A Microwell Array Cytometry System for High Throughput Single Cell Biology and Bioinformatics

by

Kenneth L. Roach

S.B., Electrical Engineering and Computer Science, MIT (2001)

Submitted to the Division of Health Sciences and Technology in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Biomedical Engineering at the MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2009

© Massachusetts Institute of Technology 2009. All rights reserved.
A Microwell Array Cytometry System for High Throughput Single Cell Biology and Bioinformatics

by

Kenneth L. Roach

Submitted to the Division of Health Sciences and Technology on September 28, 2008, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Biomedical Engineering

Abstract

Recent advances in systems biology and bioinformatics have highlighted that no cell population is truly uniform and that stochastic behavior is an inherent property of many biological systems. As a result, bulk measurements can be misleading even when particular care has been taken to isolate a single cell type, and measurements averaged over multiple cell populations in a tissue can be as misleading as the average height at an elementary school. Unfortunately, there are relatively few experimental systems available at present that can provide a combination of single cell resolution, large cell populations, and the ability to track individual cells over multiple time points. Those systems that do exist are often difficult to automate and require extensive user intervention simply to generate the raw data sets for later analysis. The goal of this thesis project was to develop a powerful, inexpensive, and easy-to-use system that meets the above requirements and can serve as a platform for single cell bioinformatics.

Our current system design is composed of two basic parts: 1) a customizable PDMS device consisting of one or more microwell arrays, each with associated alignment and identification features, and 2) a suite of custom software tools for automated image processing and data analysis. The system has a number of significant advantages over competing technologies such as flow cytometry and standard image cytometry. Unlike flow cytometry, the cells are not in suspension, and individual cells can be tracked across multiple time points or examined before and after a treatment. Unlike most image cytometry approaches, the cells are arranged in a spatially defined pattern and physically separated from one another, greatly simplifying the required image analysis. The automated analysis tools require only a minimal amount of user intervention and can easily generate multi-channel fluorescence time courses for tens of thousands of individual cells in a single experiment. For visualization purposes, tools are provided to annotate the original fluorescence images or movies with the results of later analysis, and several quality control routines are available to identify improperly seeded wells or debris.

The microwell array cytometry platform has allowed us to investigate a number of biological problems that would be difficult or impossible to tackle with standard techniques. Our earliest work focused on correlating pre-stress cell states with post-stress outcomes, with a major focus on the cryopreservation of primary hepatocytes. In particular, we wanted to know whether cell survival was dominated by extrinsic factors such as ice crystal nucleation, or intrinsic factors such as the energetic state of the cell. In one set of studies, we found that cells with a high initial mitochondrial content or mitochondrial membrane potential,
as measured by Rh123 or JC-1 staining, were significantly less likely to survive the freezing process. This demonstrated that intrinsic cell factors do play a major role in cryopreservation survival, but perhaps more importantly demonstrated the power and versatility of the microwell system by tracking individual cells across a treatment as extreme as freezing the entire device. In another set of cryopreservation experiments, cells were transiently transfected with a GFP-tagged protective protein and the resulting cell population, with its range of expression levels, was used to generate dose response curves with single cell resolution for the protein's protective effect.

More recently, our efforts have focused on generating single cell fluorescence time courses and using bioinformatics techniques such as hierarchical and k-means clustering to visualize the data and extract interesting features. More specifically, the behavior of primary hepatocytes under oxidative stress and protective metabolic manipulation was examined using a combination of mitochondrial and free radical sensitive dyes. The resulting time courses could not only be compared between the treatment groups, but a number of distinct response patterns could be identified within each treatment group. This variation in response patterns represent potentially important information that would be missed using bulk techniques or flow cytometry. In addition, membership in each response cluster was correlated between multiple dyes and with the initial state of each cell. Using a live / dead methodology, dose response curves, survival curves, and survival time distributions were also generated for each treatment condition and further subdivided based on the initial cell state and cluster assignments.

We believe that our microwell array cytometry platform will have general utility for a wide range of questions related to cell population heterogeneity, biological stochasticity, and cell behavior under stress conditions. We have really just begun exploring rich data sets of this type, and with additional work there is a great potential for groundbreaking results in many areas of biology and bioinformatics. Though we have applied techniques from gene expression analysis, there are a number of significant differences between the type of data generated by gene chips and that generated in high-throughput single cell experiments. These differences also make single cell biology a fruitful area for the development of novel bioinformatics techniques and theories.

Thesis Supervisor: Mehmet Toner
Title: Professor of Surgery and Health Sciences and Technology
Acknowledgments

I would like to start by thanking my thesis advisor, Mehmet Toner, for getting me started on such an interesting thesis topic and guiding me through the PhD process. His advice was invaluable, not just for my thesis work, but for my life in general.

I would also like to thank my thesis committee members Martin Yarmush and Isaac Kohane. Their advice and comments were a great help in my thesis research, and they really stood by me when things got tough towards the end.

Next, I would like to thank Kevin King, Heidi Elmoazzan, Anthony Chang, and the other members of the Cryopreservation group. I really appreciate all the time they spent helping me in lab, discussing my work, and simply being themselves. Working with such wonderful people made this thesis project a lot more fun and a lot less frustrating than it could have been. Without all of their help, I would probably still be trying to get my first few experiments running.

I would also like to thank our collaborators, Steve Hand and Mike Menze, for their help with the LEA protein experiments and deep insights on mitochondrial biology and anhydrobiotic organisms.

To avoid leaving someone out, I would like to thank everyone else in the Center for Engineering in Medicine. Even if I never met with some of them directly, I certainly gained a lot from all of their work. I hope my work will be just as valuable to them in the future.

On a more personal note, I would like to thank my wife Leah for putting up with my never ending thesis work, always knowing how to cheer me up, and being such a wonderful person in general. I think she will be even more excited than me when this thesis is finally signed and submitted.

Above all I would like to thank my family, especially my parents and grandparents for always believing in me, my sister Kaitie for always being a good friend, my sister Kristin for making me laugh, and my sister Kara for being such a sweet girl. Without their love and support I would certainly not be where I am today. I would not trade them for the world.
# Contents

1 Introduction ............................................................................. 13
   1.1 Thesis Structure ............................................................... 13
   1.2 Motivation ........................................................................ 15
   1.3 Background ...................................................................... 17
       1.3.1 Limitations of Bulk Measurements ......................... 17
       1.3.2 Time Courses and Biological Stochasticity .............. 19

2 Design of the Microwell Array Cytometry System ................. 23
   2.1 Microwell Array Device ................................................... 23
       2.1.1 Device Fabrication ................................................. 23
       2.1.2 Assembly and Coating ............................................ 23
       2.1.3 Hepatocyte Seeding ................................................. 25
   2.2 Probes and Reporters ....................................................... 25
       2.2.1 Mitochondrial Mass and Activity ............................... 27
       2.2.2 Reactive Oxygen Species ....................................... 28
       2.2.3 Viability Assays ....................................................... 28
       2.2.4 GFP Reporter Cells .................................................. 28
   2.3 Basic Protocols .................................................................. 29
       2.3.1 Mitochondrial Staining ............................................. 29
       2.3.2 Fluorescence Microscopy ......................................... 29
       2.3.3 Cryopreservation and Recovery ................................. 30
       2.3.4 Viability Assays ....................................................... 30
   2.4 Image Analysis .................................................................. 30
       2.4.1 Image Processing ..................................................... 30
2.4.2 Quality Control .................................................. 32

3 Testing and Validation of the Microwell Array Cytometry System 35
   3.1 Overlay Images ................................................. 35
   3.2 Array Loading Efficiency .................................... 36
   3.3 Reproducibility of Fluorescence Values ....................... 38
   3.4 Focus Variation .................................................. 39
   3.5 Spatial Homogeneity ............................................ 40

4 Microwell Array Cytometry as a Novel Tool in Biopreservation 45
   4.1 Introduction .................................................. 45
      4.1.1 LEA Proteins ............................................. 46
      4.1.2 Mitochondria in Cryopreservation ....................... 47
   4.2 Methods and Materials ........................................ 49
      4.2.1 Assembly, Coating, and Seeding ......................... 49
      4.2.2 Mitochondrial Staining .................................. 49
      4.2.3 LEA Protein Transfection ................................ 49
      4.2.4 Fluorescence Microscopy ................................ 50
      4.2.5 Cryopreservation and Recovery ......................... 50
      4.2.6 Viability Assays .......................................... 51
      4.2.7 Image Processing and Statistical Analysis ............. 51
   4.3 Results .......................................................... 51
      4.3.1 JC-1 and Rh123 Fluorescence Distributions ............. 51
      4.3.2 Live-Dead Assays ........................................ 53
      4.3.3 Mitochondrial Membrane Potential and Cryopreservation Outcomes 53
      4.3.4 Multidimensional Analysis ................................ 56
      4.3.5 LEA Protein Dose Response .............................. 58
   4.4 Discussion ..................................................... 58

5 High Throughput Single Cell Bioinformatics 63
   5.1 Introduction .................................................. 63
   5.2 Methods and Materials ........................................ 65
      5.2.1 Assembly, Coating, and Seeding ......................... 65
<table>
<thead>
<tr>
<th>5.2.2</th>
<th>Dyes and Treatments</th>
<th>65</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2.3</td>
<td>Fluorescence Microscopy</td>
<td>66</td>
</tr>
<tr>
<td>5.2.4</td>
<td>Image Processing and Data Extraction</td>
<td>66</td>
</tr>
<tr>
<td>5.2.5</td>
<td>Clustering of Fluorescence Time Courses</td>
<td>66</td>
</tr>
<tr>
<td>5.3</td>
<td>Results</td>
<td>67</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Single Fluorescence Time Courses</td>
<td>67</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Fluorescence Time Course Modeling</td>
<td>70</td>
</tr>
<tr>
<td>5.3.3</td>
<td>K-Means Clustering of Single Cell Time Courses</td>
<td>70</td>
</tr>
<tr>
<td>5.3.4</td>
<td>Hierarchical Clustering for Single Cell Dose Response Experiments</td>
<td>72</td>
</tr>
<tr>
<td>5.4</td>
<td>Discussion</td>
<td>76</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Conclusion</td>
<td>78</td>
</tr>
</tbody>
</table>

### 6 Summary and Conclusions

<table>
<thead>
<tr>
<th>6.1</th>
<th>Summary of Results</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1.1</td>
<td>System Design and Validation</td>
<td>80</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Biopreservation Studies</td>
<td>80</td>
</tr>
<tr>
<td>6.1.3</td>
<td>High Throughput Single Cell Bioinformatics</td>
<td>82</td>
</tr>
<tr>
<td>6.2</td>
<td>Future Work</td>
<td>83</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Microfluidic Integration</td>
<td>83</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Device Design and Seeding</td>
<td>83</td>
</tr>
</tbody>
</table>

### A Morphological Analysis of Intracellular Fluorescence Distributions

<table>
<thead>
<tr>
<th>A.1</th>
<th>Morphological Analysis Approaches</th>
<th>85</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.2</td>
<td>Zernike Moments</td>
<td>87</td>
</tr>
<tr>
<td>A.2.1</td>
<td>Definitions and Properties</td>
<td>87</td>
</tr>
</tbody>
</table>
List of Figures

1-1 Overview of the microwell array cytometry system .................. 18
1-2 Limitations and artifacts of bulk measurements ...................... 20
1-3 Lack of temporal synchronization across cell populations .......... 21
2-1 Microwell device and array layout ................................ 24
2-2 Coating procedure for the microwell array devices .................. 26
2-3 Alignment of array images ........................................ 31
2-4 Gridding of array images .......................................... 32
2-5 Quality control of well loading .................................... 33
3-1 Overlay images for rapid visual inspection of the data .............. 36
3-2 Loading efficiency for a sample of 64 microwell arrays ............. 37
3-3 Reproducibility of fluorescence values ............................. 38
3-4 Fluorescence intensity variations with changes in focus .......... 39
3-5 Spatial distribution of hepatocyte subpopulations seen with JC-1 staining . 41
3-6 Spatial variation of initial fluorescence intensities, cell survival times, and cluster assignments .................................. 42
3-7 Statistical tests of spatial homogeneity ............................. 43
4-1 Microwell array cytometry as a novel tool in biopreservation .......... 47
4-2 Biopreservation experiment array images ............................ 52
4-3 Data analysis after image processing ................................ 54
4-4 Viability after cryopreservation as a function of JC-1 and Rh123 intensity . 55
4-5 Viability across the JC-1 fluorescence landscape ..................... 57
4-6 Viability as a function of LEA protein transfection level ........... 59
5-1  High throughput single cell bioinformatics ......................................... 65
5-2  Fluorescence images of primary hepatocytes stained with rhodamine 123 and
dihydroethidium ................................................................. 68
5-3  Single cell fluorescence time courses .................................................. 69
5-4  Modeling of dihydroethidium time courses ...................................... 71
5-5  K-means clustering of timecourses .................................................. 73
5-6  Figures of merit for kmeans clustering ............................................ 74
5-7  Hierarchical clustering of timecourses ............................................. 75

6-1  Microwell designs for cell shape control and cell-cell interaction studies ... 84

A-1  Zernike polynomials up to order six .................................................. 89
Chapter 1

Introduction

Recent advances in systems biology and bioinformatics have highlighted that no cell population is truly uniform and that stochastic behavior is an inherent property of many biological systems. As a result, bulk measurements can be misleading even when particular care has been taken to isolate a single cell type, and measurements averaged over multiple cell populations in a tissue can be as misleading as the average height at an elementary school. Unfortunately, there are relatively few experimental systems available at present that can provide a combination of single cell resolution, large cell populations, and the ability to track individual cells over multiple time points. Those systems that do exist are often extremely difficult to automate and require extensive user intervention simply to generate the raw data sets for later analysis. The goal of this project was to develop a powerful, inexpensive, and easy-to-use system that meets the above requirements and can serve as a platform for single cell bioinformatics.

1.1 Thesis Structure

This thesis document is divided into a number of chapters, starting with motivations and background, working through a technical description of the system, and then validating its behavior. Several sets of experiments are then presented to demonstrate the power of the system. The first correlates initial cell states with experimental outcomes, the second examines single cell fluorescence time courses, and a third looks at the spatial distribution of fluorescence within each cell.

The remainder of this chapter provides an overview of existing technologies for single
cell biological measurements and past efforts at single cell bioinformatics. The advantages and disadvantages of these techniques are discussed and used to motivate the development of the system presented in this thesis. Chapter 2 provides detailed descriptions of both the physical microwell array device and the custom software tools used for image processing and data analysis. Protocols are included for device assembly and coating, cell seeding, and imaging. Chapter 3 describes the results of several experiments used to evaluate the array cytometry system and validate the generated fluorescence intensity data. These tests includes measures of array filling rates, the reproducibility of fluorescence values, the spatial homogeneity of cell responses, and quantifying the effects of nonideal focussing.

