

DNA Encoded Biotechnologies for Informative Cancer Diagnostics

Thesis by
Gabriel Abner Kwong

In Partial Fulfillment of the Requirements

For the degree of
Doctor of Philosophy



CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2009

(Defended May 28, 2009)

© 2009

Gabriel Abner Kwong

All Rights Reserved

Here, then are the true characteristics of objectivity... Objectivity does not demand that we estimate man's significance in the universe by the minute size of his body, by the brevity of his past history or his probable future career. It does not require that we see ourselves as a mere grain of sand in a million Saharas. It inspires us, on the contrary, with the hope of overcoming the appalling disabilities of our bodily existence, even to the point of conceiving a rational idea of the universe which can authoritatively speak for itself. It is not a counsel of self-effacement, but the very reverse—a call to the Pygmalion in the mind of man.

- Michael Polanyi, 1958, *Personal Knowledge*

Acknowledgements

I am truly in debt to a number of people with whom I have interacted and worked with during my stay here at Caltech. First, I would like to thank my advisor Professor Jim Heath. Of the many things that I have learned from him, I hold in high regard his intensity and infectious passion for science. I would like to thank him for the academic freedom he grants his students, and for the incredible resources he makes available for his group, including the organization of the Nanosystems Biology Cancer Center (NSBCC) between Caltech, UCLA and the Institute for Systems Biology (ISB). I would like to offer my gratitude to many senior collaborators within the NSBCC, including Professor Caius Radu (UCLA), Professor Antoni Ribas (UCLA), Professor Paul Mischel (UCLA), Dr. Leroy Hood (ISB), Professor Mike Phelps (UCLA) and Professor Owen Witte (UCLA) all of whom have helped during my graduate career either in providing access to the resources in their labs or project guidance. I would like to thank Professor Ton Schumacher (NKI) for many insightful email exchanges regarding the NACS project. I would like to thank Professor Pamela Bjorkman, Professor Sarkis Mazmanian and Professor Chin-Lin Guo for agreeing to serve on my thesis committee with little prior notice.

I had the pleasure of interacting with many intelligent and talented co-workers within the Heath group. I would like to thank Dr. Ryan Bailey with whom I co-developed the DEAL platform with. I wish him the best in his lab at UIUC. I would like to thank Heath group graduate students Kiwook Hwang and Chao Ma for helping during the latter stages of the NACS project. Thanks to Ophir Vermesh for being so down-to-

earth—you'll be a very successful doctor! I would like to thank Diane Robinson for processing all necessary paperwork (reimbursements!) and for all her direct statements and questions. Thanks to Kevin Kan for managing the group facilities and providing financial advice.

Thanks to my good friends from the Pierce lab, Harry Choi (whom I've known and worked with since my undergraduate days at UC Berkeley) and Suvir Venkataraman (the smoothest graduate student on campus) for their help in designing the orthogonal DNA sequences.

Because of the highly collaborative nature of the NSBCC, I would like to thank many collaborators at UCLA. In particular, I would like to thank Jenny Shu for help with protein expression and the Witte lab for their hospitality. I would like to thank Begonya Comin-Anduix, Richard Koya and Thinle Chodon for preparing T cell samples for NACS. I would like to thank my good friends from the Mischel Lab; Tiffany Huang for work on the GBM project, providing assistance with flow cytometry and conversing about people, dates, family history; Shawn Sarkaria for taking the leadership role in the GBM project, being a sounding board for scientific ideas, and his meticulous attention to detail.

Thanks to Bruz Marzolf at ISB for printing DNA microarrays and his promptness in response to problems and questions.

Lastly, I would like to acknowledge my family whose influence on my life I have yet to fully understand but has made me learn more about myself, and as a result lead a more genuine life.

Abstract

This thesis describes the development of DNA-encoded, multi-parametric, sensing platforms for informative cancer diagnostics. In the first part of this thesis, I will present a technology called “DNA-encoded antibody library (DEAL).” In this approach, computationally derived, orthogonal ssDNA sequences are conjugated to antibodies specific for protein targets and cell surface markers. The resulting collection of conjugates is applied to a biological sample of interest, binds to their cognate antigens, and is detected after the complexes are hybridized to a glass substrate printed with spatially distinct complementary DNA sequences. By using DNA assembly, the DEAL platform enables the simultaneous detection of the major classes of biological molecules, namely nucleic acids, proteins and cells.

The second part of this thesis focuses on the development of a cell sorting platform that can detect antigen-specific T cells called “Nucleic Acid Cell Sorting (NACS).” In NACS, ssDNA encoding is used to assemble peptide major histocompatibility complexes (p/MHC) on glass substrates by hybridization to cDNA microarrays. These assembled peptide/MHC microarrays are then used to sort mixed populations of antigen-specific T cells. This spatially encoded scheme addresses the widespread desire for methods that allow the multiplexed detection of antigen-specific T cells. The sensitivity and selectivity of NACS is similar to flow cytometry, demonstrated in key experiments with T cells derived from multiple sources, including endogenous and TCR-engineered T cells collected from cancer patients. Finally, this platform is used to

monitor the persistence of cancer-specific T cells in peripheral blood collected from a patient undergoing T cellular immunotherapy.

Lastly, a scheme for the detection of cell surface markers is presented. In this approach, DEAL and NACS conjugates prepared with UV labile ssDNA oligonucleotides are allowed to bind to target cell samples in solution. The ssDNA tags are released in solution by UV-induced photocleavage. The presence and expression of the cognate antigen is determined by collecting the pool of reporter ssDNA tags followed by exponential amplification by PCR. A DEAL conjugate specific for the oncogene EGFR was used to determine the expression level of EGFR in a low-passage brain tumor primary cell line. The feasibility of using ssDNA-p/MHC complexes for detecting unique TCRs was also demonstrated. Finally an experimental flow is described for integration with second generational high-throughput sequencing platforms for global and quantitative surface-ome profiling.

Table of Contents

Acknowledgements	iv
Abstract	vi
Table of Contents	viii
List of Figures and Tables	xii
Chapter 1: Introduction	1
1.1 Complexity of human cancers	1
1.2 Diagnostic challenges presented by cancer	3
1.3 DNA as an encoding element	5
1.4 Thesis overview	6
1.5 References	10
Chapter 2: DNA-Encoded Antibody Libraries: A Unified Platform for Multiplexed Cell Sorting and Detection of Genes and Proteins	16
2.1 Introduction	16
2.2 Experimental Methods	20
2.2.1 DNA sequences for spatial encoding	20
2.2.2 DNA antibody conjugation	21
2.2.3 Microarray fabrication	23
2.2.4 Fabrication of microfluidic devices	23
2.2.5 1 ^o antibody microarray generation and DEAL-based immunoassays	24
2.2.6 Microfluidic-based assay procedures	24

2.2.7	Microfluidic Au amplification methods	25
2.2.8	Analysis of DNA-encoded antibodies by flow cytometry	25
2.2.9	Cell capture, separation, and sorting methods	26
2.3	Results and Discussion	28
2.3.1	<i>In silico</i> design of orthogonal DNA oligonucleotides	28
2.3.2	Generation of DNA-antibody conjugates	29
2.3.3	Multiplexed protein detection by DEAL	33
2.3.4	Detection of multiple proteins within a single microfluidic channel..	36
2.3.5	Multiplexed sorting of immortalized and primary immune cells	40
2.3.6	Single environment detection of specific cDNAs, proteins and cells	43
2.4	Conclusions	44
2.5	References	46
2.6	Appendix A: Computation derivation of orthogonal DNA oligomers	52
2.6.1	Computational analysis of sequences A1, B1, and C1	53
2.6.2	Computing orthogonal sequences A, B, and C constrained by a polyA ₁₀ header: input file	54
2.6.3	PolyA ₁₀ header computational results	55
2.6.4	Computing a fourth sequence: input code	58
2.6.5	Results of fourth strand computation	59
2.7	Appendix B: FPLC of DEAL conjugates	60
Chapter 3: Modular Nucleic Acid Assembled p/MHC Microarrays for Multiplexed Sorting of Antigen-Specific T cells		61
3.1	Introduction	61
3.2	Experimental Methods	65

3.2.1	Microarray fabrication	65
3.2.2	Production of ssDNA-SAC conjugates	66
3.2.3	Preparation of T cells	67
3.2.4	T cell sorting methods	69
3.3	Results and Discussion	70
3.3.1	Rational design of ssDNA-encoded p/MHC tetramers	70
3.3.2	Performance of p/MHC arrays produced via DNA immobilization and direct spotting	74
3.3.3	NACS specificity and limit of detection	76
3.3.4	Selective release of immobilized T cells with restriction endonucleases	80
3.3.5	NACS sorting of endogenous primary human T cells	82
3.3.6	Persistence of MART-1 specific, TCR-engineered human T cells <i>in vivo</i>	85
3.3.7	Homogeneous platform for cell sorting and functional analysis	87
3.4	Conclusions	90
3.5	References	92
3.6	Appendix A: Protein sequences	98
3.7	Appendix B: Chromatography	104
3.7.1	Iminobiotin purification of SAC	104
3.7.2	FPLC of ssDNA-SAC conjugates	105
Chapter 4: Detection of Cell Surface Markers with Encoded ssDNA Reporters: Towards Global Cell Surface-ome Profiling		106
4.1	Introduction	106

4.2	Experimental Methods	108
4.2.1	DNA sequences and production of conjugates	108
4.2.2	Detection of surface markers with PCR	109
4.3	Results and Discussion	110
4.3.1	Detection of differential cell surface expression of EGFR by PCR ..	110
4.3.2	Cellular barcoding limits of detection	112
4.3.3	Detection of antigen-specific T cells using ssDNA-p/MHC tetramers	113
4.4	Conclusions and Future Directions	115
4.5	References	119

List of Figures and Tables

Chapter 1

Figure 1.1	Numerous components of tumor biology must be queried for informative diagnostics	3
------------	--	---

Chapter 2

Figure 2.1	Illustration of the DEAL method for cell sorting and co-detection of proteins and cDNAs (mRNAs)	18
Figure 2.2	DNA sequences derived from computation minimize cross hybridization	29
Figure 2.3	Illustration of the two step coupling strategy utilized to prepare DEAL antibodies	30
Figure 2.4	Optimization of DNA loading of DEAL antibodies for cell surface marker recognition	32
Figure 2.5	Spatially encoded protein array	34
Figure 2.6	Illustration of the resistance of the DEAL approach towards non-specific protein absorption	35
Figure 2.7	Protein array assembled in microfluidics in 10 minutes	36
Figure 2.8	DNA-templated protein immunoassays executed within microfluidic channels	38
Figure 2.9	Optimization and use of DEAL for multiplexed cell sorting	42
Figure 2.10	Microscopy images demonstrating simultaneous cell capture at spot B1 and multiparameter detection of genes and proteins, at spots A1 and C1, respectively	44
Figure 2.11	Fast protein liquid chromatography of DEAL conjugates	60
Table 2.1	DNA sequences for spatial encoding	21

Chapter 3

Figure 3.1	Self-assembled ssDNA-p/MHC tetramer arrays for multiplexed sorting of antigen-specific cells	64
Figure 3.2	An engineered variant of streptavidin expressing C-terminal cysteine residues has superior biotin binding capacity compared to native streptavidin post conjugation with ssDNA	72
Figure 3.3	T cell capture efficiency is optimal when utilizing ssDNA-SAC conjugates to generate NACS p/MHC tetramers	73
Figure 3.4	Comparison of NACS versus spotted p/MHC arrays	75
Figure 3.5	Comparison of the performance of p/MHC arrays produced by NACs and by spotting	76
Figure 3.6	Multiplexed nucleic acid cell sorting of antigen-specific T cells	77
Figure 3.7	NACS limit of detection	79
Figure 3.8	Programmed release of sorted T cells by endonuclease cleavage	81
Figure 3.9	NACS sorting of endogenous primary human T cells specific for Epstein-Barr virus and Cytomegalovirus	83
Figure 3.10	The T cell specificities of PBMCs from patients NRA11 and NRA13	84
Figure 3.11	Limit of detection of endogeneous human T cells with NACS ...	85
Figure 3.12	Monitoring the presence of infused MART-1-specific, TCR-engineered T cells	87
Figure 3.13	Functional profiling of TCR triggered activation of captured antigen-specific T cells	88
Figure 3.14	Dynamic cytokine profiling of NACS sorted murine OT1 lymphocytes	89

Figure 3.15	SAC purification with an iminobiotin agarose support column ..	104
Figure 3.16	Fast protein liquid chromatography of ssDNA-SAC conjugates	105
Table 3.1	Orthogonal DNA sequences for spatial encoding of p/MHC tetramers	65

Chapter 4

Figure 4.1	Schematic of cellular barcoding	108
Figure 4.2	Antibody-ssDNA and p/MHC-ssDNA tetramers stain cells in solution	111
Figure 4.3	GBM1600 EGFR detection by cellular barcoding	112
Figure 4.4	Limits of detection	113
Figure 4.5	Antigen-specific T cell detection by cellular barcoding	115
Figure 4.6	Schematic of global “surface-ome” profiling with Solexa sequencing	117
Figure 4.7	DNA sequence design considerations	118
Table 4.1	Cellular barcoding DNA sequences.....	109

Chapter 1

Introduction

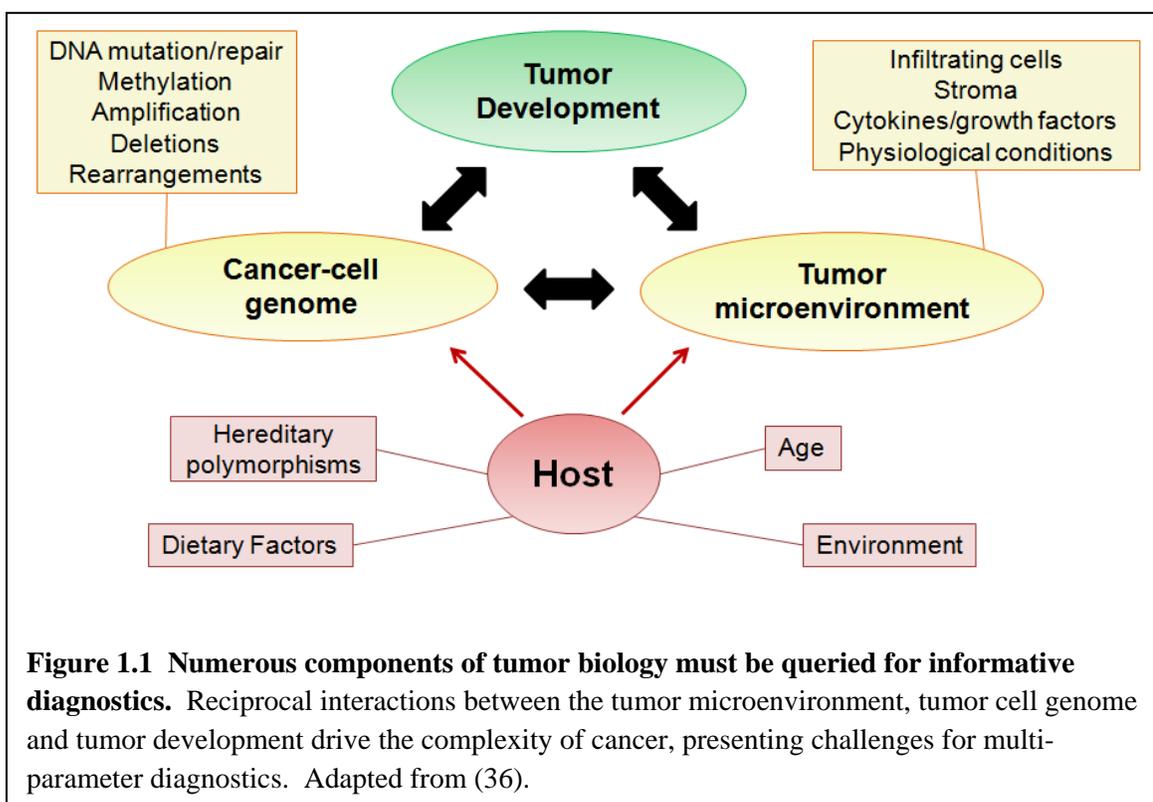
1.1 Complexity of human cancers

Our current understanding of the molecular basis of cancer has been shaped by initial landmark studies demonstrating that cancer was caused by the deregulation of a few oncogenes or tumor suppressor genes (1, 2). Identification of these genes and their encoded protein products has had a tremendous impact on molecular medicine, leading to effective targeted therapies for certain cancers. For example, patients with chronic myelogenous leukemia (CML), a bone marrow derived cancer, express the fusion oncogene BCR-ABL that is effectively targeted by the kinase inhibitor imatinib mesylate (3–7). In breast cancer, elevated levels of the membrane-bound receptor oncogene Her2/neu (ERBB2) predict response to the humanized monoclonal antibody therapeutic trastuzumab (8, 9). The clinical effectiveness of targeted molecular therapies against certain types of cancer is a testament to the progress in cancer research and medicine.

It is becoming increasingly clear, however, that for the majority of metastatic cancers, a single or few genes is insufficient in predicting tumor behavior and does not represent all necessary and sufficient targets for molecular therapy. Tumors are highly complex, dynamic, heterogeneous biological systems driven by series of genomic and epigenetic mutations that affect the genes controlling cell proliferation, survival, invasiveness and motility. Typically, multiple primary genomic alterations are present within a single tumor with secondary mutations arising through genetic instability, enabling malignancies to develop resistance to drugs. For example, the majority of lung adenocarcinomas that respond to epidermal growth factor receptor (EGFR) inhibitors (erlotinib and gefitinib) eventually develop resistance through selection of tumor cells expressing an EGFR variant containing a single amino acid mutation that confers drug resistance (10–12). Pre-malignant tissue evolve in a multi-step process into tumors, each step highly sensitive to numerous interacting extracellular stimuli, including growth factors (13, 14), extracellular matrix proteases (15, 16), physiological stresses (e.g., low oxygen levels (17, 18)) and cytokines (19). Many sources contribute to the tumor milieu, including non-tumor, stromal cells found in the local architecture (20, 21) and immune cells that have infiltrated the tumor microenvironment (19, 22, 23). Thus for a typical malignancy, a large number of genomic, proteomic, physiological conditions, and supporting cell types are involved.

1.2 Diagnostic challenges presented by cancer

One of the goals of cancer diagnostics is to detect and monitor biological elements that are representative of the interactions between the various components involved in a malignancy, and with the information be able to characterize the state of the system and/or generate network models with predictive features that would be of fundamental or therapeutic value (**Figure 1.1**). The biological elements targeted are



either directly involved in the malignancy (e.g. over expression of EGFR), or are byproducts of the process (e.g. circulating tumor DNA resulting from necrosis). Individually, they are referred to as biomarkers. Currently, most FDA approved cancer diagnostic platforms focus on the detection of a single or few biomarkers and have been effective for several cancers for staging, monitoring, and prognostication (24). This pauci-parameter approach, however, is being revised because of the molecular

heterogeneity of cancer—the accuracy and coverage of diagnostic assays will be improved by adopting a multi-parameter systems approach, namely measure as many different biological elements as possible. The motivation is that a collection of biomarkers would produce a molecular signature that would provide a higher level of sensitivity and specificity for staging, treatment and prognostication.

A number of studies reported in the literature have shown that global profiling approaches are effective in characterizing cancer. Differential gene expression studies have shown that diseased complex expression networks can be differentiated from their healthy counterparts and can be used to predict patient survival and response to cancer therapies (25–28). Many of the proteins, genes and small molecules identified in these studies have spear-headed further investigations, including mechanistic studies (29) and the use of these targets as potential biomarkers (30–32). Other global profiling approaches investigating genomic alterations (33–34) and proteomic alterations (30, 35) have also been reported, although for the latter, studies have been impeded by the lack of proteome wide capture agents and lack of technologies to capture all the proteomic functions concomitantly (36) (e.g. interactions involving protein-protein, protein-ligand, post-translational modifications).

The adoption of a multi-parameter paradigm is compounded by the fact that most biological tissues collected from patients are small, making it challenging to execute large-scale analyses of the different classes of biological molecules (e.g. cells, genes, proteins, metabolites). This has driven the miniaturization of many diagnostic assays into integrated biochips. The benefits of developing biochips are severalfold. First, integrated biochips are able to handle small amounts of tissue. This is important in

interfacing with tumor tissue derived from skinny-needle biopsies and other minimally invasive diagnostic procedures. Second, on-chip measurements can be highly parallel and multiplexed which is important in providing statistical certainty by repeated sampling. Third, chip-based devices can be manufactured with well developed processes borrowed from mature engineering disciplines (e.g. semiconductor fabrication) and are typically compatible with a host of common instrumentation (e.g. fluorescent microscopy), lowering costs and increasing accessibility.

An example of a cancer diagnostic that has benefited from miniaturization is circulating tumor cell (CTC) detection. CTCs are cells shed from primary tumors and are in circulation at low levels in blood. The abundance of CTCs can be used to monitor the efficacy of cancer therapeutics. Current technologies for CTC detection rely on bead-based assays (37) but are plagued with low recoveries and low purity. Recently, two chip-based, microfluidic devices have been reported which significantly increase recoveries and purities of CTC capture (38, 39). The robustness of one biochip enabled the correlation of CTCs levels with decreased or increased tumor burden and the identification of mutations through genomic sequencing of CTCs collected from the device (40).

1.3 DNA as an encoding element

The utility of DNA as a chemical material has reached into many areas of active research, including nanomaterials (41–43), DNA computing (44), automated machines (45), and molecular electronics (46). In the arena of biological sensing involving mainly protein detection, technologies and platforms that have been developed include immuno-

PCR (47), rolling circle amplification (48), proximity ligation (49), and nanoparticle-based assays (50). DNA-based, multiplexed assays in which multiple detection events are encoded and then decoded simultaneously include immunophenotyping using transcription (51) and self-assembled chemical libraries (52).

There are several reasons that makes DNA an attractive material and encoding element. First, the exquisite specificity and favorable energetics of DNA base-pairing provides self-assembling properties, permitting complex 2-D and 3-D structures to be designed and predicted from linear primary sequences. Second, analogous to its endogenous function, unique DNA sequences can be used to store information, scaling exponentially to the length N of the polymer. Third, in its natural environment, DNA interfaces with a host of other biological elements which can be exploited for molecular level control. For example, DNA can be specifically truncated with restriction endonucleases. Fourth, the production of DNA oligonucleotides is trivial, owing in large part to chemical automation. A wide variety of unnatural chemical handles (e.g. primary amines, thiols), readouts (e.g. fluorophores), and modified bases (e.g. locked nucleic acids) are available for incorporation into a primary sequence, expanding the utility and applications of the polymer. Lastly, platforms to characterize DNA are robust, including sequencing, PCR, microarrays, and bioinformatic algorithms.

1.4 Thesis overview

This thesis presents the development of chip-based, DNA-encoded technologies to address the current multi-parameter challenges associated with *in vitro* cancer diagnostics. In Chapter 2, I will begin by introducing the development of an approach

called DNA-Encoded Antibody Libraries (DEAL) in which computationally derived, orthogonal ssDNA tags are conjugated to an antibody library where every antibody-specificity is uniquely encoded with a distinct ssDNA sequence. A library of DEAL conjugates is exposed to a biological sample, bind to their cognate antigen, and decoded spatially on a glass substrate printed with the complementary DNA sequences. I demonstrate the DEAL technique for; (1) the rapid detection of multiple proteins within a single microfluidic channel, and with the additional step of electroless amplification of gold-nanoparticle labeled secondary antibodies, establish a detection limit of 10 fM for the protein IL-2; (2) the multiplexed, on-chip sorting of both immortalized cell lines and primary immune cells; and (3) the co-detection of ssDNAs, proteins, and cell populations on the same platform. By using a common assembly molecule, DEAL meets the diagnostic need for multi-parameter platforms able to manipulate and detect major subtypes of biological molecules (genes, proteins, cell membrane-bound markers). Moreover, this technique is fully integrable with fluidics, enabling the processing of small tissue samples. Chapter 2 has been taken in part from © *J. Am. Chem. Soc.* **2007**, *129(7)*, 1959–67.

Chapter 3 extends and develops the concept of cell sorting with DNA tags in the context of an important immunological problem with clinical implications for cancer therapy, namely that of detecting antigen-specific T cells. Through genetic recombination, T cells are capable of recognizing and engaging to cells presenting processed antigenic fragments, including antigens presented by cancer cells. The ability of T cells to cull target populations of cells has been exploited as an experimental cancer therapy by infusing patients with augmented and activated T cells specific for cancer

associated antigens (53) with documented T cell-induced tumor regression in subsets of patients with metastatic cancers (54, 55).

Despite these promising trials, T cellular immunotherapy is difficult to characterize fully because of technological bottlenecks that do not allow the multiparametric analysis of different antigen-specific T cells in small numbers. With the goal of highly multiplexed T cell detection, Chapter 3 will introduce technique called "Nucleic Acid Cell Sorting (NACS)" in which single-stranded DNA oligomers conjugated site-specifically to p/MHC tetramers are employed to immobilize p/MHC tetramers via hybridization to a complementary-printed substrate. Fully assembled p/MHC arrays are used to detect and enumerate T cells captured from cellular suspensions, including primary human T cells collected from cancer patients. Importantly, T cell array binding is optimized by utilizing cysteine-engineered streptavidin (SAC) for ssDNA-p/MHC tetramer production, resulting in NACS p/MHC arrays that outperform conventional spotted arrays assessed by performance standards such as reproducibility and homogeneity. The versatility of using DNA tags is also exploited to enable selective detachment of T cells with restriction endonucleases. Demonstrative experiments regarding NACS sensitivity, multiplexing and limit of detection are performed with cell lines and with T cells isolated from cancer patients. Lastly, I show an important clinical application of this technology by monitoring the presence and abundance of a cancer-specific T cell population collected from a melanoma patient receiving cellular immunotherapy. Chapter 3 has been taken in part from © *J. Am. Chem. Soc.*, **2009**, *in press*.

Cell surface proteins constitute an important subset of the cellular proteome, as these proteins are frequently targeted for molecular therapy, staging and for directing therapies. Chapter 4 presents current work on detecting cell membrane-bound proteins by amplification of encoded ssDNA tags. In this approach, DEAL/NACS conjugates are synthesized with photolabile ssDNA tags. Whereas in previous demonstrations, the ssDNA tags were employed as molecular tethers, in the cellular barcoding approach, the ssDNA tags are used as reporter molecules. After staining a population of cells with a library of DEAL/NACS conjugates, the ssDNA tags are released into solution via UV cleavage and collected for analysis by PCR. The quantity of each unique ssDNA tag directly correlates to the presence and expression of the targeted cell surface markers. I demonstrate that this concept by detecting EGFR over-expression from a low-passage brain cancer cell line (GBM1600) relative to EGFR null Jurkat human T cells. In addition, I show that different TCR specificities can be differentiated by this technique by using NACS conjugates to detect the presence of a TCR specific for a melanoma-associated cancer antigen, MART-1. Lastly, experimental details will be presented to interface with second generation sequencing platforms for high throughput and quantitative analysis of the reporter barcodes for global cell surface-ome profiling.

1.5 References

1. Stehelin, D.; Varmus, H.E.; Bishop, J.M.; Vogt, P.K. DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. *Nature* **1976**, *260*, 170–173.
2. Comings, D.E. A general theory of carcinogenesis. *Proc. Natl. Acad. Sci. USA* **1973**, *70*, 3324–3328.
3. Melo, J.V.; Barnes, D.J. Chronic Myeloid leukaemia as a model of disease evolution in human cancer. *Nature Rev. Cancer* **2007**, *7*, 441–453.
4. Lugo, T.G.; Pendergast, A.M.; Muller, A.J.; Witte, O.N. Tyrosine kinase activity and transformation potency of bcr-abl oncogene products. *Science* **1990**, *247*, 1079–1082.
5. Daley, G.Q.; Van Etten, R.A.; Baltimore, D. Induction of chronic myelogenous leukemia in mice by the P210bcr/abl gene of the Philadelphia chromosome. *Science* **1990**, *247*, 824–830.
6. Druker, B.J.; et al. Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. *Nature Med.* **1996**, *2*, 561–566.
7. Druker, B.J.; et al. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N. Engl. J. Med.* **2001**, *344*, 1031–1037.
8. Hudis, C.A. Trastuzumab—Mechanism of action and use in clinical practice. *N. Engl. J. Med.* **2007**, *357*, 39-51.
9. Slamon, D.J.; Clark, G.M.; Wong, S.G.; Levin, W.J.; Ullrich, A.; McGuire, W.L. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* **1987**, *235*, 177–182.
10. Yun, C.; Mengwasser, K.E.; Toms, A.V.; Woo, M.S.; Greulich, H.; Wong, K.K.; Meyerson, M.; Eck, M.J. The T790M mutation in EGFR kinase causes drug

- resistance by increasing the affinity for ATP. *Proc. Natl. Acad. Sci. USA* **2008**, *105*(6), 2070–2075.
11. Paez, J.G.; et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* **2004**, *304*, 1497–1500.
 12. Pao, W.; et al. EGF receptor gene mutations are common in lung cancers from “never smokers” and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 13306–13311.
 13. Yang, F. et al. Nf1-dependent tumours require a microenvironment containing Nf1^{+/-} and c-kit-dependent bone marrow. *Cell* **2008**, *135*, 437–448.
 14. Massague, J. TGFβ in cancer. *Cell* **2008**, *134*, 215–230.
 15. Egeblad, M.; Werb, Z. New functions for the matrix metalloproteinases in cancer progression. *Nature Rev. Cancer* **2002**, *2*, 161–174.
 16. Freije, J. M.; et al. Matrix metalloproteinases and tumor progression. *Adv. Exp. Med. Biol.* **2003**, *532*, 91–107.
 17. Denko, N.C. Hypoxia, HIF1 and glucose metabolism in the solid tumor. *Nature Rev. Cancer* **2008**, *8*, 705–713.
 18. Milosevic, M.; Fyles, A.; Hedley, D.; Hill, R. The human tumor microenvironment: invasive (needle) measurement of oxygen and interstitial fluid pressure. *Semin. Radiat. Oncol.* **2004**, *14*, 249–258.
 19. Seruga, B.; Zhang, H.; Bernstein, L.J.; Tannock, I.F. Cytokines and their relationship to the symptoms and outcome of cancer. *Nature Rev. Cancer* **2008**, *8*, 887–899.
 20. Brabletz, T.; et al. Variable b-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10356–10361.

21. Franci, C.; et al. Expression of snail protein in tumor—stroma interface. *Oncogene* **2006**, *25*, 5134–5144.
22. Balkwill, F.; Charles, K.A.; Mantovani, A. Smoldering and polarized inflammation in the initiation and promotion of malignant disease. *Cancer Cell* **2005**, *7*, 211–217.
23. Yang, J.; Weinberg, R.A. Epithelial—mesenchymal transition: at the crossroads of development and tumor metastasis. *Dev. Cell* **2008**, *14*, 818–829.
24. Ludwig, J.A.; Weinstein, J.N. Biomarkers in cancer staging, prognosis and treatment selection. *Nature Rev. Cancer* **2005**, *5*, 845–856.
25. Ntzani, E.E.; Ioannidis, J.P. Predictive ability of DNA microarrays for cancer outcomes and correlates: an empirical assessment. *Lancet* **2003**, *362*, 1439–1444.
26. Pomeroy, S.; et al. Prediction of central nervous system embryonal tumor outcome based on gene expression. *Nature* **2002**, *415*, 436–442.
27. Lossos, I.; Czerwinski, D.; Alizadeh, A.; Wechsler, M.; Tibshirani, R.; Botstein, D.; Levy, R. Prediction of survival in diffuse large b-cell lymphoma based on the expression of six genes. *New Engl. J. Med.* **2004**, *350*, 1828–1837.
28. Freije, W.; Castro-Vargas, F.; Fang, Z.; Horvath, S.; Cloughesy, T.; Liao, L.; Mischel, P.; Nelson, S. Gene expression profiling of gliomas strongly predicts survival. *Cancer Res.* **2004**, *64*, 6503–6510.
29. Lamb, J.; Ramaswamy, S.; Ford, H.L.; Contreras, B.; Martinez, R.V.; Kittrell, F.S.; Zahnow, C.A.; Patterson, N.; Golub, T.R.; Ewen, M.E. A mechanism of cyclin D1 action encoded in the patterns of gene expression in human cancer. *Cell* **2003**, *114*, 323–334.
30. Chen, G.; Gharib, T.G.; Wang, H.; Huang, C.C.; Kuick, R.; Thomas, D.G.; Shedden, K.A.; Misek, D.E.; Taylor, J.M.G.; Giordano, T.J.; Kardia, S.L.R.; Iannettoni, M.D.; Yee, J.; Hogg, P.J.; Orringer, M.B.; Hanash, S.M.; Beer, D.G.

- Protein profiles associated with survival in lung adenocarcinoma. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 13537–13542.
31. Sreekumar, A.; et al. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature* **2009**, *457*, 910–915.
 32. Zhang, Z.; et al. Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. *Cancer Res.* **2004**, *64*, 5882–5890.
 33. TCGA, Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **2008**, *455*, 1061–1068.
 34. Parsons, D.W. et al. An integrated genomic analysis of human glioblastoma multiforme. *Science* **2008**, *321*, 1807–1812.
 35. Hanash, S. Disease proteomics. *Nature* **2003**, *422*, 226–232.
 36. Hanash, S. Integrated global profiling of cancer. *Nature Rev. Cancer* **2004**, *4*, 638–643.
 37. Balic, M.; Dandachi, N.; Hofmann, G.; Samonigg, H.; Loibner, H.; Obwaller, A.; van der Kooi, A.; Tibbe Arjan, G.J.; Doye Derald, V.; Terstappen Leon, W.M.M.; Bauernhofer, T. *Cytometry, Part B* **2005**, *68*, 25–30.
 38. Adams, A.A.; Okagbare, P.I.; Feng, J.; Hupert, M.L.; Patterson, D.; Gottert, J.; McCarley, R.L.; Nikitopoulos, D.; Murphy, M.C.; Soper, S.A. Highly efficient circulating tumor cell isolation from whole blood and label-free enumeration using polymer-based microfluidics with an integrated conductivity sensor. *J. Am. Chem. Soc.* **2008**, *130*(27), 8633–8641.
 39. Nagrath, S.; et al. Isolation of rare circulating tumour cells in cancer patients by microchip technology. *Nature* **2007**, *450*, 1235–1239.
 40. Maheswaran, S.; et al. Detection of mutations in EGFR in circulating lung-cancer cells. *N. Engl. J. Med.* **2008**, *359*, 366–377.

41. Kallenbach, N.R.; Ma, R.I.; Seeman, N.C. An immobile nucleic acid junction constructed from oligonucleotides. *Nature* **1983**, *305*, 829–831.
42. Rothmund, P.W.K. Folding DNA to create nanoscale shapes and patterns. *Nature* **2006**, *440*, 297–302.
43. Winfree, E.; Liu, F.; Wenzler, L.A.; Seeman, N.C. Design and self-assembly of two-dimensional DNA crystals. *Nature* **1998**, *394*, 539–544.
44. Benenson, Y.; Gil, B.; Ben-Dor, U.; Adar, R.; Shapiro, E. An autonomous molecular computer for logic control of gene expression. *Nature* **2004**, *429*, 423–429.
45. Shin, J.S.; Pierce, N.A. A synthetic DNA walker for molecular transport. *J. Am. Chem. Soc.* **2003**, *125*, 10834–10835.
46. Robinson, B.H.; Seeman, N.C. The design of a biochip—a self-assembling molecular-scale memory device. *Protein Eng.* **1987**, *1*, 295–300.
47. Sano, T.; Smith, C.L.; Cantor, C.R. Immuno-PCR: very sensitive antigen detection by means of specific antibody-DNA conjugates. *Science* **1992**, *258*, 120–122.
48. Zhou, H.; et al. Two-color, rolling-circle amplification on antibody microarrays for sensitive, multiplexed serum-protein measurements. *Genome Biol.* **2004**, *5*, R28.
49. Fredriksson, S.; et al. Multiplexed protein detection by proximity ligation for cancer biomarker validation. *Nat. Methods* **2007**, *4*, 327–329.
50. Nam, J.M.; Thaxton, C.S.; Mirkin, C.A. Nanoparticle-based bio-bar codes for the ultrasensitive detection of proteins. *Science* **2003**, *301*, 1884–1886.
51. Kattah, M.G.; Coller, J.; Cheung, R.K.; Oshidary, N.; Utz, P.J. HIT: a versatile proteomics platform for multianalyte phenotyping of cytokines, intracellular proteins and surface molecules. *Nature Med.* **2008**, *14*, 1284–1289.

