

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