In Chapter 4, the power of the system to correlate pre-treatment cell states with post-treatment outcomes is demonstrated in a series of single cell biopreservation studies. More specifically, the system is used to examine whether the initial metabolic and energetic state of primary hepatocytes affects the outcome of cryopreservation. Another set of experiments examines whether the level of transfection with a presumed protective factor correlates with cryopreservation or dessication survival. Chapter 5 expands on this idea and demonstrates the ability of the system to generate high throughput single cell fluorescence time courses. Mitochondrial polarization and free radical generation are examined in primary hepatocytes under oxidative stress, with or without preconditioning with metabolic modulators. A number of bioinformatics techniques are used to visualize the collected data, identify significant features, and correlate behaviors between multiple dyes and multiple time points. Appendix A describes recent extensions to the analysis software to calculate additional fluorescence distribution parameters such as the Zernike moments.

Chapter 6 summarizes the major results of the previous chapters and discusses a number of possible directions for future work. Most notably, these include integrating the arrays into a combinatorial microfluidic device and using the system to study the stochastic nature of gene expression through GFP reporter constructs. This thesis provides a firm foundation for future work in single-cell studies in biology and bioinformatics using microwell array cytometry. The results presented in this thesis show the micowell array cytometry will be an invaluable tool for detailed studies of biological stochasticity and cellular responses to stress. Future device development will make the system even easier to use and expand the range of cell types and experimental systems that can be studied. Hopefully, this work will also help motivate the development of the new bioinformatics techniques necessary for
single cell analysis.

### 1.2 Motivation

Measurements averaged over the population of cells in a tissue can be as misleading as the average height in an elementary school, but bulk measurements can still be misleading even when particular care has been taken to isolate a single cell type. Recent advances in systems biology and bioinformatics have highlighted that no cell population is truly uniform and that stochasticity is an inherent property of many biological systems [1, 2, 3, 4]. Population heterogeneity has long been observed in primary cells such as hepatocytes, which are well known to specialize within the tissue based on anatomical and environmental cues [5, 6]. More surprising is that significant cell-to-cell variation can be present even in presumably uniform populations like clonal cell lines, and is a natural consequence of positive feedback in the mechanisms controlling gene expression [7, 8]. There is also growing evidence that this population heterogeneity has a meaningful biological role, such as the behavior of mammalian macrophages in response to pathogen exposure [9].

As a result, there has been much recent interest in high-throughput tools that can probe biological systems at the single-cell and subcellular levels [10, 11, 12]. Such tools will allow researchers to study the implications of heterogeneity in cell populations and better understand the structure and function of signaling pathways, metabolism, and gene expression. But beyond simply characterizing cellular heterogeneity and stochasticity, single-cell experimental systems can be used to harness this natural variation to help answer difficult biological questions. For exploratory studies, rather than driving cells into artificial testing states with an applied treatment, which is often complicated by homeostatic mechanisms and unintended side effects, the natural variation in cell behavior across the population can provide an entire spectrum of built-in test cases. When treatments are later applied, single cell approaches will provide a much richer appraisal of the effects.

In many cases, the variation observed at the cellular level is closely linked to the stochastic nature of underlying biological systems. Studies of ion channels, signaling pathway dynamics, transcription factor networks, and differentiation have also shown that the behavior of a biological system is often much more stochastic and discrete than would be assumed from population averages. At the level of individual molecules and even cells, biological
processes tend to have probabilistic rather than continuous response curves and undergo rapid shifts between metastable states [13, 14, 15]. What appears as a smooth transition at the bulk level may be anything but smooth at the level of each participant. These issues can pose a significant challenge for traditional experimental approaches that rely on bulk assays, since the average values obtained may not be representative of any given cell in the population. In many cases, high-throughput, single-cell approaches will greatly increase the accuracy, content, and statistical power of the collected biological data.

For many years, flow cytometry has been the tool of choice for fluorescence-based assays at the single cell level. Flow cytometry is a well established technique and can provide single-time-point measurements of multiple parameters, including cell size and granularity. Fluorescent dyes, labeled antibodies, and other tags are easily obtainable for a huge range of applications. Data analysis tools are also widely available. The major disadvantages of flow cytometry are that the cells must be kept in suspension and that there is no way to track individual cells across multiple time points. With digital cameras, rapid increases in computational power, and the development of better analysis techniques, several forms of image cytometry have become viable alternatives [11, 16]. Unlike flow cytometry, image cytometry is generally used with attached cells, though suspended cells can also be used in some systems. A significant advantage is that time courses can be generated for individual cells. Unfortunately, advanced image processing and data analysis techniques are often needed to extract useful results, making the systems much harder to use and automate than flow cytometry. The information content is potentially higher, but this comes at a price in complexity.

One way to reduce this complexity is by better controlling the spatial arrangement of the cell population being studied. Much of the difficulty in image cytometry comes from interpreting random seeding patterns and identifying cell boundaries. These issues can both be addressed by seeding the cells in a regular pattern and isolating them from their neighbors. Microfabrication technologies such soft lithography have led to the development of several widely-used cell patterning techniques based on the deposition of adhesion molecules and/or adhesion inhibitors in precisely controlled arrangements [17, 18]. Microfluidics, dielectrophoretic traps, and laser capture microscopy have also been used to control cell patterning in novel ways [11]. A recent addition to the cell patterning arsenal is the physical entrapment of individual cells in high density microwell arrays [19, 20, 21].
A variant of this approach is used here to simplify the image processing needed for effective image cytometry.

The microwell array system described in this paper was designed as part of our effort to develop high-throughput and high-content tools for the study of biopreservation and stress biology. Even with the best preservation protocols, some proportion of the preserved cells invariably die or function poorly. To some extent, this can be explained by the stochastic nature of ice crystal nucleation and other biophysical processes, but cell-to-cell variation in the initial population almost certainly plays a role [22]. Position in the cell cycle, metabolic state, and membrane composition could all potentially affect preservation outcome, but few good tools exist to study these effects at the single cell level. To meet this need we have developed a flexible, inexpensive, and easy-to-use platform that will allow us to correlate pre-stress cell phenotypes and post-stress outcomes with single cell resolution. To demonstrate the system, we have examined primary hepatocytes to see whether a cell’s initial mitochondrial membrane potential (ΔΨ) affects its probability of surviving cryopreservation.

1.3 Background

Recent advances in systems biology and bioinformatics have highlighted the extreme complexity of biological systems and a demonstrated a growing need for tools to probe this complexity in a detailed, systematic, and high-throughput manner. Though many systems have been developed for the high-throughput study of transcription, proteomics, signaling, and metabolism at the bulk level, an emerging challenge is to scale these technologies to the level of individual cells, thereby increasing the accuracy, content, and statistical power of the collected data. Single cell measurements will also allow us to explore the extent and impact of cell-to-cell variation within nominally uniform populations and better understand the stochastic nature and network properties of signaling pathways, metabolism, and other biological processes.

1.3.1 Limitations of Bulk Measurements

Cell-to-cell variation is often ignored or treated as a source of noise in classic biological assays. Most experiments in biology implicitly assume that the cell population being tested
Figure 1-1: Individual cells are trapped in spatially defined arrays of microwells with associated alignment and identification features, greatly simplifying the image processing needed to identify and segment the images for single cell fluorescence intensity and distribution calculations. Cells are easily tracked across multiple time points, even when those time points are interrupted by treatments as extreme as freezing or drying the entire device. Automated image analysis software can be used to produce simple data products similar to those of flow cytometry or single cell time courses for more detailed analysis using bioinformatics techniques.
is uniform in phenotype, at least with respect to the stimulus being tested. As a result, the vast majority of experiments are designed to generate bulk assay values representing the average over a large cell population. The use of bulk measurements does have its advantages, including increased signal strength, built-in noise reduction, and simplified protocols. Under certain circumstances, however, bulk measurements will give statistically weak or misleading results. Several such situations are illustrated in Figure 1-2.

A key problem with bulk experiments is the implicit assumption that the underlying cell population is uniform with respect to the current assay. By ignoring the distribution of values present within the cell population, potentially meaningful variation is treated the same as meaningless measurement noise and averaged out. The simplest error is the assumption that the distribution of single-cell measurements follow a normal distribution centered on the bulk value, or that a ‘typical’ cell will have a value close to the population average. This often is the case, but ignores the possible complication of there being discrete subpopulations. As shown in Panel A of Figure 1-2, the value for any given cell could in fact be quite far from the bulk value.

A similar problem is that bulk measurements of multiple parameters ignore the correlation structure between those parameters. As shown in Panel B of Figure 1-2, it is impossible to distinguish between cell populations with wildly different parameter patterns using only the bulk values. Bulk measurements can also mask meaningful results, particularly when there is a mixture of responsive and unresponsive cells in the population. Panel C illustrates a situation where a small fraction of the population has a strong response to a stimulus, but does not change the bulk value enough to be detected. Once the technical limitations of actually making single cell measurements are overcome, the results are generally more useful than bulk measurements. If bulk values are needed, it is a simple task to average over the single cell values, but this will generally represent a loss of potentially useful information.

1.3.2 Time Courses and Biological Stochasticity

Studies of individual ion channels, signaling pathway dynamics, transcription factor networks, and differentiation have shown that the underlying behavior of biological systems are often more stochastic and discrete than would be assumed from bulk assays. At the level of individual molecules and even cells, biological processes tend to have probabilistic rather than continuous response curves and undergo rapid shifts between metastable states.
Figure 1-2: Bulk measurements have several important limitations when dealing with heterogeneous cell populations. Panel A illustrates that the averaging implicit in bulk measurements may lead to a bulk value far from that of any individual cell. In such cases, it is not meaningful to think of the bulk value as corresponding an “average” cell. Panel B shows how the same bulk values could result from measurements on cell populations with very different joint distributions. Panel C illustrates how the behavior of an important but small subset of the cell population may be missed using only bulk measurements.

When responses are blindly averaged together, much useful information about the system can be obscured.

Even in a relatively uniform cell population, individual cells can exhibit significant temporal differences in behavior. This poses a major problem for time course experiments using bulk measurements across the population. Many biological processes are stochastic in nature and the timing of each cell’s response to a stimulus can vary quite significantly. At the extreme, some cells may never respond to a stimulus, even if many or most cells in the population have a rapid and strong response. Because cell responses occurring at different time points are averaged together, bulk measurements can easily obscure the true dynamics of a process. Panel A of Figure 1-3 shows an example in which each individual cell undergoes a rapid and stereotyped transition to a new state after a stimulus, but the timing of the transition varies across the population. Because of this temporal variation, the transition appears to occur much more slowly in the bulk time course than it actually occurs for each cell. Situations like this where cells have rapid transitions between states with variable timing can easily arise in biological system due to positive and negative feedback mechanisms. As seen in Panel B of Figure 1-3, the problem is even worse when a cell population shows several distinct response patterns or the response of each cell varies significantly in timing and magnitude. In this case, averaging eliminates many of the most
Figure 1-3: Bulk measurements can obscure the dynamics of a biological process when used in time course experiments. Panel A shows the effect of averaging when there is significant variation in the timing of each cell's response to a stimulus. The transition appears to occur slowly in the averaged time course even though the transition is quite rapid for each individual cell. Panel B shows the effect of averaging when there is significant diversity in both the magnitude and timing of each cell's response. The data shown in this panel comes from a time course experiment in Chapter 5. Many interesting features of the individual time courses are obscured by the averaging and absent in the bulk response.
Chapter 2

Design of the Microwell Array Cytometry System

2.1 Microwell Array Device

2.1.1 Device Fabrication

The devices used in this study were fabricated by PDMS replica molding. A 4 inch silicon wafer was first spin-coated with 25 μm SU-8 negative photoresist and patterned by photolithography to generate a reusable master. The most recent design included 512 microwell arrays arranged in groups of 4 and surrounded by cutting guides. Each microwell array consisted of a 32 by 32 grid of 25 μm diameter cylindrical wells along with alignment and identification features to aid later analysis (Figure 2-1). For efficient imaging, the layout was designed such that each array filled the 4x visual field of our microscopy system. Replicas were fabricated by mixing PDMS prepolymer with initiator and pouring the mixture onto the master in a 4 mm layer, briefly degassing in a vacuum chamber, and curing at 75°C overnight. The cured PDMS sheet was then peeled from the mold and cut as needed into individual devices.

2.1.2 Assembly and Coating

Individual devices containing 8 or 16 arrays each were cut from the cast PDMS and reversibly mounted well-side-up on 2.5 x 1.0 cm miniature slides for easier handling. The original size of the slides was chosen to fit in 1.5 ml Nalgene cryovials for biopreservation
A reusable mold for casting the PDMS devices is fabricated on a 4 inch silicon wafer using photolithography. As seen in Panel A, each wafer contains 512 microwell arrays arranged in groups of four and separated by cutting guides. Each PDMS sheet is typically divided into 32 devices measuring 1.6 x 0.8 cm for easy mounting on mini glass slides. As shown in Panel B, each device contains 16 individual arrays with an empty region at the center for pipetting. In some experiments, half sized devices with 8 arrays were used to allow for faster imaging. Panel C shows the layout of a typical array, including a regular 32 x 32 grid of 25 μm wells designed to hold single cells, alignment features at each corner, and unique binary identification features on each side.

experiments, but also proved to be convenient for device handling and seeding purposes. The devices were sterilized by either autoclaving after assembly or rinsing in 100% ethanol before assembly. The outer surfaces of the arrays were then blocked by applying a droplet of 2% Bovine Serum Albumin (BSA, Sigma), incubating for 10 min at 25°C, and rinsing well with PBS. Because the PDMS is hydrophobic, the BSA solution does not enter the microwells and serves to reduce cell adhesion outside of the wells.