52. Melkko, S.; Scheuermann, J.; Dumelin, C.E.; Neri, D. Encoded self-assembling chemical libraries, *Nat. Biotech.* **2004**, *5*, 568–574.
53. Blattman, J.; Greenberg, P. Cancer immunotherapy: a treatment for the masses. *Science* **2004**, *305*, 200–205.
54. Schumacher, T.N. T-Cell-receptor gene therapy. *Nat. Rev. Immunol.* **2002**, *2*, 512–519.
55. Morgan, R.A.; Dudley, M.E.; Wunderlich, J.R.; Hughes, M.S.; Yang, J.C.; Sherry, R.M.; Royal, R.E.; Topalian, S.L.; Kammula, U.S.; Restifo, N.P.; et al. Cancer Regression in Patients After Transfer of Genetically Engineered Lymphocytes. *Science* **2006**, *314*, 126–129.

Chapter 2

DNA-Encoded Antibody Libraries: A Unified Platform for Multiplexed Cell Sorting and Detection of Genes and Proteins

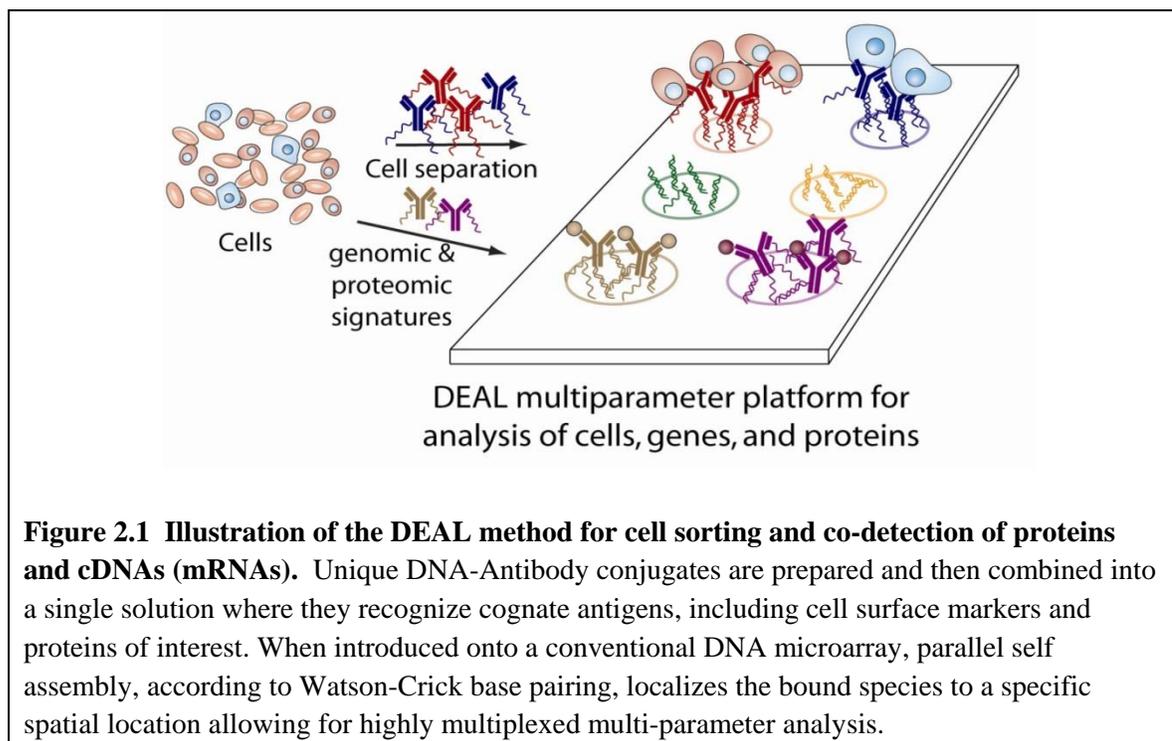
2.1 Introduction

Global genomic and proteomic analyses of tissues are impacting our molecular-level understanding of many human cancers. Particularly informative are studies that integrate both gene expression and proteomic data. Such multi-parameter data sets are beginning to reveal the perturbed regulatory networks which define the onset and progression of cancers (1–5). This new picture of cancer, and the emergence of promising new cancer drugs (6, 7) are placing new demands on clinical pathology (8). For example, traditional pathology practices (i.e. microscopic analysis of tissues) do not distinguish potential responders from non-responders for the new cancer molecular therapeutics (9). Recent examples exist in which pauciparameter molecular measurements are being employed to identify potential responders to at least two

therapeutics (10–13). However, it is unlikely that single-parameter measurements will be the norm. Instead, the coupling of molecular diagnostics with molecular therapeutics will eventually require measurements of a multi-parameter (e.g. cells, mRNAs and proteins) biomarker panel that can be used to direct patients to appropriate therapies or combination therapies.

Currently, the measurement of a multi-parameter panel of biomarkers from diseased tissues requires combinations of microscopic analysis, microarray data (14), immunohistochemical staining, western blots (8), and other methods. The collected data is integrated together within some model for the disease, such as a cancer pathway model (15), to generate a diagnosis. Currently, performing these various measurements requires a surgically resected tissue sample. The heterogeneity of such biopsies can lead to significant sampling errors since various measurements of cells, mRNAs, and proteins are each executed from different regions of the tissue.

In this chapter, the DNA-encoded antibody library, or DEAL, approach (**Figure 2.1**), is described as an important step towards executing a true multi-parameter analysis (cells, mRNAs and proteins) from the same microscopic region of tissue. We report on several key demonstrations for achieving this goal, including the rapid detection of proteins and protein panels over a broad dynamic range and with a detection limit of <10 femtoM; the sorting of immortal and primary lymphocyte populations; the co-detection of cells, cDNAs, and proteins on the same platform, and the integration of our multi-parameter platform with microfluidic techniques.



A key issue involved with a microfluidics-based multi-parameter assay is that the measurement of different classes of biomolecules (or cells) typically requires different surface chemistries, and not all of them are compatible with each other or the fabrication steps associated with building the microfluidics circuitry. Conventional antibody arrays for protein detection or for panning cells (16) require immobilization of the antibody onto aldehyde, epoxy, maleimide, or hydrophobic solid supports (17–20). It is often difficult to preserve folded (active) antibody conformations due to surface induced denaturation which depends on many variables including pH, ionic strength, temperature and concentration (21–23). This has spurred the development of alternative approaches to preserve the native conformation of proteins including 3-dimensional matrixes like hydrogels, and polyacrylamide (24, 25), cutinase-directed antibody immobilization onto SAMs (26), and the coupling of biotinylated antibodies onto streptavidin coated surfaces

(27). In addition, the arrays need to remain hydrated throughout the entire manufacturing process in order to prevent protein denaturation (18). DNA microarrays, on the other hand, are typically electrostatically absorbed (via spotting) onto amine surfaces. One option for detecting both DNA and proteins on the same slide would be to pattern both functional groups used to immobilize DNA and protein onto the same substrate, although this would significantly increase the complexity and engineering of the system. Alternatively, a compatible surface may be an activated ester glass slide to which amine-DNA and proteins can both covalently attach. However, we have found that the loading capacity of these slides for DNA is diminished, resulting in poor signal intensity when compared with DNA printed on conventionally prepared amine slides. In addition, unreacted esters are hydrolyzed back to carboxylic acids, which are negatively charged at normal hybridization buffers (pH 7), electrostatically reducing the DNA interaction. Moreover, to interrogate cells and proteins, the best surface to reduce non specific binding of cells while maintaining full antibody functionality is acrylamide (28, 29), which is incompatible with DNA.

By using DNA as a common assembly strategy for cells, cDNAs, and proteins, we are able to optimize the substrate conditions for high DNA loading onto the spotted substrates, and for complementary DNA loading on the antibodies. This leads to highly sensitive sandwich assays for protein detection, as well as high efficiency cell sorting (compared with traditional panning). We also find that non-selective binding (biofouling) of proteins to DNA-coated surfaces is reduced. Importantly, DNA coated surfaces can be dried out, stored or heated (overnight at 80° C), thus making them compatible with robust microfluidics fabrication.

DNA-labeled antibodies have been previously used to detect proteins (30–32), largely with the pendant oligomers serving as immuno-PCR tags (33, 34). DNA-tags have been used to direct the localization of proteins allowing assays to take advantage of spatial encoding, via several different read-out strategies (35–37). Conventional multi-well ELISA assays are capable of quantitating multiple proteins, but typically require separate sample volumes for each parameter. Optical multiplexing can expand this, but is limited by the number of non-spectrally overlapping chromophores. Spatial multiplexing, such as is used with DEAL, allows for the execution of many measurements on a small sample, since the number of different measurements is limited only by the patterning method utilized to prepare the cDNA array. Spotted antibody arrays (18), while potentially useful for protein detection and/or cell sorting, are not easily adaptable towards microfluidics-based assays, since the microfabrication process for preparing robust microfluidics devices often involves physical conditions that will damage the antibodies. Complementary DNA arrays are robust to such fabrication conditions.

2.2 Experimental Methods

2.2.1 DNA sequences for spatial encoding

All DNA strands were purchased with a 5'-amino modification from the Midland Certified Reagent company. Sequences for sequences A1, B1, C1 and their respective complements A1', B1' and C1' are given in Table 2.1. Computationally derived sequences, designated as A3, B3, C3 and their respective complements A3', B3' and C3' were designed following the paradigm outlined by Dirks et al. (38). Example input files

and output sequences can be found in the Appendix A. The sequences are reported in Table 2.1.

Name	Sequence
A1	5'-NH ₂ -AAAAAAAAAACGTGACATCATGCATG-3'
A1'	3'-GCACTGTAGTACGTACAAAAAAAAA-NH ₂ -5'
B1	5'-NH ₂ -AAAAAAAAAAGGATTCGCATACCAGT-3'
B1'	3'-CCTAAGCGTATGGTCAAAAAAAAAA-NH ₂ -5'
C1	5'-NH ₂ -AAAAAAAAATGGACGCATTGCACAT-3'
C1'	3'-ACCTGCGTAACGTGTAAAAAAAAA-NH ₂ -5'
A3	5'-NH ₂ -AAA AAA AAA A AT CCT GGA GCT AAG TCC GTA
A3'	5'-NH ₂ - AAA AAA AAA ATA CGG ACT TAG CTC CAG GAT
B3	5'-NH ₂ -AAA AAA AAA AGC CTC ATT GAA TCA TGC CTA
B3'	5'-NH ₂ -AAA AAA AAA AGC ACT CGT CTA CTA TCG CTA
C3	5'-NH ₂ -AAA AAA AAA AGC ACT CGT CTA CTA TCG CTA
C3'	5'-NH ₂ -AAA AAA AAA ATA GCG ATA GTA GAC GAG TGC

2.2.2 DNA antibody conjugation

AlexaFluor 488, 594, and 647-labeled polyclonal Goat anti-Human IgGs were purchased from Invitrogen. Monoclonal Rabbit anti-Human Interleukin-4 (clone: 8D4-8), non-fluorescent and APC-labeled Rabbit anti-Human Tumor Necrosis Factor- α (clones: MAb1 and MAb11, respectively), and non-fluorescent and PE-labeled Rabbit anti-Human Interferon- γ (clones: NIB42 and 4S.B3, respectively) were all purchased from eBioscience. Non-fluorescent and biotin-labeled mouse anti-Human Interleukin-2

(clones: 5344.111 and B33-2, respectively) were purchased from BD Biosciences. Prior to use, all antibodies were desalted, buffer exchanged to pH 7.4 PBS and concentrated to ~ 1mg/ml using 3000 MWCO spin filters (Millipore). Succinimidyl 4-hydrazinonicotinate acetone hydrazone in DMF (SANH, Solulink) was added to the antibodies at variable molar excess of (1000:1 to 5:1) of SANH to antibody. In this way the number of hydrazide groups introduced to the antibodies was varied. Separately, succinimidyl 4-formylbenzoate in DMF (SFB, Solulink) was added at a 20-fold molar excess to 5'aminated 26mer oligomers in PBS. This ratio of SFB to DNA ensured complete reaction of the 5' amine groups to yield 5' aldehydes. No further improvement in yield was observed for either the antibody and oligonucleotide coupling reactions after 4 hours at room temperature. Excess SANH and SFB were removed and samples buffered exchanged to pH 6.0 citrate buffer using protein desalting spin columns (Pierce). A 20-fold excess of derivatized DNA was then combined with the antibody and allowed to react overnight at room temperature. Non-coupled DNA was removed with size exclusion spin columns (Bio-Gel P-30, Bio-Rad) or purified using a Pharmacia Superdex 200 gel filtration column at 0.5 ml/min isocratic flow of PBS. The synthesis of DNA-antibody conjugates was verified by non-reducing 7.5% Tris-HCl SDS-PAGE at relaxed denaturing conditions of 60°C for 5 minutes, and visualized with a Molecular Imager FX gel scanner (Bio-Rad). Conjugation reactions involving fluorescent antibodies or fluorescent oligonucleotides were imaged similarly using appropriate excitation and emission filters.

2.2.3 Microarray Fabrication

DNA microarrays were printed via standard methods by the microarray facility at the Institute for Systems Biology (ISB—Seattle, WA) onto amine-coated glass slides. Typical spot size and spacing were 150 and 500 μm , respectively. Poly-lysine slides were made in house. Blank glass slides were cleaned with IPA and water in a sonication bath for 10 minutes each. They were then treated with oxygen plasma at 150 W for 60 sec., and then quickly dipped into DI water to produce a silanol terminated, highly hydrophilic surface. After drying them with a nitrogen gun, poly-L-lysine solution (Sigma P8920, 0.1% w/v, without dilution) was applied to the plasma treated surfaces for 15 minutes, and then rinsed off with DI water for several seconds. Finally, these treated slides were baked at 60°C for 1hr. These slides were then sent to ISB and printed as described above.

2.2.4 Fabrication of Microfluidic Devices

Microfluidic channels were fabricated from polydimethylsiloxane (PDMS) using conventional soft lithographic techniques. The goal was to fabricate robust microfluidics channels that could be disassembled after the surface assays were complete for optical analysis. Master molds were made photolithographically from a high resolution transparency mask (CadArt) so that the resulting fluidic network consisted of 20 parallel channels each having a cross-sectional profile of 10 x 600 μm and were 2 cm long. This corresponds to channel volumes of 120 nl. A silicone elastomer (Dow Corning Sylgard 184) was mixed and poured on top of the mold. After curing, the PDMS was removed from the mold and sample inlet and outlet ports punched with a 20 gauge steel pin

(Technical Innovations). The microfluidic channels were then aligned on top of the microarray and bonded to the substrate in an 80°C oven overnight.

2.2.5 1° Antibody Microarray Generation and DEAL-Based Immunoassays

Antibody microarrays were generated by first blocking the DNA slide with 0.1% BSA in 3x SSC for 30 minutes at 37°C. The slides were washed with dH₂O and blown dry. A 30 µl solution containing DNA-antibody conjugates (3x SSC, 0.1% SDS, 0.1% BSA, 15 ng/µl of each conjugate) was sandwiched to the array with a microscope slide, and incubated at 37°C for 4 hours. Arrays were then washed first in 1x SSC, 0.05% SDS at 37°C with gentle agitation, then at 0.2x SSC, then finally at 0.05x SSC. The slides were blown dry and scanned with a Gene Pix 4200 A two-color array scanner (Axon Instruments). For immunoassays, the DNA-encoded 1° antibody (15 ng/µl), antigen (3 ng/µl) and fluorescent 2° antibody (0.5 ng/µl) were combined in a single tube. After 2 hour incubation at 37°C, the formed antibody-antigen-antibody complexes were introduced to the microarrays as described above. Subsequent wash steps and visualization were identical.

2.2.6 Microfluidic-based assay procedures

Microfluidic devices were interfaced with 23 gauge steel pins and Tygon tubing to allow pneumatically controlled flow rates of ~0.5 µl/min. The assays were performed in Tris Buffered Saline (TBS), which was found to be better than 1x SSC and PBS. Each channel was blocked with 1.0% BSA in TBS prior to exposure to DNA-antibody conjugates or immunoassay pairs for 10 minutes under flowing conditions. After a 10 minute exposure to conjugates or antigens under flowing conditions, channels were

washed with buffer for 2 minutes and the microfluidics disassembled from the glass slide in order to be scanned. Immediately prior to imaging, the entire slide was briefly rinsed in TBS, blown dry and imaged on an array scanner as described above. For the human IL-2 concentration series, primary DNA-antibody conjugates were laid down first on the surface, before exposure to antigen and secondary antibody. This was necessary because at lower concentrations of antigen, the signals decrease, due to the high ratio of antigen-unbound primary antibody competing with antigen-bound primary for hybridization to the DNA array. By first exposing the array to the primary DNA-antibody conjugate, excesses were washed away before subsequent exposure to antigen and secondary antibody, increasing signal.

2.2.7 Microfluidic Au amplification methods

Microfluidics-based Au amplification experiments were performed in a similar manner, with the notable exception that a biotin-secondary antibody was used instead of a fluorescently labeled antibody. Subsequently, Au-streptavidin (Nanoprobes) was introduced into each channel (3ng/ μ l) for 10 minutes, after which the channels were thoroughly rinsed with buffer. After removal of the PDMS, the entire slide was then amplified with gold enhancer kit (Nanoprobes) according to manufacturer's protocol.

2.2.8 Analysis of DNA-encoded antibodies by flow cytometry

VL3 and A-20 cells were incubated for 20 min. on ice with 0.5 μ g of FITC-conjugated Rat Anti-Mouse CD90.2 (Thy1.2, BD Pharmingen, clone 30-H12, catalog # 553012) in 100 μ L PBS-3% FCS. Cells were also incubated with equimolar amounts of α -CD90.2/FITC-DNA conjugates characterized by various FITC-DNA loadings. Cells

were washed once with PBS-3% FCS and then were analyzed by flow cytometry on a BD FACSCanto™ instrument running the BD FACSDiva™ software.

2.2.9 Cell capture, separation, and sorting methods

Two murine cell lines, VL-3 T cells (thymic lymphoma line (39)) and A20 B cells (mouse B cell lymphoma (40), purchased from ATCC) were engineered to express mRFP and EGFP, respectively, using standard retroviral transduction protocols. Antibodies against surface markers for each of these cell lines, α -CD90.2 for VL-3 and α -B220 for A20 (eBioscience), were encoded as described above with DNA strands A1' and B1', respectively.

For sorting experiments, cells were passaged to fresh culture media [RPMI 1640 (ATCC) supplemented with 10 % fetal bovine serum, 0.1 mM non-essential amino acids and 0.05 mM β -mercaptoethanol at a concentration of 10^6 cells/100 μ l media and incubated with DNA-antibody conjugate (0.5 μ g/100 μ l) for 30 minutes on ice. Excess conjugate was removed from the supernatant after centrifugation, after which cells were resuspended in fresh media. Prior to cell incubation the microarray slide was passivated, to reduce non-specific cell adhesion, by reaction of the residual amine groups with methyl-PEO₁₂-NHS ester (Pierce) 10 mM in pH = 7.4 PBS for 4 hours at room temperature. Cells were spread evenly across the microarray surface and allowed to localize for one hour on ice. After this period, non-adherent cells were removed with gentle washing with room temperature Tris-buffered saline solution including 1 mM MgCl₂. Cell enrichment experiments were performed identically except that all incubation steps were performed in the presence of a 1:1 mixture of both T- and B-cells (each at 10^6 /100 μ l).

Primary CD4⁺ and CD8⁺ T cells were purified from EGFP and dsRed transgenic mice (obtained from Jackson Laboratories), respectively, using standard magnetic bead negative selection protocols and the BD IMag™ cell separation system. Prior to DEAL-based fractionation, the purity of these populations was analyzed by FACS and found to be greater than 80%.

Simultaneous cell, gene and protein experiments were performed similarly to those as previously described on a PEGylated microarray substrate. Briefly, GFP-expressing B cells ($10^6/100 \mu\text{l}$) were located on B1 spots after labeling with $\alpha\text{-B220-B1}'$ ($0.5 \mu\text{g}/100\mu\text{l}$). Following removal of non-adherent cells, a TNF- α ELISA pair with C1'-encoded 1^o and APC-labeled 2^o antibodies were introduced along with $0.5 \text{ ng}/\mu\text{l}$ FITC-labeled A1' and allowed to hybridize for a period of 30 minutes at room temperature. The slide was then rinsed with TBS+MgCl₂ and visualized via brightfield and fluorescence microscopy.

Homogeneous and panning cell experiments were performed in parallel. For the homogenous cell capture process, 5×10^6 Jurkats (ATCC) suspended in 1 ml of RPMI media along with $5 \mu\text{g}$ of $\alpha\text{-CD3/C3}'$ conjugates and incubated on ice for 1 hour. Excess conjugates were removed by centrifugation and the Jurkats were resuspended into $200 \mu\text{l}$ of fresh media before exposure to the DNA microarray. After 1 hour incubation on ice, the slides were rinsed gently with TBS. The cell panning experiments were performed in parallel; $5 \mu\text{g}$ of $\alpha\text{-CD3/C3}'$ conjugate in 1 ml RPMI media was incubated on a microarray for 1 hour on ice before rinsing in 0.5x PBS, then deionized water. The slide was not blown dry, but gently tapped on the side to remove the majority of the excess

solution, keeping the array hydrated. Jurkats ($5 \times 10^6/200 \mu\text{L}$) were immediately placed on the array for one hour on ice. Subsequent wash and visualization steps are identical.

2.3 Results and Discussion

2.3.1 *In silico* design of orthogonal DNA oligonucleotides

The pendant DNA oligonucleotides were designed *de novo* in order to minimize inter- and intra-strand cross hybridization. We followed the paradigm outlined by Dirks et al. (38) to computationally derive a set of orthogonal 30mers. These sequences were designed with a polyA₁₀ sequence followed by a variable 20mer encoding region. The polyA₁₀ stretch was incorporated to provide molecular flexibility and to prevent steric hindrance between the 20mer encoding region and the antigen binding domains of the antibody after conjugation. Three sequences, designated A3–C3 were generated using this approach and were tested empirically. Identical cDNA arrays printed with A3, B3, and C3 were probed with fluorescent complements A3' (green), A3'+B3' (red), and A3'+C3' (red). Minimal noise was observed between the probe sequences and noncomplementary spots (**Figure 2.2**). We performed initial DEAL experiments with sequences A1–C1 before determining that a rational design of the sequences was necessary to minimize noise. Therefore the majority of the experiments outlined in this chapter are presented using sequences A1–C1.

One advantage of using DNA oligonucleotides as molecular addresses is modularity. One working orthogonal set of sequences that has been experimentally validated can be used interchangeably with distinct sets of antibody libraries without modification of the underlying cDNA microarray. This feature allows

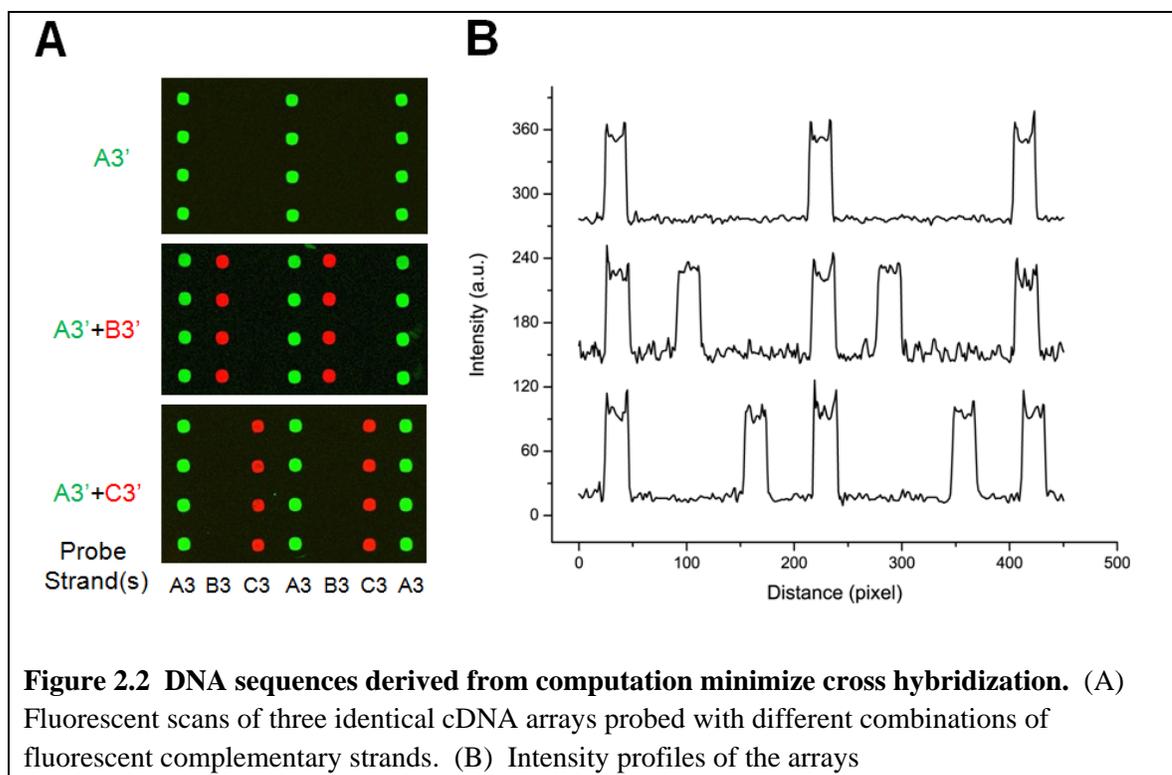
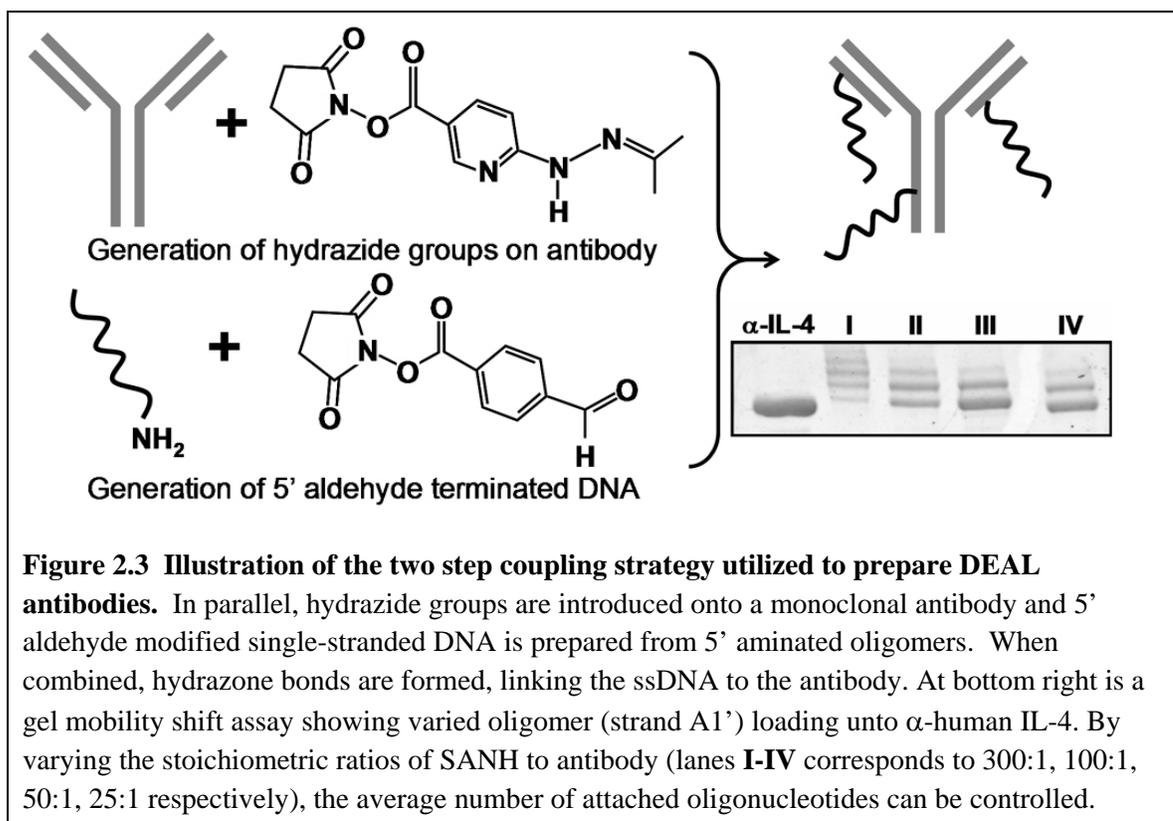


Figure 2.2 DNA sequences derived from computation minimize cross hybridization. (A) Fluorescent scans of three identical cDNA arrays probed with different combinations of fluorescent complementary strands. (B) Intensity profiles of the arrays

2.3.2 Generation of DNA-Antibody conjugates

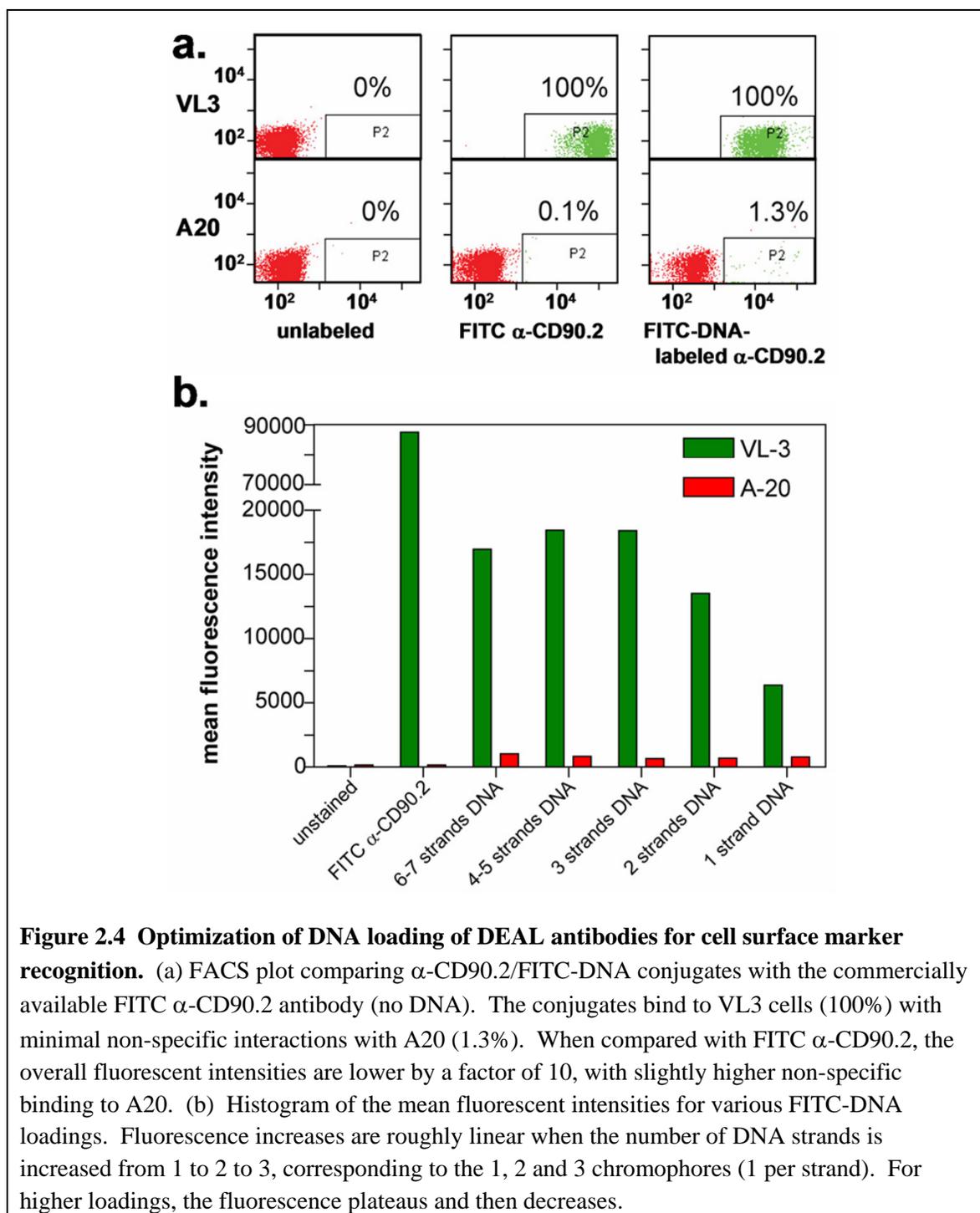
Chemically modified antibodies to aid in protein immobilization and/or detection are nearly universal for use in immunoassays. Such labeling introduces the risk of detrimentally affecting antibody function; however, that risk can be reduced by minimizing the size, and thus the steric hindrance, of the pendant moieties. With this in mind, we employed a covalent conjugation strategy in which 5'-aminated single-stranded oligonucleotides were coupled to antibodies via a hydrazone linkage (31), as shown in Scheme 2. Using commercially available reagents, an aldehyde functionality was introduced to the 5'-aminated oligonucleotide via succinimide chemistry. Similarly, a hydrazide moiety was introduced via reaction with the lysine side chains of the respective antibody. DNA-antibody conjugate formation was then facilitated via stoichiometric

hydrazone bond formation between the aldehyde and hydrazide functionalities. Conjugate formation and control over DNA-loading (41) was verified by PAGE electrophoresis, as shown in **Figure 2.3**.



