered, one-step chip for rapid, quantitative and multiplexed detection of proteins from pinpricks of whole blood. *Lab Chip* **10**, 3157–3162 (2010).
19. Findlay, J. W. A. & Dillard, R. F. Appropriate calibration curve fitting in ligand binding assays. *AAPS J* **9**, E260–7 (2007).
20. Levine, R. D. Information theory approach to molecular reaction dynamics. *Annu Rev Phys Chem* **29**, 59–92 (1978).
21. Levine, R. D. Information Theoretical Approach to Inversion Problems. *Journal of Physics a-Mathematical and General* **13**, 91–108 (1980).
22. Remacle, F. & Levine, R. D. The Elimination of Redundant Constraints in Surprisal Analysis of Unimolecular Dissociation and Other Endothermic Processes. *J Phys Chem A* **113**, 4658–4664 (2009).
23. Jaynes, E. T. Probability theory: The logic of science. *Cambridge University Press* (2004).
24. Mayer, J. E. & Mayer, M. G. *Statistical mechanics*. (Wiley, 1977).
25. Lotka, A. J. Note on moving equilibria. *P Natl Acad Sci USA* **7**, 168–172 (1921).
26. Bellman, R. *Introduction to Matrix Analysis*. (McGraw-Hill, 1970).
27. Wall, M. E., Rechtsteiner, A. & Rocha, L. M. in *Approach to Microarray Data Analysis* 91–109 (Kluwer Academic Publishers, 2003). doi:10.1007/0-306-47815-3_5
28. Jolliffe, I. T. *Principal Component Analysis*. (Springer Verlag, 2002).
29. Roger, T., David, J., Glauser, M. P. & Calandra, T. MIF regulates innate immune responses through modulation of Toll-like receptor 4. *Nature* **414**, 920–924 (2001).
30. Calandra, T., Bernhagen, J., MITCHELL, R. A. & Bucala, R. Macrophage Is an Important and Previously Unrecognized Source of Macrophage-Migration Inhibitory Factor. *J Exp Med* **179**, 1895–1902 (1994).
31. Murphy, K. P., Travers, P., Walport, M. & Janeway, C. *Janeway's immunobiology*. (Garland Pub, 2008).

32. Levine, R. D. How large is 'large' for a thermodynamic-like behavior. *Physica E* **9**, 591–599 (2001).
33. Callen, L. *Thermodynamics and an Introduction to Thermostatistics*. (Wiley, 1985).
34. Alhassid, Y. & Levine, R. D. Experimental and Inherent Uncertainties in the Information Theoretic Approach. *Chemical Physics Letters* **73**, 16–20 (1980).