The inner surfaces of the wells were then coated with a protein that enables cell adhesion such as collagen or fibronectin. The optimal choice of adhesion protein is highly dependent on the cell type being used. Primary hepatocytes were adhered well to both collagen and fibronectin, and both were used depending on the experiment. H35 hepatoma cells adhered to both proteins, but the adherence to fibronectin was faster and tighter. C3A cells, on the other hand, had almost no adherence to fibronectin, but very strong adherence to collagen. The choice of adhesion molecules is important for optimal array seeding, and the seeding protocol needs to be optimized for each cell type. Poorly adherent coatings or short seeding times will lead to cells being washed out of the wells, while excessively adherent coatings or long seeding times will lead to massive cell attachment outside of the wells.
Stock solutions of collagen or fibronectin were either purchased or prepared as described elsewhere. The 1.0 mg/ml stock solution of fibronectin was purchased from Sigma and diluted 1:40 in phosphate buffered saline (PBS) immediately before use. The 1.25 mg/ml stock solution of Type I collagen was prepared as described elsewhere and diluted 1:50 in distilled water before use [23]. Each device was then covered with a droplet of the either the fibronectin or collagen solution and placed under vacuum for 30 min to promote microwell filling. The devices were then incubated for another 30 min at 25°C, rinsed with media, and kept wet until seeding.

2.1.3 Hepatocyte Seeding

Primary rat hepatocytes were freshly isolated from female Lewis rats and resuspended at 1x10^6 cells/ml in C+H hepatocyte culture medium [23]. Particular care was taken with the pipetting to gently break up clumps of cells and achieve a single cell suspension. Droplets of the suspension were then placed on each device and periodically recirculated by gentle pipetting to improve filling efficiency and reduce cell attachment outside of the microwells. After 5 to 10 minutes, excess cells were removed by rinsing with fresh media, and the devices were briefly examined under the microscope. If loading was poor after the first round of seeding, it was repeated a second time. Once fully loaded, each device was transferred to a dish of fresh medium and gently agitated to remove any cells remaining on the surface of the device.

The seeding procedure was optimized to maximize the number of filled wells while minimizing the attachment of clumped cells or debris outside of the wells, but even under ideal conditions, some amount of extraneous material was present in small amounts. This debris is harmless when positioned away from the wells, but can interfere with later analysis if it covers or surrounds one. The software filtering routines described later accurately detected any wells obscured by bubbles, debris, and poorly seeded cells and was used to eliminate these wells from later analysis.

2.2 Probes and Reporters

Mitochondria have many important functions that can be investigated using fluorescent probes. Several experiments in this thesis focus on aspects of the mitochondrial state. Ex-
Array device is cut, mounted on a slide, and sterilized

Blocking solution droplet does not enter wells due to the hydrophobicity of the PDMS

Device is rinsed, leaving outer surfaces of array coated with the blocking solution

Collagen or fibronectin solution fills wells when device is placed in a vacuum chamber

Device is rinsed again, leaving a highly adhesive coating for the cells on the inner well surfaces

Figure 2-2: The coating procedure for the microwell array devices is illustrated above. The red droplet coats the outer surface of the device and represents a blocking solution such as albumin or PEG that prevents cell attachment. The green droplet coats the inner surfaces of the microwells and represents a solution containing a cell adhesion molecule such as collagen or fibronectin. The coating procedure is simple to perform and does not require any special equipment or complex preparation of reagents.
peripheral targets include the mitochondrial mass, membrane potential, metabolic rate, free radical production, and apoptotic pathways. The biopreservation experiments of Chapter 4 focus on how the initial mitochondrial membrane potential of a cell affects cryopreservation survival. The single cell time of Chapter 5 examine the temporal dynamics of mitochondrial membrane potential and free radical production in cells under oxidative stress.

2.2.1 Mitochondrial Mass and Activity

The most important measurements of a cell’s mitochondrial state are total mitochondrial mass and energy production. Since energy production is difficult to measure with fluorescent probes, mitochondrial membrane potential will be used as a rough proxy [24]. The mitochondria content of a cell could influence cryopreservation several ways. During recovery, having more mitochondria could give a cell a greater capacity for energy production, ion gradient homeostasis, and repair processes. Cells with more mitochondria might also be less affected by the loss of mitochondria due to intracellular damage. However, having more mitochondria could also be a liability since mitochondria are major regulators of apoptosis, generate reactive oxygen species, and contain large calcium stores that may leak into the cytoplasm if the mitochondrial membranes are damaged [25, 26]. Very little is currently known about the role this may play in cryopreservation. The mitochondrial dye TMRE [27] can be used to estimate the mitochondrial content of each cell and determine if having more or less mitochondria is protective.

The proton pumps of the mitochondrial respiratory chain generate a significant electric potential and proton gradient across the inner mitochondrial membrane, with the mitochondrial matrix depleted of protons and highly negative relative to the cytoplasm. The mitochondrial membrane potential (MMP) and proton gradient act as the driving force for the mitochondrial ATP synthase, metabolite transport proteins, and mitochondrial ion channels. The MMP also regulates the opening of the mitochondrial permeability transition pore (MPTP), a key player in the apoptotic pathways [26]. Initial work with the microwell array system has focused on the MMP-sensitive dyes JC-1 and Rhodamine 123 [28]. Both are lipophilic, cationic dyes that can cross the mitochondrial membrane and accumulate in the negatively charged matrix in a potential dependent manner according to the Nernst equation. JC-1 can also be used in a ratiometric fashion to correct for the effect of plasma membrane potential, cell size, and mitochondrial mass. At high concentrations, JC-1 forms
J-aggregates with altered fluorescence properties, allowing mitochondrial accumulation to be distinguished from other accumulation sites [29, 30].

2.2.2 Reactive Oxygen Species

Reactive oxygen species (ROS) are known to be a significant damage source during preservation and other cell stress states. In most cell types, mitochondria are the major source of free radical generation, though other production sites are also possible. Superoxide production can be monitored using dihydroethidium (DHE), which is converted to the fluorescent dye ethidium in the presence of superoxide radicals [31]. Another choice is MitoSOX Red, which is a modified version of DHE that accumulates in the inner mitochondrial membrane and is relatively insensitive to other free radicals [32]. Other ROS species such as peroxide will can be monitored using CM-H₂DCFDA, which accumulates in the cytoplasm and is sensitive to a wider range of free radicals than DHE [33]. In Chapter 4, these assays will be used both before preservation to measure initial ROS generation rates and after preservation to monitor changes in production resulting from mitochondrial damage and other mechanisms. In Chapter 5, these assays will be used in conjunction with mitochondrial dyes to examine the dynamics of cellular responses to oxidative stress.

2.2.3 Viability Assays

Since only a limited number of fluorescence channels are available in most microscopy systems, viability probes need to be chosen carefully to avoid interfering with the mitochondrial and free radical dyes already in use. Given the many viability stains available, this is not be a major problem in most cases. Immediate or necrotic cell death can be monitored using a standard Live/Dead assay system. For the live part of the assay, Calcein-AM or one of its many color variants is used to identify intact cells. If needed, JC-1 or Rh123 can also fulfill this function. For the dead part of the assay, a membrane impermeant dsDNA stain such as ethidium homodimer or Sytox Blue is used to identify cells with a disrupted plasma membrane.

2.2.4 GFP Reporter Cells

As part of the Microfluidic Living Cell Array (mLCA) project, a number of GFP reporter cells have been developed in our research group [34, 35, 36]. The current reporter cells
are derived from H35 rat hepatoma cells and maintain a fairly high level of hepatospecific function. Each reporter cell line expresses a destabilized GFP protein in response to activation of a particular transcription factor. Several of the available lines respond to transcription factors with a close link to cell stress pathways or metabolism and would be particularly interesting to study either alone or in conjunction with a mitochondrial dye. These include heat shock factor 1 (HSF1), hypoxia inducible factor 1α (HIF1α), peroxisome proliferator-activated receptor (PPAR), sterol regulatory element binding protein (SREBP), and carbohydrate response element binding protein (ChREBP). Like some of the dyes, the GFP expression levels of the cells could be used as either initial states or outcome measurements. Preliminary experiments were carried out with several of these cell lines and will be an important part of future work with the microwell array system.

2.3 Basic Protocols

2.3.1 Mitochondrial Staining

The fluorescent dyes JC-1 and rhodamine 123 (Rh123) can be used to assess mitochondrial membrane potential (ΔΨ) in the cells [28, 24]. Stock JC-1 solution was prepared by dissolving 1 mg lyophilized JC-1 in 1 ml DMSO. The stock solution was diluted 1:200 in C+H culture medium before use for a final concentration of 5 μg/ml. Stock Rh123 solution was prepared by dissolving 10 mg lyophilized Rh123 in 1 ml DMSO. In this case, the stock was diluted 1:1000 in C+H before use for a final concentration of 10 μg/ml. Both stock solutions were stored at -20°C. Seeded devices were incubated for 15 min at 37°C in the diluted dye solution, briefly rinsed to remove excess dye, and incubated at 37°C in fresh media for an additional 30 min before imaging.

2.3.2 Fluorescence Microscopy

Brightfield and fluorescence images were captured on a Zeiss 200 Axiovert microscope with an AxioCam MRm digital camera, typically using a 2.5x objective and 1.6x optovar for full images of a single array. JC-1 fluorescence was measured using Zeiss #38 and #31 filter sets for the green and red channels. Rh123 fluorescence was measured using a Zeiss #38 filter set. DHE fluorescence was measured in the red channel using a Zeiss #31 filter set. Exposure times were selected to maximize the dynamic range of the resulting images. A
brightfield image was taken along with each set of fluorescence images for alignment, quality control, and display purposes.

2.3.3 Cryopreservation and Recovery

After imaging, cells were allowed to recover for 30 min in a 37°C incubator. Freezing solutions were prepared using HypoThermosol-Base (BioLife Solutions) with 10% DMSO added [37]. The devices were quickly transferred into 1.5 ml Nalgene cryovials filled with the freezing solution, placed on ice for a short time for cell equilibration with the cryoprotectant, and placed in a controlled-rate freezer (Planer KRYO 10). The devices were first cooled to -6°C, at which time extracellular ice was seeded by the applying cold forceps to the outside of each vial. This was followed by a 10 minute holding period to allow for ice growth and temperature equilibration. The devices were then cooled at a rate of -5°C/min to -80°C, plunged into liquid nitrogen, and stored for up to a week. After the storage period, the vials were thawed rapidly by gentle agitation for 2 min in a 37°C water bath. The devices were then rinsed in warm C+H, transfered to fresh medium, and allowed to recover at 37°C for 1 hr.

2.3.4 Viability Assays

Stock solutions of 1 mM Sytox Blue (Molecular Probes) were diluted 1:1000 in C+H to a final concentration of 1 µM and used in combination with JC-1 stained cells. Stock solutions of 2 mM ethidium homodimer-1 (Molecular Probes) were diluted 1:1000 in C+H to a final concentration of 2 µM, and used in combination with Rh123 stained cells. After recovering from cryopreservation, devices were transferred into 35 mm dishes containing one of the diluted dyes and incubated at 37°C for 5 min before imaging. Sytox Blue fluorescence was measured using a Zeiss # 5 filter set, while EthD-1 was measured using a Zeiss # 31 filter set. JC-1 or Rh123 fluorescence images were usually taken at this time as well.

2.4 Image Analysis

2.4.1 Image Processing

Custom image analysis software was developed based on the Insight Segmentation and Registration Toolkit (ITK) and was used for general image processing tasks, automated
alignment of array images, and calculation of pixel intensity statistics for each well [38].

Alignment was performed in three stages. Brightfield images of the arrays were first normalized, thresholded, and flood filled to generate simplified black and white images. A rough alignment was then performed between these images and a centered template image of the entire array. From these results, the approximate locations of the alignment features were calculated and used as the seed positions for fine alignment of each feature. In both cases, alignment was performed using a mean squares distance metric and a one-plus-one evolutionary optimizer. The expected coordinates of each well were then calculated based on the identified positions of the four alignment features.

Figure 2-3: Alignment of the array images is carried out two stages. Coarse alignment of the entire array image is first performed using a template image such as the one shown in Panel A. Rough feature positions are then calculated as shown in Panel B and used as the starting point for fine alignment of each feature as shown in Panel C. Highly accurate alignments can be achieved with no user intervention as shown in Panel D. The process is aided by preprocessing raw array images like the one shown in panel D to generate simplified images like the one seen in Panel C.

A number of processing steps were applied to the fluorescence images before computing fluorescence statistics for each well. The images were first corrected for stray background
illumination and spatial variation in lamp intensity using black and white reference images [39]. In some cases, morphological image filters were used to create a background image and eliminate local variation arising from device orientation and nonuniformities in the device or surrounding media. The resulting pixel intensities were finally renormalized to a standard range using either a fixed multiplier or by setting black and white pixel percentages. Pixel intensity statistics including the mean, median, mode, and standard deviation were calculated for the pixels belonging to each well and exported as a spreadsheet file. Further analysis and plotting was performed using Python and the R statistical package [40].

Figure 2-4: Total pixel intensities and other statistics are calculated for each well in the array. The position of each well is calculated using a perspective transform based on the positions of the four alignment features. The use of the perspective transform enables accurate well positioning even when the array is tilted out of the focal plane or the alignment feature positions are skewed.

2.4.2 Quality Control

A simple filtering algorithm was used to automatically identify empty and poorly seeded wells and tag them as such so that they could be excluded from later analysis. Empty wells were identified by settings fluorescence intensity thresholds. Obstructed wells were identified by first examining brightfield images of the annular region immediately surrounding each
well as shown in Figure 2-5. Using the thresholded images generated for alignment purposes, the number of black pixels in each annular region was counted. Properly seeded wells typically contained no more than a few such pixels, and wells were considered obstructed if the percent of black pixels in the region exceeded 5%. This value was chosen as a strict threshold based on histograms of annular pixel counts and manual validation with a subset of the arrays. A similar procedure was carried out with each fluorescence channel to identify additional wells with improper seeding. Wells identified as empty or obstructed were excluded from later plotting and data analysis.

Figure 2-5: The annular ring surrounding each well is examined in the thresholded bright-field images and optionally one or more of the fluorescence images. In the brightfield images, black pixels in this region represent debris, cells, or bubbles that may interfere with obtaining accurate fluorescence intensity values. Wells with more than a strict threshold of black pixels or fluorescence intensity in this region are excluded from later analysis.
Chapter 3

Testing and Validation of the Microwell Array Cytometry System

A number of experiments were run in order to physically test the array devices and validate the results obtained with the automated analysis tools. The validation experiments included calculations of the well loading efficiency for a sample of typical devices, reproducibility testing of the fluorescence intensities obtained with repeated imaging in different positions, examining how variations in focus affect the reported intensity values, and determining whether cell behavior is spatially homogeneous within each array. Software tools were also developed to generate overlay images of the arrays for rapid visual inspection of the data and troubleshooting. The results of these experiments show that the system can achieve high filling rates with minimal cell attachment outside of the wells, that the fluorescence values obtained with the system are highly reproducible and robust against variations in device orientation and minor focusing errors, and that the position of a cell within an array has little to no effect on the fluorescence values obtained or the cell’s behavior.