Clearly the adverse steric effects of tagging antibodies with oligonucleotides are of concern when performing various assays, such as the immunoassays and cell sorting/capture experiments described herein. For this reason, we investigated the ability of DNA-encoded antibodies to retain recognition of cell surface markers, as visualized by fluorescence activated cell sorting (FACS). By using a fluorophore covalently tagged to the DNA, but *not* the antibody, FACS was used to optimize DNA-loading for the DEAL conjugates. For the analysis, 5' aminated, 3' FITC-labeled DNA as tagged unto α -CD90.2 antibodies at various stoichiometric ratios of SANH to antibody (5:1, 25:1, 50:1, 100:1, 300:1). This produced, on average, conjugates with 1, 2, 3, 4–5 and 6–7 strands of

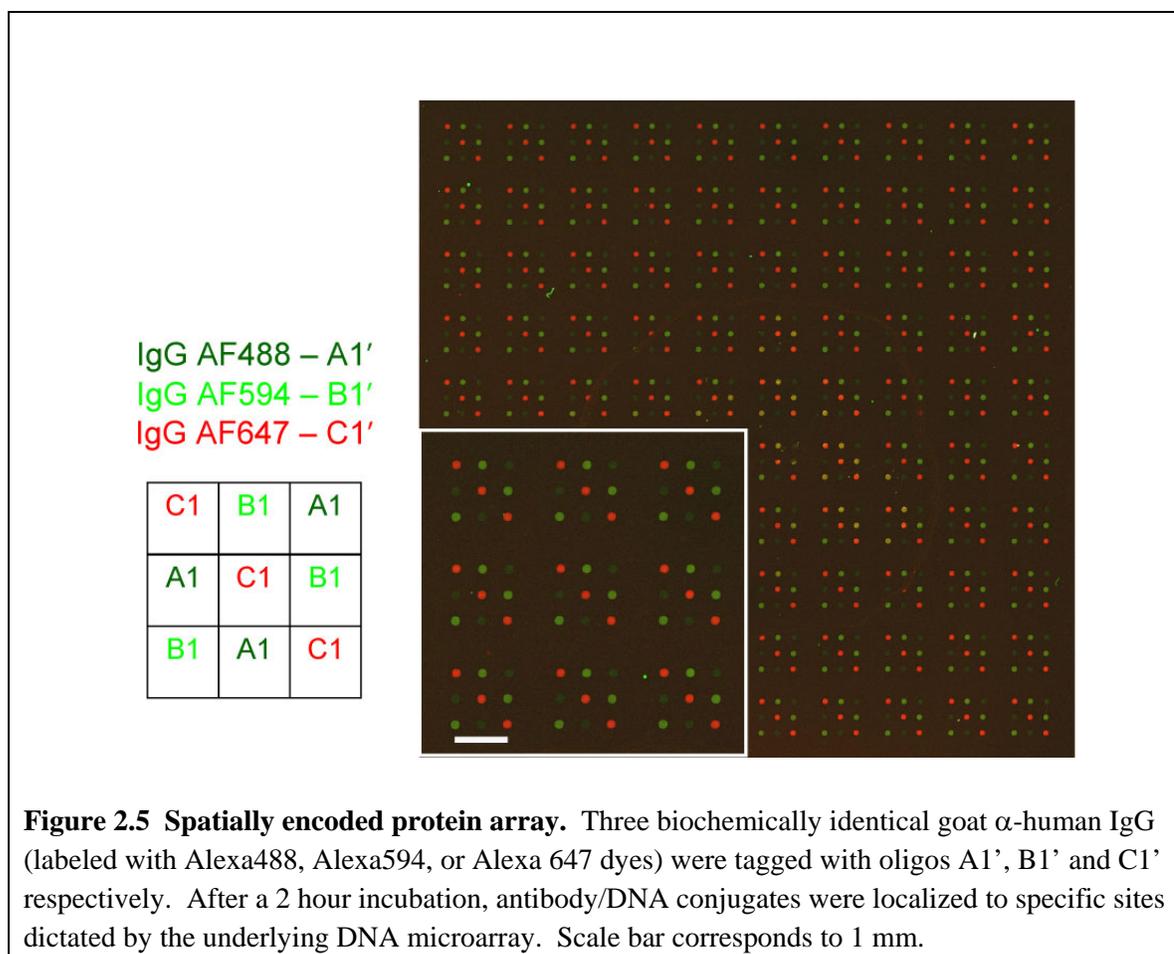
FITC-DNA respectively, as measured by gel mobility shift assays (**Figure 2.3**). These conjugates were tested for their ability to bind to the T cell line VL3 (CD90.2 expressing), by monitoring the FITC fluorescence with the flow cytometer. The B cell line A20 (CD90.2 negative) was used as a negative control. The performance of the conjugates was also compared with commercially available FITC α -CD90.2. The results are shown in **Figure 2.4**. The histogram of the mean fluorescent intensities for various FITC-DNA loadings shows that fluorescence increases are roughly linear when the number of DNA strands is increased from 1 to 2 to 3, corresponding to 1, 2 and 3 chromophores (1 per strand). At higher loadings, the increase in fluorescence first plateaus (4–5 oligomers) and then decreases up to the highest loading (6–7 oligomers). Thus, excess DNA labels (4–7 oligomers) did sterically reduce the ability of antibodies to recognize cell surface markers. Optimal loading for cell surface marker recognition was achieved with antibodies synthesized with the 50:1 SANH:antibody ratio, corresponding to approximately three DNA strands per antibody. Subsequent cell sorting experiments were performed in consideration of this observation. When compared with the FITC α -CD90.2 control, the DNA antibody conjugates had reduced fluorescence by a factor of 10 and slightly higher nonspecific binding to A20 cells.



This could be due to a couple of reasons. A likely factor is that the stoichiometric ratio of fluorophore to antibody for the DEAL conjugates versus the commercial antibody is different. For the DEAL conjugates, each strand of DNA is attached to one fluorophore only (i.e. conjugates with one DNA strand has a fluorophore to antibody ratio of 1:1) whereas the commercial antibodies generally have more than one fluorophore per antibody (i.e. fluorescent antibodies have a fluorophore to antibody ratio >1). Thus the factor of 10 less fluorescence should not be strictly interpreted as a 10x reduction in the binding affinity of the DEAL conjugates, although it is possible that the oligomer steric effects discussed earlier do account for some reduction in relative fluorescence intensity. Direct measurement of the affinity of the DEAL conjugate compared with the corresponding unmodified antibody using methods like Surface Plasmon Resonance (SPR) will be more conclusive.

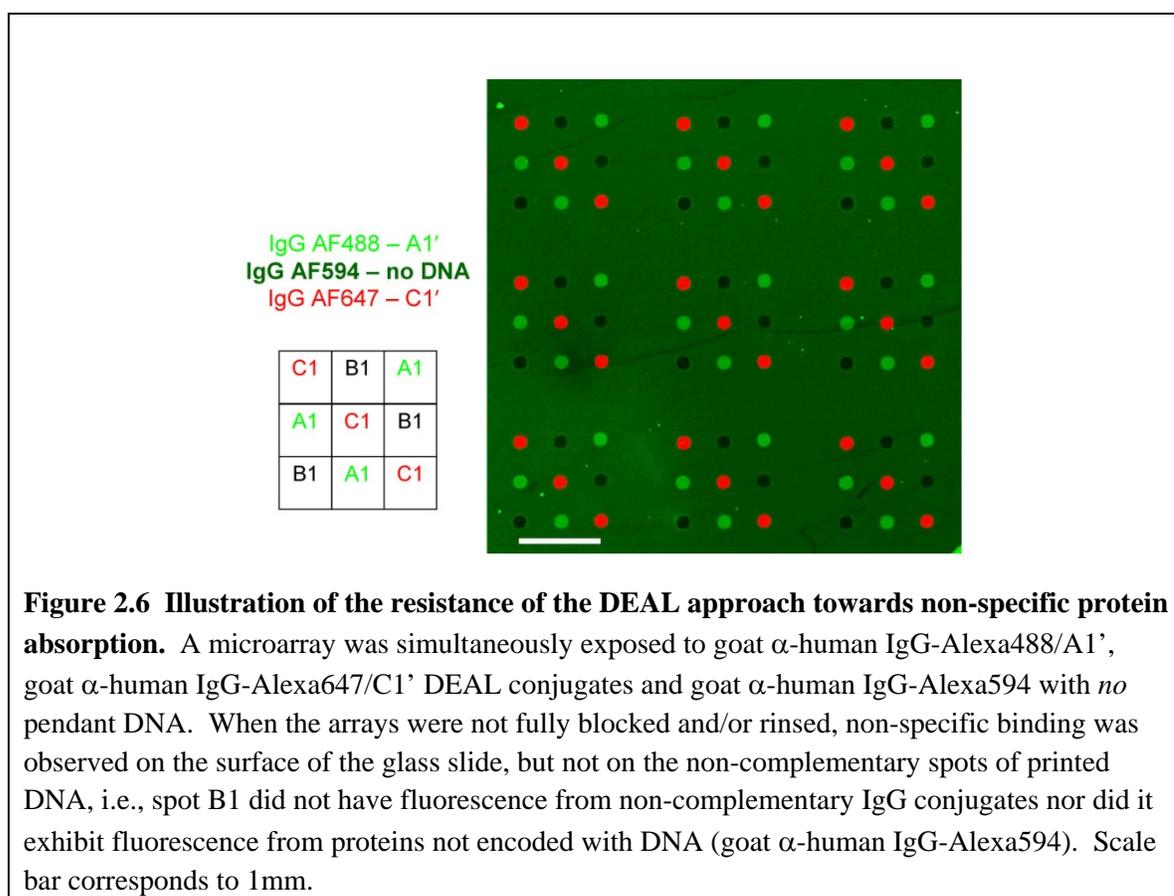
2.3.3 Multiplexed protein detection by DEAL

We demonstrated the DEAL concept for spatially localizing antibodies using three identical goat anti-human IgGs, each bearing a different molecular fluorophore and each encoded with a unique DNA strand. A solution containing all three antibodies was then introduced onto a microarray spotted with complementary oligonucleotides. After a two hour hybridization period and substrate rinse, the antibodies self-assembled according to Watson-Crick base-pairing, converting the >900 spot complementary DNA chip into a multi-element antibody microarray (**Figure 2.5**). This observation implied that quite large antibody arrays can be assembled in similar fashion.



The ultimate size of any protein array, however, will likely be limited by interference from non-specific binding of proteins. In an effort to visualize the contributions of non-specific binding, three antibodies were similarly introduced onto a microarray: two antibodies having complementary DNA-labeling spotted oligonucleotides and a third unmodified antibody (**Fig. 2.6**). For demonstration purposes, the slide was not thoroughly rinsed following hybridization and accordingly a high background signal due to non-specific adsorption of non-encoded fluorescently-labeled antibody was observed. The spotted nucleotide regions, to which no antibody was chemically encoded, displayed much less non-specifically attached protein, implying that

DNA greatly diminishes active area biofouling. Such retardation of biofouling is reminiscent of substrates that are functionalized with polyethyleneglycol (PEG) (41–43). By analogy with postulated mechanisms associated with PEG (44–46), we hypothesize that the hydrophilic nature of the spotted oligonucleotides minimizes interactions with hydrophobic portions of proteins often exposed during non-specific adsorption. Conjugate hybridization experiments were also carried out within 5 degrees of the calculated duplex melting temperatures, taking advantage of Watson-Crick stringencies and thus diminishing non-complementary DNA interactions. In any case, this reduced biofouling means that the DEAL method can likely be harnessed to detect reasonably large panels of proteins within a single environment.



2.3.4 Detection of multiple proteins within a single microfluidic channel

Microfluidic-based assays offer advantages such as reduced sample and reagent volumes, and shortened assay times (47). For example, under certain operational conditions, the surface binding assay kinetics are primarily determined by the analyte (protein) concentration and the analyte/antigen binding affinity, rather than by diffusion (48). We evaluated a microfluidics-based DEAL approach by bonding a polydimethylsiloxane (PDMS)-based microfluidic channel on top of a DNA microarray (**Figure 2.7A**). We initially performed a multiplexed antibody localization experiment, similar to that described above. The antibody conjugates self-assembled at precise spatial locations encoded by the pendant oligonucleotide in <10 minutes (**Figure 2.7B**), consistent with the time scales reported on DNA hybridization in microfluidics (49–51).

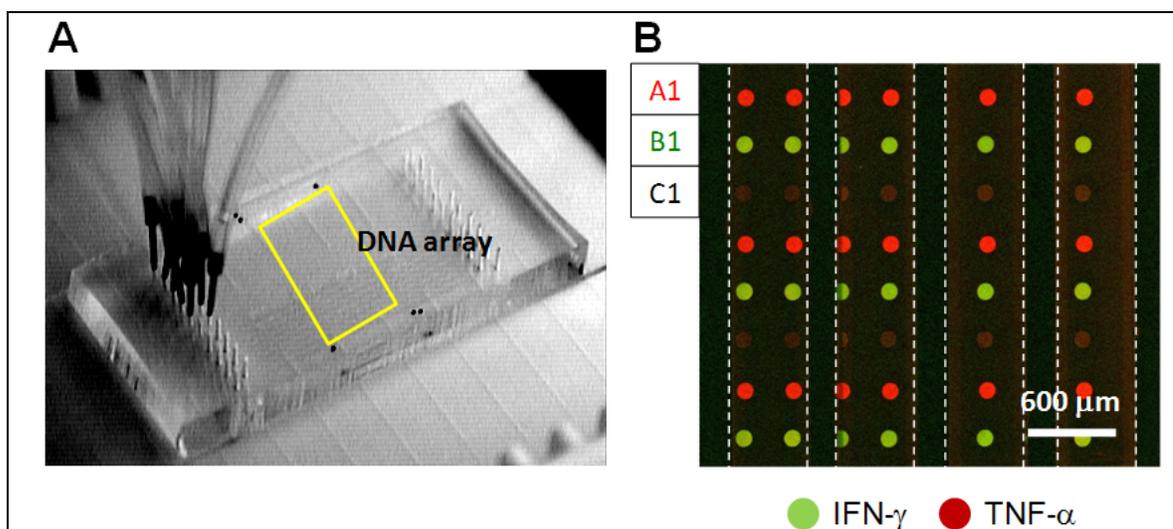
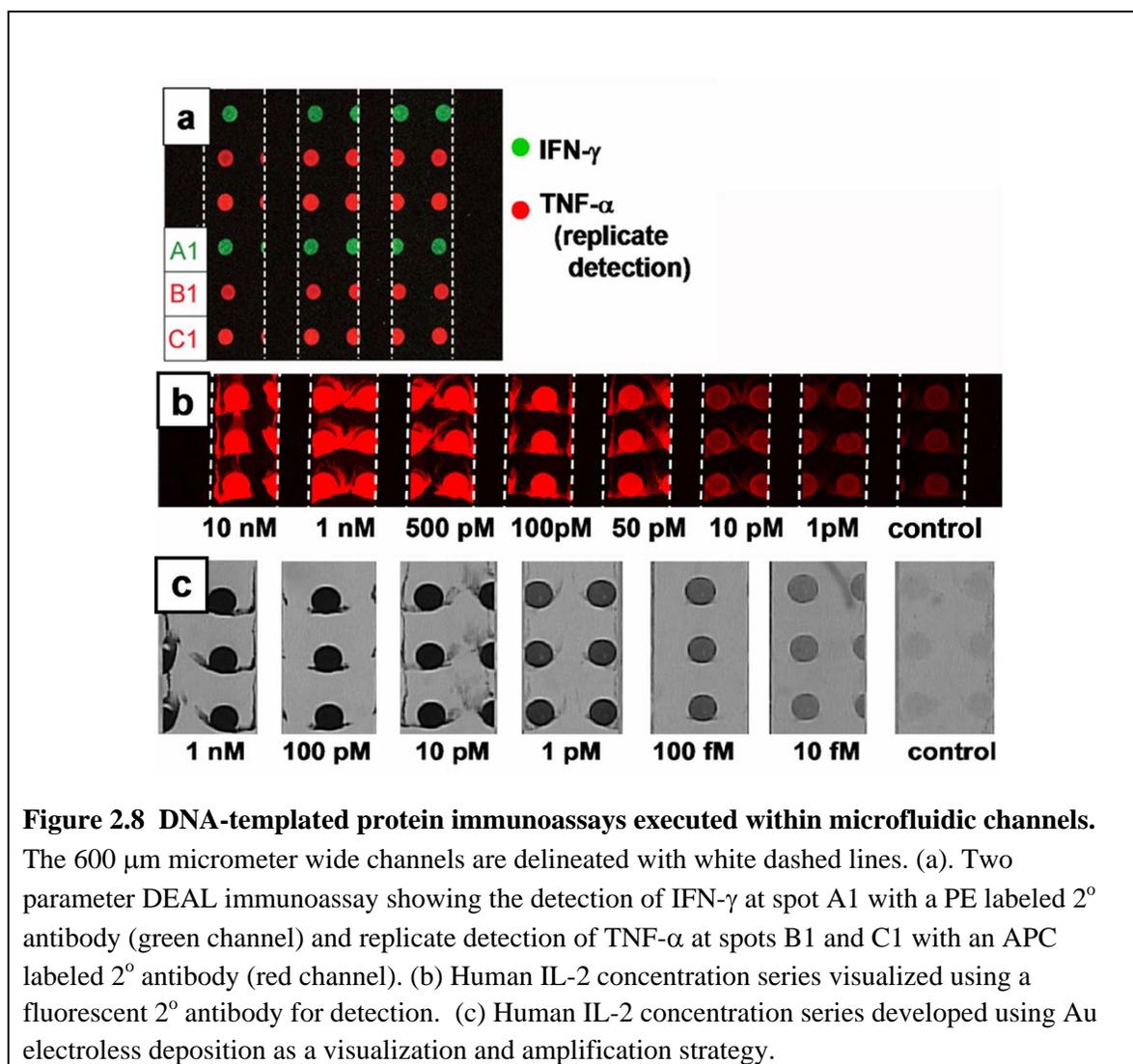


Figure 2.7 Protein array assembled in microfluidics in 10 minutes. (A) Picture of PDMS microfluidic device encapsulated a DNA array (yellow box) (B) Two goat α -human IgG (labeled with Alexa594 or Alexa 647) were tagged with oligos A1' and B1' respectively and introduced into a microfluidic device bonded on top of a DNA microarray with corresponding complementary strands A1 and B1 along with non-complementary strand C1. No DEAL conjugate encoded to spot C1 was added. After flowing at $\sim 0.5 \mu\text{l}/\text{min}$ for 10 minutes, the

microfluidic PDMS slab was removed and the glass slide imaged. The dashed lines delineate separate microfluidic channels of 600 μm width.

To validate the DEAL strategy for protein detection, we utilized encoded antibodies to detect cognate antigens in a variant of standard immunoassays. In a standard immunoassay (52), a primary antibody is adsorbed onto a solid support, followed by the sequential introduction and incubation of the antigen-containing sample and secondary labeled “read-out” antibody, with rinsing steps in between. In order to simplify this conventional five step immunoassay, we reasoned that the encoding power of the DEAL antibodies could serve to position the entire sandwich complex to the appropriate location for multiplexed readout, reducing the assay to a single step. To test this concept, in the same solution, a non-fluorescent, DNA-encoded 1^o antibody was combined with antigen and a fluorescently-labeled (no DNA) 2^o antibody. Under these conditions, a fluorescent signal will be spatially encoded only if an antibody-antigen-antibody sandwich is successfully formed in homogeneous solution and localized onto the microarray. Upon introduction of DNA-encoded antibodies against two cytokines, human IFN- γ and TNF- α , cognate antigens and fluorescently-labeled 2^o antibodies, the DEAL sandwich assays self-assembled to their specific spatial locations where they were detected, as shown in **Figure 2.8**. This multi-protein immunoassay also took 10 minutes to complete.



We explored the sensitivity limits of a microfluidics, DEAL-based sandwich immunoassay, using a third interleukin, IL-2. Using a fluorescent readout strategy, the assay peaked with a sensitivity limit of around 1 nM on slides printed at saturating concentrations of 5 μM of complementary DNA. Several strategies were employed to increase the sensitivity. First, we reasoned that increasing the loading capacity of the glass slide for DNA will increase the density of DEAL conjugates localized and therefore, increase the number of capture events possible. Conventional DNA microarrays are printed on primary amine surfaces generated by reacting amine-silane

with glass (53). DNA strands are immobilized through electrostatic interactions between the negative charges on the phosphate backbone of DNA and the positive charges from the protonated amines at neutral pH conditions. To increase the loading capacity of the slide, we generated poly-lysine surfaces, increasing both the charge density as well as the surface area of interaction with DNA. By adopting these changes, it became possible to print complementary DNA at saturating concentrations of 100 μM on the glass slides. Correspondingly, the sensitivity of the fluorescent based assays increased to 10 pM (**Figure 2.8b**). In addition, we chose to employ Au nanoparticle-labeled 2^o antibodies, followed by electroless metal deposition (54), to further amplify the signal and transform a fluorescence based read out to an optical one. This is possible since spatial, rather than colorimetric multiplexing, is utilized. Adopting these improvements, the presence of IL-2 interleukin can be readily detected at a concentration limit less than 10 fM (**Figure 2.8c**), representing at least a 1000-fold sensitivity increase over the fluorescence based microfluidics immunoassay. In comparison, this method is 100–1000-fold more sensitive than conventional ELISA (55), and 150 times more sensitive than the corresponding human IL-2 ELISA data from the manufacturer (56).

In performing these experiments, the idea of a 1 step immunoassay was revised. The sensitivity of the assays was reduced when performing a 1 step immunoassay, especially at lower concentrations of antigen. This is most likely due to competitive binding between DEAL conjugates with and without cargo for hybridization onto the underlying DNA microarray. By sequentially exposing the array to DEAL conjugate, antigen, and then secondary antibody, the sensitivities were increased. This is a clear trade off between convenience and sensitivity. It should still be stressed however, that

maximum signal is still reached under microfluidic flowing conditions within 10 minutes for each step. Thus in a fully automated device, a complete microfluidic immunoassay with sensitivities down to 10 fM can be obtained in 1 hour (including a 30 minute step for Au amplification).

In addition to the sample size and time-scale benefits that accompany this type of microfluidics immunoassay, there are other advantages. For example, since the entire assay is performed in solution prior to read-out, protein denaturation (a concern for spotted antibody microarrays) does not reduce binding efficiency. In addition, any assay that involves substrate-supported antibodies, would not have survived microfluidic chip assembly (which involved an extended bake at 80°C). That procedure was designed to yield robust PDMS microfluidics channels that could then be disassembled for the optical readout step. Another benefit of performing solution phase assays is that the orientational freedom enjoyed by both the antigens and antibodies ensures that the solid support will not limit the access of analytes to the binding pocket of the capture agent. We explore this issue in further detail below in the section of cell sorting. Other improvements, such as reducing the DNA spot size (57), and removing spot redundancy are currently being investigated to further lower detection limits.

2.3.5 Multiplexed sorting of immortalized and primary immune cells

We extended the DEAL technique for multiplexed cell sorting. The most common method for cell sorting is FACS, which is well-suited for many applications. Unfortunately, cells separated by conventional FACS are not immediately available for post-sorting analysis of gene and/or protein expression. In addition, FACS is also limited by the number of spectrally distinct fluorophores that can be utilized to label the cell

surface markers used for the sorting. FACS, however, is robust in sorting cells according to multiple cell surface markers. Amongst other alternative cell sorting strategies, the traditional panning method, in which cells interact with surface marker-specific antibodies printed onto an underlying substrate (58), is particularly relevant. Panning is capable of separating multiple cell populations, but has the same limitations as conventional spotted protein microarrays, namely that antibodies are not always oriented appropriately on a surface, and they can also dry out and lose functionality. DEAL overcomes this limitation, by keeping all reagents in solution.

We compared DEAL-based cell sorting with panning by evaluating homogeneous cell capture (solution phase cell capture) and heterogeneous capture of cells (surface confined cell capture). The homogeneous DEAL method exhibited higher cell capture efficiency as shown in **Figure 9a,b**. The increase in capture efficiency can be attributed to several factors. In homogeneous cell capture, the DEAL conjugates are allowed to properly orient and bind to the cell surface markers in solution. Cell capture is not driven by antibody to cell surface marker interactions, but rather by the increased avidity of the multivalent DEAL conjugates for the complementary DNA strands on the microarray through cooperative binding, greatly increasing capture efficiency. Similar trends have been reported for nanoparticle, DNA hybridization schemes (59). With this process, it is typical to see a DNA spot entirely occupied by a confluent layer of cells. With panning methods, which are analogous to our (heterogeneous) DEAL defined arrays, the capture agents are restricted to adopt a random orientation on the surface. The activity of the antibodies is reduced, simply because of improper orientation for interaction with the cell

surface markers, decreasing maximum avidity and cooperation with neighboring antibodies.

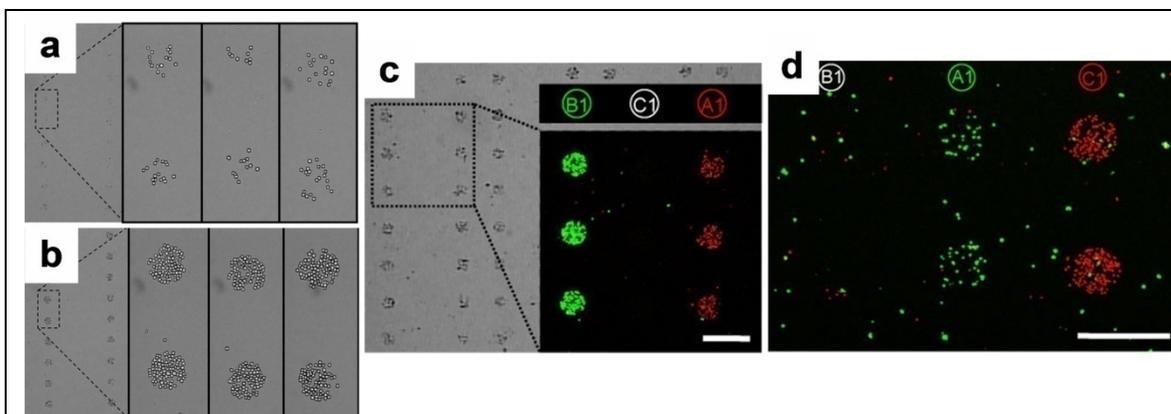


Figure 2.9 Optimization and use of DEAL for multiplexed cell sorting. Panels a and b are brightfield images showing the efficiency of the homogeneous DEAL cell capture process. (a) DEAL labeled antibodies are first assembled onto a spotted DNA array, followed by introduction of the cells. This heterogeneous process is similar to the traditional panning method of using surface bound antibodies to trap specific cells. (b) A homogeneous assay in which DEAL labeled antibodies are combined with the cells, and then the mixture is introduced onto the spotted DNA array microchip. This process is clearly much more efficient. Brightfield and fluorescence microscopy images of multiplexed cell sorting experiments where a 1:1 mixture of mRFP-expressing T cells (red channel) and EGFP-expressing B cells (green channel) is spatially stratified onto spots A1 and B1, corresponding to the encoding of α -CD90.2 and α -B220 antibodies with A1' and B1', respectively. (c) Fluorescence micrograph of multiplexed sorting of primary cells harvested from mice. A 1:1 mixture of CD4+ cells from EGFP transgenic mice and CD8+ cells from dsRed transgenic mice are separated to spots A1 and C1 by utilizing DEAL conjugates α -CD4-A1' and α -CD8-C1', respectively.

We also investigated the use of DEAL for multiplexed cell sorting. Two unique DNA strands were conjugated to antibodies raised against the T cell marker CD90.2 (Thy1.2) and the B cell marker CD45R (B220), respectively. Multiplexed DEAL-based cell sorting was demonstrated by spatially separating a 1:1 mixture of monomeric Red fluorescent protein (60) (mRFP)-expressing T cells (VL-3, murine thymic lymphoma) and EGFP-expressing B cells (mouse B cell lymphoma). This mixture was incubated

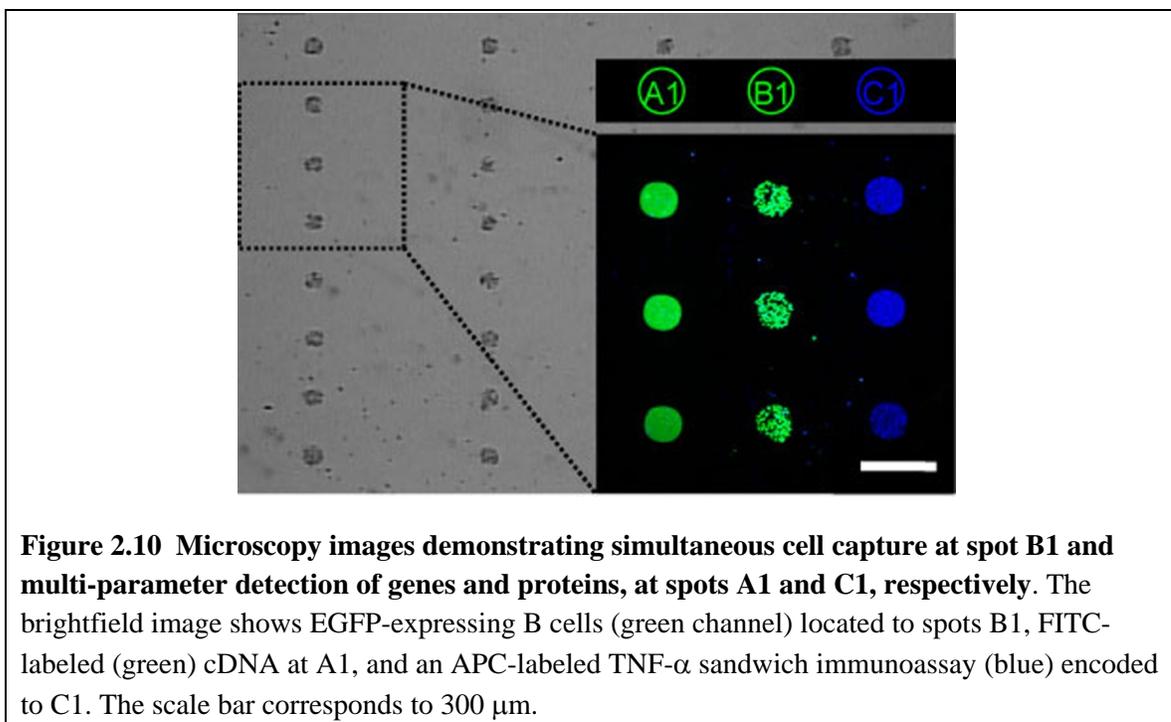
with uniquely encoded DNA-antibody conjugates against both T and B cell markers and introduced to an appropriately spotted microarray. **Figure 2.9c** shows both brightfield and false color fluorescence micrographs demonstrating that the mRFP-expressing T cells are enriched at spots A1 and EGFP-expressing B-cells located at B1, consistent with the DNA-encoding of the respective antibodies.

Primary cells are usually more fragile than established cell lines. This is due to the fact that they have to be extracted (usually by enzymatic digestions) from the surrounding tissues, a process that can lead to decreased viability. Moreover, the culture process often selects for clones characterized by greatly increased viability as well as proliferation potential. A generalized cell sorting technology must therefore also work on primary cells with minimal sample manipulation. To demonstrate the utility of DEAL for primary cell sorting, a synthetic mixture of CD4⁺ and CD8⁺ T cells was isolated via magnetic negative depletion from EGFP- and dsRED- transgenic mice, respectively. The mixture was stratified using α -CD4 and α -CD8 DNA-antibody conjugates. As shown in **Figure 2.9d**, the two cell types were separated to different spatial locations according to the pendant DNA encoding.

2.3.6 Single environment detection of specific cDNAs, proteins and cells

To highlight the universal diversity of this platform, GFP-expressing B cells were tagged with B1' DNA-encoded antibody conjugates and spatially located onto spots (B1) encoded with the complementary oligonucleotide. Post cell localization, FITC-labeled A1' DNA and a C1'-encoded TNF- α immunosandwich, were combined and introduced to the same microarray platform. The resulting brightfield and fluorescence microscopy

images, shown in **Figure 2.10**, demonstrate the validity of the DEAL platform for simultaneously extending across different levels of biological complexity.



2.4 Conclusions

By utilizing DNA as a universal linkage we have demonstrated a platform capable of simultaneous cell sorting, ssDNA and protein detection. DEAL represents a promising approach for the large scale, multi-parameter analysis of biological samples. We are currently applying DEAL towards the separation of highly complex primary cell mixtures such as whole mouse spleen and whole mouse thymus extracts. In addition, microfluidics-based DEAL immunoassays arrays are currently being harnessed for the analysis of protein biomarker panels from mouse whole-blood. We are particularly interested in integrating DEAL with advanced, on chip tissue handling tasks followed by simultaneous quantitation of mRNAs and proteins, because this is where DEAL can

potentially assist in pathological analysis of cancerous tissues. From a more fundamental cancer biology perspective, a near-term targeted application is the capture and functional evaluation of tumor-specific cytotoxic lymphocytes (28, 61). Such an application requires both rare cell capture, cell activation, and the subsequent detection of secreted proteins. For such problems, DEAL has the potential to eliminate any adverse effects of sample dilution and can thus greatly simplify the analysis of the biological system.

2.5 References

1. Lin, B.; White, J. T.; Lu, W.; Xie, T.; Utleg, A. G.; Yan, X.; Yi, E. C.; Shannon, P.; Khretbukova, I.; Lange, P. H.; Goodlett, D. R.; Zhou, D.; Vasicek, T. J.; Hood, L. Evidence for the presence of disease-perturbed networks in prostate cancer cells by genomic and proteomic analyses: a systems approach to disease. *Cancer Res.* **2005**, *65*, 3081–3091.
2. Kwong, K. Y.; Bloom, G. C.; Yang, I.; Boulware, D.; Coppola, D.; Haseman, J.; Chen, E.; McGrath, A.; Makusky, A. J.; Taylor, J.; Steiner, S.; Zhou, J.; Yeatman, T. J.; Quackenbush, J. Synchronous global assessment of gene and protein expression in colorectal cancer progression. *Genomics* **2005**, *86*, 142–158.
3. Huber, M.; Bahr, I.; Kratzchmar, J. R.; Becker, A.; Muller, E.-C.; Donner, P.; Pohlenz, H.-D.; Schneider, M. R.; Sommer, A. Comparison of proteomic and genomic analyses of the human breast cancer cell line T47D and the antiestrogen-resistant derivative T47D-r. *Molec. Cell. Proteomics* **2004**, *3*, 43–55.
4. Tian, Q.; Stepaniants, S. B.; Mao, M.; Weng, L.; Feetham, M. C.; Doyle, M. J.; Yi, E. C.; Dai, H.; Thorsson, V.; Eng, J.; Goodlett, D.; Berger, J. P.; Gunter, B.; Linseley, P. S.; Stoughton, R. B.; Aebersold, R.; Collins, S. J.; Hanlon, W. A.; Hood, L. E. Integrated genomic and proteomic analyses of gene expression in Mammalian cells. *Molec. Cell. Proteomics* **2004**, *3*, 960–969.
5. Chen, G.; Gharib, T. G.; Huang, C.-C.; Taylor, J. M. G.; Misek, D. E.; Kardia, S. L. R.; Giordano, T. J.; Iannettoni, M. D.; Orringer, M. B.; Hanash, S. M.; Beer, D. G. Discordant protein and mRNA expression in lung adenocarcinomas. *Molec. Cell. Proteomics* **2002**, *1*, 304–313.
6. Prados, M.; Chang, S.; Burton, E.; Kapadia, A.; Rabbitt, J.; Page, M.; Federoff, A.; Kelly, S.; Fyfe, G. *Proc. Am. Soc. Clin. Oncology* **2003**, *22*, 99.
7. Rich, J. N.; Reardon, D. A.; Peery, T.; Dowell, J. M.; Quinn, J. A.; Penne, K. L.; Wikstrand, C. J.; van Duyn, L. B.; Dancey, J. E.; McLendon, R. E.; Kao, J. C.; Stenzel, T. T.; Rasheed, B. K. A.; Tourt-Uhlig, S. E.; Herndon, J. E.; Vredenburgh, J. J.; Sampson, J. H.; Friedman, A. H.; Bigner, D. D.; Friedman, H. S. Phase II trial of gefitinib in recurrent glioblastoma. *J. Clin. Oncology* **2004**, *22*, 133–142.
8. Mellinghoff, I. K.; Wang, M. Y.; Vivanco, I.; Haas-Kogan, D. A.; Zhu, S.; Dia, E. Q.; Lu, K. V.; Yoshimoto, K.; Huang, J. H. Y.; Chute, D. J.; Riggs, B. L.; Horvath, S.; Liau, L. M.; Cavenee, W. K.; Rao, P. N.; Beroukhi, R.; Peck, T.