3.1 Overlay Images

For rapid visual inspection of the data, annotated images showing the classification of each well were also automatically generated for each array based on filling, quality control, and
outcome information. Two typical examples are shown in Figure 3-1. The image in Panel A was generated for one of the biopreservation experiments described in Chapter 4. In this experiment, primary hepatocytes were seeded and stained with the mitochondrial dye JC-1. The color of the circle surrounding each well represents the status of the cell in that well at the end of the experiment. Green indicates a cell that survived preservation, red indicates that the cell died, and yellow indicates that the cell was lost from the array. Dark blue and cyan indicate wells that were poorly seeded or obscured by debris at the beginning or end of the experiment respectively.

The image in Panel B was generated for one of the time course clustering experiments in Chapter 5. The color of each circle corresponds to the cluster assignment for that well using the k-means algorithm. Wells without a circle were either empty or left out for quality control reasons. These images proved particularly valuable in troubleshooting the system and identifying potential patterns in the data. The overlay tool was designed to be flexible and can also be used to generate a wide variety of related images.

Figure 3-1: Overlay images can be generated for rapid visual inspection of the data. Panel A highlights the fate of each cell in a biopreservation experiment. Panel B shows the k-means cluster assignment for each cell in a time course experiment.

3.2 Array Loading Efficiency

Loading efficiencies were determined for 64 microwell arrays on four separate devices. The devices were seeded, stained with either Rh123 or JC-1, and imaged. For each well, the presence of a cell was determined by gating on total Rh123 or JC-1 fluorescence. No
additional dye was needed since essentially all cells stained to a level well above background. In experiments with a significant population of non-staining cells, an additional dye such as DAPI could be added for this purpose. Automated cell counting was also attempted using the brightfield or phase images, but was much less reliable because of cell-to-cell variation in well filling patterns. Poorly seeded and debris obstructed wells were identified using brightfield images of the arrays as described in Section 2.4.2.

Figure 3-2 shows the loading efficiency of the 64 analyzed array, with the percentage of empty, obstructed, and properly filled wells. For this set of arrays, the percentage of properly filled wells was 70.4 ± 14.5 % (mean ± s.d.). The percentage of debris obstructed wells was 5.0 ± 2.2 % (mean ± s.d.). Array position within each 16-array device was a minor factor in the loading efficiency, with arrays toward the center of the device tending to have slightly higher filling rates. The debris rates did not appear to show such a pattern. The observed spatial variation is likely due to the shape of the cell suspension droplet during seeding and the flow patterns that develop during subsequent pipetting.

Figure 3-2: Loading efficiency and debris rates were determined for a total of 64 microwell arrays from four seeded devices. The percentage of properly filled wells was 70.4 ± 14.5 % (mean ± s.d.). The percentage of debris obstructed wells was 5.0 ± 2.2 % (mean ± s.d.). This corresponds to an average of more than 700 individual cells per array and almost 12,000 cells per device.
3.3 Reproducibility of Fluorescence Values

In order to test the reproducibility of the intensity values obtained with the system, repeated images were taken of the same array positioned at various points within the visual field. Each image was processed independently, empty and poorly seeded wells were filtered out, and the resulting time courses were plotted for each well. With each image taken, a slight global decrease in well intensity values was observed across the entire array, most likely due to photobleaching. This systematic global variation could largely be eliminated by normalizing the well intensities in a given array image by the median well intensity for that image. After this correction, the well intensities were remarkably consistent from image to image. A sample of the results for JC-1 red fluorescence are shown in Figure 3-3. The coefficient of variation (CV%) across the time course was calculated individually for each well. For JC-1 Red, JC-1 Green, and Rh123 channels, the CV% values (mean ± SD) were 1.6 ± 1.0%, 2.0 ± 0.8%, and 1.1 ± 0.5% respectively with very few outliers.

Figure 3-3: Fluorescence values obtained using the analysis routines were very reproducible. After a global normalization to account for photobleaching, there was little variation in the fluorescence values obtained for each well with repeated imaging. Panel A shows a sample of normalized JC-1 red fluorescence for eight independently measured and processed images of the same array. Panel B shows the distribution of CV% values obtained for these wells.
3.4 Focus Variation

Due to variations in PDMS thickness and device positioning, some parts of an array may be in better focus than others during imaging. This could potentially lead to systematic spatial or temporal variation in the measured intensity values. A simple experiment was carried out to evaluate the potential impact of focus variation on fluorescence intensity values. Primary hepatocytes were seeded in the arrays and stained with Rh123. Fluorescence images of the array were taken with the cells in good focus and at other focal positions ranging from +60 to -60 μm. The maximum distances tested are much further out of focus than what is seen in a typical experiment. Because taking each image causes some degree of photobleaching, well-focused images were taken between each out-of-focus image in the series to generate an intensity decay curve and cancel out the effect of photobleaching. Focus variations were found to have only a minimal effect on the measured intensity values for each well. Figure 3-4 shows the fluorescence values measured for a sample of cells with the microscope focused at various distances above and below the optimal position. Even at the maximum distances tested, the out-of-focus values were less than 3% smaller on average than the values measured with good focusing.

Figure 3-4: Focus variations have only a minimal effect on the measured intensity values for each well. Primary hepatocytes were stained with Rh123 and imaged at various positions in and out of focus. Panels A and B show the measured intensity values for a random sample of cells with the microscope focused either above or below the ideal position. The values at 60 μm out-of-focus were less than 3% smaller on average.
3.5 Spatial Homogeneity

Spatial variation and nonuniformity are major concerns with many high throughput technologies and have been demonstrated for both the cDNA and oligonucleotide arrays used in gene expression analysis [41]. In microwell array cytometry, fluorescence intensities and cell behaviors can vary systematically across the arrays for several different reasons. One potential cause of spatial variation is a nonuniform lamp intensity across the microscope’s field of view. The direct effect of this on fluorescence intensities can largely be corrected for by calibrating with white and black reference images, but such calibration will not eliminate differences in cell behavior caused by nonuniform light exposure across the array. Another potential source of spatial variation in cell behavior is the presence of different cell densities across or between the arrays due to edge effects and nonuniform cell seeding. Cells at the outer margin of each array or in regions of sparse seeding will have fewer neighbors than those at the center of each array or in regions of dense seeding. This could potentially lead to differences in cell behavior through paracrine effects, nutrient availability, or oxygen tension. Several different qualitative and quantitative techniques were used to evaluate the spatial homogeneity of both fluorescence values and cell behavior. For most of the experimental systems tested, systematic spatial effects were found to be minimal and played little role in determining experimental outcomes.

One qualitative technique used to examine these spatial effects is the generation of overlay images displaying cell classification data for each well on top of the original fluorescence or brightfield images. As shown in Panel A of Figure 3-5, primary hepatocytes stained with JC-1 can be divided into two subpopulations based on fluorescence intensity in the green channel. Overlay images displaying the subpopulation assignment of each cell were generated for each array to show that this was not simply an artifact of lamp intensity, seeding density, variation between arrays, or other spatial effects. As seen in Panel B of Figure 3-5, the cells in each subpopulation appear to be randomly distributed across the array, at least qualitatively. Similar results were seen for the other arrays, and each array was found to contain a similar proportion of cells from each subpopulation. This largely rules out spatial artifacts as the origin of the two subpopulations.

Another qualitative technique is to represent the array as a 2D image where each square in the grid corresponds to one of the wells and the color is determined by some parameter
Figure 3-5: Primary hepatocytes stained with JC-1 can be divided into two subpopulations based on their fluorescence intensity in the green channel. The uniform spatial distribution of these cells both within and between the arrays largely rules out spatial artifacts from lamp intensity variation or image processing.

measured or calculated for that well. This was used in a number of experiments to evaluate the spatial uniformity of parameters such as initial fluorescence values, cell survival times, and cluster assignments. The panels of Figure 3-6 illustrate the spatial distribution of these parameters for one of the arrays in a time course experiment of primary hepatocytes stained with Rh123 and exposed to 25 µM menadione. Once again, little systematic spatial variation can be seen across the arrays.

A number of statistical tests have been developed to examine the spatial distribution of geographical and ecological data [42]. Several of these techniques were used to quantitatively assess whether spatial factors had a significant impact on our microwell array cytometry results. Spatial homogeneity across the arrays was evaluated quantitatively using statistical tests such as Geary’s C [43] and Moran’s I [44]. For continuous parameters, a Mantel test was also performed to check for correlations between the spatial distance and parameter difference matrices for the wells in each array [45]. A summary of the test results are shown in Figure 3-7 for several of the parameters discussed qualitatively earlier in the section. For most of the parameters, no statistically significant spatial correlation using any of the statistical test.
Figure 3-6: Spatial nonuniformity is minimal with the microwell array cytometry system. Initial cell fluorescence values, cell death time, and cluster assignments are plotted as heat map representing the 32x32 grid of microwells in an array from an experiment in Chapter 5. All four parameters are distributed in a relatively uniform manner across the array. Blank spaces represent empty or dropped wells.
The one exception was the initial JC-1 red fluorescence in the Chapter 4 biopreservation experiments. The small p-values in the three tests result from negative spatial correlations rather than positive correlations as would be expected for artifacts like lamp intensity variation. This represents the observation that cells with high JC-1 red intensity tend to be spatially separated from each other, perhaps due to local nutrient or oxygen depletion. In the current experiments, spatial effects were neither desired nor expected, and based on the statistical tests, spatial factors had only a small impact on the results. Though not used here, the spatial control offered by microwell array cytometry is a great opportunity for future studies where spatial effects are known to be important, such as in paracrine signaling or the innate immune response to viral DNA.

<table>
<thead>
<tr>
<th>Biopreservation:</th>
<th>Geary's C</th>
<th>Moran's I</th>
<th>Mantel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rh123 Initial</td>
<td>0.9533</td>
<td>0.9120</td>
<td>0.261</td>
</tr>
<tr>
<td>JC-1 Green Initial</td>
<td>0.1355</td>
<td>0.2006</td>
<td>0.710</td>
</tr>
<tr>
<td>JC-1 Red Initial</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Oxidative Stress:</th>
<th>Geary's C</th>
<th>Moran's I</th>
<th>Mantel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rh123 Initial</td>
<td>0.0960</td>
<td>0.1385</td>
<td>0.512</td>
</tr>
<tr>
<td>DHE Initial</td>
<td>0.0996</td>
<td>0.0404</td>
<td>0.315</td>
</tr>
<tr>
<td>Cell Death Time</td>
<td>0.4372</td>
<td>0.5828</td>
<td>0.186</td>
</tr>
</tbody>
</table>

Figure 3-7: Several different statistical tests were carried out to evaluate whether significant spatial patterns are present in the data from each array. Shown are a sample of the p-values obtained with each test when carried out on a number of different parameters such as initial fluorescence intensity values. All tests were two-sided. The biopreservation and oxidative stress experimental systems are described in Chapters 4 and 5 respectively. Significant p-values were obtained for the JC-1 red fluorescence channel, which had a negative spatial autocorrelation.
Chapter 4

Microwell Array Cytometry as a Novel Tool in Biopreservation

4.1 Introduction

There is a significant need for high-throughput and high-content tools in biopreservation and stress biology, particularly for the investigation and discovery of new ways to improve preservation outcomes [11]. In biopreservation, even with highly optimized protocols, the extreme stress of freezing or desiccation almost always leads to dysfunction or death in a significant fraction of the cell population. It is well known that cell fates are determined in part by the stochastic nature of ice crystal nucleation and other biophysical processes [22]. Less well understood is the extent to which metabolic, bioenergetic, transcriptional, and genetic variation in the initial cell population plays a significant role. By understanding how the initial state of each cell affects its response to preservation, it will be possible to develop new targets for intervention.

Cell-to-cell variation can be quite significant, particularly with primary cells that have been isolated from their native tissue environment. Primary hepatocytes, for instance, vary widely in their metabolic behavior, with either enhanced detoxification or ammonia removal functions depending on their locations in the sinusoids [46]. Mitochondrial activity, metabolic state, energy reserves, position in the cell cycle, and membrane composition could all potentially affect preservation outcomes, but few good tools exist to examine these effects at the single cell level.
Microwell array cytometry is a powerful experimental technique that can be used to correlate pre-stress cell phenotypes and post-stress outcomes with single cell resolution. The system discussed in this paper was designed in our research group as part of a general effort to develop high-throughput and high-content tools for single cell biology. Because the cells are seeded in rugged, high density grids of cell-sized microwells, thousands of individual cells can be tracked and imaged through manipulations as extreme as freezing or drying. Automated image analysis software makes working with the data as straightforward as in flow cytometry.

Microwell array cytometry also has a number of significant advantages over conventional approaches. Unlike flow cytometry, measurements can be made at multiple time points for the same set of cells. Unlike image cytometry with random cell seeding, image analysis is greatly simplified by arranging the cells in a spatially defined pattern and physically separating them from one another. To demonstrate the utility of microwell array cytometry in the field of biopreservation, we have employed it in several different experimental systems. The first set of experiments investigated the role of mitochondrial membrane potential (ΔΨ) in the cryopreservation of primary hepatocytes. The second set of experiments investigated whether the expression level of GFP-tagged protective proteins transfected into the cells correlated with the probability of surviving cryopreservation or desiccation. Experiments like these demonstrate how a sources of biological variation like initial cell state or transfection efficiency can be transformed from unwanted sources of noise that complicate analysis into useful sources of diversity that can be leveraged to study difficult biological problems.

4.1.1 LEA Proteins

LEA (Late Embryogenesis Abundant) proteins are a heterogeneous class of proteins that were first identified in plants, where they have been shown to protect against dehydration stress and dessication damage [47, 48]. Possible mechanisms for this protection include stabilizing sugar glasses and preventing the denaturation and aggregation of cellular proteins [49, 50, 51, 52]. LEA proteins are expressed at particularly high levels in developing seeds, which can reach a very low water content and remain viable. LEA protein homologues have also been identified in a number of dessication tolerant animals, such as several species of nematodes, arthropods, and yeast [53, 54, 55]. Two such proteins, AfrLEA1 and AfrLEA2, have recently been isolated from the anhydrobiotic brine shrimp *Artemia franciscana* and
Figure 4-1: Microwell array cytometry has a number of powerful applications in the field of biopreservation. Most importantly, it can be used to correlate pre-stress cell states with post-stress outcomes at the single cell level. Microwell devices are first prepared and seeded with primary hepatocytes or another cell type. The cells are then stained with one or more fluorescent dyes and imaged to determine the initial cell state. After imaging, the cells are cryopreserved or exposed to a related stress, allowed to recover, and checked for viability with a live-dead assay. The alignment and identification features of the array design allow for easy automation and tracking of each cell through the extreme physical stress of the preservation process.

cloned into vectors for transfection into mammalian cells [54]. In combination with small molecule protectants such as trehalose [56, 57, 58, 59] or 3-O-methylglucose (3OMG) [60, 61], LEA protein expression could potentially improve the viability of mammalian cells after freezing or drying [62, 52]. In order to test this idea at the single cell level, GFP fusion proteins were developed for each of the isolated LEA proteins. Using the microwell array cytometry system, the expression of the fusion proteins could then be correlated to cell survival after freezing or dessication.

4.1.2 Mitochondria in Cryopreservation

Resistance to one or more extreme stresses can be found in many organisms, though such resistance is by no means common or universal. Even closely related organisms can have widely different tolerance levels. The adaptations allowing for tolerance are subtle and often the result of many small changes to metabolic and stress response systems rather than the production of a single protective element [63]. Even in cases where such a protective element obviously exists, such as trehalose in rotifers and brine shrimp or circulating glucose in wood frogs, the use of these compounds to protect otherwise intolerant cells has been only
modestly successful [64]. In other cases, the presence of a protective compound has been shown to be neither necessary nor sufficient for tolerance in the original organism [65]. It has become clear that further progress in biopreservation will require a better understanding of metabolic, bioenergetic, and transcriptional systems, both the protective ones employed by the cells of tolerant organisms and the detrimental ones activated in other cell types. Given the complexity of these systems, bioinformatics and systems biology approaches will be invaluable in this pursuit.

Mitochondria are known to play a key role in metabolism, bioenergetics, calcium homeostasis, and cell death pathways [26]. Naturally tolerant organisms are also known to possess a number of mitochondrial adaptations that are critical for survival in the stressed state [66, 67, 68]. Many such organisms enter a state of metabolic stasis and reduced mitochondrial activity in anticipation of environmental stress [69, 70]. Brine shrimp embryos, for instance, reduce their oxygen consumption by up to 92% upon entering diapause [71]. These observations suggest that having relatively quiescent mitochondria and reduced mitochondrial activity might also be protective for mammalian cells during biopreservation. If significant correlations exist between mitochondrial states and cryopreservation outcomes, novel preservation methods and treatments can be developed to adjust these states prior to cryopreservation.

The experiments of this paper are intended to investigate these ideas in a powerful new way. The goal is to use microwell array cytometry system to precisely determine the extent to which the natural variation in mitochondrial membrane potential across a cell population helps determine preservation outcomes. To date, most biopreservation studies have focused on either bulk assays or small numbers of cells. The experiments described here are some of the first to use high throughput, single cell technologies in biopreservation research. By identifying the most favorable mitochondrial states before and after preservation, it may be possible to improve overall outcomes by pushing cells towards these states with various modulators of mitochondrial activity.
4.2 Methods and Materials

4.2.1 Assembly, Coating, and Seeding

Standard glass slides were cut into 2.5 x 1.0 cm miniature slides intended to fit in 1.5 ml Nalgene cryovials. Individual devices containing 16 arrays each were cut from replica molded polydimethylsiloxane (PDMS) sheets and reversibly mounted well-side-up on a miniature slide. The devices were sterilized by either autoclaving after assembly or rinsing in 100% ethanol before assembly. A 1.25 mg/ml stock solution of Type I collagen was prepared as described elsewhere and diluted 1:50 in distilled water before use [23]. Each device was covered with a droplet of collagen solution and placed under vacuum for 15 min to improve microwell filling. The devices were then incubated for another 30 min at 25°C, rinsed with media, and kept wet until seeding. Hepatocytes were seeded as described in Section 2.1 and kept at 37°C in C+H culture media until ready for staining.

4.2.2 Mitochondrial Staining

The fluorescent dyes JC-1 and rhodamine 123 (Sigma Aldrich) were used to assess mitochondrial membrane potential (ΔΨ) in the cells [28, 24]. Stock JC-1 solution was prepared by dissolving 1 mg lyophilized JC-1 in 1 ml dimethylsulphoxide (DMSO). The stock solution was diluted 1:200 in C+H culture medium before use for a final concentration of 5 µg/ml. Stock rhodamine 123 (Rh123) solution was prepared by dissolving 10 mg lyophilized Rh123 in 1 ml DMSO. In this case, the stock was diluted 1:1000 in C+H before use for a final concentration of 10 µg/ml. Both stock solutions were stored at -20°C. Seeded devices were incubated for 15 min at 37°C in the diluted dye solution, briefly rinsed to remove excess dye, and incubated at 37°C in fresh media for an additional 30 min before imaging.

4.2.3 LEA Protein Transfection

The AfrLEA-1 and AfrLEA-2 genes isolated from the anhydrobiotic brine shrimp *Artemia franciscana* were cloned into mammalian expression vectors as fusion proteins with DSRed or GFP respectively. C3A hepatoma cells, a derivative of the HepG2 hepatoma cell line, were seeded in 35 mm culture dishes at a density of 300,000 cells per dish and cultured in serum-free Opti-MEM media. The next day, the plated cells were transfected with either the AfrLEA-1 or AfrLEA-2 expression vectors using the FuGENE HD transfection reagent.
(Roche) according to the package instructions. Expression vector DNA was used at 2 μg per 35 mm dish. The cells were then incubated for 24 to 48 hours before trypsinization and seeding in the microwell devices. The devices were coated with collagen as described above. Unlike primary hepatocytes or H35 hepatoma cells, fibronectin should not be used with the C3A cells due to a lack of attachment efficiency.

4.2.4 Fluorescence Microscopy

Brightfield and fluorescence images were captured on a Zeiss 200 Axiovert microscope with an AxioCam MRm digital camera, typically using a 2.5x objective and 1.6x optovar for full images of a single array. JC-1 fluorescence was measured using Zeiss # 38 and # 31 filter sets for the green and red channels. The green fluorescence channel represents the loading of JC-1 into the cell, while the red channel represents the shift in fluorescence that occurs with the potential dependent formation of J-aggregates at high JC-1 concentrations in the mitochondrial membranes [29, 30]. Rh123 fluorescence was measured using a Zeiss # 38 filter set. GFP fluorescence was also measured using a Zeiss # 38 filter set. DSRed could not be imaged with the available filter sets. Exposure times were selected to maximize the dynamic range of the resulting images. A brightfield image was taken along with each set of fluorescence images for alignment, quality control, and display purposes.

4.2.5 Cryopreservation and Recovery

After imaging, cells were allowed to recover for 30 min in a 37°C incubator. Freezing solutions were prepared using HypoThermosol-Base (BioLife Solutions) with 10% DMSO added [37]. The devices were quickly transferred into 1.5 ml Nalgene cryovials filled with the freezing solution, placed on ice for a short time for cell equilibration with the cryoprotectant, and placed in a controlled-rate freezer (Planer KRYO 10). The devices were first cooled to -6°C, at which time extracellular ice was seeded by the applying cold forceps to the outside of each vial. This was followed by a 10 minute holding period to allow for ice growth and temperature equilibration. The devices were then cooled at a rate of -5°C/min to -80°C, plunged into liquid nitrogen, and stored for up to a week. After the storage period, the vials were thawed rapidly by gentle agitation for 2 min in a 37°C water bath. The devices were then rinsed in warm C+H, transfered to fresh medium, and allowed to recover at 37°C for 1 hr.
4.2.6 Viability Assays

Stock solutions of 1 mM Sytox Blue (Molecular Probes) were diluted 1:1000 in C+H to a final concentration of 1 μM and used in combination with JC-1 stained cells. Stock solutions of 2 mM ethidium homodimer-1 (Molecular Probes) were diluted 1:1000 in C+H to a final concentration of 2 μM, and used in combination with Rh123 stained cells. After recovering from cryopreservation, devices were transferred into 35 mm dishes containing one of the diluted dyes and incubated at 37°C for 5 min before imaging. Sytox Blue fluorescence was measured using a Zeiss # 5 filter set, while EthD-1 was measured using a Zeiss # 31 filter set. JC-1 or Rh123 fluorescence images were usually taken at the same time.

4.2.7 Image Processing and Statistical Analysis

A set of custom software tools were developed for automated image processing, quality control, and preparing convenient data sets for later analysis. This software is described in detail elsewhere and available upon request. Plotting and statistical analysis were performed using the R programming environment [40].

4.3 Results

Primary hepatocytes were seeded in the microwell arrays, stained with a ΔΨ-sensitive dye, imaged, cryopreserved, and assessed for viability. The result of each processing step is shown for a typical array in Figure 4-2. Data analysis centered on correlating the initial mitochondrial membrane potential, as represented by JC-1 or Rh123 fluorescence, with short term survival after cryopreservation.

4.3.1 JC-1 and Rh123 Fluorescence Distributions

The hepatocytes were stained with either JC-1 or Rh123 and imaged before cryopreservation. Using the microwell array cytometry system, it is easy to create data products similar to those generated by flow cytometry. A scatter plot of the observed red and green JC-1 fluorescence intensities is included in Panel A of Figure 4-3. The green channel was found to have a bimodal distribution, representing two hepatocyte subpopulations. Cell counts and array overlay images showed that cells from the two subpopulations were uniformly distributed both between the arrays and spatially across each array. The significance of these
Figure 4-2: The microwell array design. Panel A is a brightfield image taken after seeding hepatocytes in the 32-by-32 grid of microwells. Alignment features and identification features are included on each side. In Panel B the cells have been stained with the mitochondrial dye JC-1 before cryopreservation and imaged in the red and green fluorescence channels. Panel C shows the results of Live/Dead staining after cryopreservation, with red representing polarized mitochondria in live cells and blue representing nuclear staining in dead cells. Panel D was generated using the viability data and displays the eventual fate of each cell on top of the original JC-1 fluorescence image for rapid visual inspection. Green = survival, Red = death, Yellow = lost from well, Magenta = uncertain, Blue or Cyan = debris around well.
subpopulations is not clear from this experiment alone, but may represent cells from different regions of the liver. Significant differences in gene expression and function have been found between hepatocytes in the centrilobular and periportal regions [5, 6]. Despite the bimodal peak of the JC-1 green channel, and a clear difference between the mean JC-1 red intensity in the two subpopulations, the JC-1 red and Rh123 intensity distributions did not resolve into multiple peaks. Since these channels are primarily sensitive to ΔΨ, high variability in this factor may overwhelm any systematic differences between the subpopulations [24].

4.3.2 Live-Dead Assays

Live-dead assays were performed after cryopreservation using retained Rh123 or JC-1 for the live part of the assay and a membrane impermeant nucleic acid stain for the dead part of the assay. For fluorescence channel compatibility, ethidium homodimer was used with the Rh123 stained cells and Sytox Blue was used with the JC-1 stained cells. The results of these viability assays are plotted in the remaining panels of Figure 4-3. In both cases there is good separation between the live and dead cell populations. The rectangular regions in Panels C and D represent the gates that were used to designate live cells, dead cells, and cells lost during thawing. The trimodal distribution seen with the nucleic acid stains in Panel B reflect the fact that primary hepatocytes can be either mono or binucleated. Overall viabilities after cryopreservation for the JC-1 and Rh123 experiments were 56.4% and 47.2% respectively, with 4.8% and 4.6% of the cells lost from the array. These results are somewhat lower than those of similar bulk experiments with suspended cells [37].

4.3.3 Mitochondrial Membrane Potential and Cryopreservation Outcomes

The effect of initial ΔΨ on cryopreservation outcomes was investigated by combining the pre-preservation Rh123 or JC-1 intensity data with the post-preservation viability data. Using the viability data in each experiment, the hepatocytes were separated into three subpopulations: live cells, dead cells, and cells lost from the device during preservation. Histograms of the Rh123 and JC-1 intensity data were then plotted for each of these populations and compared.

The resulting histograms are shown in the top row of Figure 4-4 for JC-1 red, JC-1 green, and Rh123. Subtle differences can be observed in the intensity distributions for each
Figure 4-3: Data analysis after image processing. Panel A shows the distribution of JC-1 red and green channel fluorescence intensities for approximately 10,000 cells in a 16-array device. Panel B shows the intensity histogram of a membrane impermeant nuclear stain used in the viability assays. The two peaks correspond to mono and binucleated hepatocytes. Panels C and D illustrate the gating used to identify live and dead cells after cryopreservation.
Figure 4-4: Hepatocyte viability after cryopreservation was plotted as a function of JC-1 and Rh123 fluorescence intensity as a proxy for mitochondrial membrane potential. Hepatocytes were separated into three classes: cells that survived, cells that died, and cells that were lost from the array. Top row: cell counts in each class for each of three fluorescence channels. Bottom row: percent viability plotted as a function of fluorescence intensity. The error bars represent the standard error of the percent viability given the number of cells in that bin.
population. Most strikingly, the JC-1 red and Rh123 intensity distributions are shifted to the left for cells that go on to survive preservation. Though the difference appears small, the number of cell samples is large and unpaired t-tests showed that the mean JC-1 red and Rh123 intensities are significantly smaller in cells that went on to survive preservation compared to those that did not survive preservation (p < 0.01 in both cases). This trend was not observed in the JC-1 green channel and the t-test did not show a significant difference in the mean intensity between the groups (p = 0.195).

The effect of the intensity shift can be better appreciated by plotting percent viability as a function of initial fluorescence intensity as shown in the bottom row of Figure 4-4. In the case of JC-1 red and Rh123, the viability is approximately 60% at low fluorescence intensities and drops to approximately 40% at high intensities. Logistic regression of cell survival on initial fluorescence intensity showed highly significant negative correlations for JC-1 red and Rh123 (p < 0.01), but was not significant for JC-1 green (p = 0.189). The significant negative correlation with viability suggests that having a high mitochondrial membrane potential before cryopreservation is detrimental to the cells and leads to worse outcomes. However, the relatively small magnitude of the effect suggests that other factors also play important roles.