- C.; Lee, J. C.; Sellers, W. R.; Stokoe, D.; Prados, M.; Cloughesy, T. F.; Sawyers, C. L.; Mischel, P. S. Molecular Determinants of the Response of Glioblastomas to EGFR Kinase Inhibitors. *N. Engl. J. Med.* **2006**, *353*, 2012–2024.
9. Betensky, R. A.; Louis, D. N.; Cairncross, J. G. Influence of unrecognized molecular heterogeneity on randomized clinical trials. *J. Clin. Oncology* **2002**, *20*, 2495–2499.
 10. Hughes, T.; Branford, S., 2003. *Semin Hematol.* 2 Suppl 2, 62–68.
 11. Lamb, J.; Crawford, E. D.; Peck, D.; Modell, J. W.; Blat, I. C.; Wrobel, M. J.; Lerner, J.; Brunet, J. P.; Subramanian, A.; Ross, K. N.; Reich, M.; Hieronymus, H.; Wei, G.; Armstrong, S. A.; Haggarty, S. J.; Clemons, P. A.; Wei, R.; Carr, S. A.; Lander, E. S.; Golub, T. R. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **2006**, *313* (5795), 1929–1935.
 12. Martin, M. Molecular biology of breast cancer. *Clin. Transl Oncol.* *8* (1), 7–14.
 13. Radich, J. P.; Dai, H.; Mao, M.; Oehler, V.; Schelter, J.; Druker, B.; Sawyers, C. L.; Shah, N.; Stock, W.; Willman, C. L.; Friend, S.; Linsley, P. S. Gene expression changes associated with progression and response in chronic myeloid leukemia. *Proc. Natl. Acad. Sci.* **2006**, *103*(8), 2794–2799.
 14. Mischel, P. S.; Cloughesy, T. F.; Nelson, S. F. DNA-microarray analysis of brain cancer: molecular classification for therapy. *Nature Rev. Neuroscience* **2004**, *5*, 782–794.
 15. Weinberg, R. A., *Cancer Biology*. Garland Science: 2006.
 16. Wysocki, L. J.; Sato, V. L. "Panning" for lymphocytes: a method for cell selection. *Proc. Natl. Acad. Sci.* **1978**, *75*(6), 2844–2848.
 17. Liu, X.; Wang, H.; Herron, J.; Prestwich, G. Photopatterning of antibodies on biosensors. *Bioconjugate Chem.* **2000**, *11*, 755–761.
 18. Macbeath, G.; Schreiber, S. L. Printing proteins as microarrays for high-throughput function determination. *Science* **2000**, *289*, 1760–1763.
 19. Pal, M.; Moffa, A.; Sreekumar, A.; Ethier, S.; Barder, T.; Chinnaiyan, A.; Lubman, D. Differential phosphoprotein mapping in cancer cells using protein microarrays produced from 2-D liquid fractionation. *Anal. Chem.* **2006**, *78*, 702–710.

20. Thirumalapura, N. R.; Morton, R. J.; Ramachandran, A.; Malayer, J. R. Lipopolysaccharide microarrays for the detection of antibodies. *Journal of Immunological Methods* **2005**, *298*, 73–81.
21. Seigel, R. R.; Harder, P.; Dahint, R.; Grunze, M.; Josse, F.; Mrksich, M.; Whitesides, G. M. On-line detection of nonspecific protein adsorption at artificial surfaces. *Anal. Chem.* **1997**, *69*, 3321–3328.
22. Ramsden, J. J. Puzzles and paradoxes in protein adsorption. *Chem. Soc. Rev.* **1995**, *24*, 73–78.
23. Fainerman, V. B.; Lucassen-Reynders, E.; Miller, R. Adsorption of surfactants and proteins at fluid interfaces. *Colloids Surf. A* **1998**, *143*, 141.
24. Arenkov, P.; Kukhtin, A.; Gemmel, A.; Voloshchuk, S.; Chupeeva, V.; Mirzabekov, A. Protein microchips: use for immunoassay and enzymatic reactions. *Anal. Biochem.* **2000**, *278*, 123–131.
25. Kiyonaka, S.; Sada, K.; Yoshimura, I.; Shinkai, S.; Kato, N.; Hamachi, I. Semi-wet peptide/protein array using supramolecular hydrogel. *Nature Materials* **2004**, *3*, 58–64.
26. Kwon, Y.; Han, Z.; Karatan, E.; Mrksich, M.; Kay, B. K. Antibody arrays prepared by cutinase-mediated immobilization on self-assembled monolayers. *Anal. Chem.* **2004**, *76*, 5713–5720.
27. Peluso, P.; Wilson, D.; Do, D.; Tran, H.; Venkatasubbaiah, M.; Quincy, D.; Heidecker, B.; Poindexter, K.; Tolani, N.; Phelan, M.; Witte, K.; Jung, L.; Wagner, P.; Nock, S. Optimizing antibody immobilization strategies for the construction of protein microarrays. *Anal. Biochem.* **2003**, *312*, 113–124.
28. Chen, D. S.; Soen, Y.; Stuge, T. B.; Lee, P. P.; Weber, J. S.; Brown, P. O.; Davis, M. M. Marked differences in human melanoma antigen-specific T cell responsiveness after vaccination using a functional microarray. *PLoS Medicine* **2005**, *2*, 1018–1030.
29. 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 Biology* **2003**, *1*(3), 429–438.
30. Boozer, C.; Ladd, J.; Chen, S.; Yu, Q.; Homola, J.; Jiang, S. DNA directed protein immobilization on mixed ssDNA/oligo(ethylene glycol) self-assembled monolayers for sensitive biosensors. *Anal. Chem.* **2004**, *76*, 6967–6972.

31. Kozlov, I. A.; Melnyk, P. C.; Stromborg, K. E.; Chee, M. S.; Barker, D. L.; Zhao, C. Efficient strategies for the conjugation of oligonucleotides to antibodies enabling highly sensitive protein detection. *Biopolymers* **2004**, *73*, 621–630.
32. Adler, M.; Wacker, R.; Booltink, E.; Manz, B.; Niemeyer, C. M. Detection of femtogram amounts of biogenic amines using self-assembled DNA-protein nanostructures. *Nature Methods* **2005**, *2*, 147–149.
33. Sano, T.; Smith, C. L.; Cantor, C. R. Immuno-PCR: very sensitive antigen detection by means of specific antibody-DNA conjugates. *Science* **1992**, *258*, 120–122.
34. Niemeyer, C. M.; Adler, M.; Wacker, R. Immuno-PCR: high sensitivity detection of proteins by nucleic acid amplification. *TRENDS Biotech* **2005**, *23*, 208–216.
35. Wacker, R.; Niemeyer, C. M. DDI-microFIA--A readily configurable microarray-fluorescence immunoassay based on DNA-directed immobilization of proteins. *ChemBioChem* **2004**, *5*, 453–459.
36. Becker, C. F. W.; Wacker, R.; Bouschen, W.; Seidel, R.; Kolaric, B.; Lang, P.; Schroeder, H.; Muller, O.; Niemeyer, C. M.; Spengler, B.; Goody, R. S.; Engelhard, M. Direct readout of protein-protein interactions by mass spectrometry from protein-DNA microarrays. *Angew. Chem.* **2005**, *44*, 7635–7639.
37. Boozer, C.; Ladd, J.; Chen, S.; Jiang, S. DNA-Directed Protein Immobilization for Simultaneous Detection of Multiple Analytes by Surface Plasmon Resonance Biosensor. *Anal. Chem.* **2006**, *78*, 1515–1519.
38. Dirks, R. M.; Lin, M.; Winfree, E.; Pierce, N. A. Paradigms for computational nucleic acid design. *Nucleic Acids Research* **2004**, *32(4)*, 1392–1403.
39. Groves, T.; Katis, P.; Madden, Z.; Manickam, K.; Ramsden, D.; Wu, G.; Guidos, C. J. In vitro maturation of clonal CD4+CD8+ cell lines in response to TCR engagement. *J. Immunol.* **1995**, *154*, 5011–5022.
40. Kim, K. J.; Langevin, C. K.; Merwin, R. M.; Sachs, D. H.; Asfsky, R. Establishment and characterization of BALB/c lymphoma lines with B cell properties. *J. Immunol.* **1979**, *122*, 549–554.
41. This approach to conjugate synthesis is expected to result in a distribution of DNA loadings for each antibody, however, we feel as this effect is exaggerated in preparation for PAGE analysis. We observed that normal conditions for the heat-induced denaturation proceeding gel electrophoresis (100° for 5 minutes) reduced the number of DNA-strands visualized, presumably by breaking the hydrazone

linkage between the DNA and the protein. By relaxing the denaturing conditions, a sample heated at 60° for 5 minutes (minimum required for good gel) showed up to 7 discrete bands, whereas the same sample heated at 100° for 5 minutes showed no pendant oligonucleotides.

42. Prime, K. L.; Whitesides, G. M. Self-assembled organic monolayers: model systems for studying adsorption of proteins at surfaces. *Science* **1991**, *252*, 1164–1167.
43. Prime, K. L.; Whitesides, G. M. Adsorption of proteins onto surfaces containing end-attached oligo(ethylene oxide): a model system using self-assembled monolayers. *J. Am. Chem. Soc.* **1993**, *115*(23), 10714 – 10721.
44. Jeon, S. I.; Lee, J. H.; Andrade, J. D.; De Gennes, P. G. Protein-surface interactions in the presence of polyethylene oxide. I. Simplified theory. *Journal of Colloid and Interface Science* **1991**, *142*(1), 149–158.
45. Jeon, S. I.; Andrade, J. D. Protein-surface interactions in the presence of polyethylene oxide. II. Effect of protein size. *Journal of Colloid and Interface Science* **1991**, *142*(1), 159–166.
46. Andrade, J. D.; Hlady, V. Protein adsorption and materials biocompatibility: a tutorial review and suggested hypotheses. *Advances in Polymer Science* **1986**, *79*, (1–63).
47. Breslauer, D. N.; Lee, P. J.; Lee, L. P. Microfluidics-based systems biology. *Mol. Biosyst.* **2006**, *2*, 97–112.
48. Zimmermann, M.; Delamarche, E.; Wolf, M.; Hunziker, P. Modeling and optimization of high-sensitivity, low-volume microfluidic-based surface immunoassays. *Biomedical Microdevices* **2005**, *7*(2), 99–110.
49. Erickson, D.; Li, D.; Krull, U. Modeling of DNA hybridization kinetics for spatially resolved biochips. *Anal. Biochem.* **2003**, *317*, 186–200.
50. Bunimovich, Y.; Shin, Y.; Yeo, W.; Amori, M.; Kwong, G.; Heath, J. Quantitative real-time measurements of DNA hybridization with alkylated nonoxidized silicon nanowires in electrolyte solution. *J. Am. Chem. Soc.* **2006**, *128*(50), 16323–16331.
51. Wei, C.; Cheng, J.; Huang, C.; Yen, M.; Young, T. Contextual interactions determine whether the *Drosophila* homeodomain protein, Vnd, acts as a repressor or activator. *Nucleic Acids Research* **2005**, *33*(8), 1–11.

52. Engvall, E.; Perlmann, P. O. Enzyme-linked immunosorbent assay, Elisa. 3. Quantitation of specific antibodies by enzyme-labeled anti-immunoglobulin in antigen-coated tubes. *J. Immunol.* **1972**, *109*, 129–135.
53. Pirrung, M. How to make a DNA chip. *Angew. Chem. Int. Ed.* **2002**, *41*, 1276–1289.
54. Hainfeld, J. F.; Powell, R. D., Silver- and Gold-Based Autometallography of Nanogold. In *Gold and Silver Staining: Techniques in Molecular Morphology*, Hacker, G. W.; Gu, J., Eds. CRC Press: Washington, DC, 2002; pp 29–46.
55. Crowther, J. R., ELISA; Theory and Practice. In *Methods in Molecular Biology*, Humana Press Inc.: Totowa, New Jersey, 1995.
56. <http://www.bdbiosciences.com/ptProduct.jsp?prodId=6725>.
57. Thibault, C.; Le Berre, V.; Casimirius, S.; Tervisiol, E.; Francois, J.; Vieu, C. Examination of Cholesterol oxidase attachment to magnetic nanoparticles. *Journal of Nanobiotechnology* **2005**, *3*(7), 1–12.
58. Cardoso, A. A.; Watt, S. M.; Batard, P.; Li, M. L.; Hatzfeld, A.; Geneviev, H.; Hatzfeld, J. An improved panning technique for the selection of CD34+ human bone marrow hematopoietic cells with high recovery of early progenitors. *Exp. Hematol.* **1995**, *23*, 407–412.
59. Taton, T. A.; Mirkin, C. A.; Letsinger, R. L. Scanometric DNA array detection with nanoparticle probes. *Science* **2000**, *289*, 1757–1760.
60. Campbell, R. E.; Tour, O.; Palmer, A. E.; Steinbach, P. A.; Baird, G. S.; Zacharias, D. A.; Tsien, R. Y. A monomeric red fluorescent protein. *Proc. Natl. Acad. Sci.* **2002**, *99*, 7877–7882.
61. 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 Biology* **2003**, *1*, 429–438.

2.6 Appendix A: Computational derivation of Orthogonal

DNA oligomers

There are several applications of the computational algorithm developed by Dirks et al. (38). First it can accept a list of sequences (A, B, C, ... n) and return with an exhaustive file listing the relative interaction strengths between any two DNA sequences. This value, reported as $n(s^*)$, roughly represents the orthogonality of the two sequences that are being compared. As an example, sequences A1, B1 and C1 (inputs as A, B, and C respectively) were analyzed and the results are listed in **Appendix 2.6.1**. Here, the interaction strength of sequence B with B, representing intra-strand interactions, has the lowest $n(s^*)$ value of 7.525723, and thus is the most orthogonal pair. In comparison, the interaction strength of A_n with A_n (n represents the complement operator) has the highest $n(s^*)$ value of 16.406083 and thus is the least orthogonal pair. A good measure of the global orthogonality of a set of sequences is determined by the set with the lowest $\sum n(s^*)$.

Besides analysis, a set of sequences can be generated by inputting a set of constraints (e.g. sequence length, defined sequences, etc.) and the program will return with a set of sequences ranked according to $n(s^*)$ adhering to the given constraints. An example of this is given by the input file shown in **Appendix 2.6.2**, where the input code asks for 3 orthogonal sequences (A, B, and C) such that each sequences begins with a polyA₁₀ header before a variable 20mer region. The set of sequences (truncated to show only 3 sets of 10 total) with the lowest $n(s^*)$ value of 149.225 was taken and defined to be A3, B3 and C3, the sequences used in this chapter (**Appendix 2.6.3**). This

process can be repeated iteratively to increase the number of orthogonal sequences. Example input (**Appendix 2.6.4**) and output (**Appendix 2.6.5**) is given for the computation of a fourth DNA sequence.

2.6.1 Computational analysis of sequence A1, B1, and C1

Results sorted by $n(s^*)$ sum

```

A_A:
AAAAAAAAAACGTGACATCATGCATG+AAAAAAAAAACGTGACATCATGCATG 15.893007
.....+..... <- target
.....(((((((+.....))))))))) <- predicted
An_An:
AAAAAAAAAACATGCATGATGTCACG+AAAAAAAAAACATGCATGATGTCACG 16.406083
.....+..... <- target
.....(((((((.....+.....)))))))))..... <- predicted
An_Bn:
AAAAAAAAAACATGCATGATGTCACG+AAAAAAAAAACTGGTATGCGAATCC 9.679928
.....+..... <- target
.....+..... <- predicted
An_Cn:
AAAAAAAAAACATGCATGATGTCACG+AAAAAAAAAATGTGCAATGCGTCCA 11.736368
.....+..... <- target
.....+..... <- predicted
An_B:
AAAAAAAAAACATGCATGATGTCACG+AAAAAAAAAAGGATTTCGCATAACAGT 10.011576
.....+..... <- target
.....((((.....+.....))))..... <- predicted
An_C:
AAAAAAAAAACATGCATGATGTCACG+AAAAAAAAAATGGACGCATTGCACAT 13.267340
.....+..... <- target
.....+..... <- predicted
B_B:
AAAAAAAAAAGGATTTCGCATAACAGT+AAAAAAAAAAGGATTTCGCATAACAGT 7.525723
.....+..... <- target
.....+..... <- predicted
Bn_Bn:
AAAAAAAAAACTGGTATGCGAATCC+AAAAAAAAAACTGGTATGCGAATCC 8.707963
.....+..... <- target
.....((.....+.....))..... <- predicted
Bn_Cn:
AAAAAAAAAACTGGTATGCGAATCC+AAAAAAAAAATGTGCAATGCGTCCA 9.372607
.....+..... <- target
.....+..... <- predicted
Bn_A:
AAAAAAAAAACTGGTATGCGAATCC+AAAAAAAAAACGTGACATCATGCATG 9.717578
.....+..... <- target
.....((((.....+.....))))..... <- predicted
Bn_C:

```

```

AAAAAAAAAACTGGTATGCGAATCC+AAAAAAAAAATGGACGCATTGCACAT 12.147723
.....+..... <- target
.....((((.....+.....))))..... <- predicted
C_C:
AAAAAAAAAATGGACGCATTGCACAT+AAAAAAAAAATGGACGCATTGCACAT 14.639235
.....+..... <- target
.....((((.....+.....))))..... <- predicted
Cn_Cn:
AAAAAAAAAATGTGCAATGCGTCCA+AAAAAAAAAATGTGCAATGCGTCCA 14.117287
.....+..... <- target
.....((((.....+.....))))..... <- predicted
Cn_A:
AAAAAAAAAATGTGCAATGCGTCCA+AAAAAAAAAACGTGACATCATGCATG 12.522849
.....+..... <- target
.....+..... <- predicted
Cn_B:
AAAAAAAAAATGTGCAATGCGTCCA+AAAAAAAAAAGGATTTCGCATACCAGT 15.107010
.....+..... <- target
.....(((((((.....+.....))))..))))..... <- predicted
D:
C 0.000000
. <- target
. <- predicted
Total n(s*) = 180.852

```

2.6.2 Computing orthogonal sequences A, B, and C constrained by a polyA₁₀

header: input file

```

R: 10, 20, 20, 20      %%Defining R to contain 10, or 20 bases
R1: AAAAAAAAAA        %%Defining the first 10 bases to be polyA10

A_A: 1, 2, 1, 2       %%Sequence A is composed of R1 = polyA10, and R2
                       which is variable 20 mer
.....+.....
An_An: 1, -2, 1, -2
.....+.....
An_Bn: 1, -2, 1, -3
.....+.....
An_Cn: 1, -2, 1, -4
.....+.....
An_B: 1, 2, 1, 3
.....+.....
An_C: 1, 2, 1, 4
.....+.....
B_B: 1, 3, 1, 3
.....+.....
Bn_Bn: 1, -3, 1, -3
.....+.....
Bn_Cn: 1, -3, 1, -4
.....+.....
Bn_A: 1, -3, 1, 2

```

```

.....+.
Bn_C: 1, -3, 1, 4
.....+.
C_C: 1, 4, 1, 4
.....+.
Cn_Cn: 1, -4, 1, -4
.....+.
Cn_A: 1, -4, 1, 2
.....+.
Cn_B: 1, -4, 1, 3
.....+.

```

2.6.3 PolyA₁₀ header computational results

Results sorted by n(s*) sum

```

A_A:
AAAAAAAAAACGTGCCTACGGATCATTCTA+AAAAAAAAAACGTGCCTACGGATCATTCTA 14.752380
.....+. <- target
.....(((.....)))......+......(((.....)))...... <- predicted
An_An:
AAAAAAAAAATAGAATGATCCGTAGGCACG+AAAAAAAAAATAGAATGATCCGTAGGCACG 12.191140
.....+. <- target
.....+. <- predicted
An_Bn:
AAAAAAAAAATAGAATGATCCGTAGGCACG+AAAAAAAAAATACGAGCTACTAAGTGTCGG 10.821610
.....+. <- target
.....+. <- predicted
An_Cn:
AAAAAAAAAATAGAATGATCCGTAGGCACG+AAAAAAAAAATACCGAGTCAGGACCGTCGC 11.704642
.....+. <- target
.....(((.....)))......+...... <- predicted
An_B:
AAAAAAAAAATAGAATGATCCGTAGGCACG+AAAAAAAAAACGGACACTTAGTAGCTCGTA 12.387949
.....+. <- target
.....((((.....+......)))...... <- predicted
An_C:
AAAAAAAAAATAGAATGATCCGTAGGCACG+AAAAAAAAAAGCGACGGTCCTGACTCGGTA 13.792005
.....+. <- target
.....(((.....((.....+......))).))...... <- predicted
B_B:
AAAAAAAAAACGGACACTTAGTAGCTCGTA+AAAAAAAAAACGGACACTTAGTAGCTCGTA 15.189105
.....+. <- target
.....+. <- predicted
Bn_Bn:
AAAAAAAAAATACGAGCTACTAAGTGTCGG+AAAAAAAAAATACGAGCTACTAAGTGTCGG 10.362398
.....+. <- target
.....+. <- predicted
Bn_Cn:
AAAAAAAAAATACGAGCTACTAAGTGTCGG+AAAAAAAAAATACCGAGTCAGGACCGTCGC 11.009212
.....+. <- target
.....(((.....)))......+...... <- predicted
Bn_A:
AAAAAAAAAATACGAGCTACTAAGTGTCGG+AAAAAAAAAACGTGCCTACGGATCATTCTA 11.155809
.....+. <- target
.....+. <- predicted

```

```

Bn_C:
AAAAAAAAAATACGAGCTACTAAGTGTCCG+AAAAAAAAAAGCGACGGTCCTGACTCGGTA 12.599213
.....+..... <- target
.....((((.....+.....))))... <- predicted
C_C:
AAAAAAAAAAGCGACGGTCCTGACTCGGTA+AAAAAAAAAAGCGACGGTCCTGACTCGGTA 22.259374
.....+..... <- target
.....(.(((.....(((.....(((.....(.+.)))))))).)) <- predicted
Cn_Cn:
AAAAAAAAAATACCGAGTCAGGACCGTCGC+AAAAAAAAAATACCGAGTCAGGACCGTCGC 15.521820
.....+..... <- target
.....(.(.+.....).)..... <- predicted
Cn_A:
AAAAAAAAAATACCGAGTCAGGACCGTCGC+AAAAAAAAAACGTGCCTACGGATCATTCTA 12.952999
.....+..... <- target
.....((((.....+.....))))... <- predicted
Cn_B:
AAAAAAAAAATACCGAGTCAGGACCGTCGC+AAAAAAAAAACGGACACTTAGTAGCTCGTA 13.708535
.....+..... <- target
.....((((.....+.....))))... <- predicted
Total n(s*) = 200.408

...

A_A:
AAAAAAAAAATCAAGTAGCGAATGGACTA+AAAAAAAAAATCAAGTAGCGAATGGACTA 7.851744
.....+..... <- target
.....+..... <- predicted
An_An:
AAAAAAAAAATAGTCCATTCGCTACTTGAT+AAAAAAAAAATAGTCCATTCGCTACTTGAT 10.114856
.....+..... <- target
.....+..... <- predicted
An_Bn:
AAAAAAAAAATAGTCCATTCGCTACTTGAT+AAAAAAAAAATACCGTGAAGCGTCCTAGTA 11.335990
.....+..... <- target
.....((((.....+.....))))... <- predicted
An_Cn:
AAAAAAAAAATAGTCCATTCGCTACTTGAT+AAAAAAAAAATAGATACGATGGCTCAGTGC 11.862416
.....+..... <- target
.....(.+.....).)..... <- predicted
An_B:
AAAAAAAAAATAGTCCATTCGCTACTTGAT+AAAAAAAAAATACTAGGACGCTTCACGGTA 11.058647
.....+..... <- target
.....+..... <- predicted
An_C:
AAAAAAAAAATAGTCCATTCGCTACTTGAT+AAAAAAAAAAGCACTGAGCCATCGTATCTA 10.700609
.....+..... <- target
.....+..... <- predicted
B_B:
AAAAAAAAAATACTAGGACGCTTCACGGTA+AAAAAAAAAATACTAGGACGCTTCACGGTA 14.394146
.....+..... <- target
.....((((.....(.+.....))))..((.....))... <- predicted
Bn_Bn:
AAAAAAAAAATACCGTGAAGCGTCCTAGTA+AAAAAAAAAATACCGTGAAGCGTCCTAGTA 13.083925
.....+..... <- target
.....+..... <- predicted
Bn_Cn:
AAAAAAAAAATACCGTGAAGCGTCCTAGTA+AAAAAAAAAATAGATACGATGGCTCAGTGC 12.305284
.....+..... <- target
.....((((.....+.....))))...((.....).) <- predicted

```

```

Bn_A:
AAAAAAAAAATACCGTGAAGCGTCCTAGTA+AAAAAAAAAATCAAGTAGCGAATGGACTA 9.982085
.....+..... <- target
.....(.....+.....).. <- predicted
Bn_C:
AAAAAAAAAATACCGTGAAGCGTCCTAGTA+AAAAAAAAAAGCACTGAGCCATCGTATCTA 10.427249
.....+..... <- target
.....+..... <- predicted
C_C:
AAAAAAAAAAGCACTGAGCCATCGTATCTA+AAAAAAAAAAGCACTGAGCCATCGTATCTA 8.990198
.....+..... <- target
.....+..... <- predicted
Cn_Cn:
AAAAAAAAAATAGATACGATGGCTCAGTGC+AAAAAAAAAATAGATACGATGGCTCAGTGC 9.885630
.....+..... <- target
.....((.....(.....)+.....(.....)).....(.....) <- predicted
Cn_A:
AAAAAAAAAATAGATACGATGGCTCAGTGC+AAAAAAAAAATCAAGTAGCGAATGGACTA 9.740523
.....+..... <- target
.....+..... <- predicted
Cn_B:
AAAAAAAAAATAGATACGATGGCTCAGTGC+AAAAAAAAAATACTAGGACGCTTCACGGTA 10.394821
.....+..... <- target
.....+.....((.....))..... <- predicted
Total n(s*) = 162.128

...

A_A:
AAAAAAAAAATCCTGGAGCTAAGTCCGTA+AAAAAAAAAATCCTGGAGCTAAGTCCGTA 11.081137
.....+..... <- target
.....((.....(.....+.....)).....)..... <- predicted
An_An:
AAAAAAAAAATACGGACTTAGCTCCAGGAT+AAAAAAAAAATACGGACTTAGCTCCAGGAT 11.417329
.....+..... <- target
.....(.....+.....)..... <- predicted
An_Bn:
AAAAAAAAAATACGGACTTAGCTCCAGGAT+AAAAAAAAAATAGGCATGATTCAATGAGGC 9.031869
.....+..... <- target
.....+..... <- predicted
An_Cn:
AAAAAAAAAATACGGACTTAGCTCCAGGAT+AAAAAAAAAATAGCGATAGTAGACGAGTGC 9.887457
.....+..... <- target
.....+..... <- predicted
An_B:
AAAAAAAAAATACGGACTTAGCTCCAGGAT+AAAAAAAAAAGCCTCATTGAATCATGCCTA 9.893760
.....+..... <- target
.....+..... <- predicted
An_C:
AAAAAAAAAATACGGACTTAGCTCCAGGAT+AAAAAAAAAAGCACTCGTCTACTATCGCTA 9.416199
.....+..... <- target
.....+..... <- predicted
B_B:
AAAAAAAAAAGCCTCATTGAATCATGCCTA+AAAAAAAAAAGCCTCATTGAATCATGCCTA 10.372686
.....+..... <- target
.....+..... <- predicted
Bn_Bn:
AAAAAAAAAATAGGCATGATTCAATGAGGC+AAAAAAAAAATAGGCATGATTCAATGAGGC 11.143822
.....+..... <- target
.....((.....+.....))..... <- predicted

```

```

Bn_Cn:
AAAAAAAAAATAGGCATGATTCAATGAGGC+AAAAAAAAAATAGCGATAGTAGACGAGTGC 9.725931
.....+..... <- target
.....+..... <- predicted
Bn_A:
AAAAAAAAAATAGGCATGATTCAATGAGGC+AAAAAAAAAATCCTGGAGCTAAGTCCGTA 9.849003
.....+..... <- target
.....+..... <- predicted
Bn_C:
AAAAAAAAAATAGGCATGATTCAATGAGGC+AAAAAAAAAAGCACTCGTCTACTATCGCTA 10.070369
.....+..... <- target
.....+..... <- predicted
C_C:
AAAAAAAAAAGCACTCGTCTACTATCGCTA+AAAAAAAAAAGCACTCGTCTACTATCGCTA 8.656702
.....+..... <- target
.....+..... <- predicted
Cn_Cn:
AAAAAAAAAATAGCGATAGTAGACGAGTGC+AAAAAAAAAATAGCGATAGTAGACGAGTGC 9.905828
.....+..... <- target
.....+..... <- predicted
Cn_A:
AAAAAAAAAATAGCGATAGTAGACGAGTGC+AAAAAAAAAATCCTGGAGCTAAGTCCGTA 9.691171
.....+..... <- target
.....(((.....+.....))). <- predicted
Cn_B:
AAAAAAAAAATAGCGATAGTAGACGAGTGC+AAAAAAAAAAGCCTCATTGAATCATGCCTA 9.081401
.....+..... <- target
.....+..... <- predicted
Total n(s*) = 149.225

```

2.6.4 Computing a fourth sequence: input code

```

>Test #6
R: 10, 20, 20, 20, 20
R1: AAAAAAAAAA
R2: ATCCTGGAGCTAAGTCCGTA
R3: GCCTCATTGAATCATGCCTA
R4: GCACTCGTCTACTATCGCTA

D_D: 1, 5, 1, 5
.....+.....
Dn_A: 1, -5, 1, 2
.....+.....
Dn_B: 1, -5, 1, 3
.....+.....
Dn_C: 1, -5, 1, 4
.....+.....
Dn_An: 1, -5, 1, -2
.....+.....
Dn_Bn: 1, -5, 1, -3
.....+.....
Dn_Cn: 1, -5, 1, -4
.....+.....
Dn_Dn: 1, -5, 1, -5
.....+.....

```

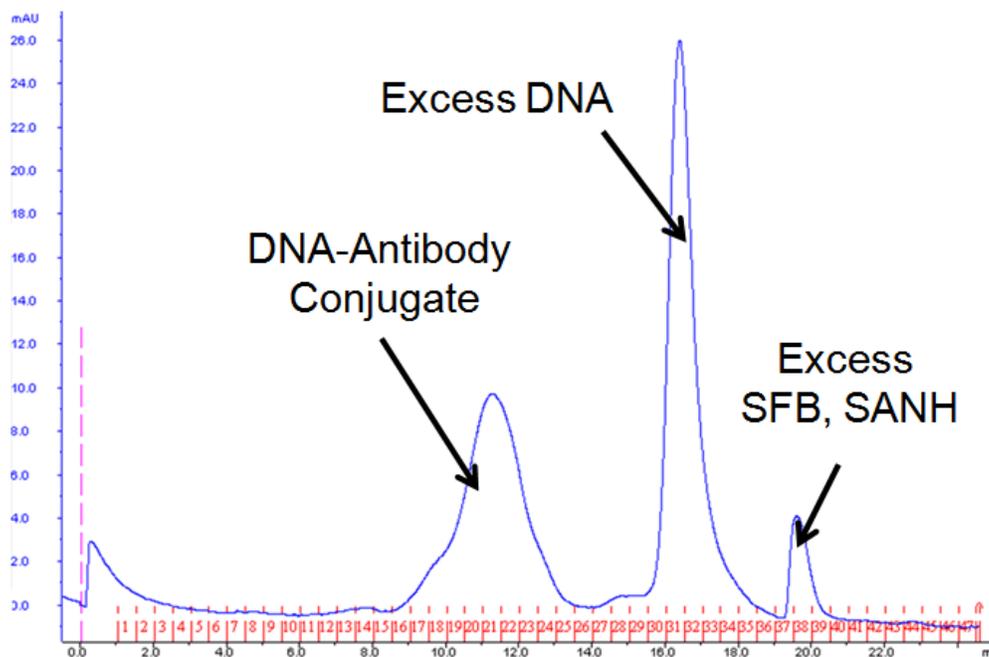
2.6.5 Results of fourth strand computation

```

D_D:
AAAAAAAAAAATGGTCGAGATGTCAGAGTA+AAAAAAAAAAATGGTCGAGATGTCAGAGTA 10.604201
.....+..... <- target
.....((.....+.....))..... <- predicted
Dn_A:
AAAAAAAAAAATACTCTGACATCTCGACCAT+AAAAAAAAAAATCCTGGAGCTAAGTCCGTA 7.984174
.....+..... <- target
.....+..... <- predicted
Dn_B:
AAAAAAAAAAATACTCTGACATCTCGACCAT+AAAAAAAAAAAGCCTCATTGAATCATGCCTA 6.890737
.....+..... <- target
.....+..... <- predicted
Dn_C:
AAAAAAAAAAATACTCTGACATCTCGACCAT+AAAAAAAAAAAGCACTCGTCTACTATCGCTA 6.684232
.....+..... <- target
.....+..... <- predicted
Dn_An:
AAAAAAAAAAATACTCTGACATCTCGACCAT+AAAAAAAAAAATACGGACTTAGCTCCAGGAT 7.998166
.....+..... <- target
.....+..... <- predicted
Dn_Bn:
AAAAAAAAAAATACTCTGACATCTCGACCAT+AAAAAAAAAAATAGGCATGATTCAATGAGGC 9.261694
.....+..... <- target
.....+..... <- predicted
Dn_Cn:
AAAAAAAAAAATACTCTGACATCTCGACCAT+AAAAAAAAAAATAGCGATAGTAGACGAGTGC 11.865802
.....+..... <- target
.....(((.....+.....)))..... <- predicted
Dn_Dn:
AAAAAAAAAAATACTCTGACATCTCGACCAT+AAAAAAAAAAATACTCTGACATCTCGACCAT 6.770436
.....+..... <- target
.....((.....+.....))..... <- predicted
Total n(s*) = 68.059

```

2.7 Appendix B: FPLC of DEAL conjugates.



*0.5 mL/min isocratic flow of PBS

Figure 2.11 Fast protein liquid chromatography of DEAL conjugates. A typical successful conjugation reaction will yield three distinct peaks, corresponding to the DEAL conjugates, excess unreacted ssDNA, and excess small molecules.

Chapter 3

Modular Nucleic Acid Assembled p/MHC Microarrays for Multiplexed Sorting of Antigen- Specific T Cells

2.1 Introduction

T cells constitute an important part of the acquired immune system. They recognize a diversity of antigens through the highly variable, hetero-dimeric T cell receptor protein (TCR), with approximately 10^7 different antigen specificities (1). The initiation of the T cell immune response is triggered by the engagement of the TCR with processed antigenic peptides (e.g. from a bacterial pathogen) that are bound to Major Histocompatibility Complex (p/MHC) molecules presented on the surface of antigen presenting cells (APCs), leading to downstream T cell proliferation and maturation into effector populations. After pathogen clearance, a subset of the activated T cells transition into memory cells, providing the immune system with the capacity for rapid response

towards previously encountered pathogens. As a consequence, an individual's collection of T cells and their antigen specificities, collectively called the T cell repertoire, is an evolving, extensive repository of cellular immune responses against self and foreign antigens. It is of fundamental and therapeutic importance to detect and survey these T cell populations.

The development of soluble p/MHC tetramers for labeling antigen-specific T cells has enabled the direct phenotypic analysis of antigen-specific T cell populations with flow cytometry (2). Conventionally, p/MHC tetramers are prepared by mixing enzymatically biotinylated p/MHC molecules with preparations of streptavidin (SA)-fluorophore conjugates. While p/MHC monomers have low affinities (2, 3), their tetramer counterparts exhibit much higher avidity, permitting T cell detection via flow cytometry to become a standard assay. However, because p/MHC tetramer-stained T cell populations are encoded optically (i.e. one unique fluorophore required per p/MHC specificity), the number of antigen-specificities that can be interrogated simultaneously within a population is limited by spectral overlap. In addition, serial flow cytometry detection of distinct antigen-specific T cells is generally restricted by sample size. Efforts to increase the degree of multiplexing generally revolve around polychromatic flow cytometry utilizing quantum dots (4, 5). However, cost, sample preparation time, and color compensation complexity also increase correspondingly.

As an alternative to flow cytometry, several groups have reported microarray-based T cell detection schemes, in which collections of p/MHC complexes are printed on a supporting substrate (6–9). A population of cells is applied directly to the p/MHC array where target antigen-specific T cells bind to regions spotted with the cognate p/MHC and

are detected optically. Analogous to DNA and protein microarrays, the readout of such assays is dependent on location rather than distinct fluorescent signals, thus potentially increasing the degree of multiplexing.

A factor that needs to be addressed before p/MHC arrays are used for broader studies and applications concerns the reproducibility and robustness of p/MHC arrays produced by spotting onto treated and/or derivatized surfaces. Ideally p/MHC complexes should be immobilized such that their functional conformations are preserved. Analogous protein arrays produced via antibody adsorption to unmodified and derivatized surfaces can suffer from surface induced effects including protein denaturation and protein adsorption in inactive orientations (10–12). To circumvent such problems, customized surfaces and relatively mild chemistries for protein immobilization have been developed (13–18). However, often the surface that meets the demands of the application requires a high level of technical expertise and/or is limited in accessibility (19).