4.3.4 Multidimensional Analysis

Since both red and green channel intensities were acquired for each cell in the JC-1 experiment, it is possible to examine cell viability across the 2D landscape of initial JC-1 fluorescence values rather than look at each channel separately as done above. Figure 4-5 shows the variation in cell survival rates seen across the landscape of JC-1 intensities, with green representing an upper limit of 70% viability and red representing a lower limit of 20% viability. Live cell and total cell density estimates were computed using a Gaussian kernel and the ratio of these density estimates was plotted. Several features are apparent in the joint distribution that are not obvious in the single channel plots, such as the poor viability for cells with low JC-1 red fluorescence but high JC-1 green fluorescence. Multidimensional analyses of this type are made possible in part by the high cell counts achievable with microwell array cytometry, without which there would be insufficient data for reliable viability estimates in regions with low cell density.
Figure 4-5: Survival rates vary across the landscape of initial JC-1 fluorescence values, with green representing a high of 70% viability and red representing a low of 20% viability. The high cell counts and measurement capabilities of microwell array cytometry allow preservation outcomes to be mapped across a multidimensional landscape of cell phenotypes with a single experiment.
4.3.5 LEA Protein Dose Response

Fluorescence images of the transiently transfected C3A cells were collected prior to dessica-
tion or cryopreservation. Since the cells were transfected with a LEA-GFP fusion protein, the
fluorescence level of each cell roughly corresponds to the level of LEA protein expres-
sion. As in the mitochondrial experiments, the viability of each cell was determined after
preservation and related to the initial expression level. Figure 4-6 shows cell counts and
percent viability plotted as a function of the initial GFP fluorescence. A positive trend
was seen with cryopreservation, but did not reach the level of statistical significance. In
the dessication experiment, cell survival patterns were dominated by spatially nonuniform
drying, and no consistent correlation was found with expression of the LEA-GFP fusion
protein and cell viability.

4.4 Discussion

The goals of this project were to investigate the role of mitochondrial state in determining
cryopreservation outcomes, determine whether LEA protein expression levels correlate with
preservation survival, and introduce microwell array cytometry as a high-throughput, high-
content platform for single cell studies in the field of biopreservation. Given the promising
results of these initial experiments, we expect the system will be used in the future to
investigate a wide range of cellular features that may impact preservation outcomes.

The experiments described in this paper focused on the role of mitochondrial membrane
potential in the cryopreservation of primary hepatocytes. Based on current literature, it was
unclear how the mitochondrial state of a cell would affect the outcome of preservation. On
the one hand, cells with a high mitochondrial membrane potential ($\Delta\Psi$) have been shown
to be more resistant to a number of death-inducing stimuli [72]. On the other hand, studies
of freeze and desiccation tolerant organisms have shown an almost universal tendency to
dramatically reduce metabolic activity and oxygen consumption prior to the onset of the stress
[69, 70]. Tolerant organisms are also known to have numerous mitochondrial adaptations
that inhibit calcium leakage and apoptosis [66, 67, 68].

Our results with JC-1 and Rh123 staining suggest that a moderately suppressed mito-
chondrial membrane potential may be beneficial in cryopreservation, but is clearly not the
sole determinant of outcome. In interpreting this result, it should be noted that the mem-
Figure 4-6: Viability after cryopreservation or dessication was plotted as a function of GFP fluorescence intensity as a proxy for LEA protein expression after transient transfection. The cells were separated into two classes: those that survived and those that died. The number of lost cells could not be determined since some cells had essentially zero GFP fluorescence, and no additional dyes were used before preservation to positively identify the loaded wells. Top row: cell counts in each class for each experiment. Bottom row: percent viability plotted as a function of fluorescence intensity. The error bars represent the standard error of the percent viability given the number of cells in that bin.
brane potential is not a direct measure of mitochondrial activity. In healthy mitochondria, a low mitochondrial membrane potential can be the result of increased oxidative phosphorylation, whereas a high membrane potential can occur under resting (state 4) conditions, where most respiration is devoted to the compensation of proton leak rather than ATP synthesis [73]. In contrast, a total collapse of ΔΨ, which was not observed here, is often seen upon the opening of the mitochondrial permeability transition pore during apoptotic signaling [74]. Activation of the transition pore involves compromise of inner membrane integrity and thus of membrane potential. This phenomenon is distinct from mitochondrial outer membrane permeabilization via Bax/Bak poration, which does not necessarily alter the membrane potential [75]. Given these complexities, caution should be exercised when interpreting ΔΨ data alone in the absence of other mechanistic information. Further exploration of mitochondrial state using the microwell image cytometry system will help address these questions.

As mentioned in the results, the overall viability with the devices was somewhat lower but comparable to what we typically achieve with suspended hepatocytes. Under the right circumstances, the use of microwell arrays could actually be beneficial, since the large physical forces involved in freezing or drying can easily displace cells attached to a flat surface or crush suspended cells between growing ice crystals [22]. In many of the freezing experiments, it was observed that any cells attached to the array, but not protected by a well were either killed or ripped away from the surface during preservation. The freezing protocol used for the arrays was directly adapted from our protocol for suspended primary hepatocytes and not independently optimized. Viability on the devices was nevertheless quite good and could potentially exceed our best results for suspended cells with additional work. One significant problem that needs to be addressed is bubble formation on the device during thawing, which sometimes killed significant numbers of cells and interfered with image analysis. We are currently working on several strategies to eliminate this problem, such as adding a cell-compatible surfactant like Poloxamer-188 and adjusting the thawing protocol [76].

The LEA protein transfection experiments did not show a significant correlation between the level of expression and survival after cryopreservation or dessication. Several factors may have contributed to this negative result. In the dessication experiments, cell viability was largely determined by the local water content, with zero viability in the driest regions.
and almost full viability in more hydrated regions. These spatial factors overwhelmed any potential variation due to LEA protein expression level. This issue will be addressed in future experiments by improving the uniformity of the drying process and optimizing the dessication protocol. Another issue is that LEA proteins are known to protect against dessication damage in tolerant organisms such as plants and nematodes, but play a less well defined role in freeze tolerance [77]. Though freeze and dessication tolerance are often closely related [78, 63], the protective action of LEA proteins may not apply in a slow freezing system like the one used here. The protective action of the LEA proteins may have been hindered by fusion with the GFP tag, or may require additional adaptations or factors that are absent in mammalian cells. Though the LEA protein experiments did not achieve positive results, they served as a proof of concept demonstration of a powerful class of biopreservation experiments made possible by the microwell array system.

A number of other improvements could potentially be made to the microwell system and future experimental designs. One weakness of the current experiments was the use of short term viability as the only cryopreservation endpoint. In future experiments, this will be addressed by looking at cell viability again after one or two days and using apoptosis assays to determine the type of cell death that has occurred. Additional fluorescent dyes and GFP reporter constructs will also be used to study factors such as free radical production, calcium release, caspase activation, and stress response pathways. Unfortunately, long term endpoints are not currently feasible for several reasons. Primary hepatocytes are very sensitive to culture conditions and long term seeding in microwells without cell-to-cell contact is far from ideal. In seeding experiments without freezing, viability remains near 100% for several days, but later drops to almost zero. With cell lines, long term endpoints are complicated by cell division and movement, which disrupt the physical separation and spatial organization of the cells in the device. We are currently working on several potential ways to address these problems and extend the length of time cells can be kept in the arrays, such as improving the surface coatings of the device or embedding the microwells in a thin Matrigel layer after cell seeding, which has been shown to maintain hepatocyte function in the absence of cell-cell contact [79].

With conventional methods such as flow cytometry or manual image analysis, it is time consuming and difficult to reliably track large numbers of individual cells through the preservation and recovery process. As a result, relatively little is known about how
heterogeneity in the bioenergetic, metabolic, transcriptional, or signaling phenotypes of a cell population affects preservation outcomes. With microwell array cytometry, these difficulties are largely eliminated, transforming cell population heterogeneity from unwanted noise into a source of useful experimental variation. This variation can then be harnessed to study the phenotypic determinants that affect preservation outcomes and used to identify protective and maladaptive cell responses to target for intervention. Based on the results of this study, we believe that microwell array cytometry will be an extremely powerful and valuable tool for the study of biopreservation damage mechanisms and the development of new approaches in biopreservation.
Chapter 5

High Throughput Single Cell Bioinformatics

5.1 Introduction

The growing understanding that no cell population is truly uniform and that stochasticity plays a large role in biological systems has motivated the development of new experimental approaches for high-throughput, single cell studies in biology [1, 2, 3, 4]. Even in a relatively uniform cell population, individual cells can exhibit significant temporal differences in behavior. As discussed in Section 1.3, the cell population averaging implicit in bulk time course measurements can easily obscure interesting and important behavior occurring at the single cell level. In the worst cases, the results from bulk time courses studies can be quite misleading, with bulk results far from the actual behavior of any individual cell. In many cases, single cell time courses would provide a much more rich and detailed answer to complex biological questions.

Unfortunately, single cell measurements are technically challenging, and few good experimental techniques exist to generate single cell time courses in an automated and high-throughput manner. The microwell array cytometry system presented in this thesis combines a high density microwell array with integrated alignment and identification features. This greatly simplifies the image processing needed for effective image cytometry and allows thousands of cells to be reliably tracked across multiple time points, even when the devices undergo significant manipulation between measurements.
In the previous chapter, the initial pre-stress state of each cell was correlated with its post-stress outcome. This involved making measurements at two time points separated by an intervening treatment. In this chapter, the system is extended to measure full single-cell time courses with multiple conditions in a single experiment and up to 50 time points over 12 hours. Several bioinformatics techniques, such as k-means and hierarchical clustering, are also presented as tools for the characterization and visualization of the complex data sets resulting from these experiments. As will be discussed later, the information generated in high-throughput cytometry studies differs from that of expression microarrays in several significant ways and will require the application of new bioinformatics techniques in order to sort through it and extract meaningful results.

The development of analysis techniques for single cell biology is an active and high impact topic in the field of bioinformatics, with many applications in pharmaceutical development and drug screening [80, 81, 82]. However, most of the efforts in high-throughput single cell biology have focused on the development of techniques for automated morphological analysis and cell classification at a single or limited number of time points. Relatively little attention has been paid to the development of experimental platforms and analytical techniques for single cell time course studies. Hopefully, the microwell array cytometry system and bioinformatics approaches presented in this thesis will help move the field forward by providing one of the first major explorations of this idea.

Using the microwell system, dynamic, high-throughput, single-cell fluorescence data can be generated with high reproducibility and minimal user intervention during the image analysis phase. Integrating the system with powerful bioinformatics techniques during the data analysis phase allows for the rich characterization of cell behavior at the single cell level.

The robust microwell array cytometry platform has allowed us to investigate a number of biological problems that would be difficult to tackle with more traditional techniques. Our initial efforts focused on generating dynamic single-cell fluorescence images of primary hepatocytes in a model system of oxidative stress using a combination of mitochondrial and free radical probes. These single cell fluorescence time courses were then analyzed using bioinformatics techniques such as hierarchical and k-means clustering to visualize the data and extract interesting features.
5.2 Methods and Materials

5.2.1 Assembly, Coating, and Seeding

Microwell devices containing 8 arrays each were prepared as described in Section 2.1, seeded with primary rat hepatocytes, and kept at 37°C in C+H culture media until ready for staining. In a single experiment with the 8-array devices, it is possible to collect brightfield and fluorescence images every 15 minutes for over 5000 individual cells in each of six different treatment conditions.

5.2.2 Dyes and Treatments

Hepatocyte mitochondrial membrane potential and mitochondrial superoxide generation was monitored after treatment with the free radical generator, menadione. The fluorescent dye rhodamine 123 (Rh123) was used to qualitatively assess mitochondrial membrane potential (ΔΨ) in the cells [24, 28]. A 1000x stock Rh123 solution was prepared by dissolving 10 mg lyophilized Rh123 in 1 ml DMSO and used at a final concentration of 10 µg/ml. Mitochondrial superoxide generation was assessed using dihydroethidium (DHE), which is initially non-fluorescent but is oxidized in the presence of superoxide to form the fluo-
cent dye ethidium [31]. A 1000x stock DHE solution was prepared by dissolving 3.15 mg lyophilized DHE in 1 ml DMSO and used at a final concentration of 10 µM. A 25 mM stock menadione solution was prepared by dissolving 4.3 mg menadione in 1 ml ethanol and used at a final concentration of 0 to 100 pM. All stock solutions were protected from light and stored at -20°C in small aliquots.

Seeded microwell devices were first incubated for 15 min at 37°C in C+H culture media with Rh123 at 10 µg/ml, rinsed to remove excess dye, and incubated at 37°C in fresh media for an additional 30 min. Immediately before imaging, DHE and menadione stock solutions were added to warm C+H culture medium in the appropriate amounts, and the seeded devices were transferred into the prepared solutions. Array imaging was started within 15 min of initial exposure to the dye and treatment.

5.2.3 Fluorescence Microscopy

Brightfield and fluorescence images were captured on a Zeiss 200 Axiovert microscope with an AxioCam MRm digital camera, typically using a 2.5x objective and 1.6x optovar for full images of a single array. Rh123 fluorescence was measured in the green channel using a Zeiss # 38 filter set, and DHE fluorescence was measured in the red channel using a Zeiss # 31 filter set. Exposure times were selected to maximize the dynamic range of the resulting images. A brightfield image was taken along with each set of fluorescence images for alignment, quality control, and display purposes.

5.2.4 Image Processing and Data Extraction

Brightfield and fluorescence images were processed as described in Section 2.4 for each array and timepoint. The results from each device were then compiled into unified data files for convenient plotting and analysis.

5.2.5 Clustering of Fluorescence Time Courses

Several unsupervised clustering techniques such as k-means and agglomerative hierarchical clustering were applied to Rh123 and DHE fluorescence time courses to help visualize the data and identify distinct patterns of cell behavior. Clustering was performed in R using routines from the Bioconductor software package [83]. K-means clustering was performed multiple times for both raw and normalized time course data. The number of clusters was
optimized by examining the within cluster sum of squares and the Gap statistic [84]. Cluster assignments were also validated by running the algorithm multiple times with random starting points. Hierarchical clustering was performed using the Euclidean distance and either Ward’s method or complete linkage.

5.3 Results

Primary hepatocytes were stained and imaged with the mitochondrial dye rhodamine 123 (Rh123) in green channel and dihydroethidium (DHE) in the red channel. The accumulation of Rh123 is dependent on the mitochondrial membrane potential, and it is lost from the cell when the mitochondria depolarize. DHE is converted to a fluorescent compound by superoxide radicals, and the rate of fluorescence increase corresponds to the level of free radical generation in the cell. Example images from several time points are shown in Figure 5-2. The highlighted cells are shown in more detail in Figure 5-3. Software tools were developed to quickly and easily generate similar images and movies for arbitrary subsets of the cell population.

5.3.1 Single Fluorescence Time Courses

Combining the microwell devices with an automated microscopy system and custom image analysis software allows single cell fluorescence time courses to be obtained with minimal user intervention. Time courses of Rh123 and DHE fluorescence were generated for cells treated with varying amounts of the free radical generator menadione. Shown in Figure 5-3 is a sample of these time courses for cells treated with 25 μM menadione and colored based on the estimated time of cell death.

It is important to note that even within this single treatment condition, the cell population displayed a wide range of response patterns that would be difficult to characterize without an unbiased and high throughput experimental system. With manual image analysis, cells at the extremes of this range would likely be missed or excluded as outliers, when in fact the behavior of these rare cells may be more important than the “average” cells when trying to understand the biological processes involved.
Figure 5-2: Shown is a false color fluorescence image of primary hepatocytes stained with rhodamine 123 (green channel) and dihydroethidium (red channel) and exposed to 25 μM menadione for four hours.
Figure 5-3: Using the microwell array cytometry system, single cell fluorescence time courses can be generated with minimal user intervention. Primary hepatocytes were stained with rhodamine 123 (green channel) and dihydroethidium (red channel) and exposed to 25 µM menadione for four hours. Panel A shows the single cell time courses of Rh123 and DHE fluorescence generated from the time series, with images collected every 16 min for approximately 13 hours. The color of each time course corresponds to the survival time determined for that cell. Panel B is an abbreviated time series highlighting individual cells in the array and the diversity of fluorescence patterns observed. K-means clustering was used to identify clusters of cells with similar time courses, and the highlighted cells in each column were randomly selected from these clusters.
5.3.2 Fluorescence Time Course Modeling

The resulting time courses tend to follow characteristic patterns reflecting the mechanism and staining pattern of each dye, but the timing and magnitude of the individual time course features vary significantly from cell to cell. Since the cells were loaded with Rh123 and rinsed before any imaging occurs, the Rh123 time courses begin with a roughly exponential decrease in intensity due to photobleaching and dye leakage. There is then a spike in fluorescence as the mitochondria begin to lose polarization, releasing the dye into the cytoplasm, and reducing the amount of self quenching. Shortly thereafter, there is a complete loss of fluorescence representing plasma membrane depolarization and cell death. Mathematical modeling of the Rh123 time courses is complicated by the significant diversity observed in the shape of the fluorescence spikes and the short or nonexistent exponential regions in many of the time courses.

The DHE time courses also follow a characteristic and expected pattern with a steadily increasing baseline intensity due in part to photoactivation and nonspecific conversion of the dye. Superimposed on this baseline is a period of rapid intensity change corresponding to mitochondrial superoxide generation as nonfluorescent DHE is converted to fluorescent ethidium in the presence of superoxide radicals. The timing of this peak roughly corresponds to the spike in Rh123 fluorescence and is followed by a drop in fluorescence back to the baseline as the mitochondria lose their integrity and release trapped dye. Many of the DHE time courses can be modeled quite accurately with only seven parameters. The baseline fluorescence can be modeled as \( A + Bt + C(1 - e^{-t/D}) \), and the peak can be modeled as \( E e^{t/F} u(G - t) \) where \( u(t) \) is the unit step function. Parameter values were estimated heuristically based on time course features such as the initial slope, the final slope, and the time of the sharpest drop. The estimated parameters were then used as seed values for a nonlinear least squares optimization. The results of this fitting are shown in Figure 5-4 for a random sample of time courses from the cells treated with 25 \( \mu \)M menadione.

5.3.3 K-Means Clustering of Single Cell Time Courses

Having individual time courses for hundreds to thousands of cells opens up the possibility of using bioinformatics techniques such as k-means and hierarchical clustering to better characterize the data and explore the behavior of cell populations with distinct behaviors.
Figure 5-4: Dihydroethidium time courses were fit using a seven parameter model. Shown in the figure is a random sample of actual and modeled time courses for cells treated with 25 μM menadione. The measured values are shown in red, the modeled values are shown in blue, and the background components of the modeled values are shown in green.
With bulk measurement or low cell count experiments, the typical diversity of cell responses is treated as a source of unwanted biological noise and one is limited to generating a single average response.

K-means clustering was performed on the Rh123 and DHE time courses, both combined as a single vector and for each channel separately. For the combined clustering, the time courses in each channel were rescaled by a fixed factor such that the average Euclidean distance between two time courses would be equal to one. This was done so that the two channels would have equal weight when combined in a single vector for clustering. Shown in Figure 5-5 are the results for 9 clusters under two different treatment conditions. Clustering was performed separately for each treatment condition, but within a treatment condition, the cluster assignments are shared for the two channels and numbered according to the median time of cell death within the cluster.

In order to optimize the number of clusters chosen for the K-means algorithm, clustering was performed on Rh123 and DHE time courses for a range of cluster counts and evaluated using several figures of merit including the between-cluster variance [85] and gap statistic [84]. The between-cluster variance plot shows the fraction of the total variance explained by the separation into clusters. This fraction always increases with the number of clusters, but typically demonstrates one or more transition points where the addition of more clusters qualitatively has a smaller effect. Figure 5-6 shows the within-cluster variance plot for the combined Rh123 and DHE time courses under both the control and 25 μM menadione conditions. In each case, transitions can be seen at approximately 5 and 10 clusters, indicating that either would be a reasonable choice for the number of clusters.

5.3.4 Hierarchical Clustering for Single Cell Dose Response Experiments

Dose response curves are a common and well established way to present data from pharmacological tests with a binary endpoint such as cell death. Such curves can be generated from high throughput single cell experiments, but this does not take full advantage of the collected data. Reducing a single cell fluorescence time courses to a simple endpoint such as cell death does not account for the initial state of the cell and its behavior leading up to that endpoint. Here, we took a different approach for analyzing the retrieved data to yield a more detailed picture of hepatocyte behavior in response to menadione treatment.

Hierarchical clustering was performed on the normalized and combined Rh123 and DHE
Figure 5-5: K-means clustering was performed on Rh123 and DHE time course data under two experimental conditions. Since clustering was performed on combined data, the cluster assignments in the two channels correspond to each other. Control cells with no menadione treatment are shown in Panels A and B. Cells treated with 25 μM menadione are shown in Panels C and D. Several distinct response patterns can be identified from the clustered data. The response patterns differ in shape and timing between the two conditions, but also within each condition, reflecting the wide diversity of behaviors possible in a cell population. For each fluorescence channel, the general shape of each time course is constrained by the mechanism of the dye used, but there is considerable variation in the magnitude and timing of the responses.
Figure 5-6: The number of clusters selected for the k-means algorithm was evaluated using the between-cluster variance. Panels A and B show the results for the control and 25 μM menadione conditions respectively using combined Rh123 and DHE time courses. The steeper slope and higher values seen with the control cells indicate that the variability present in the time courses can be captured by a smaller number of clusters than in the treatment condition.

time courses using the Euclidean distance and complete linkage. The dendrograms were then reordered such that branches with the highest mean intensity appear at the top. The results are shown in Figure 5-7. The colored bars between the heat maps and the dendrograms indicate the cluster assignment for each cell using the k-means algorithm. As the menadione dosage increases, the rate of free radical generation also increases, and cell death tends to occur earlier. Even within a single treatment condition, however, there is a great diversity of cell survival times. For the 25 μM treatment condition, survival times range anywhere from 150 to 600 min. Another striking feature is the temporal correlation between the fluorescence channels, with the contours of the paired heat maps following each other very closely.

Heat maps are particularly well suited to visualizing the data from high throughput single cell dose response experiments, particularly in combination with hierarchical clustering. Placing time courses with similar profiles close together made it easier to identify major trends in the data, both within each treatment condition and between them and highlighted the temporal relationships between multiple fluorescence channels. As an alter-
Figure 5-7: Hierarchical clustering with the Euclidean distance and complete linkage was performed on combined Rh123 and DHE time course data in a dose response experiment, with dendrogram nodes reordered according to average fluorescence intensities. Each heat map row corresponds to the behavior of a single cell from one of the arrays, and the color bar on the left side indicates cluster assignments using the k-means algorithm. Presenting dose response data as a set of clustered heat maps is an information-rich alternative to classic survival curves. Though it is straightforward to represent the collected data as a set of survival curves, the heat map approach better highlights the diversity of cell responses to each stimulus and allows one to appreciate the temporal relationships between multiple fluorescence channels.
native to clustering, the time courses were sorted according to the time at which a particular end point is reached, but this tended to obscure potentially interesting trends in the data.

### 5.4 Discussion

The microwell array cytometry system presented in this paper is a novel, high-throughput, high-content platform for the study of cell behavior, cell population heterogeneity and biological stochasticity. Most importantly, the system is able to generate dynamic single cell fluorescence data for multiple vital probes and correlate pre-treatment phenotypes with post-treatment outcomes at the single cell level. The system also provides sufficiently large cell populations for meaningful analysis, easy loading and staining protocols, automated data handling, and reproducibility of the measured fluorescence values. In addition to presenting the cytometry system, a number of powerful bioinformatics techniques were adapted and applied to the study of single cell fluorescence time courses, demonstrating the advantages of this approach over classic bulk assays and analysis techniques.

The need for a microwell array cytometry platform arose in direct response to the limitations of bulk assays and currently available cytometry technologies. Bulk assays can only deal with population averages, which represents a loss of potentially useful information and can mislead the user in a number of significant ways [10]. Single cell methods such as flow and image cytometry avoid this problem by providing the entire distribution of values seen in the population, but each has its own limitations. The major limitation of flow cytometry is its inability to track individual cells across multiple time points. One can track the overall behavior of the population across time, but not the individual behavior of each cell. With array cytometry, each cell is associated with a single microwell in a rugged PDMS device and remains attached and in place through significant manipulation of the device. The inclusion of alignment and identification features around each array makes it even easier to track cells across time.

Image cytometry is limited by the complexity of identifying and segmenting fluorescence images of cells, particularly with random seeding and high cell densities. Though analysis techniques are improving, it is not trivial to extract meaningful single cell data from such images [16]. With array cytometry, image processing is greatly simplified by seeding each cell in its own microwell, which physically separates the cells and places them in known
positions relative to easily identified alignment features. This technological approach makes it easier to automate the data analysis and use the system in a high-throughput fashion. Because of its low cost and relative simplicity, microwell array cytometry has the potential to become as widely used and versatile as flow cytometry.

Bioinformatics techniques such as k-means and hierarchical clustering are a powerful set of tools that can be used to organize, analyze, and visualize the voluminous data generated by high throughput techniques such as image cytometry and microwell array cytometry. As demonstrated in this chapter, the clustering of single cell fluorescence time courses helps identify groups of cells with distinct behaviors, makes it easier to compare the results of different treatment conditions, and highlights the diversity of responses to treatment within a presumably uniform cell population. The use of heat maps with hierarchical clustering provides a data-rich complement to survival curves for dose response experiments where additional time course data is available.

Though they are superficially similar, fluorescence time course data from a microwell array cytometry experiment has a number of features that distinguish it from the gene expression data generated by cDNA and oligonucleotide microarrays. Most importantly, the physical properties and behavior of each fluorescent dye impose a great deal of structure on the resulting time courses. For a particular dye, much of the observed variation between cells relates to timing and magnitude rather than the overall shape of the curve. This poses a different set of challenges than are addressed by the time course analysis often done for gene expression data, which emphasizes the grouping of genes with similar time course shapes, but not magnitudes [86]. Another important difference is that the number of features examined is relatively small compared to the number of replicates, the opposite of most gene expression studies [87]. These differences will require the development and refinement of bioinformatics techniques suited to this type of data and represent a great opportunity for future bioinformatics research.

There are many planned enhancements to both the physical and software aspects of the system. On the software side, this includes computing additional statistics for each well such as cell size, granularity, and fluorescence colocalization. On the hardware side, the array will be combined with an existing combinatorial microfluidic device, greatly expanding the number of markers and conditions that can be tested at one time [36]. Efforts are also underway to automate the collection of higher resolution images of interesting wells based
on position and phenotype data from low resolution scout images.

5.4.1 Conclusion

As our understanding of biology grows, single-cell and subcellular analysis techniques will become increasingly important. Already it has become clear that the forced population averaging of bulk experimental techniques provides a limited view of the rich variation and stochasticity present in biological systems. Studying large populations does have its advantages, however, and high throughput experimental techniques are beginning to recover those advantages without sacrificing the richness of single cell analysis. Further advances will require the development of improved algorithms for automated data collection, physical device designs to support this automation, and new approaches in bioinformatics to analyze the results. It is expected that the system described here is a step in this direction.
Chapter 6

Summary and Conclusions

The microwell image cytometry system presented in this paper is a key part of an ongoing effort in our research group to develop high-throughput, high-content platforms for the study of stress biology and population heterogeneity in biopreservation, bioinformatics, and other biological disciplines. The primary goal of this particular project was to develop an inexpensive, high-throughput experimental system capable of generating single cell fluorescence time courses and single cell correlations of pre-stress cell phenotypes with post-stress outcomes. Additional design requirements included automated data handling, sufficiently large cell populations for meaningful statistical analysis, compatibility with existing flow cytometry software, easy loading and staining protocols, and the ability for cells to survive preservation in the device.

The goals of the project developed in direct response to the limitations of previously available techniques. Bulk assays only deal with population averages, which as described earlier represent a loss of potentially useful information and can mislead the user in a number of significant ways. Single cell methods such as flow and image cytometry avoid this problem by providing the entire distribution of values seen in the population. The main limitation of flow cytometry is the inability to track individual cells across multiple time points. In our system, however, each cell is associated with a particular well in a rugged PDMS device and remains attached and in place even when the entire device is put through a freeze-thaw cycle, making it easy to track each cell across time. Automated microscopy is limited by the complexity of interpreting fluorescence images of cells with random seeding, particularly at high cell densities. Though analysis techniques are improving, it is far from trivial to extract
meaningful single cell data from such images. In our system, image processing is greatly simplified by positioning each cell in its own microwell, making it much easier to automate the data analysis and use the system in a high-throughput fashion. Another major goal was compatibility with existing software tools and analysis techniques. Efforts were made to export data in formats suitable for use in flow cytometry software and statistical software.

6.1 Summary of Results

6.1.1 System Design and Validation

A large number of experiments were run in order to physically test the array devices and validate the results obtained with the automated analysis tools. The validation experiments included calculations of the well loading efficiency for a sample of typical devices, reproducibility testing of the fluorescence intensities obtained with repeated imaging in different positions, examining how variations in focus affect the reported intensity values, and determining whether cell behavior is spatially homogeneous within each array. Software tools were also developed to generate overlay images of the arrays for rapid visual inspection of the data and troubleshooting. The results of these experiments show that the system can achieve high filling rates with minimal cell attachment outside of the wells, that the fluorescence values obtained with the system are highly reproducible and robust against variations in device orientation and minor focusing errors, and that the position of a cell within an array has little to no effect on the fluorescence values obtained or the cell’s behavior.