We report here on the method of Nucleic Acid Cell Sorting (NACS), which is based upon the design and application of nucleic acid assembled p/MHC tetramer arrays for multiplexed sorting of antigen-specific T cells. For NACS, p/MHC tetramers of distinct specificities are conjugated to unique sequences of ssDNA in a site-specific fashion. A collection of ssDNA-tagged p/MHC complexes is then self-assembled by DNA hybridization onto a glass slide printed with the complementary DNA sequences. Fully assembled p/MHC tetramer arrays are used to sort mixed populations of antigen-specific T cells (**Figure 3.1**). This strategy of using DNA pendants as molecular linkages (20-25) is simple and highly modular. Most importantly, T cell array binding is

optimized by utilizing cysteine-engineered streptavidin (SAC) for ssDNA-p/MHC tetramer production, resulting in NACS p/MHC arrays that outperform conventional spotted arrays assessed by performance criteria such as reproducibility and homogeneity. The versatility of using DNA tags is also exploited to enable selective detachment of T cells with restriction endonucleases. Demonstrative experiments regarding NACS sensitivity, multiplexing and limit of detection are performed with cell lines and finally with T cells isolated from cancer patients.

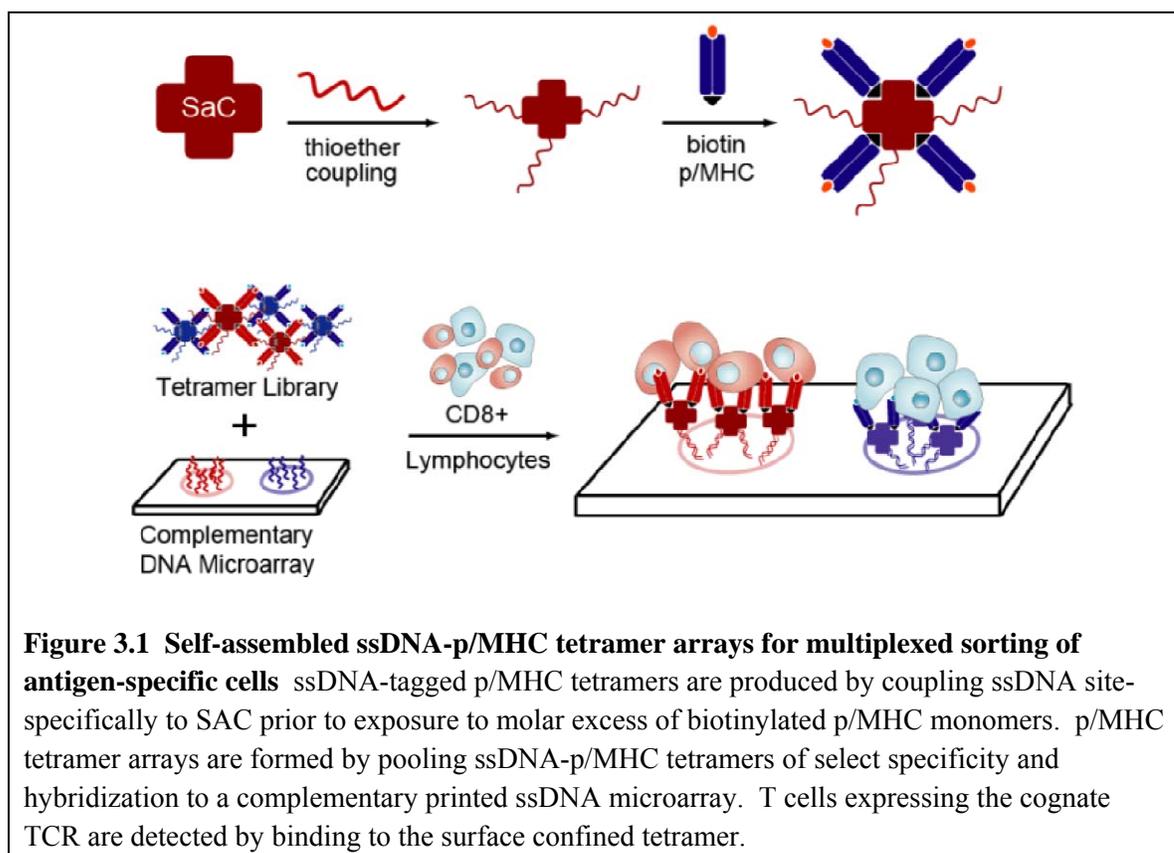


Figure 3.1 Self-assembled ssDNA-p/MHC tetramer arrays for multiplexed sorting of antigen-specific cells ssDNA-tagged p/MHC tetramers are produced by coupling ssDNA site-specifically to SAC prior to exposure to molar excess of biotinylated p/MHC monomers. p/MHC tetramer arrays are formed by pooling ssDNA-p/MHC tetramers of select specificity and hybridization to a complementary printed ssDNA microarray. T cells expressing the cognate TCR are detected by binding to the surface confined tetramer.

3.2 Experimental Methods

3.2.1 Microarray Fabrication

All DNA strands were purchased from IDT with HPLC purification. DNA microarrays were printed by the microarray facility at the Institute for Systems Biology (ISB—Seattle, WA) on amine-coated glass slides (GAPS II, Corning) in identical triplicate 12x12 arrays containing alternative rows of A, B and C spots, or A_{EcoRI} and B_{BamHI} with a SMPXB15 pin (Arrayit). Sequences for all strands can be found at **Table 3.1**.

Table 3.1 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.

3.2.2 Production of ssDNA-SAC conjugates

The pET-3a plasmid containing the SAC gene was a kind gift from Takeshi Sano (Harvard Medical School). The expression of SAC was performed according to previously published protocols (26). Briefly, transformed BL21(DE3)-pLysE cells were grown at 37°C with shaking in LB medium and selection antibiotics ampicillin and chlorophenicol. The cells were induced at OD₆₀₀ = 0.6 with IPTG and kept spinning for another 4 hours. The culture was then centrifuged at 1600g for 10 min and lysed with lysis buffer (2 mM EDTA, 30 mM Tris-HCl, 0.1% Triton X-100, pH 8.0). The insoluble inclusion bodies were then separated from the lysate by centrifugation at 39,000g for 15 min and dissolved in 6 M guanidine-HCl, pH 1.5 to the original culture volume. The SAC lysate was then refolded by dialysis in 0.2 M Sodium acetate, 10 mM β-mercaptoethanol (β-ME) pH 6.0 overnight before dialyzed against 50mM Sodium bicarbonate, 500 mM NaCl, 10 mM β-ME pH 11 in preparation for column purification. Refolded volumes of SAC were mixed 1:1 with binding buffer (50 mM Sodium bicarbonate, 500 mM NaCl, 10mM β-ME, pH 11). A gravity column packed with 1.5 ml of iminobiotin agarose resin (Pierce) was washed with 10 ml of binding buffer. The refolded mixture was then applied to the column and the eluted fractions were collected and reapplied to the column again, to maximize SAC recovery. After washing the

column with 20 ml binding buffer, SAC was eluted with pH 4 elution buffer (50 mM Sodium acetate, 10mM β -ME). Fractions containing SAC, as monitored by OD280, were collected, buffer exchanged to PBS containing 10 mM β -ME, and concentrated to 1 mg/ml final concentration using 10K mwco filters (Millipore). Immediately prior to conjugation, stock SAC was buffer exchanged to PBS containing 5mM Tris(2-Carboxyethyl) phosphine Hydrochloride (TCEP) using zeba desalting columns (Pierce). MHPH (3-N-Maleimido-6-hydraziniumpyridine hydrochloride, Solulink) in DMF was added to SAC at a molar excess of 300:1. In parallel, SFB in DMF (succinimidyl 4-formylbenzoate, Solulink) was added in a 40:1 molar excess to 5'aminated oligos. The mixtures were reacted at room temperature (RT) for 3–4 hours before buffer exchanged to citrate (50mM sodium citrate, 150 mM NaCl, pH 6.0) using zeba columns. The SFB-labeled oligos were combined in a 20:1 molar excess with the derivatized SAC and incubated overnight at RT. Unreacted oligos were removed using a Pharmacia Superdex 200 gel filtration column at 0.5 ml/min isocratic flow of PBS. Fractions containing the SAC-oligo conjugates were concentrated using 10K mwco concentration filters (Millipore).

3.2.3 Preparation of T cells

The cDNA from the alpha and beta chains of a TCR specific for tyrosinase368-376 was a kind gift from Michael I. Nishimura (Medical University of South Carolina, Charleston, SC). The alpha and beta chains were cloned into a lentiviral vector where both transgenes were linked by a 2A self-cleaving sequence as described (27). Concentrated supernatant from this lentiviral vector was used to infect Jurkat cells to generate Jurkat ^{α -Tyro} cells. The MSGV1-F5Aft2AB retroviral vector expressing the F5

MART-1 TCR was a kind gift from Steven A. Rosenberg and Richard Morgan (Surgery Branch, National Cancer Institute Bethesda, MD). The MSGV1-F5AfT2AB retroviral supernatant was used to infect Jurkat cells to generate the Jurkat^{α-MART-1} cell line. The Jurkat^{α-NY-ESO-1} cell line was a generous gift from Robert Prins (UCLA). To generate primary human T cell cultures expressing the F5 MART-1 TCR, PBMCs obtained from leukapheresis were activated for 48 hours with 50 ng/ml of OKT3 (muromonab anti-human CD3 antibody, Ortho-Biotech, Bridgewater, NJ) and 300 U/ml of IL-2 (adesleukin, Novartis, Emeryville, CA). MSGV1-F5AfT2AB retrovirus supernatant was applied to retronectin-coated wells (Takara Bio Inc., Japan). Then activated PBMC in RPMI plus 5% human AB serum supplemented by 300 IU of IL-2 were added to these wells and incubated at 37°C overnight at 5% CO₂. On the following day, PBMC are transferred to a second set of pre-coated retronectin retroviral vector tissue culture plates and incubated at 37°C overnight at 5% CO₂. Cells were subsequently washed and re-suspended in culture media described above. Frozen leukapheresis fractions from patients NRA11, NRA 13 (UCLA IRB#03-12-023) and F5-1 (UCLA IRB #08-02-020-02A) were thawed and incubated overnight in RPMI supplemented with 10% human AB serum and 1% penicillin, streptomycin, and amphotericin (Omega Scientific). F5-1 cells were used immediately following incubation. NRA11 and NRA13 samples were CD8⁺ enriched (anti-CD8 microbeads, Miltenyi Biotech) using an AutoMACS machine according to the manufacturer's instructions. Following separation, the cells were kept at in RPMI-humanAB media containing 30 U IL2/mL.

3.2.4 T Cell Sorting Methods

The HLA-A*0201 restricted MHC class I monomers loaded with tyrosinase₃₆₉₋₃₇₇ (YMDGTMSQV), MART-1₂₆₋₃₅ (ELAGIGILTV) and NY-ESO-1157-165 (SLLMWITQC) were produced in house according to previous published protocols (28). 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. Lipophilic cell membrane staining dyes DiO, DiD, and DiL were purchased from Invitrogen.

Prior to experiments, microarray slides were blocked to prevent non-specific cell binding with 1 mg/ml PEG-NHS ester (Sunbio) in PBS for 2 hours at RT. Four-fold molar excess of p/MHC monomers were combined with ssDNA-SAC at 37°C for 20 min. ssDNA-p/MHC tetramers were hybridized to DNA arrays for 1 hour at 37°C in 200 µl media and rinsed with 3% FBS in PBS. T cells (10⁶ /100 µl media) were incubated on the array at 37°C for 30 min. The arrays were rinsed with 3% FBS in PBS and cell capture visualized via brightfield (Nikon Eclipse TE2000) and/or confocal microscopy (Nikon E800). Post T cell capture p/MHC tetramer staining was done by incubating the array with 200 µl of media containing fluorescent p/MHC tetramer along with fluorescent cDNA (Cy5-A' and/or Cy3-B'). The arrays were rinsed with 3% FBS in PBS prior to imaging. For selective T cell release experiments, three identical arrays were used to immobilize cells. Treatment with EcoRI, BamHI, or DNase was in RPMI media for 1–2 hours at 37°C. DNase was purchased from Sigma, all other enzymes from NEbiolabs.

For p/MHC comparative studies, SuperEpoxy and SuperProtein (representing covalent and hydrophobic surfaces respectively) were purchased from Arrayit (Sunnyvale, CA). Amine GAPS II slides (electrostatic) were purchased from Corning. Polycarboxylate hydrogel (hydrophilic) slides were purchased from XanTec (Germany). Fluorescent MART-1 tetramers were printed according to manufacturer's instructions for each slide. Cell sorting images were quantified with ImageJ (NIH) and fitted to the Hill Function (NACS $n=2$, $R^2=0.95$, Covalent $n=2.1$, $R^2=0.97$) with Origin (OriginLab, MA).

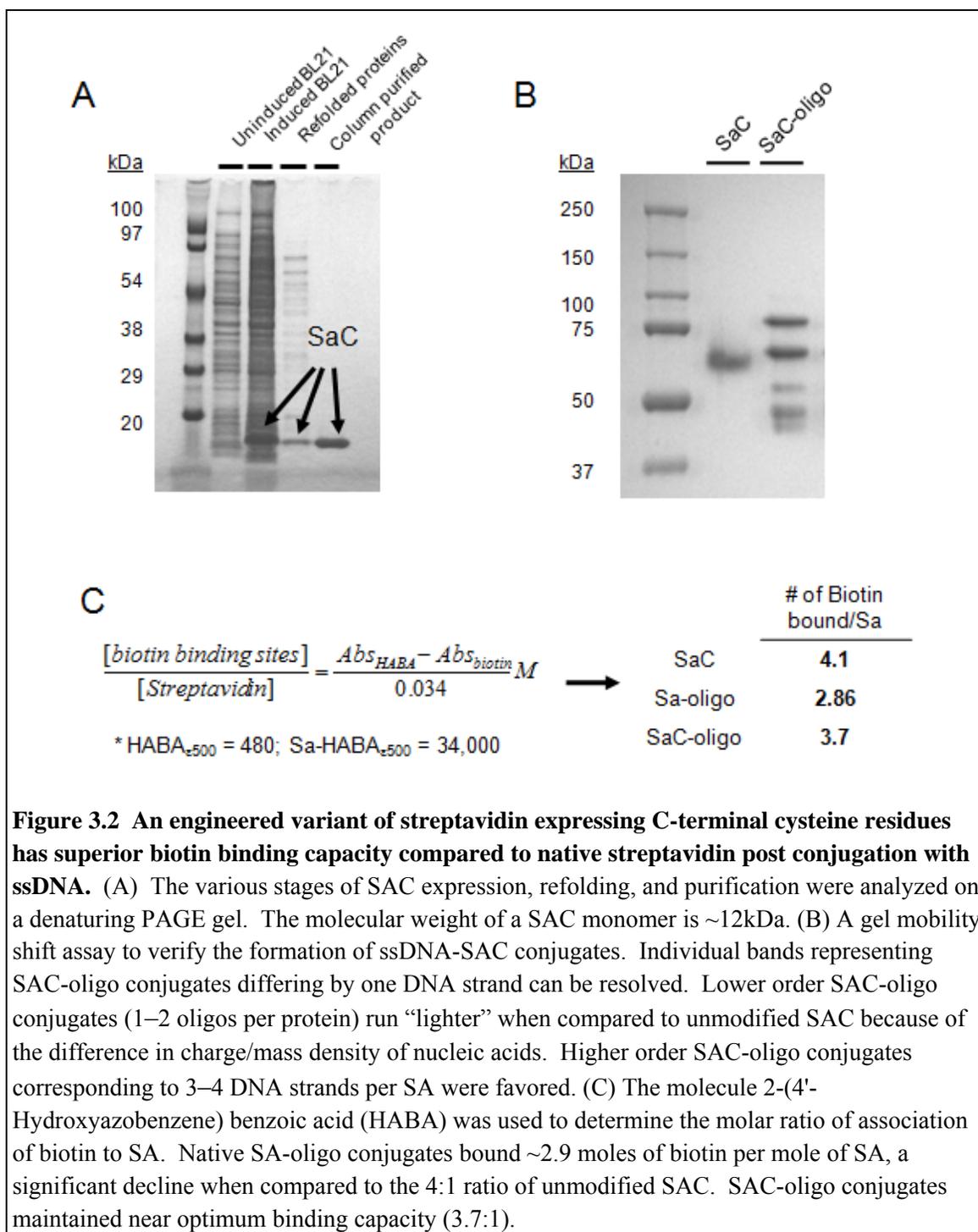
3.3 Results and Discussion

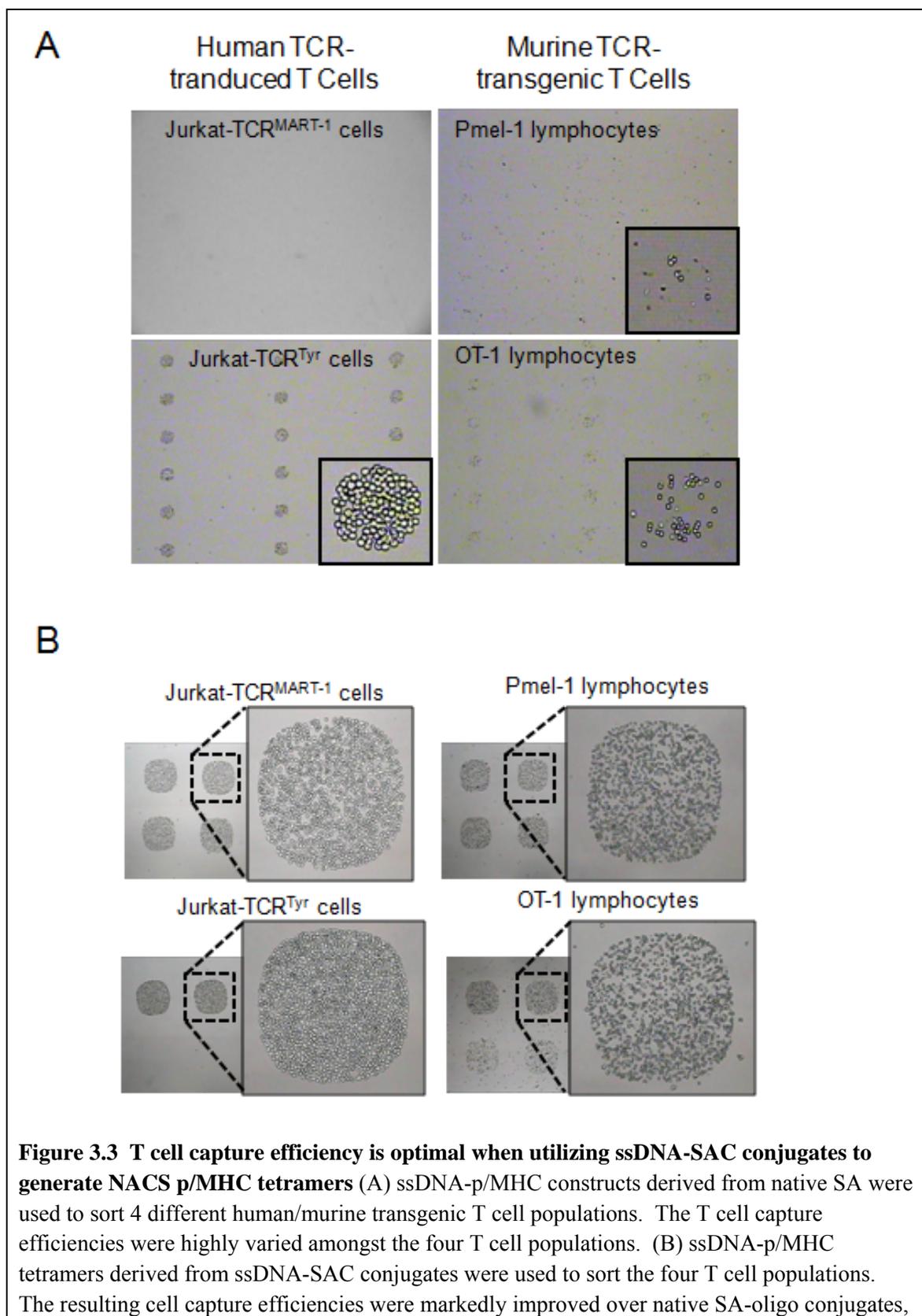
3.3.1 Rational design of ssDNA-encoded p/MHC tetramers

The standard scaffold most frequently used to assemble p/MHC monomers into tetramers is SA-phycoobiliprotein (using the protein fluorophores phycoerythrin (PE) or allophycocyanin (APC)) conjugates. Because SA-phycoobiliprotein conjugates are produced via chemical cross-linking, most functional groups are exhausted and/or modified, prohibiting the conjugation of ssDNA. In addition, the attachment of molecular fluorophores to native SA reduces the binding capacity for biotin, an effect attributed to the modification of lysine121 that occurs with amide coupling strategies. This residue is in close proximity to the ligand binding pockets (29, 30). To circumvent this, Altman and co-workers (29) employed a recombinant mutant of SA for fluorescent p/MHC tetramer preparations. This variant incorporates a cysteine residue at the carboxy-terminus (31), a site removed from the biotin binding pocket. The conjugation

of cysteine-reactive maleimide derivatives is restricted to the C-terminus because cysteine residues are absent in native SA.

We expressed SAC, coupled the protein with 5'-maleimide ssDNA, and verified the formation of conjugates with mobility shift assays (**Figure 3.2**). In parallel, ssDNA was coupled to native SA for direct comparison. To test biotin binding capacity, SAC-oligo conjugates were probed with 2-(4'-Hydroxyazobenzene) benzoic acid (HABA) (32), a molecular mimic of biotin with distinct optical density coefficients dependent on whether biotin is bound to SA or not. A biotin:SA molar ratio of association significantly below 4 in the assay would indicate a reduction in biotin binding capacity. Conjugates derived from native SA were greater than one full unit below the expected value (2.86 versus 4.0), while conjugates formed with SAC maintained near optimal (3.7) binding capacity (**Figure 3.2C**). These conjugates were then tested across 4 different monoclonal T cell populations (2 human TCR-transduced cell lines and 2 murine TCR-transgenic splenocyte cell suspensions). ssDNA-tagged SAC constructs had markedly higher cell capture efficiencies (**Figure 3.3B**) when compared with p/MHC tetramers prepared with native SA (**Figure 3.3A**). All subsequent NACS tetramers were prepared with the SAC variant.

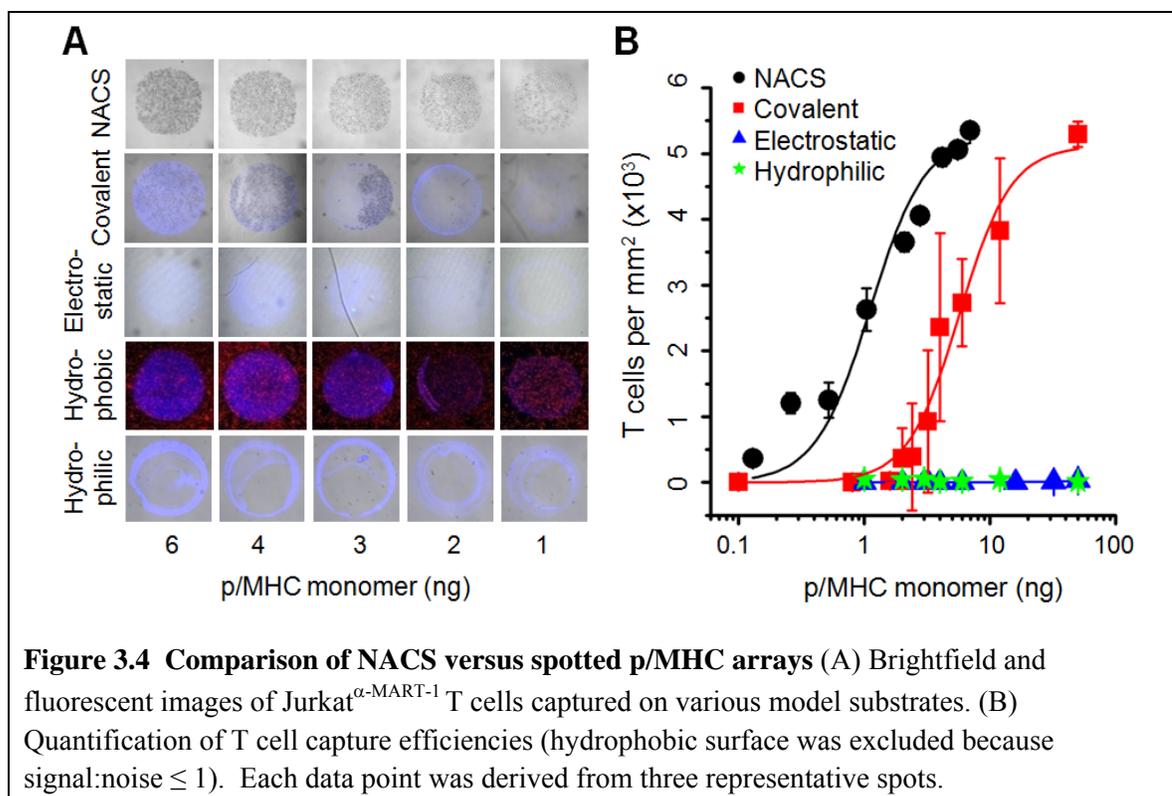




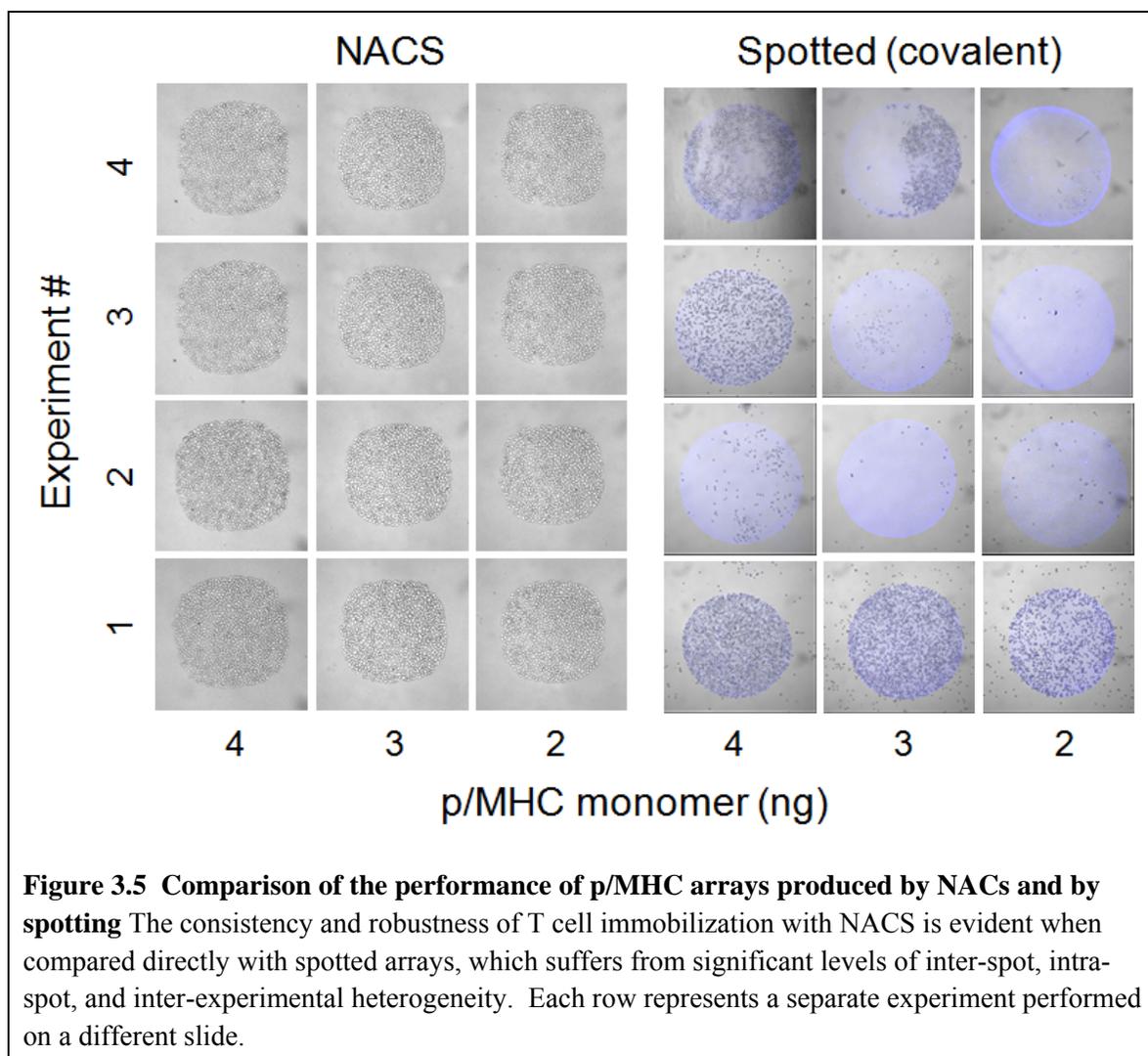
demonstrating that SAC is necessary for the production of high affinity ssDNA-p/MHC tetramers.

3.3.2 Performance of p/MHC arrays produced via DNA immobilization and direct spotting

We directly compared the performance of NACS with conventional direct spotting strategies on various model substrates. The substrates were selected to represent the spectrum of surface chemistries typically used to immobilize proteins (covalent, electrostatic, hydrophobic, and hydrophilic adsorption). Serial dilutions of a fluorescent p/MHC tetramer, MART-1, were spotted on the substrates according to manufacturer's instructions. Jurkat^{α-MART-1} T cells (the human T leukemia cell line Jurkat transduced with the F5 MART-1 TCR (**33**) specific for MART-1) were applied to the array and representative images collected (**Figure 3.4A**) and quantified (**Figure 3.4B**). We observed little to no T cell capture (electrostatic, hydrophilic) or significant noise (hydrophobic) on the majority of the surfaces investigated compared to NACS arrays immobilized with identical concentrations of p/MHC tetramers. T cell binding was observed on one surface (covalent) but cell capture was highly variable as evidenced by both intra-spot and inter-spot heterogeneity and cross experimental variation (**Figure 3.5**). Moreover, to achieve equivalent T cell capture densities, NACS p/MHC arrays required >5 times less material than covalent immobilization. (p/MHC monomer at half max $\equiv K_{1/2} = 1.1$ ng for NACS and 5.7 ng for covalent immobilization).



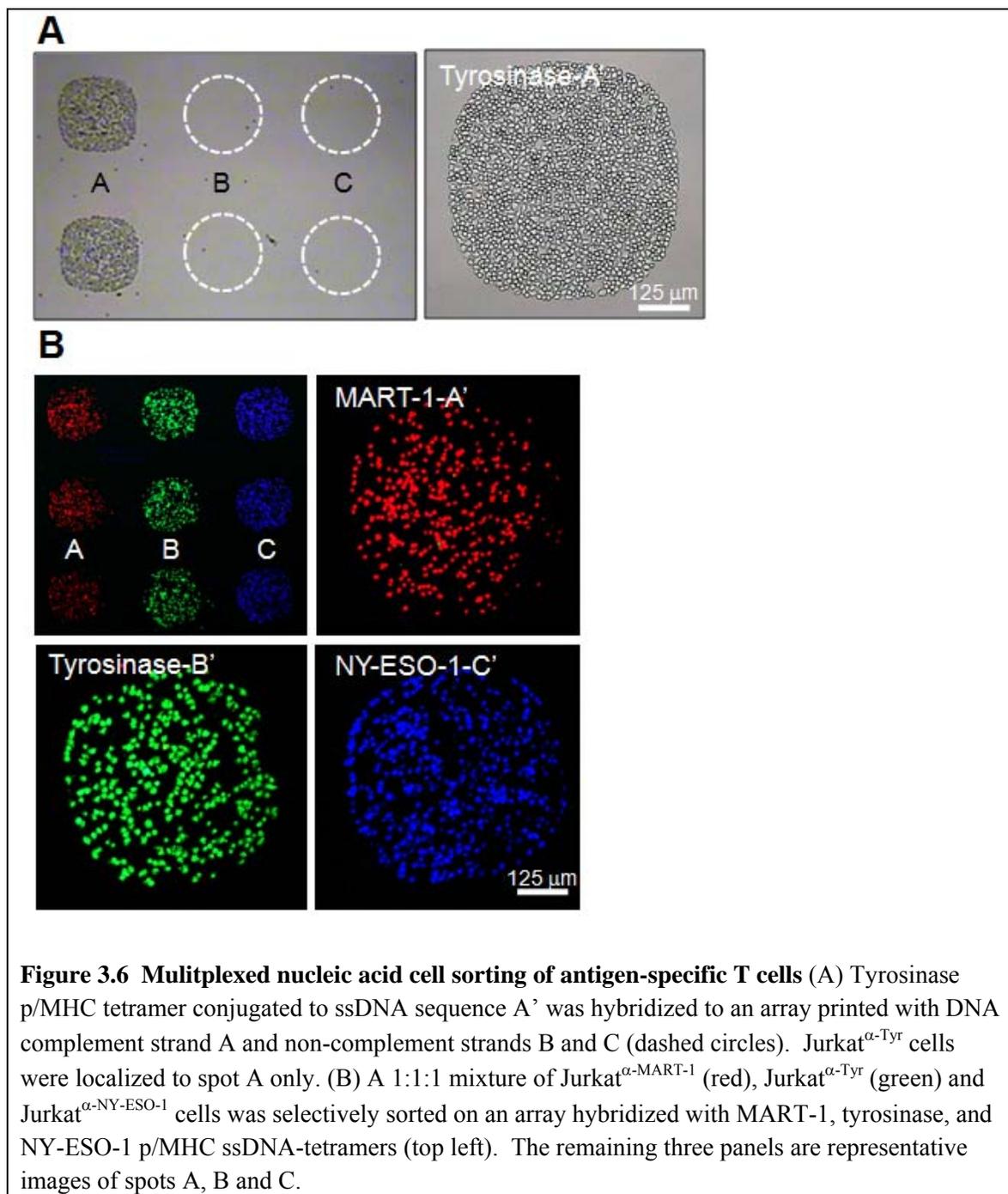
The performance and reproducibility of NACS p/MHC arrays is markedly improved and represents an integral step towards expanding array-based T cell detection schemes for broader applications. This likely has a few causes. First, surface-tethered ssDNA-p/MHC tetramers may enjoy greater orientational freedom at the surface/solution interface compared with adsorbed proteins which are required to conform to the surface. This effect may increase the density of functional protein and consequently reduce the amount of material required for array production. Second, the hydration state of the environment during the production and subsequent storage of protein arrays is an important factor for array reproducibility (10, 13, 18). This effect is minimized with NACS because DNA chips can be printed and stored dry for extended periods of time and ssDNA-tagged p/MHC tetramer arrays are self-assembled in solution immediately prior to an experiment.