6.1.2 Biopreservation Studies

Using standard methods such as flow cytometry or automated microscopy, there are few good ways to reliably track large numbers of individual cells across or through the preservation process. As a result, little is still known about how heterogeneity in the bioenergetic, metabolic, transcriptional, or signaling phenotypes of a cell population may affect preservation outcomes. With the development of the microwell array cytometry system, population heterogeneity can be transformed from a source of unwanted noise to a source of useful experimental variation and harnessed to study the cellular factors influencing preservation outcomes. The protective or maladaptive stress responses identified with the system can then be more easily targeted for intervention.
Initial work with the microwell array system focused on the role of mitochondrial function in the cryopreservation of primary hepatocytes. Based on the current literature, it was unclear whether high mitochondrial mass and function before preservation would be protective or detrimental to the outcome. On the one hand, cells with a high mitochondrial membrane potential (MMP) have been shown to be more resistant to a number of death-inducing stimuli. On the other hand, studies of freeze and desiccation tolerant organisms have shown an almost universal tendency to drastically reduce metabolic activity and oxygen consumption prior to the onset of the stress. Our results with JC-1 and Rh123 staining suggest that a moderately suppressed mitochondrial membrane potential may be beneficial in cryopreservation, but is clearly not the sole determinant of outcome. The results are complicated by the fact that the membrane potential is not a direct measure of mitochondrial activity. Further exploration of mitochondrial state using the microwell image cytometry system will help address this complication.

In a separate study, cells were transiently transfected with GFP-labelled versions of the LEA-1 or LEA-2 protein, both of which are believed to play an important role in the survival of dessication and freezing tolerant organisms. The cells were then examined after freezing or dessication to see whether the expression level of the protein was correlated with the probability of surviving preservation. Though a slight positive trend was observed for freezing survival, the LEA protein transfection experiments did not show a significant correlation between the level of expression and survival after cryopreservation or dessication. Despite the lack of firm positive results, the LEA protein experiments served as a proof of concept demonstration of a powerful class of biopreservation experiments made possible by the microwell array system.

The use of microwell arrays may be beneficial from a preservation standpoint, since the large physical forces involved in freezing or drying can easily displace cells attached to a flat surface or crush cells between growing ice crystals when in suspension. The structural stability of the PDMS wells may actually have a protective effect. At least some proportion of the cells were killed by bubble formation on the device during thawing, a process we are currently working to eliminate. In many of the freezing experiments, it was also observed that any cells attached to the array and not protected by a well were either killed or ripped away from the surface during preservation. The freezing protocol used for the arrays was directly adapted from our standard protocols for suspended primary hepatocytes. Viability
on the devices was quite good, even without optimization, and could potentially exceed our best results for suspended cells with additional work.

6.1.3 High Throughput Single Cell Bioinformatics

For this section of the thesis, the capabilities of the microwell system were extended to allow single-cell time courses to be measured for multiple treatment conditions, large number of cells, and many time points, all in a single experiment. Using the system, dynamic, high-throughput, single-cell fluorescence time course data could be generated with high reproducibility and minimal user intervention. Several bioinformatics techniques, such as k-means and hierarchical clustering, were adapted for the microwell data and presented as tools for the characterization and visualization of the complex data sets generated in the experiments. The use of powerful bioinformatics techniques during the data analysis phase allowed for the rich characterization of cell behavior at the single cell level.

The microwell array cytometry platform allowed us to investigate of a number of biological problems that would be difficult to tackle with more traditional techniques. Our initial efforts focused on generating dynamic single-cell fluorescence images of primary hepatocytes in a model system of oxidative stress using a combination of mitochondrial and free radical probes. These single cell fluorescence time courses were then analyzed using bioinformatics techniques such as hierarchical and k-means clustering to visualize the data and extract interesting features.

Bioinformatics techniques such as k-means and hierarchical clustering are a powerful set of tools that can be used to organize, analyze, and visualize the voluminous data generated by high throughput techniques such as image cytometry and microwell array cytometry. As demonstrated in this thesis, the clustering of single cell fluorescence time courses helps identify groups of cells with distinct behaviors, makes it easier to compare the results of different treatment conditions, and highlights the diversity of responses to treatment within a presumably uniform cell population. The use of heat maps with hierarchical clustering provides a data-rich complement to survival curves for dose response experiments where additional time course data is available.
6.2 Future Work

There are many potential enhancements to both the physical and software aspects of the system. In terms of hardware improvements, the array will be combined with an existing combinatorial microfluidic device, greatly expanding the number of markers and conditions that can be tested at one time. Additional modifications to the device design and well configuration will allow the microwell system to be used for cell-cell interaction studies and fine control of the cellular microenvironment.

The current suite of software tools will be updated and extended with the intention of broader use both within and beyond the research group. Significant efforts will be made to make the analysis software more user friendly and accessible to a wide range of users. As will be discussed in Appendix A, efforts are also underway to add improved support for more complex image analysis routines such as cell size and granularity estimates, subcellular localization, and colocalization of fluorescence. Efforts are also underway to automate the collection of high resolution images of interesting wells based on position and phenotype data from a low resolution scout image.

6.2.1 Microfluidic Integration

Integrating the current array design into a microfluidic device would simplify the seeding procedure, allow multiple experiments to be run in parallel, and give finer control over the microenvironment of the cells. For a preliminary trial, the microwell arrays will be integrated into the existing microfluidic device used for microfluidic Living Cell Array (mLCA) experiments reported previously by our research group [36]. This will allow for easy cell seeding and parallel testing of up to eight inputs and eight outputs in a combinatorial fashion. The GFP reporter cell lines used in the mLCA experiments were described in Section 2.2 and could potentially be used with the current array system for high-throughput single cell gene expression studies.

6.2.2 Device Design and Seeding

A wide range of experiments can be performed without significant changes to the current device design in terms of the microwell arrays or alignment features. With primary hepatocytes, high filling rates and minimal attachment outside of the wells can already be achieved
using the current well and array geometry. Minor changes may be needed for the optimal seeding of H35 reporter cells [36]. Since they tend to be smaller than primary hepatocytes, the well diameter will be reduced from 25 μm to 15 μm when using H35 cells. Image registration using the current alignment features has a very high success rate. The alignment features may be improved in future device designs, but no major changes are needed to achieve highly effective automated image analysis.

A simple but powerful modification to the device design would be to change the shape of each well in the array. Figure 6-1 shows several possible well designs for both cell shape control experiments and cell-cell interaction studies. The coating procedure could also be adjusted to include various signaling molecules or growth factors to provide fine control over the cellular microenvironment of each well.

Figure 6-1: The shape of the microwells can be changed for fine control over cell shape as shown in Panel A or for cell-cell interaction studies as shown in Panel B. Well size is easily optimized for additional cell types. The same basic coating and seeding procedures can be used for most of these designs. The top half of Panel B shows a possible well design for studying the interaction between hepatocytes and fibroblasts or endothelial cells. The large hepatocytes would be seeded first, and would attach primarily in the large center wells. The smaller fibroblasts would then be seeded and fill the wells surrounding each center well.
Appendix A

Morphological Analysis of Intracellular Fluorescence Distributions

A significant advantage of image based methods for single cell analysis is the possibility of examining cell morphology, fluorescence localization, and colocalization patterns in addition to basic parameters like the total fluorescence intensity. One of the major goals of image cytometry and high content cytometry platforms is the characterization and classification of cells based on the quantitative analysis of their morphological features [88, 89]. The rich data sets provided by these techniques have found significant use in drug discovery, mechanistic studies of pharmaceuticals, and identifying protein localization patterns [81, 82, 90]. Like expression microarrays and large scale genomic studies, the large volumes of data produced by these techniques require advanced, automated analysis techniques for any hope of extracting meaningful results. Several techniques from other areas of bioinformatics have been applied to this problem, but it remains an exciting and promising area for future research.

A.1 Morphological Analysis Approaches

Most of the approaches for automated morphological analysis in image cytometry are based on more general work in the fields of machine vision and pattern recognition. Most such
studies have relied on a three step process for cell classification. These steps consist of segmentation, feature generation, and classification. The first step is the identification and segmentation of cells based on the collected images. This step often requires significant user interaction, particularly at high cell densities. In extreme cases, the cells may be individually selected and segmented by hand. As discussed previously, a major advantage of microwell array cytometry is the way it simplifies this step in the analysis.

The second step is generating a set of quantitative features describing the morphological characteristics of the cell. At the most basic level, these features often includes characteristics such as the cell size, cell perimeter, nuclear-to-cytoplasmic ratio, position of the nucleus, and simple image moments. Some of these features are less appropriate to measure in microwell array because of the constraints placed on overall cell shape by seeding in the microwells, but many of them can and have been implemented in the image analysis software used in previous chapters of this thesis. A number of basic image moments and localization parameters can be calculated for each well in the array with minimal added computational time compared to total intensity measurements alone.

A number of more advanced features have also been developed for quantitative morphological analysis. Two of the most commonly used in the field of image cytometry are the Zernike moments and the Haralick textures [88, 91, 90]. The Zernike moment are particularly appropriate for microwell array cytometry and an efficient Zernike moment calculation routine has been incorporated into the analysis software [92]. The Zernike moments will be defined and discussed in more detail later in this chapter. The Haralick texture features are a set of statistics derived from the co-occurrence matrices of an image [93]. The values represent such features as the homogeneity of the image and the amount of local intensity variation. A Haralick texture calculation routine has not yet been incorporated into the analysis software, but is planned for future versions.

The third step in analysis is cell classification based on the generated feature sets. Several different techniques from the fields of bioinformatics and machine learning have been used for supervised or unsupervised classification of cells in image cytometry. These include neural networks, classification trees, linear discriminant analysis, support vector machines, and various clustering techniques [88, 80, 81, 82]. Cell classification based on the Zernike moments was attempted using k-means and hierarchical clustering for one of the timecourse experiments in Chapter 5 but did not proceed beyond the exploratory stage.
A.2 Zernike Moments

The Zernike polynomials are a set of orthogonal polynomials defined on the unit disk \[94\]. They were first developed by Frits Zernike to characterize the various distortions and aberrations present in optical systems \[95\]. Since then, Zernike moments have found extensive use in the fields of optics, optometry, and ophthalmology. More recently, they have been found use in image analysis and computer vision for shape recognition \[94, 96, 97, 88\]. The Zernike polynomials have a number of properties that make them attractive for use in image analysis, such as orthogonality and rotational invariance.

A.2.1 Definitions and Properties

The positive and negative Zernike polynomials \(Z_n^m(\rho, \phi)\) and \(Z_n^{-m}(\rho, \phi)\) are defined for integers \(n\) and \(m\) with \(n \geq m \geq 0\) as

\[
Z_n^m(\rho, \phi) = R_n^m(\rho) \cos(m\phi) \quad (A.1)
\]

\[
Z_n^{-m}(\rho, \phi) = R_n^m(\rho) \sin(m\phi) \quad (A.2)
\]

with the radial function \(R_n^m(\rho)\) defined as

\[
R_n^m(\rho) = \begin{cases} 
\sum_{k=0}^{(n-m)/2} \frac{(-1)^k (n-k)!}{k! \left(\frac{1}{2}(n+m)-k\right)! \left(\frac{1}{2}(n-m)-k\right)!} \rho^{n-2k} & \text{for } n-m \text{ even} \\
0 & \text{for } n-m \text{ odd}
\end{cases} \quad (A.3)
\]

Based on this definition, the first few radial Zernike polynomials are

\[
R_0^0(\rho) = 1 \quad (A.4) \quad R_3^3(\rho) = \rho^3 \quad (A.9)
\]

\[
R_1^1(\rho) = \rho \quad (A.5) \quad R_4^4(\rho) = 6\rho^4 - 6\rho^2 + 1 \quad (A.10)
\]

\[
R_2^0(\rho) = 2\rho^2 - 1 \quad (A.6) \quad R_2^2(\rho) = 4\rho^4 - 3\rho^2 \quad (A.11)
\]

\[
R_3^2(\rho) = \rho^2 \quad (A.7) \quad R_3^4(\rho) = \rho^4 \quad (A.12)
\]

\[
R_1^3(\rho) = 3\rho^3 - 2\rho \quad (A.8)
\]

The positive and negative Zernike polynomials can also be expressed together using complex numbers. In this form, the magnitude of the complex Zernike moments are rotationally invariant. This form is particularly useful for image cytometry purposes since with
random seeding or the use of round microwells, the behavior of a cell is not expected to vary depending on its absolute orientation.

\[ V_n^m(\rho, \theta) = R_n^m(\rho) \exp(jn\theta) \] (A.13)

Using the complex form of the Zernike polynomials, one can then define the Zernike moments \( A_n^m \), where \( P(x, y) \) is the intensity value of the pixel at \((x, y)\). Also important to note is that the complex conjugate of \( V_n^m(\rho, \theta) \) is used in the sum below.

\[ A_n^m = \frac{n + 1}{\pi} \sum_x \sum_y P_{x,y} |V_n^m(x, y)|^2 \] (A.14)

The change from Cartesian to polar coordinates for \( V_n^m(\rho, \theta) \) can be done with a simple change of coordinate system.

\[ \rho = \sqrt{x^2 + y^2} \] (A.15)
\[ \theta = \tan^{-1}\left(\frac{y}{x}\right) \] (A.16)

The normalized Zernike polynomials up to order six are shown in Figure A-1 with red representing positive values and green representing negative values. Many of the low order Zernike moments have a simple physical interpretation, particularly when applied to an optical system. This interpretation is less meaningful when the moments are used to generate descriptors for shape recognition and classification purposes, but it still has relevance in certain situations, such as when accounting for microscope focus variation. The zeroth order moment is equivalent to the total fluorescence intensity.

The Zernike moments are particularly well suited to characterizing cell shape and fluorescence distributions in microwell array cytometry, since the image of each cylindrical microwell has a simple and natural mapping to the unit disk. The rotational invariance of the complex Zernike moment magnitudes is also a useful property, since cell behavior is not expected to vary with the absolute rotational orientation. As a proof of concept experiment, Zernike moments up to the sixth order were calculated for each cell and time point in one of the oxidative stress data sets from Chapter 5. The resulting profiles were used to distinguish between cytoplasmic and nuclear localization of the DHE fluorescence, and attempts were made to relate the initial Zernike profile of each cell to its survival time.
Figure A-1: Zernike polynomials up to order six
and other outcome parameters. Unfortunately, this work did not proceed beyond the initial stages due to time constraints, but should prove to be a fruitful area for future research.
Bibliography