3.3.3 NACS specificity and limit of detection

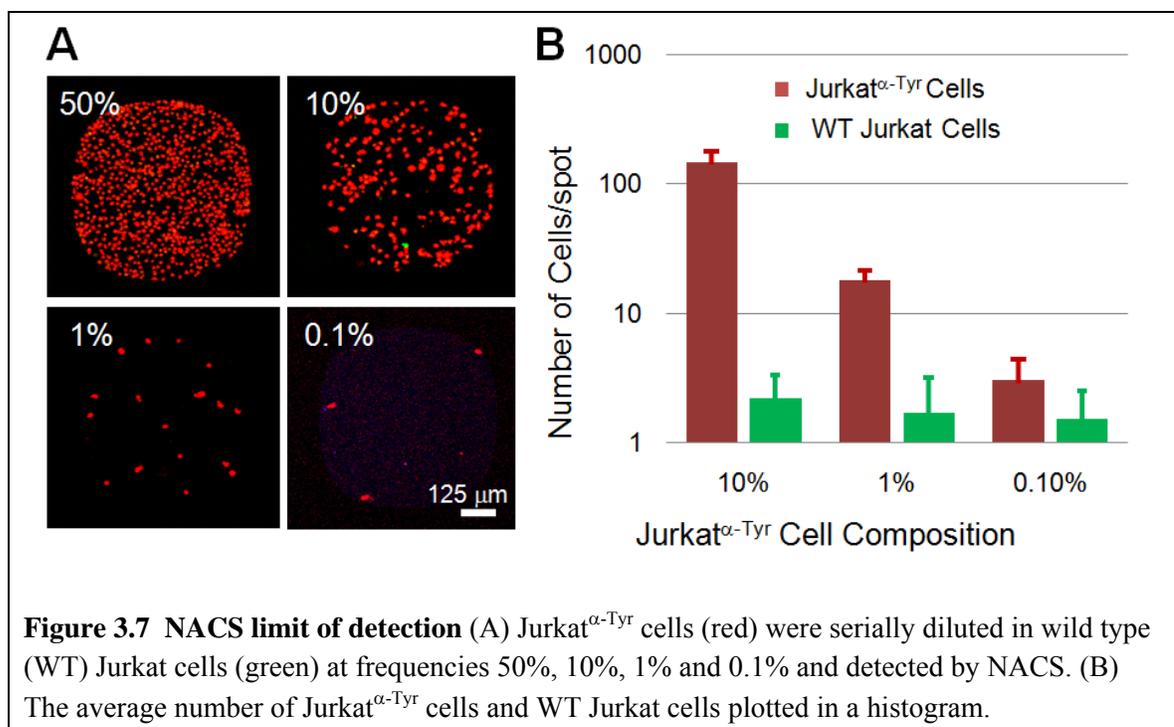
To evaluate the specificity of p/MHC array assembly and T cell sorting, a ssDNA-p/MHC tetramer, tyrosinase, with pendant DNA sequence A' was hybridized to a DNA microarray printed with the complementary strand (A) along with two additional distinct sequences (designated B and C). A homogeneous population of Jurkat^{α-Tyr} cells (Jurkat cells transduced with a TCR specific for tyrosinase) (34) was then applied to the array. Jurkat^{α-Tyr} T cells localized to the complementary spots (A) containing the hybridized

cognate p/MHC but not to spots printed with the non-complementary sequences B and C (**Figure 3.6A**). The mean binding capacity calculated from three spots ($\sim 600 \mu\text{m}$) was $\sim 1486 \pm 62$ Jurkat ^{α -Tyr} T cells.



To illustrate the multiplexing capability of NACS, MART-1, tyrosinase, and NY-ESO-1 ssDNA-p/MHC tetramers encoded to DNA sequences A, B and C respectively were combined and assembled simultaneously to a three element DNA microarray (strands A, B, and C). A 1:1:1 mixed population of Jurkat^{α-MART-1}, Jurkat^{α-Tyr} and Jurkat^{α-NY-ESO-1} (Jurkat human T cells transduced with the TCR (35) specific for NY-ESO-1) cells prestained with lipophilic dyes (red, green and blue respectively) was applied to the array and localized into alternating columns (**Figure 3.6B**). Minimal cross-reactivity was observed. The average density of spots was about a factor of three less than homogeneous sorting (440 ± 28 T cells/spot).

To determine the limit of detection, target populations of Jurkat^{α-Tyr} cells were spiked in at 10%, 1% and 0.1% into wild type (w.t.) Jurkat cells and sorted (**Figure 3.7**). The T cell capture density per spot per species for each mixture was enumerated and averaged (right panel). The number of non-specific w.t. Jurkat cells that adhered to the array was constant throughout all dilutions while the number of Jurkat^{α-Tyr} T cells captured per spot decreased linearly in relation to the fractional composition of Jurkat^{α-Tyr} cells with a detection limit that was ~ 1 in 1000 cells—a limit that corresponds well to the total number of cells that can be captured per spot. Thus, the sensitivity of this approach is strictly a geometric constraint since antigen-specific T cells that settle on inert areas cannot sample and bind to their cognate p/MHC tetramer. The sensitivity can be improved by increasing the size of the capture region (i.e. increase spot diameter and/or incorporate spot redundancy) or by reducing inert regions (i.e. increase printing density).



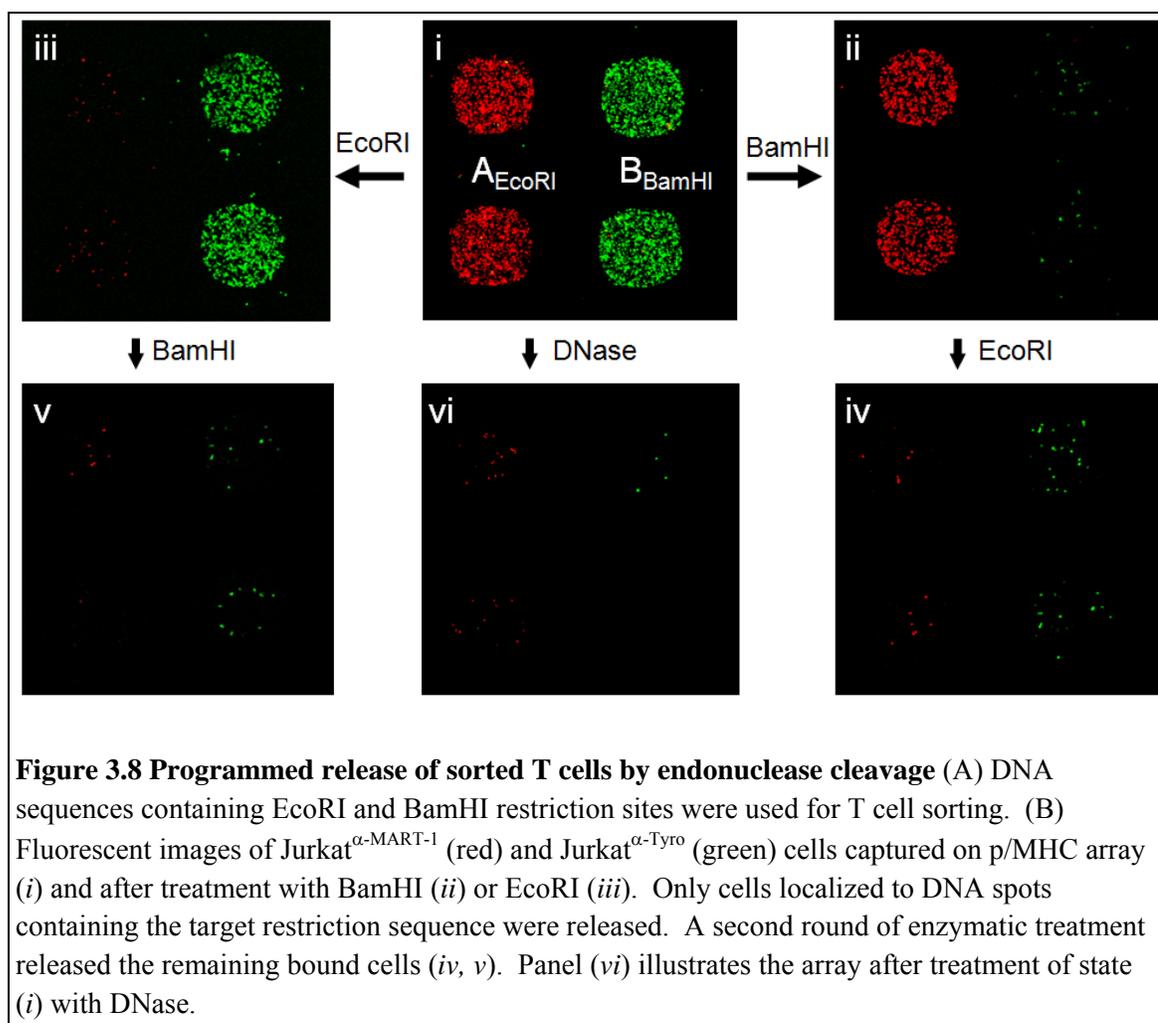
It should be noted, however, that the sensitivity of this approach cannot be increased without limiting scalability. Since spot diameters are required to be sufficiently large to detect a population at low frequency, this sets an upper bound on the number of distinct spots that can be patterned on a substrate. In the current instance, 600 μm spots are printed in 12 by 12 grids ($\sim 1 \text{ in}^2$), enabling the potential identification of 144 distinct antigen-specificities $\geq 0.1\%$ from 10^6 T cells (10^6 cells is typically required to cover a 1 in^2 region). In order to detect target populations below 0.1% without decreasing multiplexing, p/MHC-specific enrichment strategies (e.g. magnetic bead-based schemes (36)) can be implemented prior to cell sorting. Lastly, the recovery of this technique—defined as the total number of antigen-specific T cells captured as a fraction of the number of T cells applied to the array—is low relative to other sorting technologies like FACS and bead-based schemes. Typical T cell recovery for NACS is $\leq 20\%$ (37). In comparison, fluorescent-based sorting techniques like FACS have high recoveries, since

each stained T cell can be identified individually by the cytometer and sorted from the null population. The recovery (as well as sensitivity) of NACS will likely be improved by circulating T cells over the p/MHC array with agitation or with integrated microfluidic devices to allow T cells to sample the entire array.

3.3.4 Selective Release of Immobilized T cells with Restriction Endonucleases

Antigen-specific T cells immobilized onto glass are immediately available for secondary assays, since many such as immunohistochemistry (IHC), fluorescent in situ hybridization (FISH) and cytokine secretion assays (6, 8) are traditionally performed or are compatible with cells localized to a substrate. However, several other relevant assays, such as those designed to assess T cell phenotype or functional status like genomic/mRNA analysis or simply further culture for phenotypic enrichment would require a method for releasing the captured cells. Any release scheme should ideally be selective for given cell types. For NACS, we explored whether the DNA tethers could be selectively cleaved by exploiting the sequence specificity of restriction endonucleases. We integrated unique restriction sites to each DNA sequence employed for cell sorting, and found that the adhesion of different populations of antigen-specific T cells could, in fact, be independently controlled (**Figure 3.8**). Jurkat ^{α -MART-1} and Jurkat ^{α -Tyr} cells prestained with lipophilic dyes (red and green respectively) were sorted on an array printed with DNA sequences A_{EcoRI} and B_{BamHI} (**Figure 3.8i**). These oligonucleotides were modified by incorporating 6 bp restriction sites specific for endonucleases EcoRI and BamHI respectively. After T cell immobilization, the array was treated with BamHI which cleaved the B_{BamHI} spots and selectively released the bound Jurkat ^{α -Tyr} cells (**Figure 3.8ii**). Conversely, on a separate but identically cell sorted array, Jurkat ^{α -MART-1}

cells were released after treatment with EcoRI (**Figure 3.8iii**). A second round of enzymatic treatment with the complementary endonuclease (EcoRI to state *(ii)* or BamHI to state *(iii)*) removed the remaining adherent cells (*(iv, v)*). Alternatively all captured cells *(i)* could be released non-selectively in a single step with the addition of DNase *(vi)*.

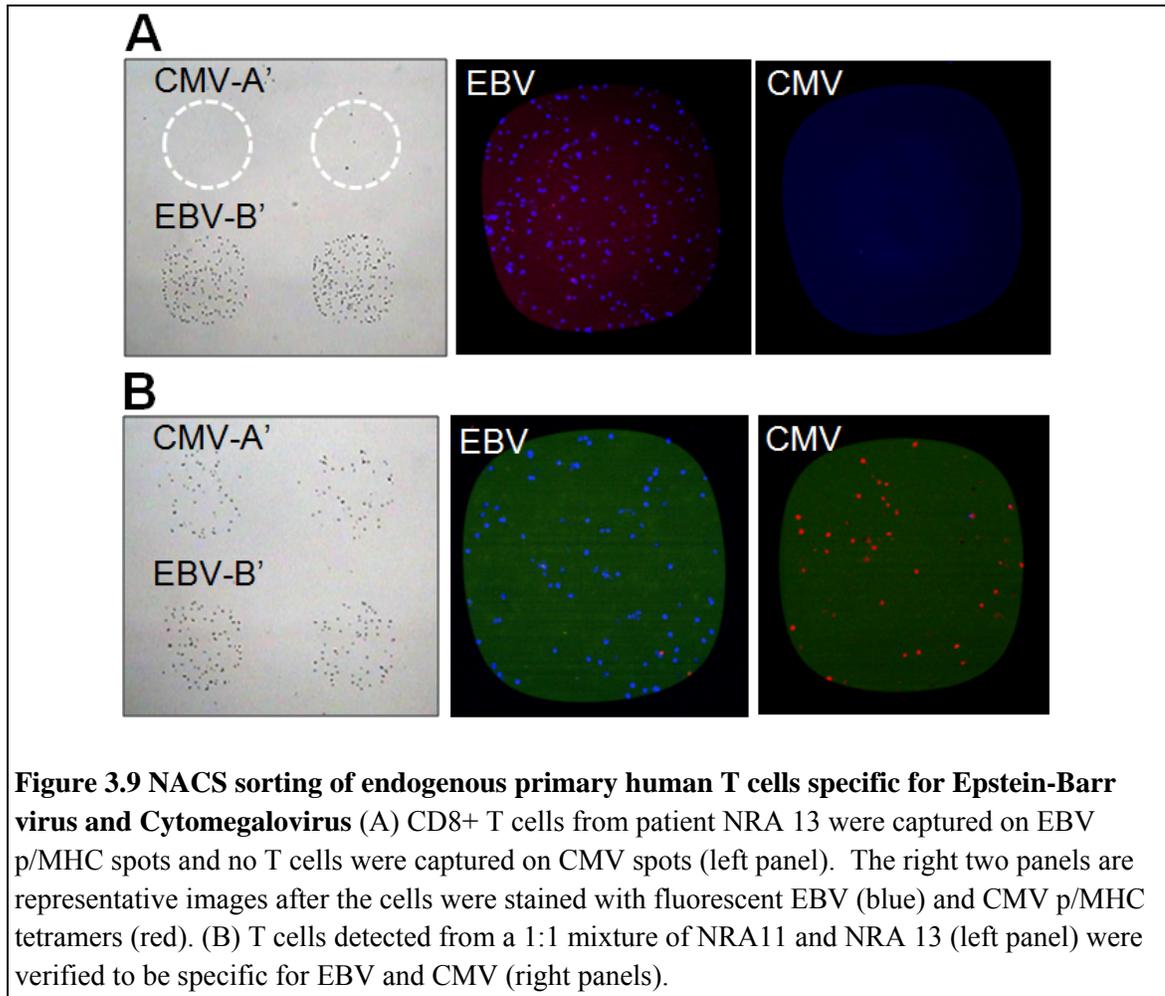


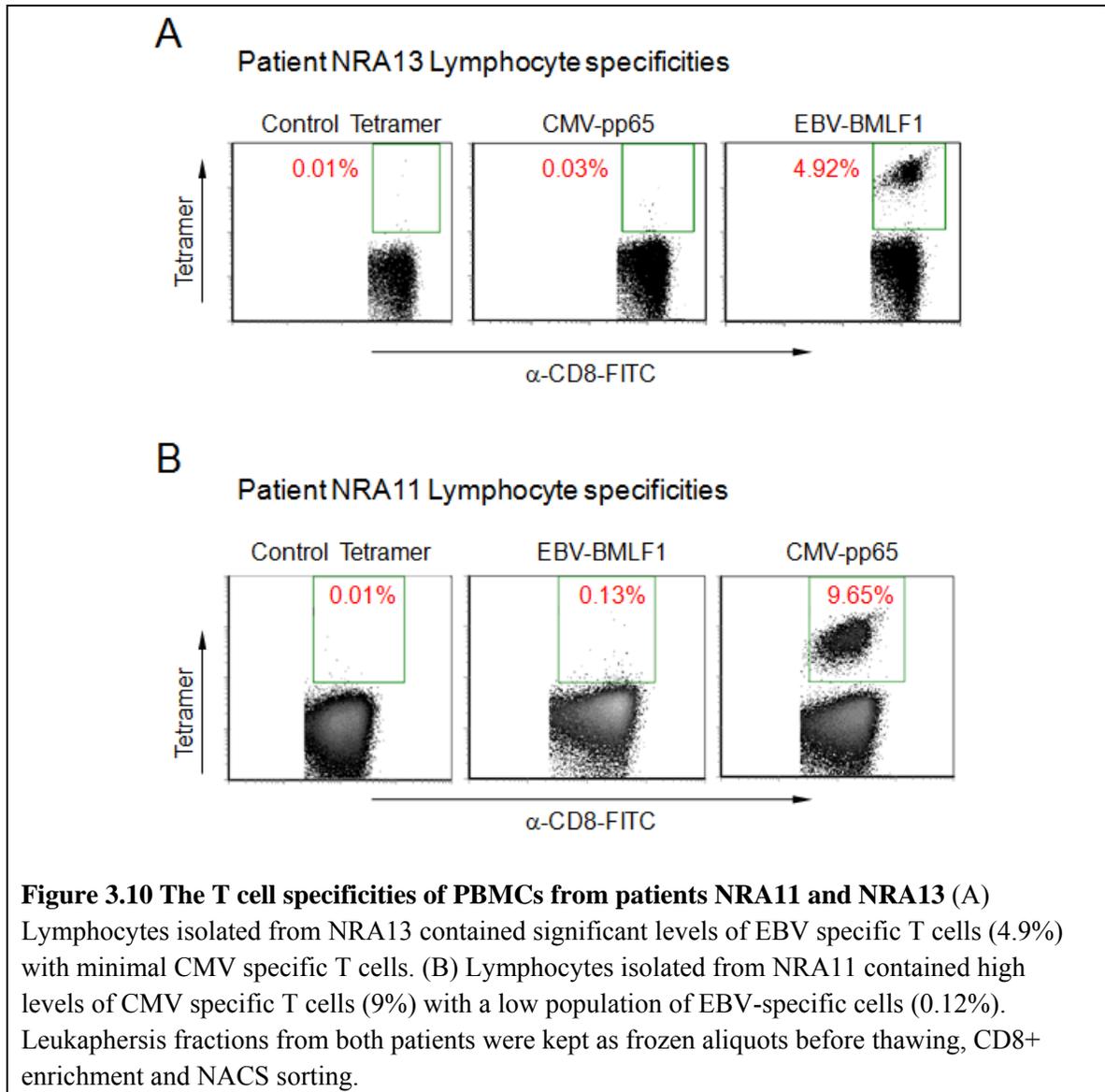
3.3.5 NACS sorting of endogenous primary human T cells

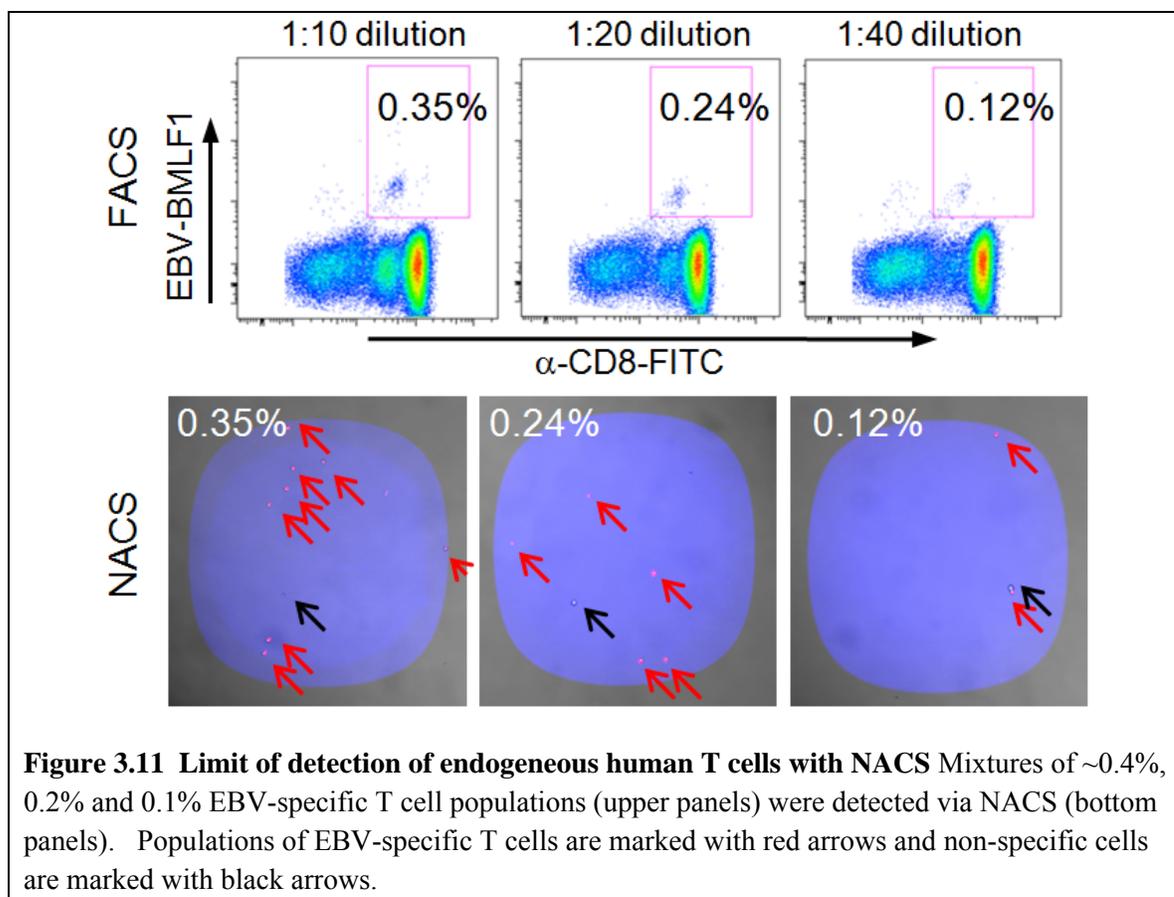
Detection of primary human T cells isolated from peripheral blood is generally more demanding than cultured cell lines because a single population of antigen-specific T cells is present within a large background of differing blood cells and of T cells expressing monoclonal and polyclonal TCRs of diverse specificities. In addition, these T cells would be expressing endogenous levels of TCR. We explored whether the same attributes of NACS that were found in the above examples would apply equally to endogenous primary human T cells. Frozen leukapheresis samples from patient NRA13 were CD8⁺ enriched and applied to a CMV and Epstein-barr virus (EBV BMLF1₂₅₉₋₂₆₇/HLA-A2.1) p/MHC array. T cells were captured only within the EBV regions only (**Figure 3.9A**). NRA13 CD8⁺ cells were verified by flow cytometry to be ~5% EBV-specific and ~0% CMV-specific (**Figure 3.10A**). The registry of the array was determined with the addition of A'-cy3 (red) and B'-cy5 (blue) conjugates.

For multiplexed detection, a 1:1 mixture of EBV-specific and CMV-specific CD8⁺ T cells was produced by combining NRA13 lymphocytes with CMV-specific T cells from patient NRA11 (**Figure 3.10B**). Following cell sorting and fluorescent p/MHC tetramer staining, the populations were complementary stained for the appropriate antigen-specificity (**Figure 3.9B**). The detection limit of antigen-specific T cells was evaluated from using serially diluted mixtures of EBV-specific T cells (~0.4%, 0.2%, and 0.1% by FACS) that were probed on an array. Isolated hits were resolved in frequencies as low as ~0.1% (**Figure 3.11**, red arrows). The number of unstained cells within the capture regions (black arrow) was constant throughout all dilutions (~1–2 cells/spot) and likely represents the level of background from non-specific interactions.

It should be noted that while we incorporated fluorescent p/MHC tetramer staining after T cell immobilization for illustrative purposes, the specificity of the captured cells could be determined solely from the registry of the array.







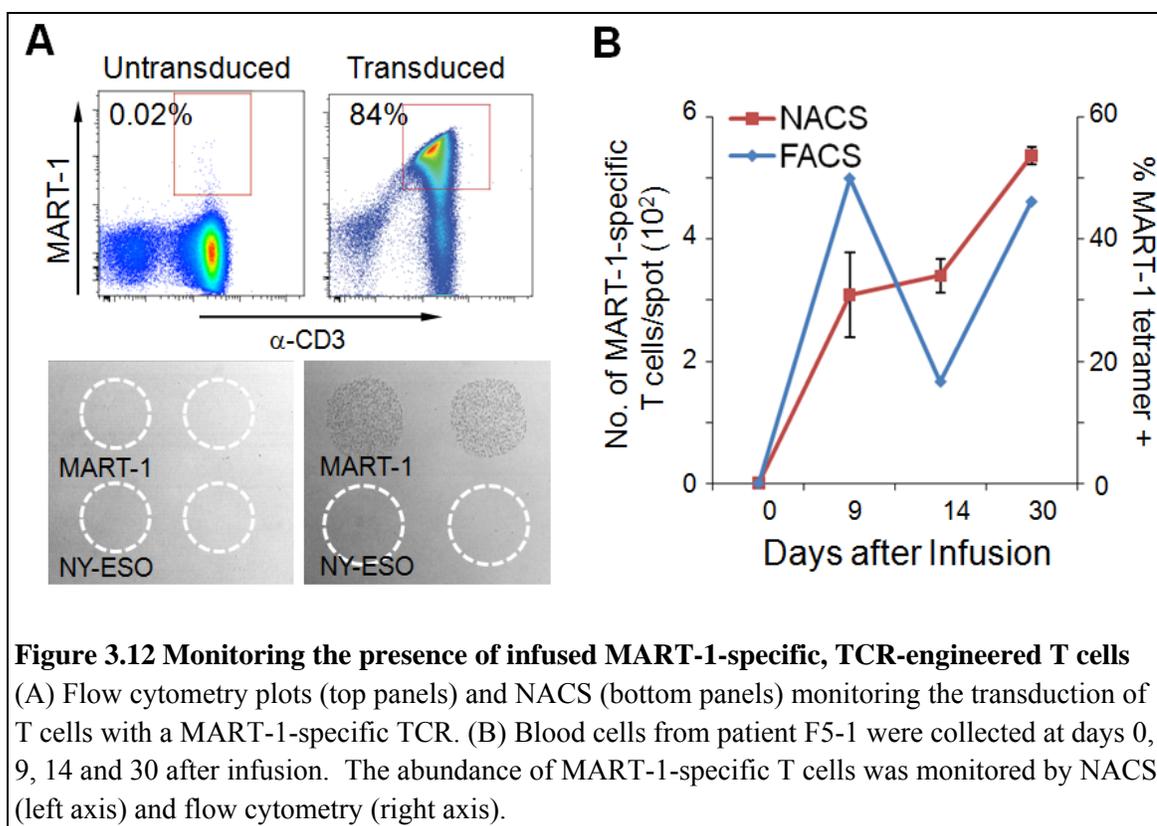
3.3.6 Persistence of MART-1 specific, TCR-engineered human T cells *in vivo*

One important potential application of NACS is to utilize the technique for monitoring cancer patients that are undergoing a particular type of immunotherapy that involves TCR engineering of peripheral blood mononuclear cells (PBMCs). This is an emerging clinical approach to rapidly generate large numbers of tumor antigen-specific T cells. In particular, a cancer patient's T cells are collected, genetically modified to express a TCR specific for a desired cancer antigen, and then they are introduced back into the patient as a cancer therapy. After re-infusion into the patient, the modified T cells can subsequently traffic to and engage with cancer cells, promoting tumor regression in a subset of patients with metastatic cancers. This type of directed cell

therapy is utilized primarily in patients with melanoma (skin cancer), but also for other cancers (38, 39). This approach is called adoptive cell transfer (ACT). We monitored the presence and abundance of TCR-engineered T cells in peripheral blood of a patient undergoing ACT. White blood cells were collected from patient F5-1 with metastatic melanoma. After in vitro expansion and TCR transduction, the cells were > 80% specific for the MART-1 p/MHC tetramer (**Figure 12A**). These cells were subsequently infused into patient F5-1 and the presence and persistence of MART-1-specific T cells in the peripheral blood was monitored by NACS at days 0 (prior to infusion), 9, 14 and 30 (**Figure 12B**). MART-1-specific T cells were not detectable in pre-infusion samples from the patient but they were detectable at all subsequent time points. This is represented as a gradual rise in the abundance of MART-1-specific T cells until day 30 (left axis). In comparison, the frequency of MART-1-specific T cells detected by flow cytometry spiked initially before decreasing and finally increasing by day 30 (right axis).

The abundance of TCR-engineered T cells, as measured by NACS, was generally correlated with parallel measurements using flow cytometry. The prospect of array-based T cell detection schemes for experimental cancer immuno-therapies such as ACT will likely increase as more cancer associated antigens are identified and targeted. For example, more than 50 different melanoma associated antigens have thus far been characterized (40, 41), and this means that more T cells types are likely to be employed for future therapies (39, 42, 43). However, the desire to probe larger sets of cancer associated antigens will be compounded by the amount of sample that can be practically collected from a patient. NACS appears to provide a feasible approach towards carrying

out the highly multiplexed cellular measurements that will eventually be required as ACT-like therapies move forward.



3.3.7 Homogeneous platform for cell sorting and functional analysis

Antigen-specific T cells that are sorted by NACS are immobilized to a solid surface and are immediately available for further studies. Traditional surface bound assays such as IHC, FISH and ELISPOT should integrate seamlessly with NACS. In addition, antigen-specific T cells have been shown to be activated upon binding to p/MHC arrays (6–8) and secrete various immunomodulatory cytokines. The cytokine secretion profiles of CD8+ T cells provide valuable information concerning the phenotype of the T cell (e.g. effector and anergic phenotypes can be differentiated by the

production of IFN- γ). Thus integrating on-chip cytokine measurements with antigen-specific T cell capture would greatly expand the diagnostic capability of NACS p/MHC tetramer arrays.

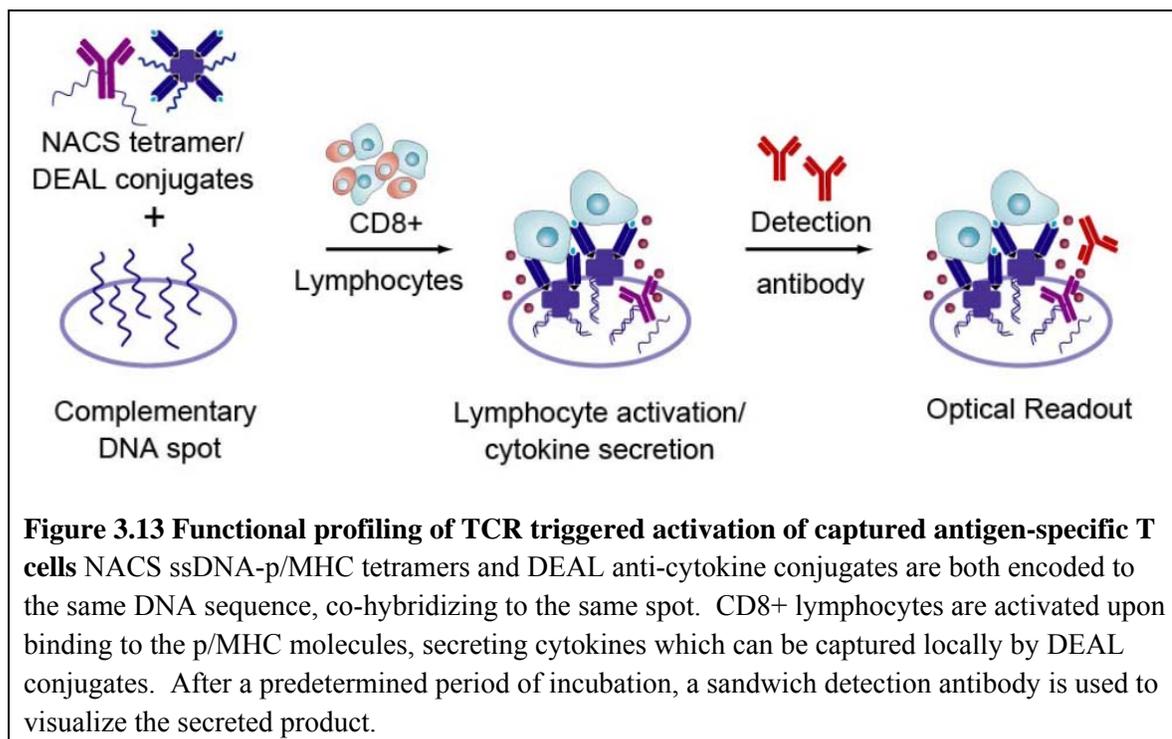
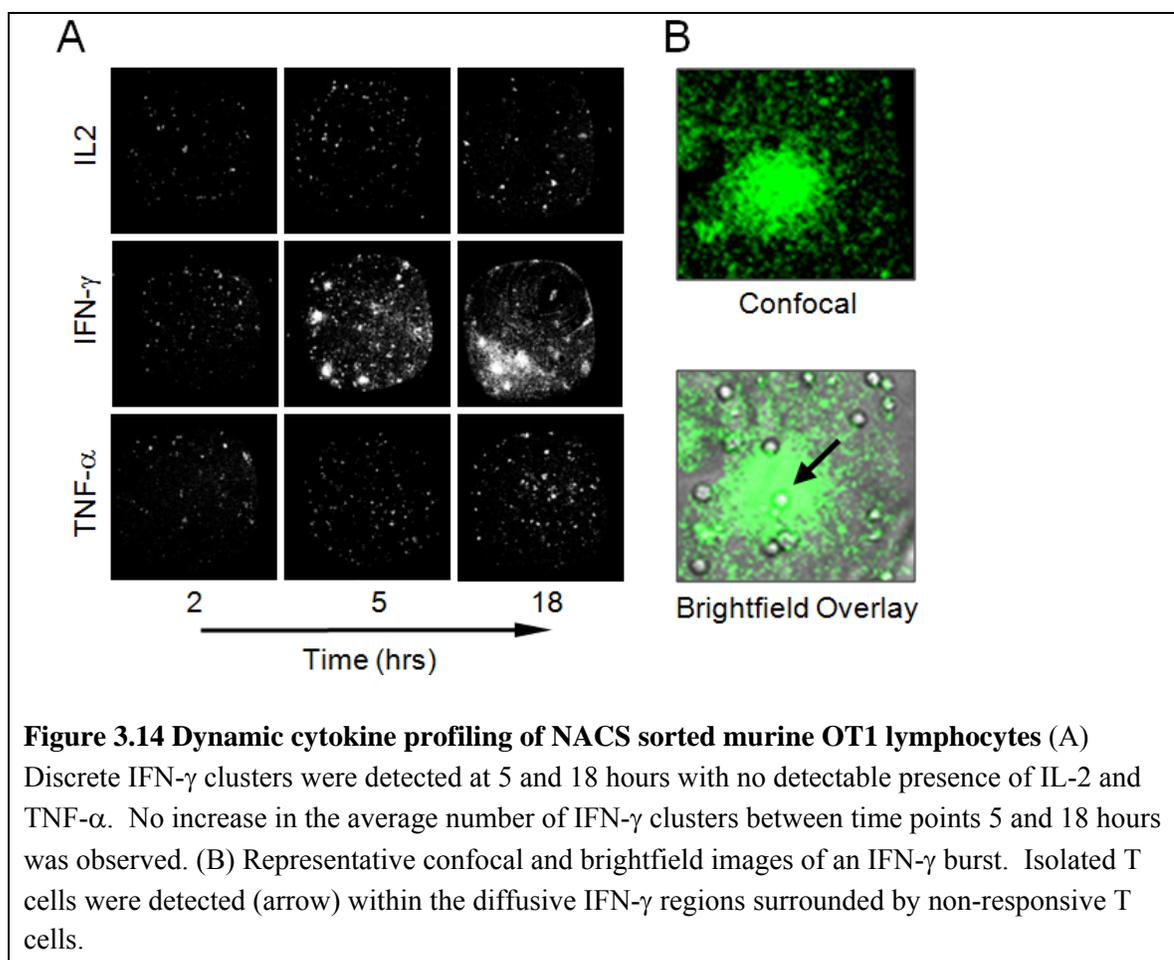


Figure 3.13 Functional profiling of TCR triggered activation of captured antigen-specific T cells NACS ssDNA-p/MHC tetramers and DEAL anti-cytokine conjugates are both encoded to the same DNA sequence, co-hybridizing to the same spot. CD8⁺ lymphocytes are activated upon binding to the p/MHC molecules, secreting cytokines which can be captured locally by DEAL conjugates. After a predetermined period of incubation, a sandwich detection antibody is used to visualize the secreted product.

We proceeded by integrating an ELISPOT-type sandwich assay with p/MHC NACS to detect cytokines produced by captured murine TCR transgenic splenocytes “on-the-spot” (**Figure 3.13**). Three murine anti-cytokine antibodies (IL-2, IFN- γ and TNF- α) were encoded with DNA strands A’, B’, and C’ respectively. H-2K^b-OVA₂₅₇₋₂₆₄ ssDNA-p/MHC tetramers were encoded to all three strands. The ssDNA-p/MHC tetramers and antibody conjugates were pooled and assembled to a microarray printed with the complementary strands A, B and C. Murine OT1 lymphocytes (derived from TCR transgenic mice in which most splenocytes are specific for the model antigen OVA₂₅₇₋₂₆₄), were then seeded on the array. Following incubation periods of 2, 5, or 18 hours, pooled cytokine detection antibodies were added and the slide imaged by confocal

microscopy (**Figure 3.14A**). The inflammatory cytokine IFN- γ was detected at time points 5 and 18, manifest as discrete diffusive clusters (~ 50 – 100 μm in diameter at 5 hrs) that increased in average diameter temporally, attributable to molecular diffusion and sustained secretion. Examination of the vicinity of each burst showed that underlying each fluorescent cluster was a single cell while neighboring cells appeared to be non-responders (**Figure 3.14B**), suggestive that each IFN- γ burst was derived from a single cell. The number of IFN- γ clusters remained constant at ~ 3 between hours 5 and 18, indicating no increase in the number of activated T cells between those hours. No significant levels of murine IL-2 and TNF- α were detected at these time points.



3.4 Conclusions

We have described a method for generating robust and modular p/MHC arrays for high efficiency T cell sorting. The inclusion of a larger set of orthogonal DNA sequences (20, 25) will enable the modular assembly of higher order p/MHC arrays for T cell screening experiments (e.g. one working set of DNA sequences can be used interchangeably to generate any combination of p/MHC arrays). This would find immediate utility in the field of TCR peptide epitope discovery where recently, novel antigen peptides were discovered via high-throughput CD8⁺ screening experiments utilizing multi-color flow cytometry in mice and humans (44, 45) (as many as 2,000 distinct p/MHC tetramers were prepared and tested). NACS arrays have the potential to streamline such experiments. Although traditional methods of producing single p/MHC monomers are time and labor intensive, recent reports using conditional peptide exchange technology enables the relatively straightforward construction of 1000 element p/MHC libraries rapidly (44–46). The integration of NACS with these peptide exchange technologies is a realistic option.

We have also demonstrated a number of advantages of the NACS platform. It significantly outperforms literature approaches that utilize surface-bound p/MHC tetramers to capture cells. It is simple and inexpensive to implement since cell sorting is performed on glass substrates prepared via traditional DNA printing technologies. In addition, sorted cells may be selectively released, which should permit for the deployment of a host of bioanalytical methods on NACS sorted cells. We envision that NACS will find uses beyond multiplexed sorting of T cells based on TCR specificity. The principal components of this platform—streptavidin-cysteine core and orthogonal

single stranded DNA sequences—were rationally developed to enable oriented coupling and spatial addressing. Thus this platform is amenable to any family of binding proteins or small molecule binders labeled with biotin. The increase in avidity of p/MHC tetramers over monomers as a consequence of the valency of SA should likewise extend to other capture agents, making it feasible to generate cellular arrays with probes ranging from high to moderate affinities like antibodies, aptamers or peptides.

3.5 References

1. Arstila, T.P.; Casrouge, A.; Baron, V.; Even, J.; Kanellopoulos, J.; Kourilsky, P. A Direct Estimate of the Human alpha beta T Cell Receptor Diversity. *Science* **1999**, *286*, 958–961.
2. Altman, J.D.; Moss, P.A.H.; Goulder, P.J. R.; Barouch, D.H.; McHeyzer-Williams, M.G.; Bell, J.I.; McMichael, A.J.; Davis, M.M. Phenotypic analysis of antigen-specific T lymphocytes. *Science* **1996**, *274*, 94–96.
3. Matsui, K.; Boniface, J.J.; Steffner, P.; Reay, P.A.; Davis, M.M. Kinetics of T-cell receptor binding to peptide/I-Ek complexes: correlation of the dissociation rate with T cell responsiveness. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 12862–12866.
4. McLaughlin, B.E.; Baumgarth, N.; Bigos, M.; Roederer, M.; DeRosa, S.C., Altman, J.D.; Nixon, D.F.; Ottinger, J.; Oxford, C.; Evans, T.G.; Asmuth, D.M. Nine-color flow cytometry for accurate measurement of T cell subsets and cytokine responses. Part I: Panel design by an empiric approach. *Cytometry A* **2008**, *73*, 400–410.
5. Chattopadhyay, P.K.; Price, D.A.; Harper, T.F.; Betts, M.R.; Yu, J.; Gostick, E.; Perfetto, S.P.; Goepfert, P.; Koup, R.A.; De Rosa, S.C.; Bruchez, M.P.; Roederer, M. Quantum dot semiconductor nanocrystals for immunophenotyping by polychromatic flow cytometry. *Nat. Med.* **2006**, *12*, 972–977.
6. Chen, D.S.; Soen, Y.; Stuge, T.B.; Lee, P.P.; Weber, J.S.; Brown, P.O.; Davis, M.M. Marked Differences in Human Melanoma Antigen-Specific T Cell Responsiveness after Vaccination Using a Functional Microarray. *PLoS Med.* **2005**, *2*, 1018–1030.
7. 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.* **2003**, *1*, 429–438.

8. 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. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 3744–3749.
9. Deviren, G.; Gupta, K.; Paulaitis, M.E.; Schneck, J.P. Detection of antigen-specific T cells on p/MHC microarrays. *J. Mol. Recognit.* **2007**, *20*, 32–38.
10. 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 Biology* **2001**, *2*, 1–13.
11. Butler, J.E. ; Ni, L.; Brown, W.R.; Joshi, K.S.; Chang, J.; Rosenberg, B.; Voss Jr, E.W. 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.* **1993**, *30*, 1165–1175.
12. Butler, J. E.; Ni, L.; Nessler, R.; Joshi, K.S.; Suter, M.; Rosenberg, B.; Chang, J.; Brown, W.R.; CantareroButler, L.A. The physical and functional behavior of capture antibodies adsorbed on polystyrene. *J Immunol. Methods*, **1992**, *150*, 77–90.
13. MacBeath, G.; Schreiber, S.L. Printing Proteins as Microarrays for High-Throughput Function Determination. *Science* **2000**, *289*, 1760–1763.
14. Lesaicherre, M. L.; Lue, Y.P.R.; Chen, G.Y.J.; Zhu, Q.; Yao, Q. Intein-Mediated Biotinylation of Proteins and Its Application in a Protein Microarray. *J. Am. Chem. Soc.* **2002**, *124*, 8768–8769.
15. Peluso, P.; Wilson, D.; Do, D.; Tran, H.; Venkatasubbaiah, M.; Quincy, D.; Heidecker, B.; Poindexter, K.; Tolani, N.; Phelan, M.; Witte, K.; Jung, L.; Wagner, P.; Nock, S. Optimizing antibody immobilization strategies for the construction of protein microarrays. *Anal. Biochem.* **2003**, *312*, 113–124.

16. Kwon, Y.; Han, Z.; Karatan, E.; Mrksich, M.; Kay, B. K. Antibody arrays prepared by cutinase-mediated immobilization on self-assembled monolayers. *Anal. Chem.* **2004**, *76*, 5713–5720.
17. Arenkov, P.; Kukhtin, A.; Gemmel, A.; Voloshchuk, S.; Chupeeva, V.; Mirzabekov, A. Protein microchips: use for immunoassay and enzymatic reactions. *Anal. Biochem.* **2000**, *278*, 123–131.
18. Kiyonaka, S.; Sada, K.; Yoshimura, I.; Shinkai, S.; Kato, N.; Hamachi, I. Semi-wet peptide/protein array using supramolecular hydrogel. *Nat. Mater.* **2004**, *3*, 58–64.
19. For example, the polyacrylamide surface employed by Davis and co-workers^{6,7} has been discontinued from the manufacturer (Perkin Elmer).
20. 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.* **2007**, *129*, 1959–1967.
21. Boozer, C.; Ladd, J.; Chen, S.F.; Jiang, S.T. DNA-directed protein immobilization for simultaneous detection of multiple analytes by surface plasmon resonance biosensor. *Anal. Chem.* **2006**, *78*, 1515–1519.
22. Niemeyer, C.M. Functional devices from DNA and proteins. *Nano Today* **2007**, *2*, 42–52.
23. 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.* **2006** *45*, 896–901.
24. Hsiao, S.C.; Crow, A.K.; Lam, W.A.; Bertozzi, C.R.; Fletcher, D.A.; Francis, M.B. DNA-coated AFM cantilevers for the investigation of cell adhesion and the patterning of live cells. *Angew. Chem. Int. Ed. Engl.* **2008**, *47*, 8473–7.

25. Fan, R.; Vermesh, O.; Srivastava, A.; Yen, B.K.H.; Qin, L.; Ahmad, H.; Kwong, G.A.; Liu, C.C.; Gould, J.; Hood, L.; Heath, J.R. Integrated barcode chips for rapid, multiplexed analysis of proteins in microliter quantities of blood. *Nat. Biotechnol.* **2008**, *26*, 1373–1378
26. Sano, T.; Cantor, C.R. Expression of a cloned streptavidin gene in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 142–146.
27. Szymczak, A.L.; Workman, C.J.; Wang, Y.; Vignali, K.M.; Dilioglou, S.; Vanin, E.F.; Vignali, D.A.A. Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector. *Nat. Biotechnol.* **2004**, *22*, 589–594.
28. 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. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 3429–3433.
29. Ramachandiran, V.; Grigoriev, V.; Lan, L.; Ravkov, E.; Mertens, S.A.; Altman, J.D. A robust method for production of MHC tetramers with small molecule fluorophores. *J. Immunol. Methods* **2007**, *319*, 13–20.
30. Cameron, T.O.; Cochran, J.R.; Yassine-Diab, B.; Sekaly, R.P.; Stern, L.J. Detection of Antigen-Specific CD4⁺ T Cells by HLA-DR1 Oligomers Is Dependent on the T Cell Activation State. *J. Immunol.* **2001**, *166*, 741–745.
31. Reznik, G.O.; Vajda, S.; Cantor, C.R.; Sano, T. A Streptavidin Mutant Useful for Directed Immobilization on Solid Surfaces. *Bioconj. Chem.* **2001**, *12*, 1000–1004.
32. Green, N.M. Spectrophotometric determination of biotin. *Methods Enzymol.* **1970**, *18*, 418–424.
33. Johnson, L.A.; Heemskerk, B.; Powell, D.J.; Cohen, C.J.; Morgan, R.A.; Dudley, M.E.; Robbins, P.F.; Rosenberg, S.A. 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.* **2006**, *177*, 6548–6559.
34. Nishimura, M.I.; Avichezer, D.; Custer, M.C.; Lee, C.S.; Chen, C.; Parkhurst, M.R.; Diamond, R.A.; Robbins, P.F.; Schwartzentruber, D.J.; Rosenberg, S.A. MHC Class I-restricted Recognition of a Melanoma Antigen by a Human CD4+ Tumor Infiltrating Lymphocyte. *Cancer Res.* **1999**, *59*, 6230–6238.
 35. Wargo, J.A.; Robbins, P.F.; Li, Y.; Zhao, Y.; El-Gamil, M.; Caragacianu, D.; Zheng, Z.; Hong, J.A.; Downey, S.; Schrump, D.S.; Rosenberg, S.A.; Morgan, R.A. Recognition of NY-ESO-1+ tumor cells by engineered lymphocytes is enhanced by improved vector design and epigenetic modulation of tumor antigen expression. *Cancer Immunol. Immunother.* **2009**, *58*, 383–394.
 36. Moon, J.J.; Chu, H.H.; Pepper, M.; McSorley, S.J.; Jameson, S.C.; Kedl, R.M.; Jenkins, M.K. Naive CD4(+) T cell frequency varies for different epitopes and predicts repertoire diversity and response magnitude. *Immunity* **2007**, *27*, 203–213.
 37. For a typical homogeneous sorting experiment, 10^6 T cells are applied to an array containing 144 spots. Each spot can bind to $\sim 1486 \pm 62$ T cells. The recovery, R, is calculated as $R = 100\% \times (144 \times 1486)/10^6 \approx 20\%$. The recovery is valid even with heterogeneous samples (e.g. 1:1 mixture of two antigen-specific T cell populations) when the entire array is used to sort a single target population ($R = 100\% \times (144 \text{ spots} \times 661)/(5 \times 10^5) \approx 20\%$). For multiplexed experiments, R is proportional to the fraction of the 144 spots dedicated for cell capture. For example, sorting 2 populations from a 1:1 mixture, $R = 100\% \times (72 \text{ spots} \times 661)/(5 \times 10^5) \approx 10\%$.
 38. Schumacher, T.N. T-cell-receptor gene therapy. *Nat. Rev. Immunol.* **2002**, *2*, 512–519.
 39. Morgan, R.A.; Dudley, M.E.; Wunderlich, J.R.; Hughes, M.S.; Yang, J.C.; Sherry, R.M.; Royal, R.E.; Topalian, S.L.; Kammula, U.S.; Restifo, N.P.; et al.

- Cancer Regression in Patients After Transfer of Genetically Engineered Lymphocytes. *Science* **2006**, *314*, 126–129.
40. Rosenberg, S.A. A New Era for Cancer Immunotherapy Based on the Genes that Encode Cancer Antigens. *Immunity* **1999**, *10*, 281–287.
 41. Rosenberg, S.A. Progress in human tumour immunology and immunotherapy. *Nature* **2001**, *411*, 380–384.
 42. Dudley M.E.; et al. Cancer Regression and Autoimmunity in Patients After Clonal Repopulation with Antitumor Lymphocytes. *Science* **2002**, *298*, 850–854.
 43. Hunder, N.N.; Wallen, H.; Cao, J.; Hendricks, D.W.; Reilly, J.Z.; Rodmyre, R.; Jungbluth, A.; Gnjjatic, S.; Thompson, J.A.; Yee, C. Treatment of Metastatic Melanoma with Autologous CD4+ T Cells against NY-ESO-1. *N. Engl. J. Med.* **2008**, *358*, 2698–2703.
 44. Toebes, M.; Coccoris, M.; Bins, A.; Rodenko, B.; Gomez, R.; Nieuwkoop, N.J.; van de Kastele, W.; Rimmelzwaan, G.F.; Haanen, J.; Ovaas, H.; Schumacher, T.N. Design and use of conditional MHC class I ligands. *Nat. Med.* **2006**, *12*, 246–251.
 45. Grotenbreg, G.M.; Roan, N.R.; Guillen, E.; Meijers, R.; Wang, J.; Bell, G.W.; Starnbach, M.N.; Ploegh, H.L. Discovery of CD8+ T cell epitopes in *Chlamydia trachomatis* infection through use of caged class I MHC tetramers. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 3831–3836.
 46. Bakker, A.H.; Hoppes, R.; Linnemann, C.; Toebes, M.; Rodenko, B.; Berkers, C.R.; Hadrup, S.R.; van Esch, W.J.; Heemskerk, M.H.; Ovaas, H.; Schumacher, T.N. Conditional MHC class I ligands and peptide exchange technology for the human MHC gene products HLA-A1, -A3, -A11, and -B7. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 3825–3830.

3.6 Appendix A: Protein Sequences

Core streptavidin + Cys in pET-3a vector (NdeI/BamHI)

NdeI

1 CATATGGGCA TCACCGGCAC CTGGTACAAC CAGCTCGGCT CGACCTTCAT
H M G I T G T W Y N Q L G S T F I

51 CGTGACCGCG GGCGCCGACG GCGCCCTGAC CGGAACCTAC GAGTCGGCCG
V T A G A D G A L T G T Y E S A V

101 TCGGCAACGC CGAGAGCCGC TACGTCCTGA CCGGTCGTTA CGACAGCGCC
G N A E S R Y V L T G R Y D S A

151 CCGGCCACCG ACGGCAGCGG CACCGCCCTC GGTTGGACGG TGGCCTGGAA
P A T D G S G T A L G W T V A W K

201 GAATAACTAC CGCAACGCCC ACTCCGCGAC CACGTGGAGC GGCCAGTACG
N N Y R N A H S A T T W S G Q Y V

251 TCGGCGGGCG CGAGGCGAGG ATCAACACCC AGTGGCTGCT GACCTCCGGC
G G A E A R I N T Q W L L T S G

301 ACCACCGAGG CCAACGCCTG GAAGTCCACG CTGGTCGGCC ACGACACCTT
T T E A N A W K S T L V G H D T F

BamHI

351 CACCAAGGTG GGTGGTTCTG GTTGCCCCTAG GGATCC
T K V G G S G C P *

Human β_2m in pET-3a vector NdeI/BamHI

NdeI

1 CATATGATCC AGCGTACTCC AAAGATTCAG GTTTACTCAC GTCATCCAGC
H M I Q R T P K I Q V Y S R H P A

51 AGAGAATGGA AAGTCAAATT TCCTGAATTG CTATGTGTCT GGGTTTCATC
E N G K S N F L N C Y V S G F H P

101 CATCCGACAT TGAAGTTGAC TTAAGTGAAGA ATGGAGAGAG AATTGAAAAA
S D I E V D L L K N G E R I E K

151 GTGGAGCATT CAGACTTGTC TTTCAGCAAG GACTGGTCTT TCTATCTCTT
V E H S D L S F S K D W S F Y L L

201 GTATTATACT GAATTCACCC CCACTGAAAA AGATGAGTAT GCCTGCCGTG
Y Y T E F T P T E K D E Y A C R V

251 TGAACCACGT GACTTTGTCA CAGCCCAAGA TAGTTAAGTG GGATCGAGAC
 N H V T L S Q P K I V K W D R D

BamHI

301 ATGTAA GGATCC
 M *

Human HLA-A2.1 (biotin tag)

1 ATGGGCTCTC ACTCCATGAG GTATTTCTTC ACATCCGTGT CCCGGCCCCG
 M G S H S M R Y F F T S V S R P G

51 CCGCGGGGAG CCCCCTTCA TCGCAGTGGG CTACGTGGAC GACACGCAGT
 R G E P R F I A V G Y V D D T Q F

101 TCGTGCGGTT CGACAGCGAC GCCGCGAGCC AGAGGATGGA GCCGCGGGCG
 V R F D S D A A S Q R M E P R A

151 CCGTGGATAG AGCAGGAGGG TCCGGAGTAT TGGGACGGGG AGACACGGAA
 P W I E Q E G P E Y W D G E T R K

201 AGTGAAGGCC CACTCACAGA CTCACCGAGT GGACCTGGGG ACCCTGCGCG
 V K A H S Q T H R V D L G T L R G

251 GCTACTACAA CCAGAGCGAG GCCGGTTCTC ACACCGTCCA GAGGATGTAT
 Y Y N Q S E A G S H T V Q R M Y

301 GGCTGCGACG TGGGGTCGGA CTGGCGCTTC CTCCGCGGGT ACCACCAGTA
 G C D V G S D W R F L R G Y H Q Y

351 CGCCTACGAC GGCAAGGATT ACATCGCCCT GAAAGAGGAC CTGCGCTCTT
 A Y D G K D Y I A L K E D L R S W

401 GGACCGCGGC GGACATGGCA GCTCAGACCA CCAAGCACAA GTGGGAGGCG
 T A A D M A A Q T T K H K W E A

451 GCCCATGTGG CGGAGCAGTT GAGAGCCTAC CTGGAGGGCA CGTGCGTGGA
 A H V A E Q L R A Y L E G T C V E

501 GTGGCTCCGC AGATACCTGG AGAACGGGAA GGAGACGCTG CAGCGCACGG
 W L R R Y L E N G K E T L Q R T D

551 ACGCCCCAA AACGCATATG ACTCACCACG CTGTCTCTGA CCATGAAGCC
 A P K T H M T H H A V S D H E A

601 ACCCTGAGGT GCTGGGCCCT GAGCTTCTAC CCTGCGGAGA TCACACTGAC
 T L R C W A L S F Y P A E I T L T

651 CTGGCAGCGG GATGGGGAGG ACCAGACCCA GGACACGGAG CTCGTGGAGA

W Q R D G E D Q T Q D T E L V E T
 701 CCAGGCCTGC AGGGGATGGA ACCTTCCAGA AGTGGGCGGC TGTGGTGGTG
 R P A G D G T F Q K W A A V V V
 751 CCTTCTGGAC AGGAGCAGAG ATACACCTGC CATGTGCAGC ATGAGGGTTT
 P S G Q E Q R Y T C H V Q H E G L
 801 GCCCAAGCCC CTCACCGGAT CCGGTGGTTC CGGTGGTTCC GCGGGTGGTG
 P K P L T G S G G S G G S A G G G
 851 GTTTGAACGA CATCTTCGAA GCTCAGAAAA TCGAATGGCA CTAA
 L N D I F E A Q K I E W H *

BirA ligase biotin sequence: GGGLNDIFEAQKIEWH

Murine β_2m

1 ATGATCCAGA AAACCCCTCA AATTCAAGTA TACTCACGCC ACCCACCGGA
 M I Q K T P Q I Q V Y S R H P P E
 51 GAATGGGAAG CCGAACATAC TGAACTGCTA CGTAACACAG TTCCACCCGC
 N G K P N I L N C Y V T Q F H P P
 101 CTCACATTGA AATCCAAATG CTGAAGAACG GGAAAAAAT TCCTAAAGTA
 H I E I Q M L K N G K K I P K V
 151 GAGATGTCAG ATATGTCCTT CAGCAAGGAC TGGTCTTTCT ATATCCTGGC
 E M S D M S F S K D W S F Y I L A
 201 TCACACTGAA TTCACCCCA CTGAGACTGA TACATACGCC TGCAGAGTTA
 H T E F T P T E T D T Y A C R V K
 251 AGCATGACAG TATGGCCGAG CCAAGACCG TCTACTGGGA TCGAGACATG
 H D S M A E P K T V Y W D R D M
 301 TGA
 *

Murine Kb (biotin tag)

NdeI
 1 CATATGGGTC CACTCTCTCT GCGCTATTTT CTTACGGCTG TTAGCCGTCC
 H M G P H S L R Y F V T A V S R P
 51 GGGTCTGGGT GAGCCGCGCT ACATGGAAGT CGGTTACGTC GACGACACCG

G L G E P R Y M E V G Y V D D T E
 101 AATTCGTGCG TTTCGACAGC GACGCGGAGA ACCCGCGTTA TGAGCCGCGT
 F V R F D S D A E N P R Y E P R
 151 GCGCGTTGGA TGGAGCAGGA AGGTCCGGAG TACTGGGAGC GTGAAACGCA
 A R W M E Q E G P E Y W E R E T Q
 201 AAAGCGAAG GGCAATGAAC AGAGCTTTTCG TGTTGATCTG CGCACTCTGC
 K A K G N E Q S F R V D L R T L L
 251 TGGGTTACTA CAACCAGAGC AAAGGTGGCA GCCATAACCAT TCAGGTGATT
 G Y Y N Q S K G G S H T I Q V I
 301 AGCGGTTGTG AAGTCGGCTC TGATGGCCGC CTGTTGCGCG GTTATCAGCA
 S G C E V G S D G R L L R G Y Q Q
 351 ATATGCATAC GACGTTGCG ACTACATTGC GCTGAATGAA GATCTGAAAA
 Y A Y D G C D Y I A L N E D L K T
 401 CGTGGACTGC GGCGGACATG GCCGCACTGA TTACCAAACA CAAGTGGGAG
 W T A A D M A A L I T K H K W E
 451 CAAGCGGGCG AAGCCGAGCG CCTGCGTGCG TATCTGGAAG GCACCTGTGT
 Q A G E A E R L R A Y L E G T C V
 501 GGAATGGCTG CGCCGCTATC TGAAGAATGG CAATGCCACG TTGCTGCGTA
 E W L R R Y L K N G N A T L L R T
 551 CGGATTCCCC GAAAGCGCAC GTGACGCACC ATAGCCGTCC TGAGGATAAA
 D S P K A H V T H H S R P E D K
 601 GTTACCCTGC GTTGCTGGGC ACTGGGCTTT TACCCGGCAG ATATCACCTT
 V T L R C W A L G F Y P A D I T L
 651 GACGTGGCAA CTGAATGGTG AAGAGCTGAT TCAGGATATG GAACTGGTGG
 T W Q L N G E E L I Q D M E L V E
 701 ARACTCGTCC GGCTGGCGAC GGTACCTTCC AGAAATGGGC ATCGGTTGTC
 T R P A G D G T F Q K W A S V V
 751 GTCCCTCTGG GTAAAGAGCA ATACTATAACC TGCCACGTTT ACCACCAAGG
 V P L G K E Q Y Y T C H V Y H Q G
 801 TCTGCCGGAG CCGCTGACCT TCGGTTGGGA GCCACCGCCG AGCACCGGCA
 L P E P L T L R W E P P P S T G S
 851 GCGGTGGTAG CGGCGGTTCC GCGGGTGGCG GTCTGAACGA CATCTTTGAG
 G G S G G S A G G G L N D I F E
 BamHI
 901 GCCCAGAAGA TCGAGTGGCA TTAAGGATCC

A Q K I E W H *

BirA ligase biotin sequence: GGGLNDIFEAQKIEWH

Murine Db (biotin tag)

1 ATGGGCCCAC ACTCGATGCG GTATTTTCGAG ACCGCCGTGT CCCGGCCCCG
M G P H S M R Y F E T A V S R P G

51 CCTCGAGGAG CCCCAGTACA TCTCTGTCGG CTATGTGGAC AACAAAGGAGT
L E E P R Y I S V G Y V D N K E F

101 TCGTGCGCTT CGACAGCGAC GCGGAGAATC CGAGATATGA GCCGCGGGCG
V R F D S D A E N P R Y E P R A

151 CCGTGGATGG AGCAGGAGGG GCCGGAGTAT TGGGAGCGGG AAACACAGAA
P W M E Q E G P E Y W E R E T Q K

201 AGCCAAGGGC CAAGAGCAGT GGTTCGAGT GAGCCTGAGG AACCTGCTCG
A K G Q E Q W F R V S L R N L L G

251 GCTACTACAA CCAGAGCGCG GCGGCTCTC ACACACTCCA GCAGATGTCT
Y Y N Q S A G G S H T L Q Q M S

301 GGCTGTGACT TGGGGTCGGA CTGGCGCCTC CTCCGCGGGT ACCTGCAGTT
G C D L G S D W R L L R G Y L Q F

351 CGCCTATGAA GGCCGCGATT ACATCGCCCT GAACGAAGAC CTGAAAACGT
A Y E G R D Y I A L N E D L K T W

401 GGACGGCGGC GGACATGGCG GCGCAGATCA CCCGACGCAA GTGGGAGCAG
T A A D M A A Q I T R R K W E Q

451 AGTGGTGCTG CAGAGCATTA CAAGGCCTAC CTGGAGGGCG AGTGCGTGGA
S G A A E H Y K A Y L E G E C V E

501 GTGGCTCCAC AGATACCTGA AGAACGGGAA CGCGACGCTG CTGCGCACAG
W L H R Y L K N G N A T L L R T D

551 ATTCCCCAAA GGCACATGTG ACCCATCACC CCAGATCTAA AGGTGAAGTC
S P K A H V T H H P R S K G E V

601 ACCCTGAGGT GCTGGGCCCT GGGCTTCTAC CCTGCTGACA TCACCCTGAC
T L R C W A L G F Y P A D I T L T

651 CTGGCAGTTG AATGGGGAGG AGCTGACCCA GGACATGGAG CTTGTGGAGA
W Q L N G E E L T Q D M E L V E T

701 CCAGGCCTGC AGGGGATGGA ACCTTCCAGA AGTGGGCATC TGTGGTGGTG
R P A G D G T F Q K W A S V V V

751 CCTCTTGGGA AGGAGCAGAA TTACACATGC CGTGTGTACC ATGAGGGGCT
P L G K E Q N Y T C R V Y H E G L

801 GCCTGAGCCC CTCACCCTGA GATGGGAGCC TCCTCCATCC ACTGGATCCG
P E P L T L R W E P P P S T G S G

851 GTGGTTCCGG TGGTTCCGCG GGTGGTGGTT TGAACGACAT CTTCGAAGCT
G S G G S A G G G L N D I F E A

901 CAGAAAATCG AATGGCACTA A
Q K I E W H *

3.7 Appendix B: Chromatography

3.7.1 Iminobiotin Purification of SAC

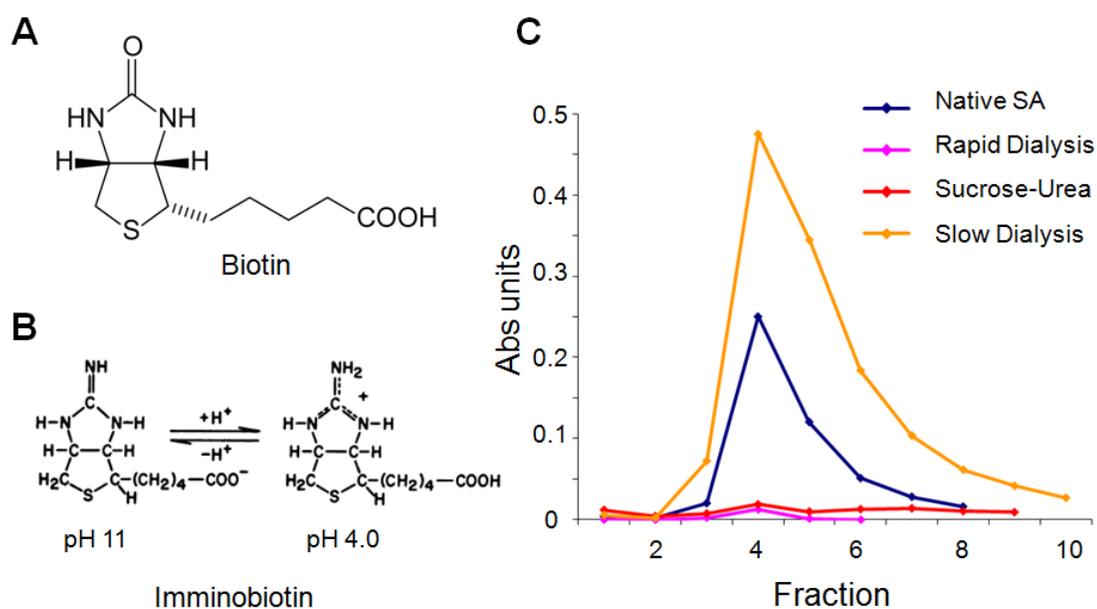
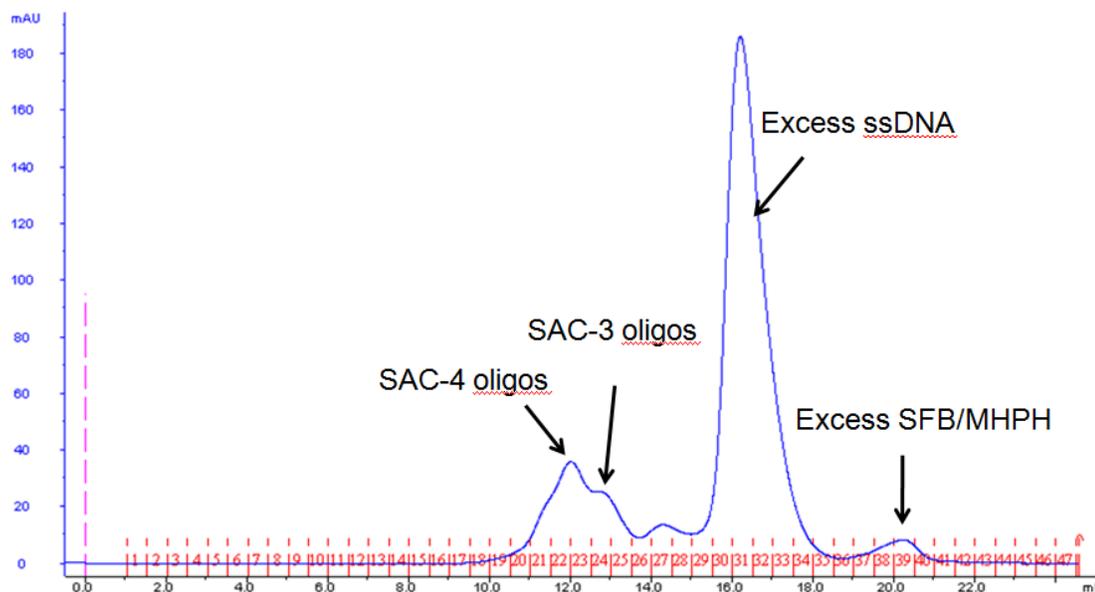


Figure 3.15 SAC purification with an iminobiotin agarose support column. (A) The structure of biotin. (B) Iminobiotin, the structural analog of biotin, has two distinct states; (1) a neutral, high affinity SAC binding state at pH 11 and a cationic state, low affinity SAC binding state at pH 4. (C) Elution profile of SA compared to SAC refolded in various buffers.

3.7.2 FPLC of SAC-ssDNA conjugates



*0.5ml/min isocratic flow of PBS

Figure 3.16 Fast protein liquid chromatography of ssDNA-SAC conjugates. A typical successful conjugation reaction will yield three distinct peaks, corresponding to the ssDNA-SAC conjugates, excess unreacted ssDNA, and excess small molecules. It is possible to resolve shoulders in the ssDNA-SAC peak, corresponding to extent of ssDNA modification.

Chapter 4

Detection of Cell Surface Markers with Encoded ssDNA Reporters: Towards Global Cell Surface-ome Profiling

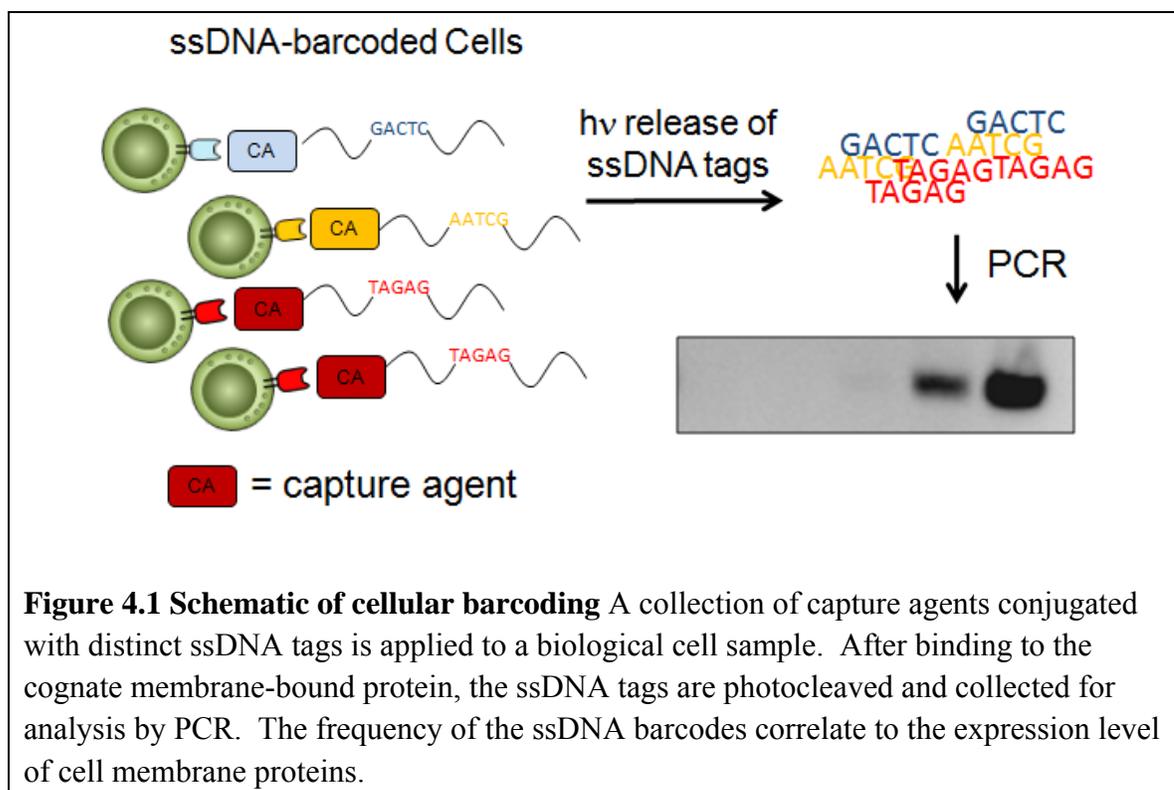
4.1 Introduction

The proteins on the surface of the cell membrane play important roles in various aspects of tumor biology. Cell surface markers are involved in cancer pathogenesis, aid in staging, and represent a large class of proteins targeted for therapy. The importance of the membrane-bound oncogenes EGFR and ERBB2 and the development of small molecule inhibitors against them have been expounded on in the previous chapters. Identification of altered or expression of cancer associated surface antigens is an area of active research. For example, the cell surface marker CD40 has been found to be expressed in many B cell malignancies and has been investigated as the possible target for anti-CD40 antibody-based cancer therapy (1). Development of methods that enable

comprehensive mapping of the cell surface proteome would provide new avenues for investigation, analogous to the effect global transcriptome expression profiling has had in providing valuable fundamental and therapeutic information for various types of cancers (2). However, there have been few documented effective strategies for high-throughput, global profiling of surface membrane proteins. One approach is to purify cellular membrane fragments by two-phase separation, but cross-contamination from cytosolic proteins is a major limitation (3). Other reports have demonstrated the feasibility of retrieving membrane fragments *in vivo* from endothelial cells using a combination of colloidal silica particles and polymers (4). Membrane proteins have also been isolated through chemical biotinylation followed by enrichment with a streptavidin column (5–7). Typically after enrichment, the membrane bound proteins are identified by the combination of 2-D gel electrophoresis and mass spectrometry. A major limitation of these studies is that the labeling strategies are non-specific and the biological samples are lysed, precluding dynamic studies.

In this chapter summarizing current work, I present an antibody-based, membrane-protein profiling approach, which uses a library of capture agents to probe membrane-bound antigens. Similar to DEAL or NACS conjugates, each capture agent is conjugated to a distinct ssDNA tag but differs in that the sequence incorporates a photolabile base. The capture agents are allowed to bind to cell surface antigens, after which the ssDNA tags are released into solution by UV-induced cleavage, collected and detected by PCR (**Figure 4.1**). This approach is called cellular barcoding. There have been quite a few studies integrating the specificity of antibodies for antigen detection with nucleic acid readouts, including immuno-PCR (8–10), immunodetection amplified

by T7 RNA polymerase (11–12), proximity ligation (13–14), and nanoparticle assays (15). However, all these assays were used to detect single proteins in idealized solutions, and were not utilized in a high throughput, multi-parameter manner for detecting and profiling cell surface proteins.



4.2 Experimental Methods

4.2.1 DNA sequences and production of conjugates

All DNA sequences were purchased with HPLC purification from Integrated DNA Technologies (www.idtdna.com) and are listed in Table 4.1.

Table 4.1 Cellular barcoding DNA sequences

Name	Sequence*
PC ¹	5' – NH ₂ – PC – ATC CTG GAG CTA AGT CCG TAG CCT CAT TGA ATC ATG CCT AGC ACT CGT CTA CTA TCG CTA
PC forward primer	5' – ATG GTC GAG ATG TCA GAG TAA TCC TGG AGC TAA GTC CGT A
PC reverse primer	5' – TAG ATA CTG CCA CTT CAC ATT AGC GAT AGT AGA CGA GTG C
A'	5' - NH ₂ - AAA AAA AAA ATA CGG ACT TAG CTC CAG GAT-cy3
* The 5' amine functional group for sequences PC and EcoRV is necessary to conjugate the oligonucleotide to SAC or antibodies.	
¹ PC = photocleavable	

Antibody-ssDNA (DEAL) conjugates were synthesized, purified and characterized according to previously published protocols (16). The HLA-A*0201 restricted MHC class I monomers loaded with MART-1₂₆₋₃₅ (ELAGIGILTV) were produced in house according to previous published protocols (17). The production of p/MHC-ssDNA constructs (NACS conjugates) were according to previous published protocols (18).

4.2.2 Detection of surface markers with PCR

Prior to all experiments, blocking buffer (1.5% BSA, 150 µg/ml salmon sperm DNA in PBS) was used to block all 1.5 ml tubes and a 96-well plate for 1 hour at room temperature (RT) before rinsing 2x with PBS. The blocked tubes and plate were used immediate for experiments. One million cells (Jurkat^{α-MART-1}, Jurkat^{α-Tyro}, or GBM1600) were transferred to 1.5 ml tubes, and resuspended in 100 µl staining buffer (HBSS supplemented with 2.5mg/ml BSA, 10mM HEPES, 0.01% azide). The cells were blocked by adding 100 µl of 2mg/ml salmon sperm DNA for 20 min. at 37°C before

resuspension in staining buffer. Capture agents tagged with ssDNA (Cetuximab, MART-1) were added (0.5 μg per 10^6 cells) to the cell suspension for 20–25 min. at 37°C. The samples were then washed 3x with staining buffer before a final wash and resuspension in 1% BSA PBS. For UV-dependent cleavage of the ssDNA tags, 5×10^5 cells in 50 μl were transferred to a 96-well plate on ice and exposed to long wave UV for 1 hr. The cells were pelleted and the supernatant containing the DNA codes was collected for PCR analysis.

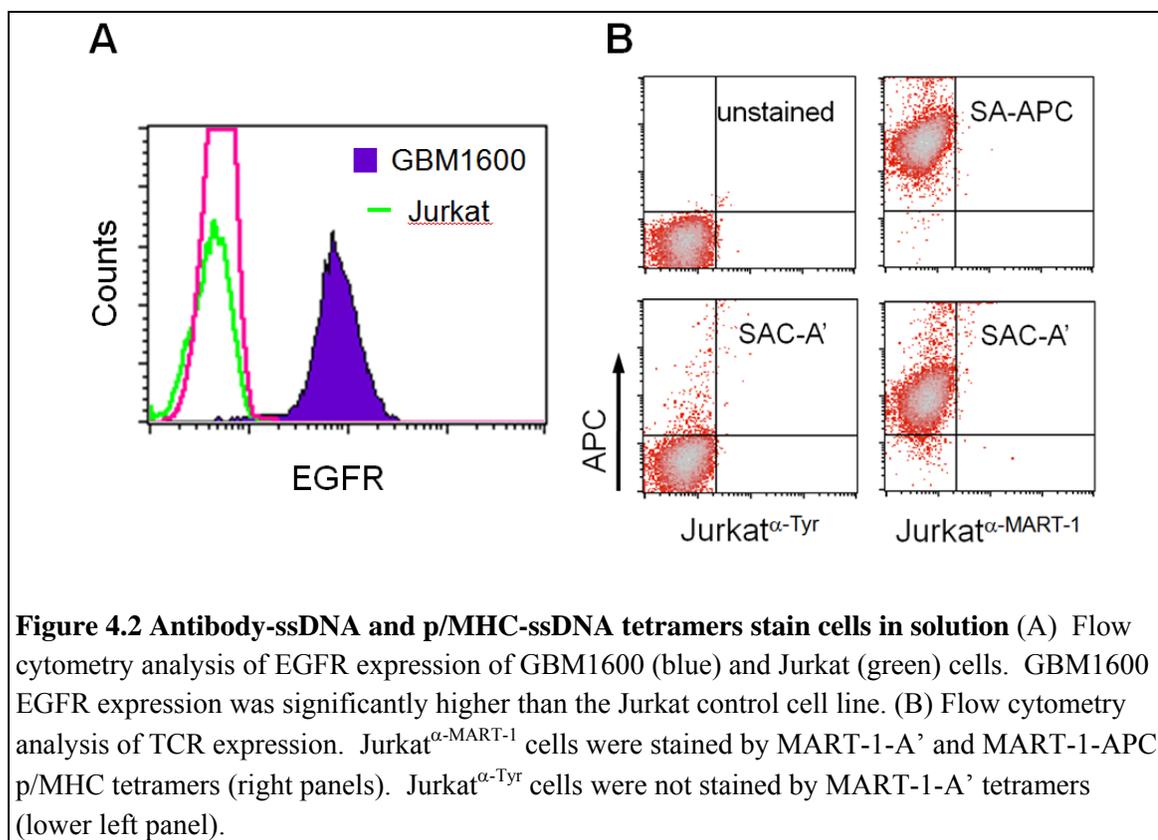
4.3 Results and Discussion

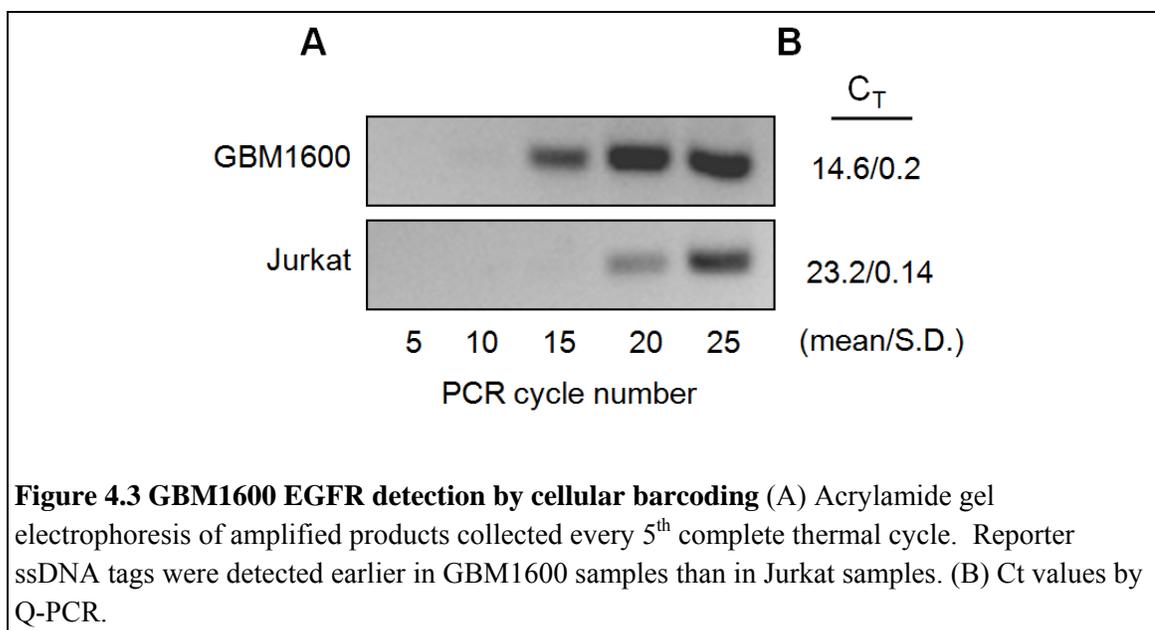
4.3.1 Detection of differential cell surface expression of EGFR

To illustrate the feasibility of using ssDNA-antibody conjugates to detect membrane-bound proteins by PCR, we selected two cell lines expressing differing levels of human epidermal growth factor receptor (EGFR). GBM1600 cells are a low-passage cell line derived from a primary brain tumor expressing high levels of EGFR. Jurkat cells are a hematopoietic derived T cell line with null expression of EGFR. The presence or absence of EGFR expression was verified flow cytometry (**Figure 4.2A**).

We prepared anti-EGFR (Cetuximab) conjugated with photocleavable ssDNA oligonucleotide (PC) and stained GBM1600 and Jurkat cells in separate tubes with the conjugate. After removing excess Cetuximab-PC molecules, the samples were treated in a 96-well plate with UV radiation for one hour. The photo-cleaved ssDNA tags were collected, amplified by PCR and visualized on a 4% agarose gel. As shown in **Figure 4.3**, the reporter tags were detected at an earlier thermal cycle during the amplification in the GBM1600 sample than in the Jurkat sample. Quantitative assessment by Q-PCR

gave a ΔCt of approximately ~ 8.6 (**Figure 4.3b**). Under ideal experimental conditions, each thermal cycle doubles the total number of amplicons. Assuming EGFR expression directly correlates with the number of ssDNA tags released in solution, the relative difference in EGFR expression between the GBM1600 and Jurkat samples is given by $2^{8.6}$, which is approximately equal to 400. In reality, since most processes are non-ideal, this likely represents the upper bound. Assuming each thermal cycle increases the number of amplicons by a factor of 1.8 instead of 2 gives a relative difference of $1.8^{8.6}$, which is approximately equal to 160.

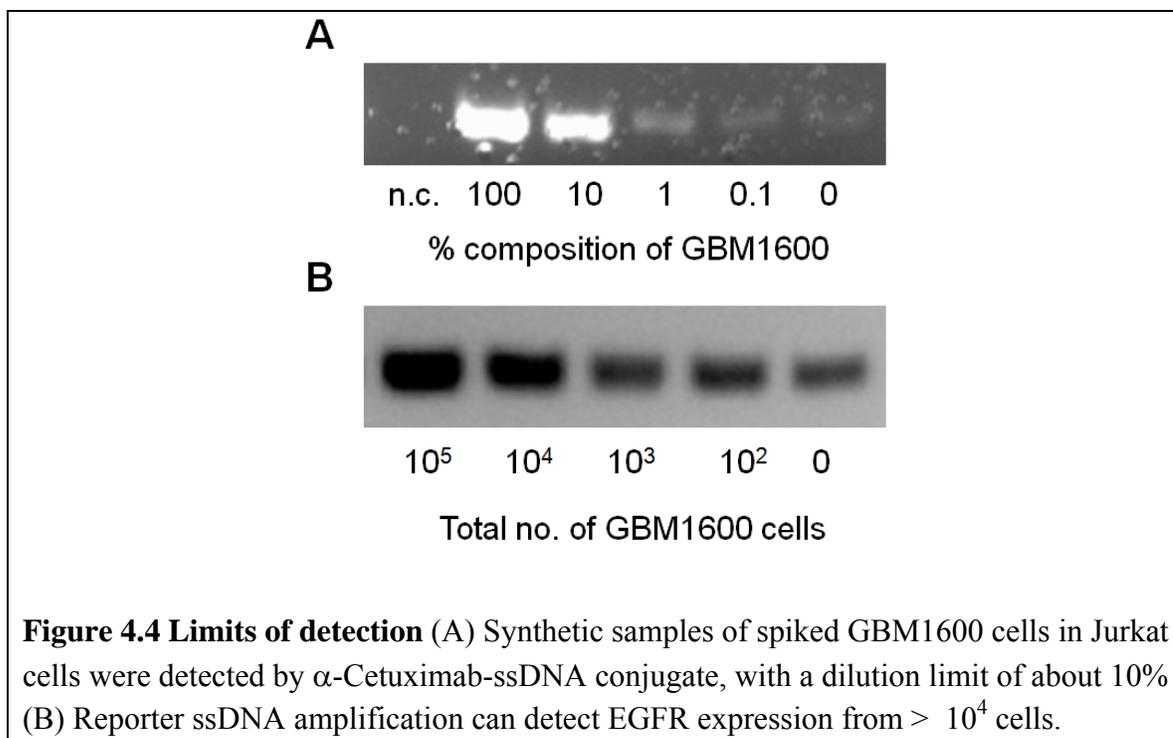




4.3.2 Cellular barcoding limits of detection

In order to determine the minimum number of cells required for surface antigen detection, two avenues were investigated. First, GBM1600 cells were spiked into Jurkat cells to produce synthetic cellular mixtures composed of 100%, 10%, 1%, 0.1%, or 0% GBM1600 cells. The total number of cells per condition was kept constant at 10^6 by increasing the number of Jurkat cells. Second, GBM1600 cells were serially diluted into separate tubes, each containing either 10^5 , 10^4 , 10^3 , 10^2 , or 0 cells per tube. No Jurkat cells were added to these vials. These samples were stained using Cetuximab-PC conjugates and the results are shown in **Figure 4.5**. For the cell mixtures, GBM1600 cells at 100% and 10% were detected significantly above the baseline while mixtures 1% and 0.1% were similar in intensity to the baseline (**Figure 4.5A**). In the serial dilutions, samples containing 10^4 cells or more were detected above background (**Figure 4.5B**). Significant signal arose from the vial without any cells, which is likely due to non-specific interactions between Cetuximab-PC conjugates and the vial. We are currently

investigating improvements in the blocking and washing steps to increase the signal to noise ratio.



4.3.2 Detection of antigen-specific T cells using ssDNA-p/MHC tetramers

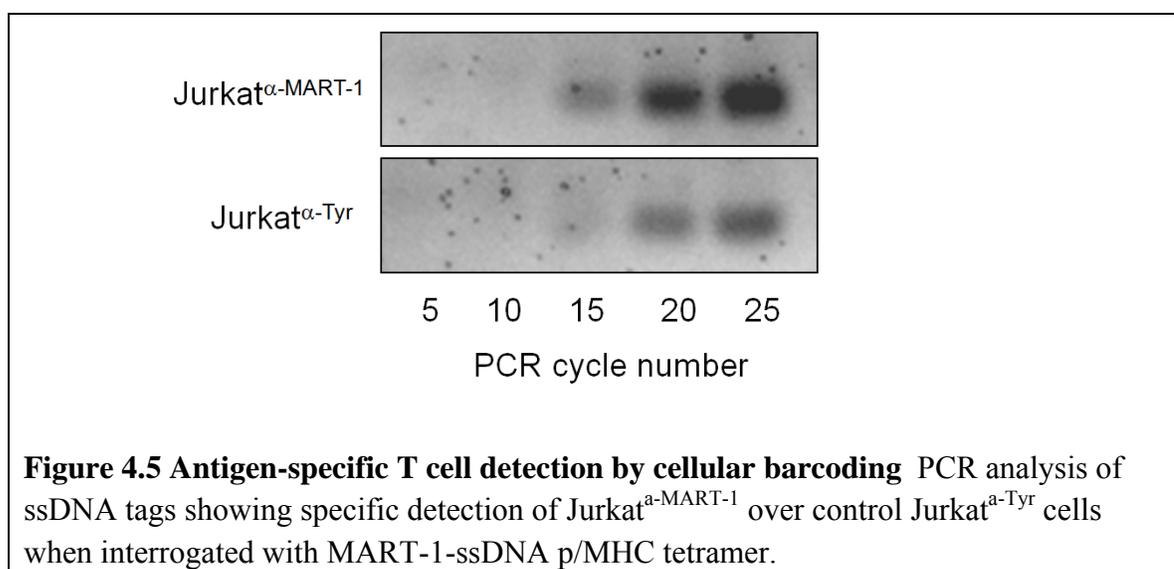
The ability to profile T cell receptors has many potential fundamental applications in immunology as well as diagnostic potential in many fields like vaccine development, immune monitoring, and cancer immunotherapy. Since the development of p/MHC tetramers (19), it has been possible to directly analyze populations of antigen-specific T cells by staining with fluorescent p/MHC tetramers. The fundamental limitation in this approach is the difficulty in multiplexing since distinct T cell specificities need to be encoded with different colors for discrimination. Therefore, there has been widespread interest in developing p/MHC protein arrays, since in a spatially encoded format, the

degree of multiplexing can be increased greatly (18, 20-23). One inherent limitation of array approaches however, is that the sensitivity and scalability is limited by the size of the spots, since the sensitivity is directly proportional to the diameter of the capture surfaces. In addition, arrays approaches are planar substrate-based detection schemes. The limited diffusive capacity of T cells prevents cells from sampling an entire array. Hence for all arrays approaches, T cell recovery is low. The majority of the cells settle on inert areas of the substrate or on non-cognate spots. With an approach like cellular barcoding, the engagement of the encoding agent with the cell surface receptor occurs in solution. Thus by using NACS conjugates, all T cells in a solution is encoded after staining and can subsequently be decoded.

To demonstrate the feasibility of detecting antigen-specific TCRs with cellular barcoding, the T cell lines Jurkat ^{α -MART-1} and Jurkat ^{α -Tyr} were chosen as the model lines. These are transduced T cells expressing TCRs specific for the antigens MART-1 and tyrosinase respectively and were described in detail in chapter 3. MART-1 p/MHC tetramers were prepared encoded with fluorescent DNA (A') or photocleavable DNA (PC). MART-1-A' p/MHC tetramers were compared with fluorescent MART-1-APC p/MHC tetramers for their ability to stain Jurkat ^{α -MART-1} cells. The results are shown in **Figure 4.2B**. Both tetramers stained Jurkat ^{α -MART-1} cells similarly (right upper, right lower panels). Importantly, cells expressing the non-cognate TCR were not stained by MART-1-A' tetramers (lower left panel), demonstrating that p/MHC tetramers appended with ssDNA pendants can engage with the cognate TCR in the suspension phase.

Both T cell lines were then stained with MART-1-PC tetramers, treated with UV and analyzed by PCR. The results are shown in **Figure 4.4**. Approximately five fewer

thermal cycles were required before PC was detected from the Jurkat^{α-MART-1} sample than from the Jurkat^{α-Tyr} sample. In comparison with EGFR detection by Cetuximab-PC, the discrimination between the two T cell lines is less significant. This may be attributed to higher affinity Cetuximab-PC conjugates relative to ssDNA-p/MHC complexes. We are currently investigating more comprehensive purification strategies that would yield the highest avidity ssDNA-p/MHC complex by separating fully tetrameric p/MHC constructs from complexes of lower valency.



4.4 Conclusions and Future Directions

In its present state, the dynamic range of membrane antigen detection with antibody-DNA conjugates is approximately 10^2 . This is sufficient in distinguishing cells that express a surface antigen from cells that do not, since most proteins are found on the cell surface at approximately 2,000–500,000 copies per cell (24). Improvements to this technique that expands the dynamic range will be beneficial as well as improvements that

lower the sensitivity of this approach, which would allow the detection of membrane proteins from small sample sizes below 10^4 cells. In addition, with heterogeneous mixtures of cells, it may be necessary to enrich for a particular phenotype prior to cellular barcoding, since a surface antigen that has been detected cannot be assigned to any particular cell type within a mixture *a priori*. To this end, it may be advantageous to integrate DEAL/NACS cell sorting prior to barcoding.

While this demonstration using conventional PCR highlights the technical feasibility of this approach, the goal is integration with high-throughput sequencing employing the Solexa platform. This second generation sequencer works by ligating cDNA libraries to the bottom of microfluidic flow cells. The flow cells are sufficiently large to ensure scattered coverage, enabling spatial resolution of individual cDNA fragments. After *in situ* amplification, the sequences are read by a sequence-by-synthesis approach. Typically, 25 base pair reads are taken from 10^7 – 10^8 possible unique sequences per flow cell (25, 26). This approach is digital and quantitative because each read is from a cDNA fragment that is spatially resolved. By integrating a library of capture agents encoded with ssDNA identifiers with Solexa sequencing, it should be possible to generate a quantitative cell surface-ome heat map. A schematic is illustrated in **Figure 4.6**.

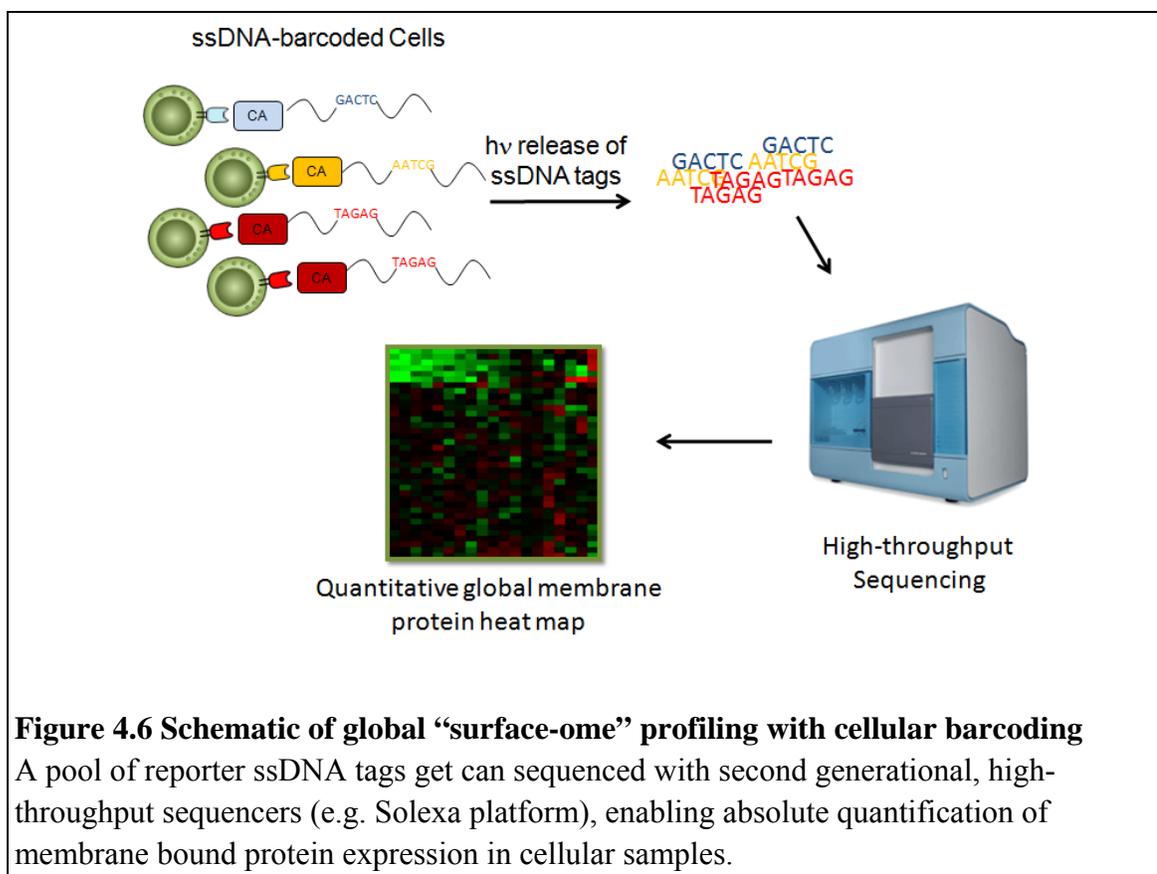
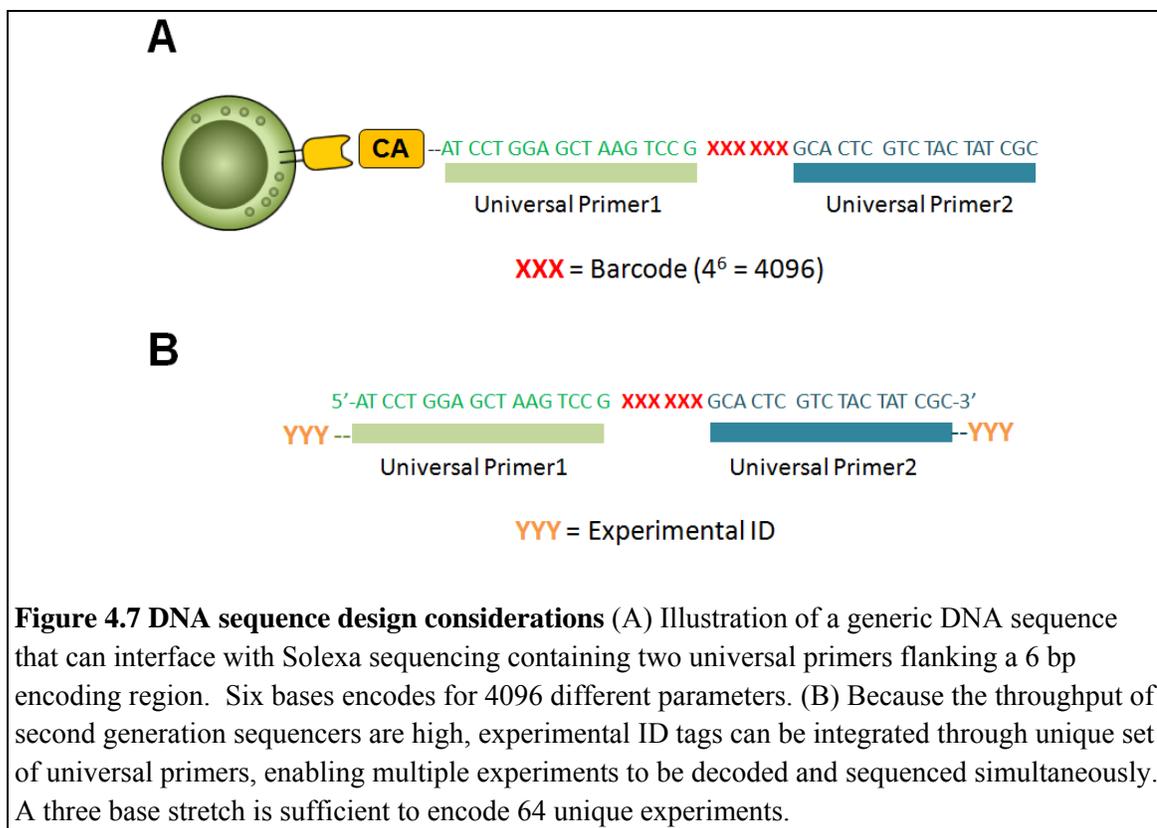


Figure 4.6 Schematic of global “surface-ome” profiling with cellular barcoding

A pool of reporter ssDNA tags get can sequenced with second generational, high-throughput sequencers (e.g. Solexa platform), enabling absolute quantification of membrane bound protein expression in cellular samples.

There are several DNA sequence design considerations to interface with Solexa sequencing. First, since 25 base pair reads are commonly employed, the barcoding segment of the DNA should lie within 25 base pairs from the 5' or 3' end. Second, the barcoding region can be small. A stretch of 6 bases will encode for $4^6 = 4096$ different parameters. Third, prior to ligation to the flow cell, small cDNA libraries are typically amplified by PCR. To avoid biasing the ssDNA tag library, universal primers can be integrated, flanking the barcode (**Figure 4.7A**). Lastly, because of the high-throughput and current cost of each sequencing run, it should be possible to utilize different sets of universal primers with 3 base pair overhangs as experimental identifiers. This would allow barcode tags from multiple experiments to be pooled together for a single

sequencing run (**Figure 4.7B**). A 3 base stretch would allow $4^3 = 64$ different experiments to be pooled together.



In conclusion, a solution phase approach has been introduced that allow cell membrane bound proteins to be detected by capture agents like antibodies or p/MHC tetramers encoded with ssDNA tags. The pendant oligonucleotides function as reporter molecules that can be detected and quantified with conventional thermal amplification approaches. The successful integration of a library of ssDNA-encoded captured agents with high-throughput sequencing for quantitative assessment would provide a global profiling tool to survey the cell surface-ome with fundamental and diagnostic implications.

4.5 References

1. Ottaiano, A.; Pisano, C.; De Chiara, A.; Ascierto, P.A.; et al. CD40 activation as potential tool in malignant neoplasms. *Tumori* **2002**, *88*, 361–366.
2. Chan, S.M.; Utz, P.J. The challenge of analyzing the proteome in humans with autoimmune diseases. *Ann. NY Acad. Sci.* **2005**, *1062*, 61–68.
3. Morre, D.J.; Morre, D.M. Mammalian plasma membranes by aqueous two-phase partition. *Biotechniques* **1989**, *7*, 946–958.
4. Jacobson, B.S.; Stolz, D.B.; Schnitzer, J.E. Identification of endothelial cell–surface proteins as targets for diagnosis and treatment of disease. *Nat. Med.* **1996**, *2*, 482–484.
5. Jang, J.H.; Hanash, S. Profiling of the cell surface proteome. *Proteomics* **2003**, *3*, 1947–1954.
6. Sostaric, E.; Georgiou, A.S.; Wong, C.H.; Watson, P.F.; Holt, W.V.; Fazeli, A. Global profiling of surface plasma membrane proteome of oviductal epithelial cells. *J. Prot. Res.* **2006**, *5*, 3029–3037.
7. Shin, B.K.; Wang, H.; Yim, A.M.; Le Naour, F.; Brichory, F.; Jang, J.H.; Zhao, R.; Puravs, E.; Tra, J.; Michael, C.W.; Misek, D.E.; Hanash, S.M. Global Profiling of the Cell Surface Proteome of Cancer Cells Uncovers an Abundance of Proteins with Chaperone Function. *J. Biol. Chem.* **2003**, *278*, 7607–7616.
8. Hendrickson, E.R.; Hatfield, T.M.; Truby, T.M.; Joerger, R.D.; Majarian, W.R.; Ebersole, R.C. High sensitivity multianalyte immunoassay using covalent DNA-labeled antibodies and polymerase chain reaction. *Nucleic Acid Res.* **1995**, *23*, 522–529.
9. Niemeyer, C.M.; Adler, M.; Wacker, R. Immuno-PCR: High Sensitivity Detection of Proteins by Means of Nucleic Acid Amplification. *Trends Biotechnol.* **2005**, *23*, 208–216.

10. Sano, T.; Smith, C.L.; Cantor, C.R. Immuno-PCR: very sensitive antigen detection by means of specific antibody-DNA conjugates. *Science* **1992**, *258*, 120–122.
11. Zhang, H.T.; Kacharina, J.E.; Miyashiro, K.; Greene, M.I.; Eberwine, J. Protein quantification from complex protein mixtures using a proteomics methodology with single-cell resolution. *Proc. Natl. Sci. USA* **2001**, *98*, 5497–5502.
12. Zhang, H.; Cheng, X.; Richter, M.; Greene, M.I. A sensitive and high-throughput assay to detect low-abundance proteins in serum. *Nat. Med.* **2006** *12*, 473–477.
13. Fredriksson, S.; et al. Multiplexed protein detection by proximity ligation for cancer biomarker validation. *Nat. Methods* **2007**, *4*, 327–329.
14. Gullberg, M.; Gústafsdóttir, S.M.; Schallmeiner, E.; Jarvius, J.; Bjarnegård, M.; Betsholtz, C.; Landegren, U.; Fredriksson, S. Cytokine detection by antibody-based proximity ligation *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 8420–8424.
15. Nam, J.M.; Thaxton, C.S.; Mirkin, C.A. Nanoparticle-based bio-bar codes for the ultrasensitive detection of proteins. *Science* **2003**, *301*, 1884–1886.
16. 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.* **2007**, *129*, 1959–1967.
17. 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. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 3429–3433.
18. Kwong, G.A.; et al. Modular nucleic acid assembled p/MHC microarrays for multiplexed antigen-specific T cell sorting. *J. Am. Chem. Soc.* **2009**, *in press*.
19. Altman, J.D.; Moss, P.A.H.; Goulder, P.J. R.; Barouch, D.H.; McHeyzer-Williams, M.G.; Bell, J.I.; McMichael, A.J.; Davis, M.M. Phenotypic analysis of antigen-specific T lymphocytes. *Science* **1996**, *274*, 94–96.

20. Chen, D.S.; Soen, Y.; Stuge, T.B.; Lee, P.P.; Weber, J.S.; Brown, P.O.; Davis, M.M. Marked differences in human melanoma antigen-specific T cell responsiveness after vaccination using a functional microarray. *PLoS Med.* **2005**, *2*, 1018–1030
21. 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.* **2003**, *1*, 429–438.
22. 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. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 3744–3749
23. Deviren, G.; Gupta, K.; Paulaitis, M.E.; Schneck, J.P. Detection of antigen-specific T cells on p/MHC microarrays. *J. Mol. Recognit.* **2007**, *20*, 32–38.
24. Rao, C.G.; et al. Expression of epithelial cell adhesion molecule in carcinoma cells present in blood and primary and metastatic tumors. *Int. J. Oncol.* **2005**, *27*, 49–57.
25. Mortazavi, A.; Williams, B.A.; McCue, K.; Schaeffer, L.; Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **2008**, *5*(7), 621–628.
26. Tang, F.; et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* **2009**, *6*(5), 377–384.